

On Phases Of Melonic Quantum Mechanics

Frank FERRARI¹ and Fidel I. SCHAPOSNIK MASSOLO^{1,2}

¹*Service de Physique Théorique et Mathématique
Université Libre de Bruxelles (ULB) and International Solvay Institutes
Campus de la Plaine, CP 231, B-1050 Bruxelles, Belgique*

²*Institut des Hautes Études Scientifiques
35 route de Chartres, 91440 Bures-sur-Yvette, France[†]*

frank.ferrari@ulb.ac.be, fidel.s@gmail.com

We explore in detail the properties of two melonic quantum mechanical theories which can be formulated either as fermionic matrix quantum mechanics in the new large D limit, or as disordered models. Both models have a mass parameter m and the transition from the perturbative large m region to the strongly coupled “black-hole” small m region is associated with several interesting phenomena. One model, with $U(n)^2$ symmetry and equivalent to complex SYK, has a line of first-order phase transitions terminating, for a strictly positive temperature, at a critical point having non-trivial, non-mean-field critical exponents for standard thermodynamical quantities. Quasi-normal frequencies, as well as Lyapunov exponents associated with out-of-time-ordered four-point functions, are also singular at the critical point, leading to interesting new critical exponents. The other model, with reduced $U(n)$ symmetry, has a quantum critical point at strictly zero temperature and positive critical mass m_* . For $0 < m < m_*$, it flows to a new gapless IR fixed point, for which the standard scale invariance is spontaneously broken by the appearance of distinct scaling dimensions Δ_+ and Δ_- for the Euclidean two-point function when $t \rightarrow +\infty$ and $t \rightarrow -\infty$ respectively. We provide several detailed and pedagogical derivations, including rigorous proofs or simplified arguments for some results that were already known in the literature.

March 18, 2019

[†]Present address.

Contents

1	Introduction and discussion	4
2	Presentation of the models	9
2.1	Large d matrix models	10
2.1.1	General setup and the large N limit	10
2.1.2	Inequivalent models	12
2.1.3	Generalized melons	13
2.2	Disordered models	16
2.3	Discussion and relation to other tensor models	18
2.3.1	Degree expansion, index expansion, melonic and non-melonic interactions	18
2.3.2	$U(N)$ -symmetric models and the tracelessness condition	20
2.3.3	Matrix models vs. tensor models, brane picture and gauging .	22
3	Euclidean time picture	23
3.1	Euclidean time two-point function	23
3.1.1	General definitions and properties	23
3.1.2	Schwinger-Dyson equations	26
3.1.3	On the solutions of Schwinger-Dyson equations and monodromies	27
3.1.4	Reparameterization invariance and standard low energy ansatz	29
3.1.5	A new low energy ansatz	33
3.2	Effective action and thermodynamics	34
3.2.1	Disordered model formulation	34
3.2.2	Matrix model formulation	36
3.2.3	Useful formulas	40
3.3	Numerical solution of Schwinger-Dyson equations	43
3.3.1	Basic numerical results for the $U(n)^2$ model	46
3.3.2	Basic numerical results for the $U(n)$ model	51
3.4	Euclidean time four-point function	54
3.4.1	General definitions and properties	54
3.4.2	Diagrammatic analysis	56
3.4.3	Algebraic proof	59

4	Phase diagrams	61
4.1	General discussion	61
4.2	Phase diagram of the $U(n)^2$ model	66
4.2.1	The (m, Q) plane and physics at fixed Q	69
4.2.2	The physics at zero temperature	71
4.2.3	The q -generalizations and critical exponents	74
4.2.4	Multi- q models, domain walls and a line of critical points . . .	79
4.3	Phase diagram of the $U(n)$ model	82
5	Real time physics	87
5.1	Real time two-point functions	87
5.1.1	General definitions and properties	88
5.1.2	Melonic real time two-point functions	91
5.1.3	Numerical solution of real time equations	96
5.1.4	Application: quasi-normal frequencies	99
5.2	Real time four-point functions	103
5.2.1	General definitions and properties	104
5.2.2	Melonic real time four-point functions	106
5.2.3	On Lyapunov exponents	108
5.2.4	Computing Lyapunov exponents	111
5.2.5	Numerical implementation	113
5.2.6	Applications	117
A	q-body interacting Hamiltonians	121
B	Algebraic derivation of the kernel formula	122
C	On the numerical errors	124

1 Introduction and discussion

Melonic theories are models for which the perturbative expansion is dominated, in some well-defined limit, by a special class of Feynman graphs called *melonic*. These graphs are relevant in a surprisingly wide range of problems. They were discovered for tensor models in the context of the discretized approach to quantum gravity [1] but also appeared implicitly in the study of disordered condensed matter systems [2]. They dominate a suitably defined large d limit of planar graphs [3–5] and thus provide an interesting non-perturbative approximation to the large n limit of matrix models. Moreover, melonic graphs can be recursively constructed and explicitly summed over using the technique of Schwinger-Dyson equations, allowing for considerable analytical progress to be made in the study of such models.

Strong interest in melonic theories, mostly fermionic quantum models, has been rekindled in recent times by the Sachdev-Ye-Kitaev (SYK) model [2, 6] and various generalizations thereof [7, 8], see also [9] and the review [10]. Within a program that has seen the convergence of very diverse approaches, new and interesting relations have been found between condensed-matter disordered systems [11–14], tensor models [15, 16] and random matrix theory [17, 18]. The emerging general picture is that melonic theories provide a quantum mechanical description of the near-horizon, low energy limit of near-extremal black holes [6, 9]. The AdS_2 factor in the near-horizon geometry is at the origin of universal properties that are also found in the infrared limit of the melonic models [19–21]. A crucial ingredient is the emergence of an approximate time reparameterization symmetry. This symmetry is spontaneously broken down to $\text{SL}(2, \mathbb{R})$, and the associated pseudo-Goldstone modes are governed by a Schwarzian action which matches the gravitational dynamics in nearly AdS_2 space-time [22]. The low-temperature thermodynamical properties of the gravitational system are reproduced, together with some correlation functions and Lyapunov chaos exponents [23].

Melonic quantum models are also extremely interesting in their own right [1, 24–27]. The class of melonic graphs is situated in between the very simple trees of bubbles which dominate the large n limit of vector models, and the intractable set of planar graphs which are relevant for the matrix models. This “midpoint” has striking properties: the models remain simple enough to be amenable to analytical study; yet they display non-perturbative features typically associated with matrix models, including the emergence of a continuous spectrum and chaos. It is even plausible that the melonic large d approximation to matrix models [3] is able to capture the correct qualitative physics associated with the full sum over planar diagrams.

The aim of this paper is to present an extensive study of two fermionic quantum mechanical models in the melonic limit, motivated by the line of research initiated in [28]. The models can be formulated either as matrix quantum mechanics, or as

disordered models. The random Hamiltonians in the disordered formulation read

$$H = Nm \chi_i^\dagger \chi^i + \sqrt{N} \lambda_{kl}^{ij} \chi_i^\dagger \chi_j^\dagger \chi^k \chi^l \quad (1.1)$$

and

$$H' = Nm \chi_i^\dagger \chi^i + \sqrt{N} (\xi_{jkl}^i \chi_i^\dagger \chi_j^\dagger \chi^k \chi^l + \xi_l^{ijk} \chi_i^\dagger \chi_j^\dagger \chi_k^\dagger \chi^l), \quad (1.2)$$

where the χ^i are complex fermions for $1 \leq i \leq N$, λ_{kl}^{ij} and ξ_{jkl}^i are random couplings and m is a mass parameter. The model based on H is often called the “complex SYK model” and has already been considered in the literature, see *e.g.* [19, 20, 29–31]. It preserves the fermion number symmetry and its matrix model version, discussed in Sec. 2, has a full $U(n) \times U(n) \times O(d)$ symmetry. For this reason, we call this model the “ $U(n)^2$ model.” The model based on H' has to the best of our knowledge never been studied before. It breaks the fermion number symmetry and its matrix model version has a reduced $U(n) \times O(d)$ symmetry; we therefore call it the “ $U(n)$ model.” The Hamiltonians (1.1) and (1.2) look very similar, but this is misleading. We shall see that they yield qualitatively very different physics.

In the natural units for which the quartic coupling strengths are set to one, the models depend on two parameters, the mass m and the temperature T . Both the large T , fixed m and large m , fixed T regions of the (m, T) -plane are weakly coupled. The matrix model formulation makes it clear that the large mass regime describes a situation akin to a system of weakly interacting and well-separated D-particles governed by open string perturbation theory. At small masses, on the other hand, the models are strongly coupled and share properties with black holes. The introduction of the mass parameter thus allows us to study in detail the interesting transition between a weakly coupled “D-brane” regime and a strongly coupled “black-hole” regime [32]. We are going to discuss at length both the physics of Euclidean time (two-point and four-point functions, thermodynamics and phase diagrams) and real time (two-point and out-of-time-ordered four-point functions, quasi-normal frequencies, chaos exponents).

Salient features and main new results for the $U(n)^2$ model

A fundamental property of the $U(n)^2$ model, first observed in [28], is the presence of a line of first order phase transitions in the strongly coupled region of the (m, T) -plane, see Fig. 21 in Sec. 4.2. The first order phase transition exists because the Schwinger-Dyson equations resumming the melonic diagrams of the theory have two distinct solutions within a finite region of the (m, T) -plane. One solution is SYK-like and describes black-hole physics, whereas the other is an analytic continuation of the perturbative solution at large mass. The first order phase transition occurs when the free energies of the two solutions match. A crucial feature is that the line of first order phase transitions terminates at a critical point $(m = m_c, T = T_c)$. Standard thermodynamical critical exponents were evaluated numerically in [28] and found to

be non-mean-field and asymmetric; for instance, at $m = m_c$, different exponents are obtained depending on whether $T \rightarrow T_c$ from above or from below.

In the present paper, we provide full derivations of the results presented in [28] and extend the analysis in two directions.

- i) First, we study the phase diagram of the q -generalizations of the model (1.1), for which the quartic interaction is replaced by a fermion number preserving term of even degree q . The physics is similar for all values of $q \geq 4$, with a line of first order phase transitions terminating at a critical point $(m_c^{(q)}, T_c^{(q)})$. We find that the values of some of the critical exponents do depend on q , see Table 2 in Sec. 4.2.3. The variations with q seem to be well above numerical error bars, especially for low values of q , providing evidence that the critical points found for different values of q are inequivalent. Moreover, we can also define “multi- q ” models for which interactions of various degrees are included simultaneously. We have initiated the study of such models, whose phase diagrams depend on the relative strengths of the couplings, see Fig. 31 in Sec. 4.2.4 for the case where the quartic and sextic couplings are both turned on. We find in this case evidence for a line of continuously varying non-trivial critical points, interpolating between the pure $q = 4$ and $q = 6$ cases.
- ii) Second, we study real time correlation functions, from which we extract the leading quasi-normal frequency ν_{qn} governing the exponential decay of two-point functions, and the Lyapunov chaos exponent η_L governing the exponential growth of the connected piece of out-of-time-ordered four-point functions. We find that the chaos bound of [33] is saturated in the zero-temperature limit for all values of the mass in the black-hole, SYK-like phase. This is true for both the particular four-point function studied in [33], but also for the physical OTOC, for which the proof of the bound presented in [33] does not apply. An interesting new result comes about when one studies the behaviour of ν_{qn} and η_L near the critical point (m_c, T_c) , see Fig. 44 and 50 in Sec. 5.1.4 and 5.2.6, respectively. We find that both ν_{qn} and η_L have a singular behaviour at the critical point. This singular behaviour leads to a new set of critical exponents, tied to the real time physics of dissipation and chaos and characterizing the critical point.

Main results for the $U(n)$ model

In contrast with the situation in the $U(n)^2$ model, for any strictly positive temperature the $U(n)$ model undergoes a smooth crossover from the weakly coupled large mass regime to the strongly coupled small mass regime. In particular, we find only one physically acceptable solution of the Schwinger-Dyson equations for each set of values of the parameters m and T . Interesting physics thus occurs at zero temperature. The mass gap m_{eff} is found to decrease from its perturbative value $m_{\text{eff}} \simeq m$ at large mass

to zero for a strictly positive critical mass $m = m_* > 0$, see Fig. 33 in Sec. 4.3. Below the quantum critical point at m_* the model is in an interesting new gapless phase.

The most novel and important feature of the model has to do with the infrared behaviour of this gapless phase found for $0 < m < m_*$. Instead of the standard conformally invariant solution, we find that *the conformal symmetry $SL(2, \mathbb{R})$, including the usual scale transformations, is spontaneously broken*. This occurs because the power-law decay of the zero-temperature Euclidean two-point function is governed by two *distinct* exponents Δ_+ and Δ_- ,

$$G(\tau) \sim \begin{cases} \frac{b_+}{|\tau|^{2\Delta_+}} & \text{if } \tau \rightarrow +\infty \\ -\frac{b_-}{|\tau|^{2\Delta_-}} & \text{if } \tau \rightarrow -\infty \end{cases}, \quad (1.3)$$

with

$$\Delta_- > \Delta_+. \quad (1.4)$$

This non-standard behaviour, with the strict inequality (1.4), is established numerically in Sec. 3.3.2 with a very high degree of confidence.

The low-temperature thermodynamical properties of the exotic gapless phase are similar to the standard behaviour also found in SYK or in the $U(n)^2$ model. The entropy can be expanded as $S/N = \sigma + cT + o(T)$, with the zero-temperature entropy σ being strictly positive for all $0 \leq m < m_*$ and vanishing at $m = m_*$. The coefficient c varies smoothly in the interval $0 \leq m < m_*$ and jumps discontinuously to zero at $m = m_*$, see Fig. 35 in Sec. 4.3. On the other hand, we find convincing numerical evidence supporting the fact that the Lyapunov exponent does not saturate the bound $\eta_L \leq 2\pi T$ when $T \rightarrow 0$ (compare Fig. 47 for the $U(n)^2$ model with Fig. 48 for the $U(n)$ model in Sec. 5.2.6). Note that there is no inconsistency here. The standard methods [9] to compute η_L in the zero-temperature limit rely on the $SL(2, \mathbb{R})$ symmetry; it is a non-trivial open problem to generalize these methods to a phase satisfying (1.3), (1.4) in which the $SL(2, \mathbb{R})$ symmetry is spontaneously broken.

Plan of the paper

In Sec. 2, we introduce and motivate the $U(n)^2$ and $U(n)$ models in detail. We explain that, to leading melonic order, they can be formulated equivalently either as large d planar matrix quantum mechanics with fundamental degrees of freedom given by d complex fermionic matrices $\psi_{\mu b}^a$, $1 \leq a, b \leq n$, $1 \leq \mu \leq d$ (Sec. 2.1), or as large N disordered models with fundamental degrees of freedom χ^i , $1 \leq i \leq N$ (Sec. 2.2). We emphasize the matrix formulation because it provides a genuine quantum mechanical theory even when the number of degrees of freedom is finite. We explain that these models are the most general single-trace matrix quantum mechanics based on the variables $\psi_{\mu b}^a$, with quartic interaction terms and the additional condition that the

fermion number violating two-point functions $\langle \text{tr} \psi_\mu \psi_\mu \rangle$ vanish at leading melonic order. We also comment in Sec. 2.3 on the relation with tensor models and the different large N limits that can be defined in these models; on subtleties associated with the case where the symmetry $U(n)^2 \times O(d)$ is broken down to $U(n) \times O(d)$, as in the $U(n)$ model; and on gauging.

In Sec. 3 we introduce the Euclidean time technology which is heavily used in the rest of the paper. In Sec. 3.1 we recall, in a detailed and pedagogical way, the fundamental properties of Euclidean two-point functions, the Schwinger-Dyson equations they satisfy, some important properties of the space of solutions of these equations (non-trivial monodromies), their reparameterization invariance and the standard low energy ansatz. We also introduce and briefly discuss the fundamental new low energy ansatz (1.3) in Sec. 3.1.5. In Sec. 3.2 we derive explicit formulas for all the thermodynamical potentials (free energy, energy and entropy). These derivations are well known in the case of the disordered models, for which an effective action can be derived through a straightforward path integral argument based on the introduction of a suitable bilocal auxiliary field. We focus on the case of the matrix (or tensor) models, for which such a simple argument does not exist. We clarify the meaning of the effective actions for these models and provide rigorous and elementary derivations of all the required identities. In Sec. 3.3 we introduce the numerical techniques used to solve the Euclidean Schwinger-Dyson equations and provide extensive tests of their precision by accurately reproducing well-known properties of the IR limit of the $U(n)^2$ model in the SYK-like phase. We then apply the same methods to the $U(n)$ model and discover that its low energy limit, for low enough mass m , is governed by the generalized ansatz (1.3), establishing unambiguously the inequality (1.4). Finally, in Sec. 3.4 we discuss the Euclidean four-point functions. As usual, their connected pieces are obtained by resumming a geometric series associated with a set of ladder diagrams. We provide a diagrammatic derivation of the associated kernels, together with a fully rigorous and general algebraic proof (see also App. B).

Sec. 4 is devoted to the study of the phase diagrams of the models, *i.e.* their properties in the (m, T) -plane. A general physical discussion is given in Sec. 4.1, where the main features are explained and motivated in simple terms. In Sec. 4.2, we focus on the $U(n)^2$ model. After recalling and extending the main results already obtained in [28] (line of first order phase transition, critical point, physics at fixed fermion number, low temperature behaviour), we study the q -generalizations of the model (see also App. A). This includes the derivation of the critical exponents for the q -generalizations of the critical points in Sec. 4.2.3, and the discussion of multi- q models with their associated domain wall solutions and line of critical points in Sec. 4.2.4. The $U(n)$ model is discussed in 4.3. The main feature of its phase diagram is the vanishing of the mass gap at a quantum critical point $m = m_* > 0$. For $0 < m < m_*$, the model is in the exotic gapless phase with a spontaneously broken scale symmetry. We find strong evidence that the zero-temperature entropy is non-vanishing in this phase and also compute the energy and heat capacity at low temperature.

Real time physics is discussed in Sec. 5. We begin in Sec. 5.1.1 with a discussion of the general properties of real time two-point functions, including Carlson’s theorem. We then explain in Sec. 5.1.2 how the Schwinger-Keldysh formalism can be used to constrain the two-point functions in the melonic limit, and how the resulting real time Kadanov-Baym equations can be solved by an elegant argument based on Carlson’s theorem, apparently yet unpublished. This provides a rigorous path from the Euclidean to the real time solution of the models. The numerical algorithm used to solve the real time equations, which is an extension of the algorithm used for the Euclidean equations, is presented in Sec. 5.1.3 and used to compute some spectral functions. We then apply our results on two-point functions to the computation of quasi-normal frequencies in Sec. 5.1.4, finding an interesting singular behaviour at the critical point of the $U(n)^2$ model, with associated non-trivial critical exponents. The problem of the real time four-point functions is tackled in Sec. 5.2. General properties (symmetries, spectral functions, KMS condition) are reviewed in Sec. 5.2.1 and the Schwinger-Keldysh formalism in the melonic case is set up in Sec. 5.2.2. We then discuss the expected behaviour of the real time four-point functions, which strongly depend on the time ordering of the operators, and introduce the chaos Lyapunov exponent η_L associated with the out-of-time-ordered correlators in Sec. 5.2.3. We carefully explain in Sec. 5.2.4 how η_L can be extracted by simplifying the exact Schwinger-Keldysh equations written down in Sec. 5.2.2. This yields the so-called real time ladder kernels. In Sec. 5.2.5, we present two methods to extract η_L numerically from the ladder kernels. The first method is the most direct and rigorous and does not rely on the assumption of any particular ansatz for the correlators. It amounts to inverting a large matrix of discretized real time ladder kernels. The second method is due to Kitaev [6] and is based on a natural ansatz for the long time behaviour of the correlators. It reduces the problem to the computation of the largest eigenvalue of a one-dimensional kernel operator. Both methods are found to yield perfectly consistent results and are applied in Sec. 5.2.6 to the explicit computation of η_L in both models. The most interesting new results we obtain concern the non-saturation of the chaos bound in the gapless phase of the $U(n)$ model, and the singular behaviour of η_L near the critical point in the $U(n)^2$ model.

2 Presentation of the models

The models we shall study can be formulated in two different ways, either as matrix or disordered models. In the matrix models there are $N = n^2 d$ degrees of freedom, organized as a vector of complex fermionic matrices $\psi_{\mu b}^a$ with $U(n)$ indices a, b and an $O(d)$ index μ . On the other hand, the disordered model variables form a vector of N complex fermions χ^i with $i = 1, 2, \dots, N$. For the purposes of the present work, we restrict ourselves to leading order in a suitably defined $N \rightarrow \infty$ limit [3–5]. It is a remarkable fact that in this limit both formulations turn out to be *strictly equivalent*,

although this equivalence breaks down if one were to consider $1/N$ corrections.

The advantage of the disordered models is that they are *a priori* technically much simpler. Being vector models, they are straightforwardly solved in the case of annealed disorder using the standard auxiliary field techniques. The main and crucial novelty, compared to the old-fashioned vector models [34], is that the auxiliary field is now bilocal [35]. By examining the leading graphs it is also easy to show that the quenched and annealed disorder averages coincide at leading large N order, see *e.g.* [36] for a detailed discussion.

The advantage of the matrix models is that they make the link with possible applications in holography much more explicit [3], since they are standard quantum mechanical matrix models with the usual string theory interpretation. Indeed, in retrospect one may see the startling equivalence with the disordered models at leading order in N as an explanation for why disordered models *à la* SYK could be related to black-hole physics.

The leading order equivalence between the two points of view deserves to be carefully justified. This essentially amounts to devising proofs that never use the auxiliary field technique, which is not applicable in the matrix case. This will be explained and exemplified in detail along the rest of the paper.

The models we consider are, in some sense to be clarified below, the most general ones based on complex fermions with quartic interaction terms and some simple additional conditions. In this Section we provide a general presentation and exposition of their basic properties.

2.1 Large d matrix models

2.1.1 General setup and the large N limit

For convenience, we work with fermionic matrices $\psi_{\mu b}^a$ normalized in such a way that the canonical anticommutator reads

$$\{\psi_{\mu b}^a, \psi_{\nu d}^{\dagger c}\} = \frac{1}{nd} \delta_{\mu\nu} \delta_d^a \delta_b^c. \quad (2.1)$$

The fermion number operator and fermion number are respectively defined by

$$\hat{Q} = \frac{1}{n} \text{tr} \psi_{\mu}^{\dagger} \psi_{\mu}, \quad Q = \langle \hat{Q} \rangle, \quad (2.2)$$

where we use matrix notation for the $U(n)$ indices. With the normalization (2.1), we have that $0 \leq Q \leq 1$.

We wish to consider the most general single-trace Hamiltonian invariant under $U(n) \times O(d)$ with a quadratic mass term and quartic interactions. It is important to distinguish two classes of interaction terms, according to the cyclic order of the $O(d)$

indices in the traces. We will say that interaction terms with alternating indices as in $\mu\nu\mu\nu$ are of type I or *tetrahedric*, while those with repeated consecutive indices such as $\mu\mu\nu\nu$ are of type II or *pillow*.¹ For example, $\text{tr} \psi_\mu^\dagger \psi_\nu \psi_\mu \psi_\nu$ is tetrahedric whereas $\text{tr} \psi_\mu^\dagger \psi_\mu \psi_\nu^\dagger \psi_\nu$ and $\text{tr} \psi_\nu^\dagger \psi_\mu \psi_\mu^\dagger \psi_\nu$ are pillows. Packing up terms of each class in h_{I} and h_{II} respectively, we can write the Hamiltonian as

$$H = nd \left(m \text{tr} \psi_\mu^\dagger \psi_\mu + \sqrt{d} h_{\text{I}} + h_{\text{II}} \right). \quad (2.3)$$

Note that the most general quadratic term depends on a unique mass parameter m , because expressions like $\text{tr} \psi_\mu \psi_\mu$ identically vanish.

We then consider the large n and large d limits defined in the following way [3–5]:

- i) First, we take $n \rightarrow \infty$ keeping d and the couplings in h_{I} and h_{II} fixed. This is the usual 't Hooft limit [37], so for example the free energy has a large n expansion of the form

$$F = \sum_{g \geq 0} n^{2-2g} F_g. \quad (2.4)$$

Here F_g is the sum over genus g vacuum Feynman graphs. We will focus on the leading order F_0 which selects the planar Feynman graphs.

- ii) Next, we take $d \rightarrow \infty$ again keeping the couplings in h_{I} and h_{II} fixed. Note that the $\sqrt{d}$ enhancement of the tetrahedric (type I) terms in (2.3) is absolutely crucial here. Without it, the $d \rightarrow \infty$ limit would exist but would be trivially vector-like [3, 4]. This second limit generates a $1/\sqrt{d}$ expansion of the sums over planar graphs, so for example for the leading-order free energy

$$F_0 = \sum_{\ell \geq 0} d^{1-\ell/2} F_{0,\ell}. \quad (2.5)$$

Once more, in what follows we focus on the leading contribution, *i.e.* $F_{0,0}$ for the free energy and similarly for other quantities.

In the limiting procedure above, the order in which one takes the limits $n \rightarrow \infty$ and $d \rightarrow \infty$ is crucial. At fixed n , due to the $\sqrt{d}$ enhancement of the tetrahedric couplings in (2.3), it is easy to find graphs having arbitrarily high powers of d . The large d limit at fixed n is thus ill-defined. On the other hand, taking the large n limit first restricts us to consider only planar graphs, and these have a power of d which is bounded above. Thus the large d limit of the sum over planar graphs is well defined even with the enhanced couplings.

For conciseness, recalling that $N = n^2 d$ we shall refer to the above successive large n and large d limits as “the large N limit” and call the leading contributions

¹A possibly confusing terminology associated to similar terms in the tensor model literature is discussed in Sec. 2.3.

in this limit “the leading large N contributions.” In this way, the terminology is also appropriate for the discussion of the disordered models presented in Sec. 2.2, so it is possible to switch back and forth between both equivalent formulations.

2.1.2 Inequivalent models

It is straightforward to check that all the single-trace interaction terms corresponding to a given cyclic ordering of the fields in the trace are equivalent at leading order in the large N limit described above, up to a possible redefinition of the mass parameter m and a c-number shift of the Hamiltonian. For example,

$$nd^{3/2} \text{tr} \psi_\mu \psi_\nu^\dagger \psi_\mu \psi_\nu^\dagger = -nd^{3/2} \text{tr} \psi_\mu^\dagger \psi_\nu \psi_\mu^\dagger \psi_\nu + n^2 \sqrt{d}, \text{ etc.} \quad (2.6)$$

Note that in some instances a double-trace term $\text{tr} \psi_\mu \text{tr} \psi_\mu^\dagger$ may be produced, but this is also sub-leading at large n . Alternatively, such terms can be altogether eliminated by imposing a tracelessness condition $\text{tr} \psi_\mu = 0$, which does not play any role at leading order.²

This remark greatly reduces the number of inequivalent terms we may consider at leading order. For the tetrahedric interactions, noting that $\text{tr} \psi_\mu \psi_\nu \psi_\mu \psi_\nu = 0$, we are left with three possibilities

$$h_{\text{I},1} = \frac{\lambda}{2} \text{tr} \psi_\mu^\dagger \psi_\nu \psi_\mu^\dagger \psi_\nu, \quad h_{\text{I},2} = \frac{\lambda'}{2} \text{tr} \psi_\mu^\dagger \psi_\nu^\dagger \psi_\mu \psi_\nu, \quad h_{\text{I},3} = \frac{\xi}{2} \text{tr} \psi_\mu^\dagger \psi_\nu \psi_\mu \psi_\nu + \text{H.c.} \quad (2.7)$$

There are a few more possibilities for pillow interactions, but those that contain the combination $\psi_\mu \psi_\mu$ or its complex conjugate cannot contribute to leading order graphs. The simplest proof of this uses the auxiliary field technique, which works for pillow interactions but not for tetrahedric interactions (an explicit example will be given in Sec. 3). Taking this into account, we are also left with the three pillows

$$h_{\text{II},1} = \frac{\kappa}{2} \text{tr} \psi_\mu^\dagger \psi_\mu \psi_\nu^\dagger \psi_\nu, \quad h_{\text{II},2} = \frac{\kappa'}{2} \text{tr} \psi_\mu \psi_\mu^\dagger \psi_\nu \psi_\nu^\dagger, \quad h_{\text{II},3} = \frac{\kappa''}{2} \text{tr} \psi_\mu^\dagger \psi_\mu \psi_\nu \psi_\nu^\dagger + \text{H.c.} \quad (2.8)$$

Treating these three terms by introducing auxiliary fields $\psi_\mu \psi_\mu^\dagger$ and $\psi_\mu^\dagger \psi_\mu$, one may show that their effect is to renormalize the mass and to shift the energy by a simple function of the fermion number Q . These changes are rather trivial and can be taken into account straightforwardly. For this reason, we shall discard the pillow terms in most of the following discussion, except for a few specific remarks.

The action of $U(n)_L \times U(n)_R \times O(d)$ is defined in matrix notation by

$$\psi'_\mu = R_{\mu\nu} U_L \psi_\nu U_R^{-1}. \quad (2.9)$$

²The tracelessness condition is useful if one wants to study the models at all orders, see Sec. 2.3.

The diagonal subgroup for which $U_L = U_R$ is simply denoted by $U(n)$, and the fermion number symmetry is the $U(1)_Q$ subgroup acting by a phase as $\psi'_\mu = e^{-i\alpha}\psi_\mu$.

All the terms we are considering are $O(d)$ -symmetric. The terms $h_{I,1}$, $h_{II,1}$ and $h_{II,2}$ are $U(n)_L \times U(n)_R$ -symmetric and thus also conserve the fermion number. The terms $h_{I,2}$ and $h_{II,3}$ violate $U(n)_L \times U(n)_R$ but preserve the fermion number, and the term $h_{I,3}$ violates both $U(n)_L \times U(n)_R$ and $U(1)_Q$.

One can also define an anti-unitary time reversal operator $\hat{T}$, acting like the identity on the Fock basis obtained by the action of the creation operators $\psi^\dagger$ on the Fock vacuum $|0\rangle$ defined by $\psi_{\mu b}^a|0\rangle = 0$. The hermiticity of the Hamiltonian implies that the couplings λ , λ' , κ and κ' must be real, and the corresponding terms are then automatically T-invariant. If complex, the couplings ξ and κ'' violate T; see, however, the discussion in Sec. 2.1.3.

2.1.3 Generalized melons

There are many ways to depict the Feynman graphs of these models. For example, we can use the standard 't Hooft ribbon graph representation, with an additional middle strand associated to the $O(d)$ index. In the case of $U(n)_L \times U(n)_R$ -invariant models the two strands making up the ribbons are associated with distinct unitary groups and can thus be colored, say in green and red. The faces of the graphs then respect such coloring. In this case a so-called “colored graph representation” is also possible, and is used in the tensor model literature (see *e.g.* [3, 4]). For more general models, in which the $U(n)_L \times U(n)_R$ symmetry can be broken down to $U(n)$, a more economic and convenient representation is as follows.

The propagator is represented by a single oriented line. We choose conventionally the orientation to represent the flow of $U(1)_Q$ charge, with $\psi_\mu^\dagger$ creating and ψ_μ destroying one unit of fermion number. At the vertices, the cyclic ordering of the lines is important: going around a vertex clockwise corresponds to the positions of operators in a single trace from left to right. Moreover, the line in the propagator may be seen as being associated with the $O(d)$ index. Tetrahedric and pillow interaction vertices are thus distinguished by the fact that the lines cross or not, respectively. The vertices associated with the terms in (2.7) and (2.8) are depicted in Fig. 1.

Following the terminology of [4], the graphs that contribute at leading order in N are called *generalized melons*. More detailed comments on these leading graphs will be given in Sec. 2.3. Here we limit ourselves to describing them.

When only tetrahedric interactions are used, the leading order graphs contributing to the two-point functions are obtained by inserting on bare propagators the generalized elementary melons depicted in Fig. 2, and repeating this process recursively. Note that on top of the “symmetric” elementary melons (λ, λ) , (λ', λ') and (ξ, ξ^*) , there exist mixed structures like (λ, ξ^*) , *etc.* These mixed structures, inserted once,

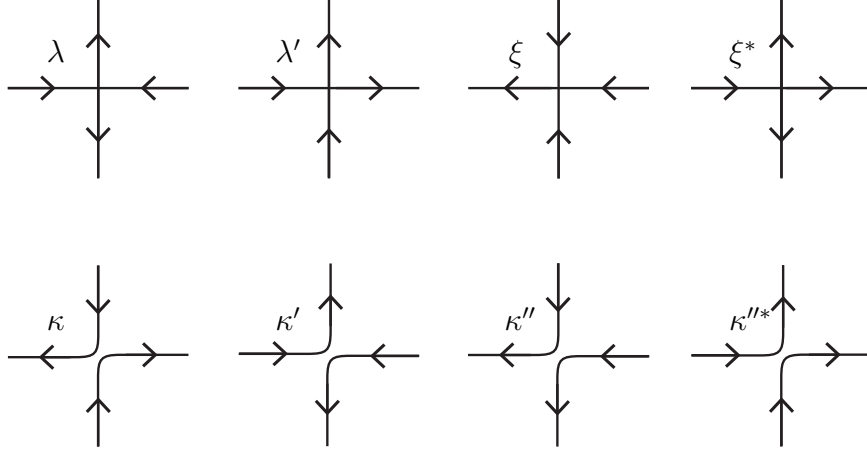


Figure 1: The tetrahedric (upper) and pillow (lower) single-trace interaction vertices (2.7) and (2.8), respectively.

can only contribute to the fermion number violating two-point functions $\langle \text{tr } \psi_\mu \psi_\mu \rangle$ and $\langle \text{tr } \psi_\mu^\dagger \psi_\mu^\dagger \rangle$, but multiple insertions do contribute to $\langle \text{tr } \psi_\mu \psi_\mu^\dagger \rangle$ as well. When pillow interactions are included, we can also insert recursively on bare propagators the elementary melonic structures depicted in Fig. 3. A generic leading order Feynman graph thus has a tree-like structure built out of bubbles attached to each other by a pillow vertex, on which arbitrary recursive insertions of generalized elementary melons may be added. Leading order graphs contributing to the vacuum amplitude, or free energy, are obtained by gluing the external legs of a graph contributing to $\langle \text{tr } \psi_\mu \psi_\mu^\dagger \rangle$ at leading order. A sketch of such a leading vacuum graph is depicted in Fig. 4.

Further simplifications result from restricting to the models for which the fermion number violating two-point functions vanish at leading order, $\langle \text{tr } \psi_\mu \psi_\mu \rangle = \langle \text{tr } \psi_\mu^\dagger \psi_\mu^\dagger \rangle = 0$. This yields one model for which we can keep the interactions $h_{\text{I},1}$ and $h_{\text{I},2}$ and another model based on $h_{\text{I},3}$. But the structure of the elementary melons implies that, at leading order in N , the only effect of adding $h_{\text{I},2}$ on top of $h_{\text{I},1}$ is to renormalize the coupling $\lambda \rightarrow \lambda + 2\lambda'$. It is thus enough to keep only the term $h_{\text{I},1}$. We are left with two basic models,

$$H_{\text{U}(n)^2} = nd \left(m \text{tr } \psi_\mu^\dagger \psi_\mu + \frac{\lambda}{2} \sqrt{d} \text{tr } \psi_\mu^\dagger \psi_\nu \psi_\mu^\dagger \psi_\nu \right), \quad (2.10)$$

$$H_{\text{U}(n)} = nd \left(m \text{tr } \psi_\mu^\dagger \psi_\mu + \frac{\xi}{2} \sqrt{d} \text{tr } \psi_\mu^\dagger \psi_\nu \psi_\mu \psi_\nu + \frac{\xi^*}{2} \sqrt{d} \text{tr } \psi_\mu^\dagger \psi_\nu^\dagger \psi_\mu^\dagger \psi_\nu \right). \quad (2.11)$$

As already noted, pillow interactions may also be added to either of these rather trivially. The models (2.10) and (2.11) are those on which we shall focus in the rest of the paper. In spite of very similar-looking Hamiltonians, their physics, in particular their phase diagrams and infrared behaviours, will turn out to be drastically different.

Note that the model (2.10) has the full $\text{U}(n)_\text{L} \times \text{U}(n)_\text{R}$ symmetry and in particular

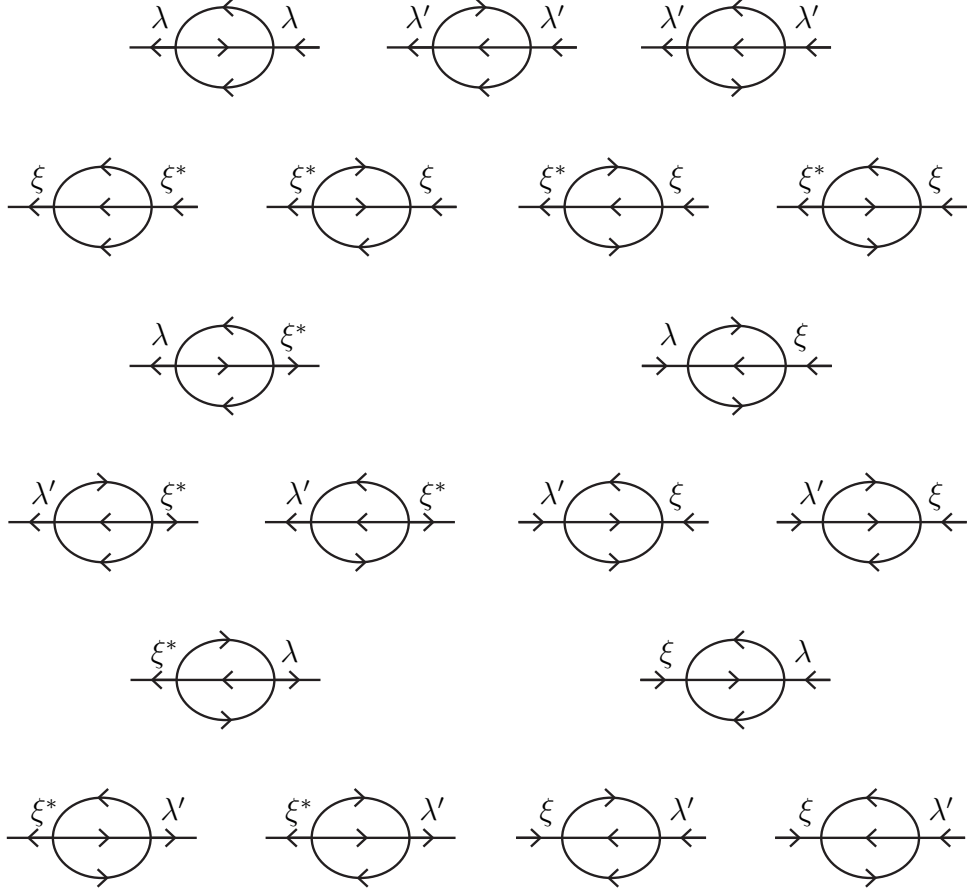


Figure 2: Generalized elementary melons built from tetrahedric (type I) vertices.

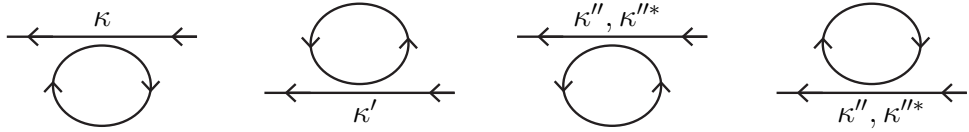


Figure 3: Elementary melons built from pillow (type II) vertices.

conserves the fermion number. The model (2.11), on the other hand, only has the diagonal $U(n)$ symmetry and violates both $U(1)_Q$ and time-reversal invariance (for a general complex coupling ξ). However, at leading order the symmetric structure of the elementary melons implies that time reversal invariance is actually restored, because only the combination $\xi\xi^* = |\xi|^2$ can enter. Moreover, for the same reasons a $\mathbb{Z}_4$ subgroup of $U(1)_Q$ acting as $\psi \mapsto i\psi$ is also preserved at leading order, explaining why $\langle \text{tr } \psi_\mu \psi_\mu \rangle$ and $\langle \text{tr } \psi_\mu^\dagger \psi_\mu^\dagger \rangle$ vanish.

Because of their symmetry properties, we shall refer to the models with Hamiltonians (2.10) and (2.11) as the $U(n)^2$ and $U(n)$ models, respectively.

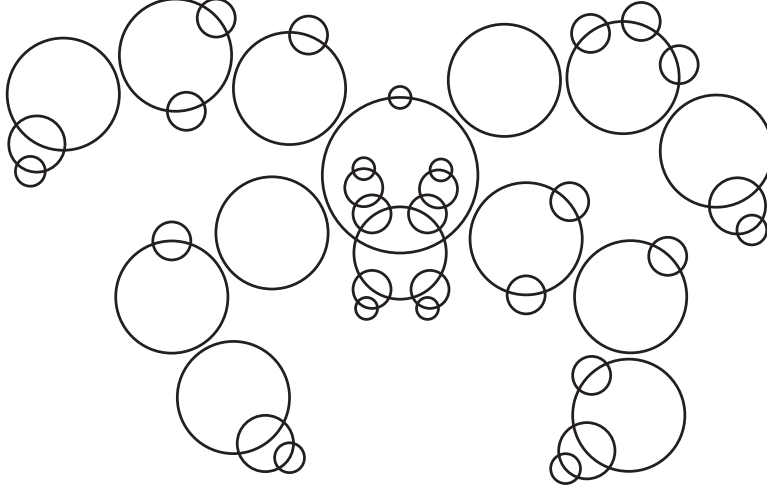


Figure 4: Sketch of a leading vacuum Feynman graph built from both tetrahedric and pillow vertices.

There are also interesting generalizations involving higher order interaction terms, similar to the q -body interacting SYK model introduced in [9] and based on the complete interaction bubble studied in [4]. We shall briefly comment on such a generalization for (2.10) in the following sections, see also App. A.

2.2 Disordered models

In the disordered model framework, the degrees of freedom are N complex fermions conveniently normalized to satisfy the anti-commutation relations

$$\{\chi^i, \chi_j^\dagger\} = \frac{1}{N} \delta_j^i, \quad (2.12)$$

and we consider a general Hamiltonian of the form

$$H = N \left(m \chi_i^\dagger \chi^i + \frac{1}{2} \kappa \chi_i^\dagger \chi^i \chi_j^\dagger \chi^j \right) + \sqrt{N} \left(\lambda_{kl}^{ij} \chi_i^\dagger \chi_j^\dagger \chi^k \chi^l + \xi_{jkl}^i \chi_i^\dagger \chi^j \chi^k \chi^l + \xi_l^{ijk} \chi_i^\dagger \chi_j^\dagger \chi_k^\dagger \chi^l + \zeta_{ijkl} \chi^i \chi^j \chi^k \chi^l + \zeta^{ijkl} \chi_i^\dagger \chi_j^\dagger \chi_k^\dagger \chi_l^\dagger \right). \quad (2.13)$$

The coupling κ mimics the pillow interactions of the matrix models. The other couplings satisfy the hermiticity conditions

$$(\lambda_{kl}^{ij})^* = \lambda_{ji}^{lk}, \quad (\xi_{jkl}^i)^* = \xi_i^{lkj}, \quad (\zeta_{ijkl})^* = \zeta^{lkji}, \quad (2.14)$$

and the symmetry conditions

$$\lambda_{kl}^{ij} = \lambda_{[kl]}^{[ij]}, \quad \xi_{jkl}^i = \xi_{[jkl]}^i, \quad \zeta_{ijkl} = \zeta_{[ijkl]}. \quad (2.15)$$

They are drawn randomly from a Gaussian distribution with weight

$$\exp\left[-\sum_{ijkl}\left(|\lambda_{kl}^{ij}|^2/\lambda^2-4|\xi_{jkl}^i|^2/\xi^2-4|\zeta_{ijkl}|^2/\zeta^2\right)\right]. \quad (2.16)$$

The large N limit is defined by taking the real couplings κ , λ , ξ and ζ fixed. In this limit, the so-called self-averaging observables, which include the thermodynamical potentials and the two- and four-point functions studied later in the paper, coincide with their average over disorder. This quenched average is usually performed by introducing replicas, but at leading large N order this turns out to be unnecessary. Indeed, the solution is given by the diagonal replica ansatz, which means that the quenched average coincides with the annealed average that is obtained by directly integrating over the random couplings as if they were dynamical variables themselves. The models can then be solved using the standard auxiliary field technique, as will be briefly reviewed in Sec. 3.

It is easy to check that at leading order in N , and for all the purposes of the present paper, the random Hamiltonian (2.13) is equivalent to (2.10) or (2.11) when turning on the coupling λ or ξ , respectively. The effect of the matrix model pillow terms are also correctly mimicked by the coupling κ in (2.13).

Note that differences do exist between the matrix model and the disordered model formulations. For instance, the coupling ζ in (2.13) has no analogue in the matrix case, because a term like $\text{tr} \psi_\mu \psi_\nu \psi_\mu \psi_\nu$ identically vanishes. We will therefore take $\zeta = 0$ in the rest of this paper. Conversely, some features of the matrix model, such as non-vanishing fermion number violating correlation functions, cannot be reproduced by (2.13) because the annealed average automatically yields an emergent $U(N)$ symmetry which implies fermion number conservation. Equivalently, the integration over the random couplings can produce “symmetric” elementary melons, similar to the ones depicted in the first two rows of Fig. 2, but cannot yield the other asymmetric structures. To achieve this, one needs to generalize the disordered model. A simple possibility is to consider two species of fermionic vectors χ^i and $\tilde{\chi}_i$ transforming in complex conjugate representations of $U(N)$, using a Gaussian integration measure over random couplings that includes terms mixing different couplings. An explicit example could be a term $\lambda_l^{ijk} \chi_i^\dagger \chi_j^\dagger \tilde{\chi}_k \chi^l + \xi_{jkl}^i \chi_i^\dagger \chi^j \chi^k \chi^l + \text{H.c.}$ in the Hamiltonian with a contribution of the form $\lambda_l^{ijk} \xi_{ijk}^l + \text{H.c.}$ in the Gaussian weight for the couplings. This does generate asymmetric melonic structures and non-zero fermion number violating two-point functions of the form $\langle \chi^i \tilde{\chi}_i \rangle$, but will not be considered further in this paper.

To summarize, the $U(n)^2$ matrix model is mimicked at leading large N order by

the random Hamiltonian³

$$H = Nm \chi_i^\dagger \chi^i + \sqrt{N} \lambda_{kl}^{ij} \chi_i^\dagger \chi_j^\dagger \chi^k \chi^l, \quad (2.17)$$

which has already been considered in the literature and is often called the “complex SYK model,” see *e.g.* [19, 20, 29–31]. As for the $U(n)$ model, it is on the other hand mimicked by

$$H' = Nm \chi_i^\dagger \chi^i + \sqrt{N} (\xi_{jkl}^i \chi_i^\dagger \chi_j^\dagger \chi^k \chi^l + \xi_l^{ijk} \chi_i^\dagger \chi_j^\dagger \chi_k^\dagger \chi^l), \quad (2.18)$$

and we are not aware of any prior work on it, even though it has obvious similarities with (2.17) and might be expected to also be relevant for condensed-matter purposes.

2.3 Discussion and relation to other tensor models

This subsection provides details and comments on the matrix model point of view that are not strictly necessary to follow the main new results of the paper, as presented in the next sections. It may thus be skipped upon first reading. The aim is to provide further motivation for our work, as well as to place it in the wider context of recent developments in tensor models. The reader can refer to [40–42] for reviews on the latter.

We shall first clarify the nature of the double large n and large d expansions we are considering, since some confusion seems to exist in the literature about certain terminology. We will discuss models with the full $U(n)_L \times U(n)_R$ symmetry in Sec. 2.3.1, and consider models with only one $U(n)$ symmetry in Sec. 2.3.2. Finally, in Sec. 2.3.3 we elaborate on the string-theory inspired interpretation of the models we are studying, as well as the relation between these and other tensor models considered in the literature.

2.3.1 Degree expansion, index expansion, melonic and non-melonic interactions

We have already mentioned in Sec. 2.1.3 that diagrams for models with $U(n)_L \times U(n)_R$ symmetry can be “colored”. As reviewed in detail for instance in [4], both interaction vertices and Feynman graphs have a representation in terms of regular colored graphs, the so-called “bubbles.” To any bubble we can associate a Gurau degree [1, 40, 43] which is in general a non-negative half integer. The original and standard definition of a *melon* is to be a *regular colored graph of Gurau degree zero*.

³One should not confuse this model with the so-called “mass-deformed SYK model,” which combines random quadratic and quartic terms (usually with Majorana fermions), see for example [38, 39].

Since interaction vertices are in one-to-one correspondence with regular colored graphs, they can be classified into two groups: the melonic and non-melonic interactions. In our case, the tetrahedric term $h_{\text{I},1}$ in (2.7) is non-melonic, whereas the pillow terms $h_{\text{II},1}$ and $h_{\text{II},2}$ in (2.8) are melonic.

To define a large N limit, one should choose the way the coupling constants scale with N . In the case of matrix models, there is a unique interesting choice first discussed by 't Hooft [37], which yields the genus expansion. In the case of tensor models, the situation is much more complex and interesting. Several inequivalent and non-trivial scalings can *a priori* be defined. In any given model, the *optimal* scaling of a coupling constant is the one for which it is impossible to further enhance that coupling and still have a well-defined expansion. Computing the optimal scaling in tensor models is in general a difficult technical problem, see *e.g.* [44].

There are two general procedures known to define scalings with well-defined large N limits for arbitrary interactions in tensor models.⁴

One is the so-called BGR scaling [45], which was used originally in the tensor model literature but can also be straightforwardly generalized to the matrix-tensors by distinguishing two sets of indices with $d \neq n$, [4]. This scaling has two fundamental properties:

- i) The Feynman graphs in the large N expansion of the tensor models are classified according to their Gurau degree. In particular, the leading graphs are the standard melons, whose structure was elucidated in [46]. Subleading orders may be studied using the results in [47].
- ii) Only melonic interaction vertices can contribute to leading order graphs. In particular, the tetrahedric terms in (2.7) do *not* contribute at leading order in this framework. This is due to the fact that the $\sqrt{d}$ enhancement that we have introduced in (2.3) is *not* present in the BGR scaling.

The physics of quantum mechanical models based on the BGR scaling is similar to the physics of the large N vector models [34] and in particular cannot describe black-hole physics (for instance, there is no continuous spectrum at large N in this scheme). Note also that, in the BGR scaling, if we distinguish $U(n)$ and $O(d)$ indices the large n and large d limits commute.

The other scaling was introduced in [4], generalizing to any tensor (or matrix-tensor) model the scaling used by Carrozza and Tanasa for a $O(n)^3$ tensor model in their seminal work [48]. The scaling of [4] is actually optimal for a large class of interactions, called maximally single-trace, which generalize the concept of single-trace matrix model interactions to general tensor or matrix-tensor models. This new scaling is the one used in the present paper, and is such that:

⁴This applies to the general “uncolored” models, in tensor model terminology.

- i) The Feynman graphs in the large N expansion of the tensor models are classified according to their *index*, which is a non-negative half-integer naturally associated to the colored graph. In particular, the leading graphs have index zero and are called *generalized melons* in [4]. Note that a graph of Gurau degree zero automatically has index zero, but the converse is not true: there are many more generalized melons than melons, and their full classification is at the moment unknown. Some special cases, which include the interactions used in the present paper, have been treated in [4, 48].
- ii) Both melonic and some non-melonic interactions can appear in the leading order graphs. In particular, both the tetrahedric $h_{\text{I},1}$ vertex in (2.7) and the pillows $h_{\text{II},1}$ and $h_{\text{II},2}$ in (2.8) contribute at leading order, which makes the physics much richer and more interesting than that of the BGR scaling.

Even though the generalized melons built out of the tetrahedric interaction vertex $h_{\text{I},1}$ in (2.7) have a strictly positive Gurau degree, they share the basic fundamental recursive structure associated with standard melons: they are built by arbitrary repeated insertions of generalized elementary melons, depicted in Fig. 2, and the sum over such graphs can be evaluated using Schwinger-Dyson equations of “melonic” type. This simple but non-trivial result is known to be valid for the interactions used in the present paper, as well as for some generalizations thereof [4, 5]. In more general cases, it is still usually very easy to show that graphs built in such a way will contribute at leading order. However, let us emphasize that the hard part of the proof is to demonstrate that no other graph can contribute; the simple “melonic” recursive structure is not *a priori* required. This hard part is often overlooked in the modern literature. The full classification of generalized melons remains a very interesting open problem.

2.3.2 $U(N)$ -symmetric models and the tracelessness condition

Almost all the tensor model literature, including [4], relies on the fact that each distinct index of the tensors (or of the matrix-tensors) is associated to a distinct symmetry group. This simplifies the analysis of the Feynman graphs considerably, because different indices of a given tensor cannot mix in a given face.

For a long time it was believed that this simplifying property was necessary to define a consistent large N limit. The fact that this is actually not the case was first pointed out in [3]. A simple argument, based on the possibility to color the faces of the relevant planar graphs with only two colors, implies that the large d limit with the enhanced scaling of [4, 48] used in the present paper is consistent and that the graphs contributing at leading order are again the generalized melons. The models containing terms that break $U(n)_{\text{L}} \times U(n)_{\text{R}}$ down to $U(n)$, like (2.11), can thus have a well-defined large N limit.

An obvious question to ask is what happens beyond the planar graphs. To some extent this question may be purely technical or academic, because for the most important physical applications we are aware of, only the planar graphs are needed. Moreover, in the matrix interpretation the higher genus terms are also subleading to all the $d^{1-\ell/2}$, $\ell > 0$, corrections that already contribute at the planar level. These planar corrections are themselves out of reach with present technology, except for the leading piece in $\sqrt{d}$ which is easy to include [48].⁵

This being said, it is an extremely interesting problem from the graph-theoretic point of view, and is not limited to the matrix-tensor models. From the purely tensorial point of view, it is very natural to impose symmetry properties on the indices by considering, for instance, completely symmetric or antisymmetric tensors. Such symmetry conditions have the same effect as the $U(n)_L \times U(n)_R$ symmetry breaking terms in the models we study, the result being that distinct indices mix up in faces and the usual tensor model technology based on colored graphs does not apply.

It was very soon realised after [3] that in these models there exists a specific class of non-planar graphs that makes the new large d limit inconsistent, see [5] for explicit examples. Similar graphs spoil the large N limit of other (anti)symmetric tensor models. An elegant way to remove these graphs, which have a very specific structure, is to impose tracelessness conditions [49]. This was shown to be correct in the case of symmetric tensors in [50], using a technical tour de force based on a set of local moves simplifying the graphs [51].

The problem remains open in the case of the matrix-tensors. The additional difficulty is that we do not have complete symmetry between all the indices in this case, and in particular the tracelessness condition is imposed only on the matrix indices.⁶ A precise conjecture for the models of interests in the present paper is as follows.

Conjecture: For the models (2.3) for which we impose the condition $\text{tr } \psi_\mu = 0$, the genus g free energy (and similarly any other observable) has a well-defined large d expansion of the form

$$F_g = \sum_{\ell \geq 0} d^{\eta_g - \ell/2} F_{g,\ell}, \quad (2.19)$$

where η_g is an upper bound on the highest possible power of d for a graph of fixed genus g . From [3], we already know that $\eta_0 = 1$. Generalizing results proven in [3, 4] for the case of $U(n)_L \times U(n)_R$ models at higher genera, for which the tracelessness condition is not required, we may further conjecture that $\eta_g = 1 + g$.

⁵The physics at finite n and the resulting mechanisms that restore unitarity are of course of great interest, see *e.g.* [17]. However, these are non-perturbative effects in the coupling $1/n^2$ and cannot be studied by computing to higher genera in the $1/n^2$ expansion.

⁶We would like to thank Sylvain Carrozza, Adrian Tanasa and Guillaume Valette for thorough discussions on this problem.

2.3.3 Matrix models vs. tensor models, brane picture and gauging

As discussed above, the models we are considering are close cousins of the fermionic tensor models introduced in [15, 25]. To make the relation more explicit, one can set $d = n$ and treat the $\psi_{\mu b}^a$ as a three-index tensor to take the large n limit.

In the present context, this procedure seems rather artificial because the indices a, b and μ are associated with different symmetry groups. We believe it is more physical and interesting to keep the matrix model point of view. In the matrix model picture, one can always interpret the indices a, b associated with the unitary symmetry as Chan-Paton factors of oriented open strings ending on some kind of D-particle. The additional orthogonal $O(d)$ symmetry is then naturally interpreted as corresponding to the rotational invariance in the directions transverse to the D-particle worldline. This picture allows interesting physical interpretations of the results, and makes the relation with string theory, holography and black-hole physics much clearer. The large d limit might also have an interesting relation with the large dimension limit of general relativity studied in [52].

At subleading order, the difference between the matrix model and the tensor model interpretations becomes quantitative.

One reason is that the corrections are organized differently in the two pictures. In the matrix models, planar contributions are always dominant over non-planar contributions, even if we consider higher order corrections in $1/\sqrt{d}$. On the other hand, if $d = n$ as in the tensor model interpretation, a given order in the large n expansion can mix up matrix model graphs having different genera.

Another reason is that the gauging of the global symmetries is done in a different way in the matrix models and in the tensor models. In the former, in line with the open string picture, we gauge only the $U(n)$ or $U(n)_L \times U(n)_R$ symmetries that act on the matrix indices. The $O(d)$ symmetry remains ungauged and is associated with the conservation of angular momentum. One then gets a standard string-like spectrum of operators. On the other hand, in the tensor model picture it is more natural to gauge the full global symmetry group, $U(n)^2 \times O(n)$ in our case, or $O(n)^3$ in well-studied examples [53].

Note, however, that both in large d matrix models and in tensor models the gauging is relevant only at next-to-next-to leading order, because the dimension of the gauge group is of order n^2 . Any contribution coming from the gauging, for instance a gauge-fixing action and a Fadeev-Popov determinant, will thus contribute at order n^2 , whereas the leading and next-to-leading orders are $n^2 d$ and $n^2 \sqrt{d}$ (or n^3 and $n^{5/2}$ in tensor models). For instance, any Hagedorn-like phase transition at weak coupling will be pushed to zero temperature at leading order (in large d matrix models, it is easy to check that the weak coupling Hagedorn temperature is of order $1/\ln d$).

3 Euclidean time picture

We now proceed to study the models (2.10) and (2.11) in the Euclidean formalism, leaving for Sec. 5 the corresponding analysis in real time. Detailed derivations of various formulas are provided for thermodynamic quantities, as well as for two- and four-point functions. The low energy solutions to the relevant Schwinger-Dyson equations are spelled out, and we also discuss the numerical methods used to solve these equations at finite coupling, exemplifying their application.

An important point we discuss is the possibility to have non-trivial monodromies in parameter space for the two-point functions. This means that the Schwinger-Dyson equations can have several distinct and consistent solutions for fixed values of their parameters. This possibility was first discovered in [28] and later also observed in [54] for a related system with two coupled SYK models. It will play a major role in Sec. 4 when we study the phase diagram of the $U(n)^2$ model (2.10).

An innovative aspect of our discussion is the introduction of a new ansatz for the infrared behaviour of two-point functions. We assume that their power-law decay is governed in general by two *distinct* exponents Δ_+ and Δ_- when the Euclidean time goes to $+\infty$ and $-\infty$ respectively. The usual ansatz used so far in the literature corresponds to $\Delta_+ = \Delta_-$, but this is not strictly required in certain cases. The possibility to have $\Delta_+ \neq \Delta_-$ will play a crucial role in elucidating the physics of the $U(n)$ model (2.11).

We also provide a careful justification of the use of effective actions for the standard bilocal fields $G(\tau_1, \tau_2)$ and $\Sigma(\tau_1, \tau_2)$ in the matrix (or tensor) models. Most of the literature has focused on disordered models, in which this effective action can be straightforwardly derived by a path integral argument, see *e.g.* [55], but this is not applicable in the case of matrix or tensor models.

Finally, we derive a general formula *à la* Bethe-Salpeter (3.133) for the kernel used in the computation of the four-point functions, which is applicable in a wide class of models. Its analytic continuation to real time will be performed in Sec. 5, allowing us to study the strength of quantum chaos as diagnosed by the Lyapunov exponent.

3.1 Euclidean time two-point function

3.1.1 General definitions and properties

We define the time-ordered finite temperature Euclidean two-point function by

$$G(\tau) = \begin{cases} \frac{1}{n} \langle \text{T tr } \psi_{\mu\text{E}}(\tau) \psi_{\mu}^{\dagger} \rangle_{\beta} & (\text{matrix models}) \\ \langle \text{T } \chi_{\text{E}}^i(\tau) \chi_i^{\dagger} \rangle_{\beta} & (\text{disordered models}). \end{cases} \quad (3.1)$$

The Euclidean time evolution of any operator is defined as usual by

$$\mathcal{O}_E(\tau) = e^{H\tau} \mathcal{O} e^{-H\tau}. \quad (3.2)$$

The equations (3.1) are valid for $-\beta < \tau < \beta$. The correlators satisfy the KMS condition and can be extended to all Euclidean times by antiperiodicity of period $\beta = 1/T$,

$$G(\tau + \beta) = -G(\tau). \quad (3.3)$$

In particular, we use the Fourier decomposition

$$G(\tau) = \frac{1}{\beta} \sum_{k \in \mathbb{Z} + \frac{1}{2}} G_k e^{-i\nu_k \tau}, \quad (3.4)$$

with Matsubara-Fourier coefficients G_k and Matsubara frequencies

$$\nu_k = \frac{2\pi}{\beta} k. \quad (3.5)$$

The tree-level Matsubara coefficients are given by

$$G_k^{(0)} = \frac{1}{m - i\nu_k}, \quad (3.6)$$

yielding the tree-level two-point function

$$G^{(0)}(\tau; m) = \begin{cases} n_F(m) e^{m(\beta - \tau)} & \text{if } 0 < \tau < \beta \\ -n_F(m) e^{-m\tau} & \text{if } -\beta < \tau < 0, \end{cases} \quad (3.7)$$

where we have introduced the Fermi distribution function

$$n_F(\omega) = \frac{1}{e^{\beta\omega} + 1}. \quad (3.8)$$

On top of (3.3), the Euclidean two-point function has a set of interesting general properties.⁷ To understand these properties, it is very useful to introduce the so-called spectral function, defined by

$$\rho(\omega) = \frac{1}{Z} \sum_{p,q} (e^{-\beta E_p} + e^{-\beta E_q}) |\alpha_{pq}|^2 \delta(\omega + E_p - E_q), \quad (3.9)$$

where

$$Z = \text{Tr } e^{-\beta H} \quad (3.10)$$

⁷See *e.g.* [56] for a thorough discussion in the bosonic case, which can be straightforwardly generalised to the fermionic case relevant for the present paper.

is the partition function, E_p are the eigenvalues of the Hamiltonian, the double sum in (3.9) is over all the eigenstates $|p\rangle$ and $|q\rangle$ of the Hamiltonian⁸ and

$$|\alpha_{pq}|^2 = \begin{cases} \sum_{\mu,a,b} |\langle p|\psi_{\mu b}^a|q\rangle|^2 & \text{(matrix models)} \\ \sum_i |\langle p|\chi^i|q\rangle|^2 & \text{(disordered models).} \end{cases} \quad (3.11)$$

The two-point function has the following spectral decomposition,

$$G(\tau) = \int_{-\infty}^{+\infty} G^{(0)}(\tau; \omega) \rho(\omega) d\omega, \quad (3.12)$$

which is equivalent to

$$G_k = \int_{-\infty}^{+\infty} \frac{\rho(\omega)}{\omega - i\nu_k} d\omega. \quad (3.13)$$

The spectral function ρ is manifestly real and positive. This yields

$$G_k^* = G_{-k}, \quad (3.14)$$

and from (3.12) and (3.13),

$$G(\tau) > 0 \quad \text{if} \quad 0 < \tau < \beta, \quad G(\tau) < 0 \quad \text{if} \quad -\beta < \tau < 0, \quad (3.15)$$

$$\text{Im } G_k > 0 \quad \text{if} \quad k > 0. \quad (3.16)$$

The moments of the spectral function,

$$\rho_n = \int_{-\infty}^{+\infty} \omega^n \rho(\omega) d\omega, \quad (3.17)$$

govern the UV behaviour of the Matsubara-Fourier coefficients (3.13),

$$G_k = \sum_{n=0}^{r-1} (-1)^n \left(\frac{i}{\nu_k} \right)^{n+1} \rho_n + \mathcal{O}(1/k^{r+1}), \quad (3.18)$$

and through (3.12) they fix the discontinuities of the two-point function and its derivatives,

$$G^{(n)}(0^+) - G^{(n)}(0^-) = (-1)^n \rho_n. \quad (3.19)$$

In particular, the canonical commutation relations (2.1) or (2.12) yield

$$G(0^+) - G(0^-) = \rho_0 = \int_{-\infty}^{+\infty} \rho(\omega) d\omega = 1. \quad (3.20)$$

Moreover, as will be demonstrated in the following (see 3.2.2), the first moment ρ_1 is exactly given by a one-loop calculation, and the one-loop contribution is subleading

⁸Note that in all of the models we consider the finite N spectrum is discrete.

in the $N \rightarrow \infty$ limit. At leading large N order, ρ_1 thus coincides with the tree-level result

$$\dot{G}(0^-) - \dot{G}(0^+) = \rho_1 = \int_{-\infty}^{+\infty} \omega \rho(\omega) d\omega = m \quad \text{at leading large } N \text{ order.} \quad (3.21)$$

A last very useful quantity that we need to introduce is the so-called resolvent R , defined by

$$R(z) = \int_{-\infty}^{+\infty} \frac{\rho(\omega)}{z - \omega} d\omega. \quad (3.22)$$

The resolvent is a holomorphic function for $\text{Im } z > 0$ and $\text{Im } z < 0$. Eq. (3.13) is equivalent to

$$R(i\nu_k) = -G_k. \quad (3.23)$$

The large z expansion of R is governed by the moments (3.17) and in particular

$$R(z) \underset{z \rightarrow \infty}{=} \frac{1}{z} + \frac{m}{z^2} + \mathcal{O}(1/z^3). \quad (3.24)$$

When the spectrum of the Hamiltonian is discrete, which is the case in these models at finite N , the only singularities of R are simple poles on the real axis at the Bohr frequencies. At infinite N , the spectrum becomes continuous and R develops a discontinuity across the real axis. The spectral function can be read off from this discontinuity as

$$\rho(\omega) = \frac{i}{2\pi} (R(\omega + i0^+) - R(\omega - i0^+)). \quad (3.25)$$

3.1.2 Schwinger-Dyson equations

Let us introduce the self-energy $\Sigma(\tau)$, which is an antiperiodic function with Matsubara coefficients

$$\Sigma_k = \frac{1}{G_k} - \frac{1}{G_k^{(0)}} = \frac{1}{G_k} - m + i\nu_k. \quad (3.26)$$

This equation is equivalent to

$$\left(\frac{d}{d\tau} + m\right)G(\tau - \tau') + \int_0^\beta \Sigma(\tau - \tilde{\tau})G(\tilde{\tau} - \tau') d\tilde{\tau} = \delta_\beta(\tau - \tau'), \quad (3.27)$$

where we have introduced the notation δ_β for the antiperiodic δ -function

$$\delta_\beta(\tau) = \sum_{n \in \mathbb{Z}} (-1)^n \delta(\tau + n\beta). \quad (3.28)$$

The self-energy is computed by resumming the one-particle irreducible (1PI) two-point graphs from which the full two-point function is obtained by the usual geometric series,

$$G_k = \frac{1}{m - i\nu_k + \Sigma_k} = \frac{1}{m - i\nu_k} - \frac{1}{m - i\nu_k} \Sigma_k \frac{1}{m - i\nu_k} + \dots. \quad (3.29)$$

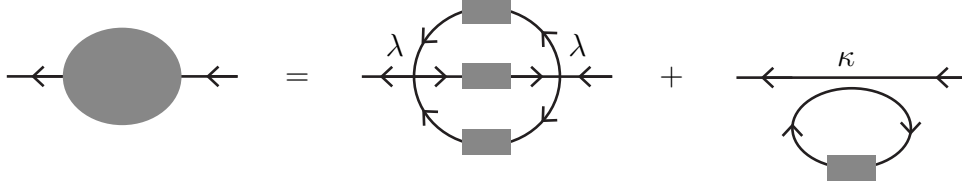


Figure 5: The Schwinger-Dyson equation for the $U(n)^2$ model with the pillow vertex $h_{II,1}$ turned on. The disk represents the self-energy Σ and the rectangle the full two-point function G . In this case, the self-energy gets contributions from the structures depicted on the upper-left corner of Fig. 2 and on the left of Fig. 3.

The recursive structure of the leading order graphs discussed in Sec. 2.1.3 implies simple self-consistency conditions, called Schwinger-Dyson equations, on the sum of the 1PI two-point graphs. For instance, for the $U(n)^2$ model (2.10) and the $U(n)$ model (2.11), we get

$$\Sigma(\tau) = \lambda^2 G(\tau)^2 G(-\tau) \quad (U(n)^2 \text{ model}), \quad (3.30)$$

$$\Sigma(\tau) = -\frac{|\xi|^2}{4} (G(\tau)^3 + 3G(\tau)G(-\tau)^2) \quad (U(n) \text{ model}). \quad (3.31)$$

The contributions for the $U(n)^2$ model come from the recursive insertions of the generalized elementary melon depicted on the top-left corner of Fig. 2. The contributions for the $U(n)$ model come from the melons on the second line of Fig. 2: the one on the left yields the term $G(\tau)^3$ and the other three the term $3G(\tau)G(-\tau)^2$.

Including contributions from the pillow vertices is very easy. For instance, if we turn on the $h_{II,1}$ term of (2.8) in the $U(n)^2$ model, the right-hand side of (3.30) gets an additional contribution $-\kappa G(0^-)\delta_\beta(\tau) = \kappa Q\delta_\beta(\tau)$ due to the possibility of inserting the structure on the left of Fig. 3. The full Schwinger-Dyson equation resulting from this addition is depicted schematically in Fig. 5. The same term would be added if $h_{II,1}$ were to be turned on in the $U(n)$ model. The term $h_{II,2}$ produces a contribution $-\kappa' G(0^+)\delta_\beta(t) = \kappa'(Q-1)\delta_\beta(\tau)$ and the term $h_{II,3}$ a contribution $(\text{Re } \kappa'')(G(0^+) + G(0^-))\delta_\beta(\tau) = (\text{Re } \kappa'')(1-2Q)\delta_\beta(\tau)$, consistently with the other structures in Fig. 3.

From the point of view of the disordered models, one does not need to analyze the leading graphs to derive the Schwinger-Dyson equations. One can instead apply a purely algebraic reasoning using the standard auxiliary field method, based on the saddle-point equations of the effective action derived from the path integral.

3.1.3 On the solutions of Schwinger-Dyson equations and monodromies

The Schwinger-Dyson equations are perturbative in nature, since they are associated with the summation of a class of Feynman graphs. However, they also contain non-perturbative information because the sum of Feynman graphs they compute has a

finite radius of convergence. Performing the analytic continuation of the perturbative series to strong coupling can yield extremely interesting and non-perturbative physics.

Morally speaking, this is very similar to what happens, for example, in some supersymmetric gauge theories [57], where one resums weakly coupled semiclassical contributions that have a finite radius of convergence to obtain non-trivial results at strong coupling via analytic continuation. In the models studied in [57], these are instanton contributions, not perturbative contributions, but the spirit is essentially the same.

In both the $U(n)^2$ and $U(n)$ models, the perturbative series in λ^2 or $|\xi|^2$ can be straightforwardly computed from the Schwinger-Dyson equations. For instance, in the $U(n)^2$ model we may plug (3.7) into (3.30) to obtain the first non-trivial order of Σ , which is $\mathcal{O}(\lambda^2)$. Inserting this approximate solution into (3.29) to get a correction for G_k , we find

$$G_k = \frac{1}{m - i\nu_k} \left[1 + \frac{e^{\beta m}}{(e^{\beta m} + 1)^2} \frac{\lambda^2}{(m - i\nu_k)^2} + \mathcal{O}(\lambda^4) \right]. \quad (3.32)$$

This procedure can be iterated to any desired order in λ . In particular, it leads to Eq. (3.21) since all the perturbative corrections to the tree-level result decay at least as fast as $1/k^3$ when $k \rightarrow \infty$.

An important feature of the models we are studying is that the perturbative expansion converges in a region of parameter space that includes two physically very different regimes.

- i) One regime is the high temperature limit at fixed mass and coupling, $\beta m \ll 1$ and $\beta \lambda \ll 1$ (in the $U(n)^2$ model) or $\beta |\xi| \ll 1$ (in the $U(n)$ model). This regime is perturbative because the fermions get a large effective thermal mass of order πT . In this limit, the leading order solution is $G_k = i/\nu_k$, which is equivalent to $G(\tau) = \frac{1}{2} \text{sign}(\tau)$. At the level of the states, this is an expansion around the maximally entropic state associated with the density matrix $2^{-N} \mathbb{I}$.
- ii) Another regime is the high mass limit $\lambda/m \ll 1$ (in the $U(n)^2$ model) or $|\xi|/m \ll 1$ (in the $U(n)$ model). This is the standard perturbative regime in particle physics, whose leading order is governed by the usual harmonic oscillator solution (3.7). In this limit, the models remain perturbative even at zero temperature and the ground state is simply the Fock vacuum $|0\rangle$.

Overall, there is an effective coupling constant $g_{\text{eff}} = \min(\lambda/T, \lambda/m)$ in the $U(n)^2$ model, or $g_{\text{eff}} = \min(|\xi|/T, |\xi|/m)$ in the $U(n)$ model. The perturbative series converges when g_{eff} is small enough.

A crucial property of the solutions of the Schwinger-Dyson equations is that they can undergo non-trivial monodromies in parameter space. This possibility was first

realized in [28], and later also found in other disordered models with quadratic terms [54]. In other words, depending on the path one uses to perform the analytic continuation from the weakly coupled region $g_{\text{eff}} \ll 1$ where the perturbative series converges to a point in the strongly coupled region beyond the radius of convergence, one may end up with *distinct* solutions. These distinct solutions are *a priori* all physically relevant, in the sense that they satisfy all the general consistency conditions listed in Sec. 3.1.1, namely (3.3), (3.14), (3.15), (3.16), (3.20) and (3.21). This property is illustrated in Fig. 6 and will be abundantly discussed in the following.

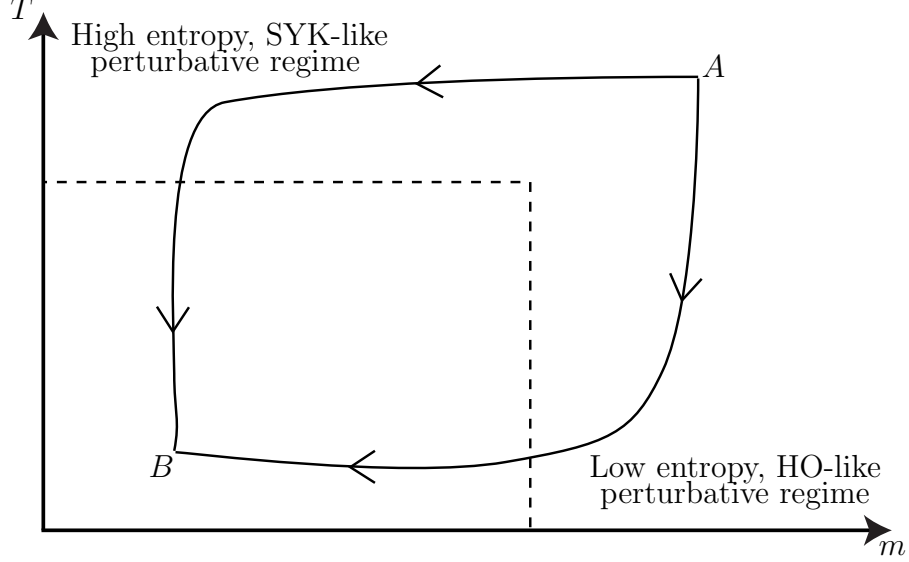


Figure 6: Non-trivial monodromy in parameter space associated with two paths along which the weakly coupled solution at A is analytically continued to B . Units are such that $\lambda = |\xi| = 1$ and the dashed line represents the separation between the weakly coupled and the strongly coupled regions. When going first through the high T , high entropy perturbative state and then to the low T , non-perturbative regime (upper path), it is natural to get a solution *à la* SYK at point B . When going first through the high mass, low T , low entropy, harmonic oscillator-like perturbative state (lower path), one may obtain a completely different, low entropy solution at B [28].

3.1.4 Reparameterization invariance and standard low energy ansatz

The $U(n)^2$ model Let us look at the $U(n)^2$ model in which the term $h_{\text{II},1}$ is being turned on as well (including other pillow vertices would change the discussion only in a trivial way). The Schwinger-Dyson equations read

$$\frac{1}{G_k} = -i\nu_k + m + \Sigma_k, \quad (3.33)$$

$$\Sigma(\tau) = \lambda^2 G(\tau)^2 G(-\tau) - \kappa G(0^-) \delta_\beta(\tau). \quad (3.34)$$

If we define

$$\hat{\Sigma}(\tau_1, \tau_2) = \Sigma(\tau_1 - \tau_2) + m\delta_\beta(\tau_1 - \tau_2), \quad G(\tau_1, \tau_2) = G(\tau_1 - \tau_2) \quad (3.35)$$

and we neglect the term $i\nu_k$ and the ultraviolet δ -function contributions on the right-hand side of (3.33) and (3.34), we can write the low energy (or long time) limit of the equations in the convenient form

$$\int_0^\beta G(\tau_1, \tau) \hat{\Sigma}(\tau, \tau_2) d\tau = \delta_\beta(\tau_1 - \tau_2), \quad (3.36)$$

$$\hat{\Sigma}(\tau_1, \tau_2) = \lambda^2 G(\tau_1, \tau_2)^2 G(\tau_2, \tau_1). \quad (3.37)$$

These equations are covariant under the transformations

$$G'(\tau'_1, \tau'_2) = \left(\frac{d\tau_1}{d\tau'_1}\right)^\Delta \left(\frac{d\tau_2}{d\tau'_2}\right)^\Delta G(\tau_1, \tau_2), \quad (3.38)$$

$$\hat{\Sigma}'(\tau'_1, \tau'_2) = \left(\frac{d\tau_1}{d\tau'_1}\right)^{1-\Delta} \left(\frac{d\tau_2}{d\tau'_2}\right)^{1-\Delta} \hat{\Sigma}(\tau_1, \tau_2), \quad (3.39)$$

for $\Delta = 1/4$ and where $\tau'(\tau)$ is an arbitrary diffeomorphism of $S_\beta^1 \rightarrow S_{\beta'}^1$,

$$\frac{d\tau'}{d\tau} > 0, \quad \tau'(\tau + \beta) = \tau'(\tau) + \beta'. \quad (3.40)$$

Invariance of the equations is achieved for diffeomorphisms such that $\beta' = \beta$. The equations are also invariant under the local multiplications

$$G'(\tau_1, \tau_2) = \frac{g(\tau_2)}{g(\tau_1)} G(\tau_1, \tau_2), \quad \hat{\Sigma}'(\tau_1, \tau_2) = \frac{g(\tau_2)}{g(\tau_1)} \hat{\Sigma}(\tau_1, \tau_2), \quad (3.41)$$

where g is an arbitrary non-vanishing function of period β . The emergence of the reparameterization symmetry (3.38)-(3.39) is a remarkable feature of the low energy limit of Schwinger-Dyson equations resumming melonic graphs, consistent with a dual gravitational description of the models [6, 9]. The invariance (3.41) is a local Euclidean version of the global fermion number symmetry of (2.10). The fact that a global symmetry becomes local is also beautifully consistent with holographic ideas.

The full low energy reparameterization symmetry (3.38)-(3.39) will always be spontaneously broken. In particular, time translation invariance implies that the physical two-point functions depend only on $\tau_1 - \tau_2$, a property which is of course not invariant under arbitrary reparameterizations. Moreover, Eq. (3.36) involves an integral over all time scales. Since it is a low energy equation, a UV cutoff is typically needed to avoid UV divergences. In practice, the reparameterization symmetry is thus also explicitly broken by this cutoff.

A fundamental property of the models, first discussed in [11], is that some solutions can preserve a $SL(2, \mathbb{R})$ subgroup of reparameterizations, or more generally a $SL(2, \mathbb{R})$

subgroup of the full set of transformations (3.38), (3.39) and (3.41) [20], at low energy. For the $U(n)^2$ model, the standard zero-temperature ansatz is

$$G(\tau) \sim \begin{cases} \frac{b_+}{|\tau|^{2\Delta}} & \text{if } \tau \rightarrow +\infty \\ -\frac{b_-}{|\tau|^{2\Delta}} & \text{if } \tau \rightarrow -\infty. \end{cases} \quad (3.42)$$

Instead of b_+ and b_- , it is often convenient to use $C > 0$ and θ defined by

$$b_+ = \frac{C}{\sqrt{\pi}} \sin(\pi/4 + \theta), \quad b_- = \frac{C}{\sqrt{\pi}} \sin(\pi/4 - \theta). \quad (3.43)$$

Note that (3.15) implies that $b_{\pm} > 0$ and thus $|\theta| < \pi/4$. Computing the Fourier transform $\tilde{G}$ of G , Eq. (3.42) and (3.43) can be rewritten as

$$\tilde{G}(\omega) \underset{\omega \rightarrow 0^+}{\sim} \frac{i}{\omega^{1-2\Delta}} C e^{-i\theta}. \quad (3.44)$$

In terms of the spectral function introduced in Sec. 3.1.1, this is equivalent to

$$\rho(\omega) \underset{\omega \rightarrow 0^{\pm}}{\sim} \frac{C}{\pi \sqrt{|\omega|}} \sin(\pi/4 \pm \theta). \quad (3.45)$$

Plugging the above ansatz in (3.36) and (3.37) yields in the standard way

$$\Delta = \frac{1}{4}, \quad (3.46)$$

and

$$C^4 \cos(2\theta) = \frac{\pi}{\lambda^2}. \quad (3.47)$$

One low energy parameter remains unconstrained, for example the angle θ . This parameter is naturally related, via a complicated non-perturbative formula, to the mass parameter m in the Hamiltonian, possibly renormalized by the pillow couplings. When $m = 0$ the model recovers a large N particle-hole symmetry which yields $G(-\tau) = -G(\tau)$, as in the standard SYK model, and thus $\theta = 0$.

Note that (3.36) is equivalent to

$$\tilde{\Sigma}(\omega) = \frac{1}{\tilde{G}(\omega)} \underset{\omega \rightarrow 0^+}{\sim} -\frac{ie^{i\theta}}{C} \omega^{1-2\Delta}, \quad (3.48)$$

and thus implies in particular that $\tilde{\Sigma}(0) = 0$. Using the first equation in (3.35), we see that the Fourier transform of the self-energy Σ at zero frequency is constrained to satisfy

$$\tilde{\Sigma}(0) = -m \quad (3.49)$$

for the $\text{SL}(2, \mathbb{R})$ -preserving solution. Satisfying this condition is therefore an indicator that the solution is indeed conformal in the IR.

The scaling form of the two-point function at non-zero temperature can then be obtained from the zero-temperature ansatz by using the mapping $\tau = \tan(\pi\tau'/\beta)$ and (3.38) to go from a solution defined on the real line ($\beta = +\infty$) to a solution defined on the interval $[-\beta/2, \beta/2]$.⁹ The resulting solution does not satisfy the antiperiodicity $G(-\beta/2) = -G(\beta/2)$, but this can be corrected by performing an additional transformation of the form (3.41), with a function $g(\tau) = (b_+/b_-)^{\tau/\beta}$. This yields

$$G(\tau) = \left(\frac{\pi}{\beta}\right)^{1/2} \frac{e^{-2\pi a\tau/\beta}}{\sqrt{\sin(\pi|\tau|/\beta)}} \left(b_+ \Theta(\tau) - b_- \Theta(-\tau)\right), \quad (3.50)$$

with a damping factor

$$a = \frac{1}{2\pi} \ln \frac{b_+}{b_-} = \frac{1}{2\pi} \ln \tan(\pi/4 + \theta). \quad (3.51)$$

Eq. (3.50) provides a good approximation to the two-point function in the scaling regime $\beta\lambda \gg 1$, with the additional condition $|\tau| \gg 1/\lambda$ which excises the UV region. By taking the Fourier transform, we get the Matsubara-Fourier coefficients

$$G_k = i \left(\frac{\beta}{2\lambda}\right)^{1/2} \frac{e^{-i\theta}}{(\pi \cos(2\theta))^{1/4}} \frac{\Gamma(1/4 + k + ia)}{\Gamma(3/4 + k + ia)}, \quad (3.52)$$

which are valid when $\beta\lambda \gg 1$ and $|\nu_k| \ll \lambda$, as well as

$$\rho(\omega) = \left(\frac{\beta}{2\lambda}\right)^{1/2} \frac{1}{(\pi^5 \cos(2\theta))^{1/4}} \text{Re} \left[e^{-i\theta} \frac{\Gamma(1/4 + ia - i\frac{\beta\omega}{2\pi})}{\Gamma(3/4 + ia - i\frac{\beta\omega}{2\pi})} \right], \quad (3.53)$$

which is valid when $\beta\lambda \gg 1$ and $|\omega| \ll \lambda$.

U(n) model A very similar analysis can be performed for the $U(n)$ model using (3.31) instead of (3.30). At low energy, the reparameterization symmetry (3.38), (3.39) is still valid, but (3.41) does not hold anymore, consistently with the fact that the fermion number is no longer conserved. The low energy ansatz (3.42), (3.43) can still be used, and plugging it into the Schwinger-Dyson equations we find that the only possibility with $b_{\pm} \neq 0$ corresponds to

$$\Delta = \frac{1}{4}, \quad \theta = 0, \quad C^4 = \frac{\pi}{|\xi|^2}, \quad b_+^4 = \frac{1}{4\pi|\xi|^2}. \quad (3.54)$$

⁹The particular mapping $\tau = \tan(\pi\tau'/\beta)$ from S^1_β to the real line is singled out by the fact that it is consistent with time translation invariance. We would like to thank Paolo Gregori for providing this argument.

Important remark: The existence of a consistent low energy ansatz solving the low energy Schwinger-Dyson equations does not ensure that there exists a physical solution to the full Schwinger-Dyson equations whose low energy limit matches the low energy ansatz. In other words, it may not always be possible to UV complete solutions that seem to be consistent at low energies. Sometimes this occurs for very elementary reasons. For instance, the low energy ansatz (3.42) can *a priori* solve the low energy Schwinger-Dyson equations for any mass parameter m in the Hamiltonian. But, at least in the perturbative regime $m \gg \lambda$, we know that such a solution cannot be UV completed: the perturbative series converges in this regime and the solution must be very near the harmonic oscillator which has a trivial IR behaviour. The lack of a UV completion is more subtle and has important consequences in the context of purely bosonic models [5, 58, 59]; see also the discussion in Sec. 3.3.2.

3.1.5 A new low energy ansatz

The standard ansatz (3.42) for the long time behaviour of the zero-temperature two-point function can be generalized to

$$G(\tau) \sim \begin{cases} \frac{b_+}{|\tau|^{2\Delta_+}} & \text{if } \tau \rightarrow +\infty \\ -\frac{b_-}{|\tau|^{2\Delta_-}} & \text{if } \tau \rightarrow -\infty, \end{cases} \quad (3.55)$$

where the exponents Δ_+ and Δ_- are *a priori* distinct. This possibility may seem exotic and, to the best of our knowledge, has never been considered in the literature so far. However, in models for which the particle-hole symmetry is broken it is a plausible and mathematically consistent possibility. One salient new result of the present paper will be to provide strong numerical evidence that this behaviour is actually realized in the $U(n)$ model.

An important new feature of the $\Delta_+ \neq \Delta_-$ case is that the two terms on the right-hand side of (3.31) scale differently at $\tau = +\infty$ and $\tau = -\infty$. Assuming, for example, that $\Delta_+ < \Delta_-$, Eq. (3.31) implies that¹⁰

$$\Sigma(\tau) \sim \begin{cases} -\frac{|\xi|^2}{4} G(\tau)^3 & \text{if } \tau \rightarrow +\infty \\ -\frac{3|\xi|^2}{4} G(\tau) G(-\tau)^2 & \text{if } \tau \rightarrow -\infty. \end{cases} \quad (3.56)$$

The $\tau \rightarrow +\infty$ term then dominates in the Fourier transform at low energy. Using the Schwinger-Dyson equations (3.36) and (3.56) as before, and keeping the leading terms at low energy, we get

$$\Delta_+ = \frac{1}{4}, \quad b_+^4 = \frac{2}{\pi|\xi|^2}, \quad (3.57)$$

¹⁰Of course, the case $\Delta_- < \Delta_+$ can be discussed along the same lines.

with no further constraints on Δ_- and b_- . In passing, we note that the relation (3.49) must also be satisfied by such a solution.

Let us remark that both the standard (3.42) and the more general (3.55) ansätze spontaneously break the reparameterization symmetry (3.38)-(3.39). However, the scale symmetry, which is preserved by (3.42), is spontaneously broken by (3.55) when $\Delta_+ \neq \Delta_-$. This is an important qualitative difference between the two classes of solutions.

It is natural to seek a finite temperature solution generalizing (3.50) using the reparameterization covariance of the Schwinger-Dyson equations. However, the usual method does not work. First, the invariance under local multiplications (3.41), which is crucially used in the $U(n)^2$ model to obtain a finite temperature two-point function with the correct antiperiodicity property, is no longer valid in the $U(n)$ model. Second, and even more crucially, the mapping $\tau = \tan(\pi\tau'/\beta)$ yields a two point function that is not time-translation invariant, except if $\Delta_+ = \Delta_- = 1/4$.¹¹ Because of these difficulties, we shall limit ourselves to the zero-temperature formula (3.55) when $\Delta_+ \neq \Delta_-$.

On the practical side, it is interesting to observe that, even though the exponent Δ_+ in the case $\Delta_+ < \Delta_-$ and the exponent Δ in the case of the standard ansatz $\Delta = \Delta_+ = \Delta_-$ are both equal to $1/4$, the two situations can be sharply distinguished even if we look only at the $\tau \rightarrow +\infty$ asymptotics of the two-point functions. Indeed, there is a factor of 8 discrepancy for b_+^4 between (3.54) and (3.57).

3.2 Effective action and thermodynamics

3.2.1 Disordered model formulation

In the case of the disordered models, one can compute the Euclidean finite temperature path integral for the Hamiltonian (2.13) straightforwardly. This is a completely standard calculation and thus we shall be brief.

Since the annealed and quenched averages coincide at leading order, as already mentioned in Sec. 2.2, we first integrate over the disorder with the Gaussian weight (2.16). This produces a vector model with bilocal interaction terms which can be treated in the standard way [34]. We introduce two bilocal auxiliary fields $\mathbf{G}$ and Σ , where Σ plays the role of a Lagrange multiplier enforcing the constraint $\mathbf{G}(\tau_1, \tau_2) =$

¹¹One can actually check that no diffeomorphism from the real line to S^1_β is compatible with time translation invariance if $\Delta_+ \neq \Delta_-$.

$\chi^i(\tau_1)\bar{\chi}_i(\tau_2)$. Integrating out the complex fermions yields the effective action

$$\begin{aligned} \frac{1}{N}S_{\text{eff}}[\mathbf{G}, \Sigma] = & -\ln(1 + e^{-\beta m}) - \ln \frac{\det \mathcal{O}_\Sigma}{\det \mathcal{O}_0} + \frac{1}{2}\kappa\beta\mathbf{G}(0^-)^2 + \int_{[0,\beta]^2} \left[\Sigma(\tau_2, \tau_1)\mathbf{G}(\tau_1, \tau_2) \right. \\ & \left. - \frac{1}{4}\lambda^2\mathbf{G}(\tau_1, \tau_2)^2\mathbf{G}(\tau_2, \tau_1)^2 + \frac{1}{4}\xi^2\mathbf{G}(\tau_1, \tau_2)^3\mathbf{G}(\tau_2, \tau_1) \right] d\tau_1 d\tau_2, \end{aligned} \quad (3.58)$$

where the operators $\mathcal{O}_\Sigma$ and $\mathcal{O}_0$, which act on antiperiodic functions, are defined by

$$\begin{aligned} \mathcal{O}_\Sigma(\tau_1, \tau_2) &= \delta_\beta(\tau_1 - \tau_2) \left[\frac{\partial}{\partial \tau_2} + m \right] + \Sigma(\tau_1, \tau_2), \\ \mathcal{O}_0(\tau_1, \tau_2) &= \delta_\beta(\tau_1 - \tau_2) \left[\frac{\partial}{\partial \tau_2} + m \right]. \end{aligned} \quad (3.59)$$

Note that, formally, the tree-level terms $-\ln(1 + e^{-\beta m})$ and $\ln \det \mathcal{O}_0$ in (3.58) cancel each other. However, we prefer to use the rigorous formula (3.58), which does not have UV divergences and takes into account that only the ratio of determinants $\det \mathcal{O}_\Sigma / \det \mathcal{O}_0$ is actually well defined.

The large N limit is governed by the saddle points of (3.58). Varying this effective action with respect to Σ and $\mathbf{G}$, denoting the saddle point values of Σ and $\mathbf{G}$ by Σ and G respectively, and using time translation invariance, we get

$$\frac{1}{G_k} = -i\nu_k + m + \Sigma_k, \quad (3.60)$$

$$\Sigma(\tau) = \lambda^2 G(\tau)^2 G(-\tau) - \frac{|\xi|^2}{4} (G(\tau)^3 + 3G(\tau)G(-\tau)^2) - \kappa G(0^-) \delta_\beta(\tau). \quad (3.61)$$

These equations generalize the Schwinger-Dyson equations written down in (3.26), (3.30), (3.31), (3.33) and (3.34) to the case where the couplings λ , ξ and κ are turned on simultaneously.¹² In particular, and not surprisingly, the on-shell auxiliary field G coincides with the Euclidean two-point function.

Moreover, by construction the on-shell value of the effective action,

$$S_{\text{eff}} = S_{\text{eff}}(G, \Sigma), \quad (3.62)$$

yields the free energy F

$$\beta F = S_{\text{eff}}. \quad (3.63)$$

From the free energy one can then obtain the entropy and the energy by using the standard thermodynamical identities,

$$S = -\frac{\partial F}{\partial T}, \quad E = F + TS. \quad (3.64)$$

¹²As explained in Sec. 2.1.3, the Schwinger-Dyson equations are more complicated in the case of the matrix model when λ and ξ are both turned on. In the present subsection, we are dealing with the simpler case of the disordered model (2.13).

The above results are very powerful: they imply that the knowledge of the Euclidean two-point function only is sufficient to derive all the thermodynamical properties of the model. In practice, we solve the Schwinger-Dyson equations numerically as described in Sec. 3.3, and plug the result into (3.58) to compute F , S and E straightforwardly. Very explicit formulas are given in Sec. 3.2.3 below.

3.2.2 Matrix model formulation

The case of the matrix models is a bit more subtle to deal with, because there is no longer an auxiliary field method available to compute the large N path integral. It is thus not immediately obvious how to compute the free energy. In particular, it is unclear whether an effective action like (3.58) could still be relevant. The aim of this subsection is to present a detailed derivation of the free energy and to carefully justify the use of effective actions like (3.58) in the context of matrix models.¹³

We explained in Sec. 2.1.3 that turning on both couplings λ and ξ in the matrix model yields complications due to non-vanishing fermion-number-violating two-point functions. For simplicity and because we won't need the more general results in the rest of this paper, we are then going to focus on the models (2.10) and (2.11). Note however that our method can be generalised straightforwardly to the models in which both couplings are turned on simultaneously. Note also that, as usual, pillow couplings may be turned on quite trivially by using the auxiliary field method.

The derivations for (2.10) and (2.11) are extremely similar. We thus present the derivation in parallel for both models.

Step one The Heisenberg Euclidean equations of motion follow from (3.2) and read

$$\frac{d\mathcal{O}_E}{d\tau} = \dot{\mathcal{O}}_E = [H, \mathcal{O}_E]. \quad (3.65)$$

In the $U(n)^2$ model, this yields

$$\dot{\psi}_\mu = [H_{U(n)^2}, \psi_\mu] = -\left(m - \frac{\lambda}{2\sqrt{d}}\right)\psi_\mu - \lambda\sqrt{d}\psi_\nu\psi_\mu^\dagger\psi_\nu. \quad (3.66)$$

Contracting with $\psi_\mu^\dagger$ on both sides and taking the trace we get

$$\left(\frac{d}{d\tau} + m - \frac{\lambda}{2\sqrt{d}}\right)G(\tau) = \begin{cases} -\frac{\lambda\sqrt{d}}{n} \langle \text{tr}[(\psi_\mu\psi_\nu^\dagger\psi_\mu)(\tau)\psi_\nu^\dagger] \rangle_\beta & \text{if } \tau > 0 \\ \frac{\lambda\sqrt{d}}{n} \langle \text{tr}[\psi_\mu^\dagger(\psi_\nu\psi_\mu^\dagger\psi_\nu)(\tau)] \rangle_\beta & \text{if } \tau < 0. \end{cases} \quad (3.67)$$

¹³The arguments apply to tensor models as well, see also [60].

In particular,

$$\dot{G}(0^+) = -\left(m - \frac{\lambda}{2\sqrt{d}}\right)G(0^+) - \frac{\lambda\sqrt{d}}{n} \langle \text{tr } \psi_\mu \psi_\nu^\dagger \psi_\mu \psi_\nu^\dagger \rangle_\beta, \quad (3.68)$$

$$\dot{G}(0^-) = -\left(m - \frac{\lambda}{2\sqrt{d}}\right)G(0^-) + \frac{\lambda\sqrt{d}}{n} \langle \text{tr } \psi_\mu^\dagger \psi_\nu \psi_\mu^\dagger \psi_\nu \rangle_\beta. \quad (3.69)$$

Let us note *en passant* that, subtracting the above two equations and using (3.20) and the canonical anticommutation relations to evaluate

$$\text{tr}(\psi_\mu \psi_\nu^\dagger \psi_\mu \psi_\nu^\dagger + \psi_\mu^\dagger \psi_\nu \psi_\mu^\dagger \psi_\nu) = \frac{n}{d}, \quad (3.70)$$

we get

$$\dot{G}(0^-) - \dot{G}(0^+) = m + \frac{\lambda}{2\sqrt{d}}. \quad (3.71)$$

This together with (3.19) yields an exact formula for the first moment of the spectral function, generalising (3.21) at finite n and d .

In the $U(n)$ model, we have¹⁴

$$\dot{\psi}_\mu = [H_{U(n)}, \psi_\mu] = -m\psi_\mu - \frac{\xi}{2}\sqrt{d}\psi_\nu\psi_\mu\psi_\nu - \frac{\xi^*}{2}\sqrt{d}(\psi_\nu^\dagger\psi_\mu^\dagger\psi_\nu - \psi_\nu^\dagger\psi_\mu\psi_\nu^\dagger + \psi_\nu\psi_\mu^\dagger\psi_\nu^\dagger). \quad (3.72)$$

Massaging a bit the traces, we get

$$\begin{aligned} \dot{G}(0^+) + mG(0^+) &= \dot{G}(0^-) + mG(0^-) \\ &= \frac{\xi\sqrt{d}}{2n} \langle \text{tr } \psi_\mu^\dagger \psi_\nu \psi_\mu \psi_\nu \rangle_\beta + \frac{3\xi^*\sqrt{d}}{2n} \langle \text{tr } \psi_\mu^\dagger \psi_\nu^\dagger \psi_\mu^\dagger \psi_\nu \rangle_\beta. \end{aligned} \quad (3.73)$$

In particular, in the $U(n)$ model

$$\dot{G}(0^-) - \dot{G}(0^+) = m, \quad (3.74)$$

and Eq. (3.21) is valid even at finite n and d .

Step two The partition function Z and free energy F are defined in general by

$$Z = \text{Tr } e^{-\beta H} = e^{-\beta F}, \quad (3.75)$$

where Tr traces over the Hilbert space of states. Then

$$\frac{\partial F}{\partial m} = nd \langle \text{tr } \psi_\mu^\dagger \psi_\mu \rangle = -n^2 d G(0^-), \quad (3.76)$$

$$\frac{\partial F}{\partial \lambda} = \frac{1}{2} n d^{3/2} \langle \text{tr } \psi_\mu^\dagger \psi_\nu \psi_\mu^\dagger \psi_\nu \rangle_\beta, \quad (3.77)$$

¹⁴At finite n and d , we use the tracelessness condition $\text{tr } \psi_\mu = 0$ to simplify the formulas. This plays no role at infinite n .

in the $U(n)^2$ model and

$$\frac{\partial F}{\partial m} = nd \langle \text{tr } \psi_\mu^\dagger \psi_\mu \rangle = -n^2 d G(0^-), \quad (3.78)$$

$$\frac{\partial F}{\partial \xi} = \frac{1}{2} n d^{3/2} \langle \text{tr } \psi_\mu^\dagger \psi_\nu \psi_\mu \psi_\nu \rangle_\beta, \quad (3.79)$$

$$\frac{\partial F}{\partial \xi^*} = \frac{1}{2} n d^{3/2} \langle \text{tr } \psi_\mu^\dagger \psi_\nu^\dagger \psi_\mu^\dagger \psi_\nu \rangle_\beta, \quad (3.80)$$

in the $U(n)$ model. From (3.68) and (3.73), neglecting the subleading $1/\sqrt{d}$ term in the $U(n)^2$ model, we thus get

$$\dot{G}(0^-) + mG(0^-) = \dot{G}(0^+) + mG(0^+) = \begin{cases} \frac{2\lambda}{n^2 d} \frac{\partial F}{\partial \lambda} & (U(n)^2 \text{ model}) \\ \frac{1}{n^2 d} \left(\xi \frac{\partial F}{\partial \xi} + 3\xi^* \frac{\partial F}{\partial \xi^*} \right) & (U(n) \text{ model}). \end{cases} \quad (3.81)$$

In particular, since G is real, in the $U(n)$ model we must have

$$\xi \frac{\partial F}{\partial \xi} = \xi^* \frac{\partial F}{\partial \xi^*}, \quad (3.82)$$

or, equivalently,

$$F(\xi, \xi^*) = F(|\xi|). \quad (3.83)$$

This result is obvious from the structure of the generalized melons in the large N limit, but we see here that it is valid even at finite n and d . In the $U(n)$ model, we can thus rewrite (3.81) as

$$\dot{G}(0^-) + mG(0^-) = \dot{G}(0^+) + mG(0^+) = \frac{4\xi}{n^2 d} \frac{\partial F}{\partial \xi} = \frac{4\xi^*}{n^2 d} \frac{\partial F}{\partial \xi^*}. \quad (3.84)$$

Finally, note that, by definition of the self-energy (3.26), we can also write

$$\dot{G}(0^-) + mG(0^-) = \dot{G}(0^+) + mG(0^+) = - \int_0^\beta G(\tau) \Sigma(-\tau) d\tau. \quad (3.85)$$

From the point of view of the Fourier transform, the first equality in the above equation follows from the fact that both G_k and Σ_k are proportional to $1/k$ at large k and thus the series $\sum_{k \in \mathbb{Z}+1/2} G_k \Sigma_k$ converges.

Step three Define the functionals

$$\begin{aligned} \frac{1}{n^2 d} S_{\text{eff}}^{\text{U}(n)^2}[\mathbf{G}, \Sigma] &= -\ln(1 + e^{-\beta m}) - \ln \frac{\det \mathcal{O}_\Sigma}{\det \mathcal{O}_0} + \int_{[0, \beta]^2} \left[\Sigma(\tau_2, \tau_1) \mathbf{G}(\tau_1, \tau_2) \right. \\ &\quad \left. - \frac{1}{4} \lambda^2 \mathbf{G}(\tau_1, \tau_2)^2 \mathbf{G}(\tau_2, \tau_1)^2 \right] d\tau_1 d\tau_2, \end{aligned} \quad (3.86)$$

$$\begin{aligned} \frac{1}{n^2 d} S_{\text{eff}}^{\text{U}(n)}[\mathbf{G}, \Sigma] &= -\ln(1 + e^{-\beta m}) - \ln \frac{\det \mathcal{O}_\Sigma}{\det \mathcal{O}_0} + \int_{[0, \beta]^2} \left[\Sigma(\tau_2, \tau_1) \mathbf{G}(\tau_1, \tau_2) \right. \\ &\quad \left. + \frac{1}{4} |\xi|^2 \mathbf{G}(\tau_1, \tau_2)^3 \mathbf{G}(\tau_2, \tau_1) \right] d\tau_1 d\tau_2, \end{aligned} \quad (3.87)$$

where $\mathbf{G}$ and Σ are antiperiodic functions of their arguments and the operator $\mathcal{O}_\Sigma$ is defined as in (3.59). Equations (3.86) and (3.87) are, at this point, nothing more than mathematical definitions. They are constructed in such a way that the saddle point equations obtained by varying $S_{\text{eff}}^{\text{U}(n)^2}$ and $S_{\text{eff}}^{\text{U}(n)}$ with respect to Σ and $\mathbf{G}$ do match the correct Schwinger-Dyson equations (3.26), (3.30) and (3.31) of the $\text{U}(n)^2$ and $\text{U}(n)$ models respectively. There is no reference to a path integral argument here: we simply use the matrix model Schwinger-Dyson equations, which are derived diagrammatically, and the mathematical statement that varying the functionals (3.86), (3.87) yields precisely these equations.

Let us denote by Σ and G the on-shell values of Σ and $\mathbf{G}$, *i.e.* solutions to (3.26) and (3.30) or (3.31). Using the fact that the functionals (3.86) and (3.87) are stationary when evaluated at $\mathbf{G} = G$ and $\Sigma = \Sigma$, we get

$$\frac{\partial S_{\text{eff}}^{\text{U}(n)^2}}{\partial \lambda} = -\frac{\lambda \beta}{2} \int_0^\beta G(\tau)^2 G(-\tau)^2 d\tau, \quad (3.88)$$

$$\frac{\partial S_{\text{eff}}^{\text{U}(n)}}{\partial \xi} = \frac{\xi^* \beta}{4} \int_0^\beta G(\tau)^3 G(-\tau) d\tau, \quad (3.89)$$

which, using the Schwinger-Dyson equations (3.30) or (3.31), as well as the antiperiodicity of G , can be rewritten as

$$\lambda \frac{\partial S_{\text{eff}}^{\text{U}(n)^2}}{\partial \lambda} = -\frac{\beta}{2} \int_0^\beta G(\tau) \Sigma(-\tau) d\tau, \quad (3.90)$$

$$\xi \frac{\partial S_{\text{eff}}^{\text{U}(n)}}{\partial \xi} = -\frac{\beta}{4} \int_0^\beta G(\tau) \Sigma(-\tau) d\tau. \quad (3.91)$$

Comparing with (3.81), (3.84) and (3.85), we get

$$\frac{\partial S_{\text{eff}}^{\text{U}(n)^2}}{\partial \lambda} = \beta \frac{\partial F}{\partial \lambda} \quad \text{and} \quad \frac{\partial S_{\text{eff}}^{\text{U}(n)}}{\partial \xi} = \beta \frac{\partial F}{\partial \xi}, \quad (3.92)$$

for the $\text{U}(n)^2$ and $\text{U}(n)$ models, respectively.

Step four Integrating (3.92) and fixing the constant of integration by using the fact that at zero coupling both models are simply a collection of $N = n^2 d$ decoupled fermionic oscillators, we obtain (3.63). This is exactly as in the case of the disordered model formulation, where S_{eff} is replaced by $S_{\text{eff}}^{\text{U}(n)^2}$ or $S_{\text{eff}}^{\text{U}(n)}$ for the $\text{U}(n)^2$ or $\text{U}(n)$ models, according to which couplings are turned on.

3.2.3 Useful formulas

Numerically, the solution of the Schwinger-Dyson equations is most naturally encoded in the Matsubara-Fourier coefficients G_k (or Σ_k). It is thus very convenient to express all the relevant physical quantities, such as fermion number, free energy, energy and entropy, as convergent sums over these coefficients.

Fermion number The fermion number (2.2) is given in terms of the two-point function

$$Q = -G(0^-) = 1 - G(0^+). \quad (3.93)$$

One cannot use immediately the expansion (3.4) to evaluate $G(0^\pm)$ because $\sum_k G_k$ diverges. This is related to the fact that the function G is discontinuous at $\tau = 0$. To go around this problem, one can use Dirichlet's trick of considering $G(\tau) + G(-\tau)$, which is continuous at $\tau = 0$ and has a convergent Fourier expansion. This yields

$$G(0^+) + G(0^-) = \frac{2}{\beta} \sum_{k \in \mathbb{Z} + \frac{1}{2}} \text{Re } G_k. \quad (3.94)$$

Using then (3.93), (3.20) and (3.14), we get

$$Q = \frac{1}{2} - \frac{2}{\beta} \sum_{k \in \mathbb{N} + \frac{1}{2}} \text{Re } G_k. \quad (3.95)$$

Subtracting the tree-level result, with tree-level coefficients $G_k^{(0)}$ given in (3.6), we also get

$$Q = n_{\text{F}}(m) - \frac{2}{\beta} \sum_{k \in \mathbb{N} + \frac{1}{2}} \text{Re}(G_k - G_k^{(0)}). \quad (3.96)$$

Formulas (3.95) and (3.96) are of course mathematically equivalent, but (3.96) is more convenient to use in numerical evaluations because the infinite series converges more quickly. This is due to the fact that $\text{Re } G_k$ and $\text{Re}(G_k - G_k^{(0)})$ decay like $1/k^2$ and $1/k^4$ respectively at large k . We shall repeatedly use a similar rewriting of the formulas in the following, separating the tree-level result and the quantum corrections.

There is also an interesting formula for the fermion number in terms of the spectral density, which is obtained from (3.93) and (3.12),

$$Q = \int_{-\infty}^{+\infty} n_F(\omega) \rho(\omega) d\omega. \quad (3.97)$$

This formula coincides with the formula for a collection of *decoupled* fermionic harmonic oscillators with a frequency distribution given by the spectral density ρ .

Free energy The free energy is obtained by evaluating the effective actions (3.86) or (3.87) on the solution for the two-point function and using (3.63). It is convenient to use the Schwinger-Dyson equations (3.30) or (3.31) to simplify the result. We get in this way

$$\frac{F}{N} = m - T \ln(1 + e^{\beta m}) + T \sum_{k \in \mathbb{N} + \frac{1}{2}} \left[\ln \left| \frac{G_k}{G_k^{(0)}} \right|^2 + \frac{3}{2} \left(1 - \operatorname{Re} \frac{G_k}{G_k^{(0)}} \right) \right]. \quad (3.98)$$

Note that this formula is valid for both the $U(n)^2$ and the $U(n)$ models.

Energy The energy is simply the expectation value of the Hamiltonian,

$$E = \langle H \rangle_\beta. \quad (3.99)$$

From the definitions (2.10) and (2.11), Eq. (3.81), (3.82) and (3.93), we get

$$\frac{E}{N} = mQ + \frac{1}{2} (\dot{G}(0^+) + mG(0^+)), \quad (3.100)$$

which is valid for both the $U(n)^2$ and the $U(n)$ models. The right-hand side of (3.100) can be evaluated using (3.95) and (3.85), yielding

$$\frac{E}{N} = \frac{m}{2} - \frac{1}{\beta} \sum_{k \in \mathbb{N} + \frac{1}{2}} \left(2m \operatorname{Re} G_k + 1 - \operatorname{Re} \frac{G_k}{G_k^{(0)}} \right). \quad (3.101)$$

Subtracting the tree-level result yields the equivalent formula

$$\frac{E}{N} = m n_F(m) - \frac{1}{\beta} \sum_{k \in \mathbb{N} + \frac{1}{2}} \left(2m \operatorname{Re}(G_k - G_k^0) + 1 - \operatorname{Re} \frac{G_k}{G_k^{(0)}} \right), \quad (3.102)$$

to be used in the numerical evaluations.

It is interesting to express the energy in terms of the spectral density, using (3.100) and (3.12) together with (3.20) and (3.21). We obtain

$$\frac{E}{N} = \frac{1}{2} \int_{-\infty}^{+\infty} (m + \omega) n_F(\omega) \rho(\omega) d\omega. \quad (3.103)$$

Note that, unlike the case of the fermion number (3.97), this formula differs slightly from the formula for a collection of decoupled fermionic harmonic oscillators, by the replacement of the frequency factor ω by $\frac{1}{2}(m + \omega)$.

Entropy The entropy is then given by the standard relation

$$S = \beta(E - F). \quad (3.104)$$

We have also checked explicitly that the result so obtained is consistent with the first relation in (3.64).

Zero temperature in the $U(n)^2$ model One can derive very interesting exact formulas at zero temperature in the $U(n)^2$ model [13, 19] when the low energy ansatz discussed in Sec. 3.1.4 is valid. The charge at zero temperature can be expressed in terms of the low energy parameter θ defined in (3.43) as

$$Q(T = 0) = \frac{1}{2} - \frac{\theta}{\pi} - \frac{1}{4} \sin(2\theta). \quad (3.105)$$

Moreover, if we denote by σ the zero-temperature entropy,

$$\sigma(m, \lambda) = \frac{1}{N} \lim_{T \rightarrow 0} S(T, m, \lambda), \quad (3.106)$$

one finds in the case $m = 0$

$$\sigma(m = 0) = \ln 2 - \int_0^{\pi/4} \left(1 - \frac{2\theta}{\pi}\right) \tan \theta \, d\theta = \frac{\text{Catalan}}{\pi} + \frac{1}{4} \ln 2 \simeq 0.46, \quad (3.107)$$

where

$$\text{Catalan} = \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k+1)^2} \quad (3.108)$$

is the so-called Catalan constant. When $m \neq 0$, it is useful to consider σ as a function of the zero-temperature charge $Q(T = 0, m)$ instead of m . One can then derive the elegant formula

$$\frac{\partial \sigma}{\partial Q} = \ln \frac{b_+}{b_-}. \quad (3.109)$$

Using (3.43), (3.105) and (3.107) as an initial condition,¹⁵ and after a little massaging, this equation can be integrated to

$$\sigma(Q) = N \int_{\theta(Q)}^{\pi/4} \left(\frac{1}{\pi} + \frac{1}{2} \cos(2\theta') \right) \ln \frac{\cos \theta' + \sin \theta'}{\cos \theta' - \sin \theta'} \, d\theta', \quad (3.110)$$

¹⁵One may also use $\sigma(Q = 0) = 0$ as an initial condition to integrate (3.109). This yields the correct answer, but is marred by an important subtlety. Indeed, as explained in [28] and reviewed in Sec. 4, the low energy ansatz discussed in Sec. 3.1.4, which is at the basis of the derivation of (3.109), seems to be valid only when the zero-temperature charge Q is greater than a strictly positive critical value.

or more explicitly

$$\begin{aligned} \sigma(\theta) = & \left(\frac{1}{2} - \frac{\theta}{\pi} - \frac{1}{4} \sin 2\theta \right) \ln \frac{\cos \theta + \sin \theta}{\cos \theta - \sin \theta} - \frac{i}{2\pi} (\text{Li}_2(ie^{2i\theta}) - \text{Li}_2(-ie^{2i\theta})) \\ & + \frac{1}{4} \ln(2 \sec 2\theta) - \text{atanh} \tan \theta - \frac{2i\theta}{\pi} \text{atan} e^{2i\theta}. \end{aligned} \quad (3.111)$$

The above zero-temperature formulas will not play a crucial role in our analysis of the model, but will be used in the next subsection to check the precision of our numerical results.

3.3 Numerical solution of Schwinger-Dyson equations

In this subsection, and in all numerical computations except otherwise stated, we always choose “natural units” such that the couplings are set to one, *i.e.*

$$\lambda = 1 \quad \text{and} \quad |\xi| = 1. \quad (3.112)$$

We want to solve numerically, for any values of the parameters, the Schwinger-Dyson equations (3.29), with self-energy (3.30) or (3.31), which we rewrite here for convenience in the units (3.112),

$$\begin{aligned} G_k &= \frac{1}{m - i\nu_k + \Sigma_k}, \\ \Sigma(\tau) &= \begin{cases} G(\tau)^2 G(-\tau) & (\text{U}(n)^2 \text{model}) \\ -\frac{1}{4}(G(\tau)^3 + 3G(\tau)G(-\tau)^2) & (\text{U}(n) \text{model}). \end{cases} \end{aligned} \quad (3.113)$$

We introduce a high-frequency cutoff k_{UV} and restrict our attention to the $2k_{\text{UV}}$ Matsubara frequencies labelled by $k = -k_{\text{UV}} + \frac{1}{2}, \dots, k_{\text{UV}} - \frac{1}{2}$. This effectively reduces (3.113) to a system of $4k_{\text{UV}}$ algebraic equations for the coefficients G_k and Σ_k . We solve this system using a standard iterative procedure which, given an approximate solution $(G_k^{[i-1]}, \Sigma_k^{[i-1]})$ to (3.113), generates another one through [9]

$$G_k^{[i]} = (1 - \alpha) G_k^{[i-1]} + \frac{\alpha}{m - i\nu_k + \Sigma_k^{[i-1]}}, \quad (3.114)$$

with $\Sigma_k^{[i]}$ calculated from $G_k^{[i]}$ through (3.113).

The weighting factor α can *a priori* be an arbitrary real number, and this choice will directly affect the convergence of the algorithm. One may refine it by making α frequency-dependent, or updating it from one iteration to the next as in [9]. For most purposes of the present paper, we found that a fixed, frequency-independent value $\alpha = 0.75$ is sufficient. We found no other solutions by using other values of α , by introducing a frequency dependence or by updating the weighting factor during the iterations.

The above algorithm is a generalization of the iterative procedure commonly used to solve fixed-point equations of the form $f(x) = x$, of which (3.113) can be thought of as being a multivariate version. Solutions x^* to $f(x) = x$ can be found from the iteration $x^{[i]} = f(x^{[i-1]})$ as long as the initial guess for the solution is close enough to x^* and $|f'(x^*)| < 1$. When this condition is satisfied, we say that the fixed point is attractive. If $|f'(x^*)| \geq 1$, on the other hand, the fixed point is repulsive and the method cannot converge. This is where the use of a weighting factor α can be helpful. Indeed, all fixed-points of $f(x)$ are also fixed-points of $f_\alpha(x) = \alpha f(x) + (1 - \alpha)x$, and it is easy to check that we can always find a non-zero value of α such that $|f'_\alpha(x^*)| < 1$. Thus, a repulsive fixed-point of the original problem $f(x) = x$ may be found as an attractive fixed point of the modified $f_\alpha(x) = x$ problem, for some suitably chosen value of α .

An important ingredient of the algorithm is the choice of the initial approximation. It may be taken to be the tree-level coefficients (3.6), $G_k^{[0]} = G_k^{(0)}$, or alternatively any other previously found Matsubara-Fourier coefficients deemed to be close to the actual solution. The iteration is then repeated until the equations are satisfied up to a prescribed tolerance ε , *i.e.*

$$\frac{1}{k_{\text{UV}}} \sum_{k=-k_{\text{UV}}+\frac{1}{2}}^{k_{\text{UV}}-\frac{1}{2}} \left| G_k^{[i]} - \frac{1}{m - i\nu_k + \Sigma_k^{[i]}} \right| < \varepsilon. \quad (3.115)$$

Since we work with 64-bit processors, we chose the tolerance to be $\varepsilon = 10^{-12}$ in the units (3.112), safely above machine-precision.

There is no guarantee that this algorithm terminates, but if convergence is achieved in $I_\alpha(\varepsilon)$ steps then the running time is $\mathcal{O}(I_\alpha(\varepsilon) \times k_{\text{UV}} \log k_{\text{UV}})$. The $k_{\text{UV}} \log k_{\text{UV}}$ factor in this expression comes from performing two Fast Fourier Transforms (FFTs) in each step, first to obtain $G^{[i]}(\tau)$ from $G_k^{[i]}$ and then to obtain $\Sigma_k^{[i]}$ from $\Sigma^{[i]}(\tau)$. Note that the calculation of $\Sigma^{[i]}(\tau)$ is itself $\mathcal{O}(k_{\text{UV}})$, or more generally $\mathcal{O}(k_{\text{UV}} \log q)$ for the q -body generalizations of (3.113) (see App. A), so it is always subleading because we shall always have $q \ll k_{\text{UV}}$.

We implemented the above algorithm in C++, and also in Mathematica for reference and double-checking purposes. In our C++ implementation, FFTs are calculated using the Cooley-Tukey algorithm [61] optimized to work with k_{UV} being a power of 2. In general, lower temperatures require higher values of k_{UV} to obtain good solutions. The reason is that the highest Matsubara frequency we include must always be much greater than the scale set by the coupling constant, *i.e.*

$$\nu_{k_{\text{UV}}-\frac{1}{2}} \equiv \frac{2\pi}{\beta} (k_{\text{UV}} - \tfrac{1}{2}) \gg 1 \quad \text{in natural units.} \quad (3.116)$$

The condition (3.116) ensures that at scale $\nu_{k_{\text{UV}}-1/2}$ we are well into the UV free asymptotic regime of the models, governed by the expansion (3.18) with (3.20) and

(3.21), for which

$$G_{k_{\text{UV}}-\frac{1}{2}} = G_{k_{\text{UV}}-\frac{1}{2}}^{(0)} + \mathcal{O}(1/k_{\text{UV}}^3). \quad (3.117)$$

After achieving convergence, one should always check that (3.117) is satisfied, as well as the basic consistency conditions (3.14), (3.15) and (3.16).

In practice, the cutoff k_{UV} also affects the precision with which we can compute thermodynamic quantities through formulas such as (3.96), (3.98) or (3.102). Indeed, introducing a high-frequency cutoff amounts to the truncation of the corresponding series, thus resulting in a systematic error in the computed values. To keep this error under control at temperatures of the order $T \sim 0.1$, one finds that $k_{\text{UV}} = 2^{10}$ is high enough, but for $T \sim 0.001$ the cutoff should be increased at least to $k_{\text{UV}} = 2^{16}$, and for $T \sim 0.0001$ we can reach $k_{\text{UV}} = 2^{21}$. See App. C for a more thorough discussion.

We shall use the above algorithm extensively in the following. For instance, when computing the phase diagrams in Sec. 4 we map the solutions that exist at each point in the space parameterized by (m, T) . In order to do this we may begin by finding numerical solutions of the Schwinger-Dyson equations in a region where an approximate solution is known to be valid, *e.g.* in one of the perturbative regimes $T \gg 1$ or $m \gg 1$. The algorithm then converges if we use as initial solution the tree-level coefficients (3.6), $G_k^{[0]} = G_k^{(0)}$. Keeping all but one of the parameters fixed, we can then find another solution close to this point using the previously found solution as an initial approximation, successively repeating this process to explore all of parameter space. While doing this, it is very useful to keep track of the thermodynamic properties of the solutions and the convergence of the algorithm, in order to adjust the step sizes accordingly.

One extremely important feature, that we have already mentioned in Sec. 3.1.3, is the possible existence of multiple solutions at a given point (m, T) , due to non-trivial monodromies in the space of solutions. This is signalled by the fact that the process described in the paragraph above can sometimes result in hysteresis curves for the thermodynamic functions. A representative example is depicted in Fig. 7. This means that the solution we find numerically depends not only on the parameters (m, T) explicitly entering the equations (3.113), but also on the initial approximation $G_k^{[0]}$ we used to start the iterative procedure. In this way, the solution “remembers” the specific sequence of steps in parameter space that was taken to reach it, and through it also the perturbative regime where the whole process typically started. This phenomenon will be discussed extensively in Sec. 4.

*Note on numerical errors and error bars*¹⁶ As discussed in detail in App. C, our numerical results are subject to two main sources of errors: systematic errors due to the introduction of cut-offs or discretizations; and methodological errors related to our fitting methods, *etc.* We do not have rigorous control over systematic errors,

¹⁶The remarks below apply also to the the real time numerics presented in Sec. 5:

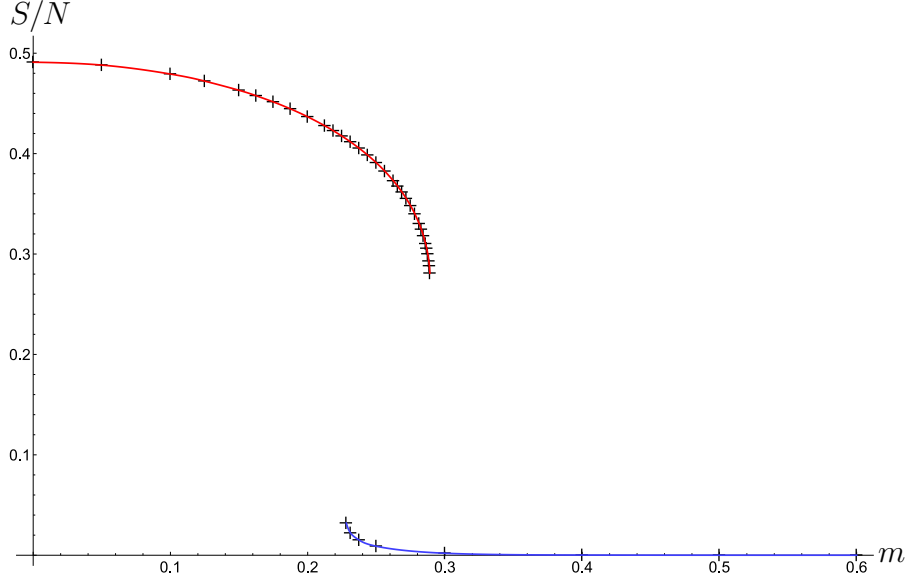


Figure 7: Hysteresis curve for the entropy of the $U(n)^2$ model as a function of mass, for fixed temperature $T = 0.035$ in natural units (data points shown with crosses for illustrative purposes). The red curve is obtained starting from the solution found at $m = 0$ and gradually increasing the mass, whereas the blue curve is obtained by decreasing the mass starting from the solution found at $m = 0.6$. In a range of masses around the value $m \sim 0.26$, two distinct solutions to the Schwinger-Dyson equations clearly coexist.

which typically may become important at very low temperatures. The data analysis methods are however under better control, and error bars can be associated with them as explained in App. C. The error bars that appear in the following therefore always refer to the methodological errors, but do not take into account the systematic errors that may appear on top of them. This implies that the actual values of some low-temperature observables may lie outside our error bar intervals. This issue is only relevant in some specific, yet important, instances and will be mentioned and discussed explicitly whenever it occurs.

3.3.1 Basic numerical results for the $U(n)^2$ model

The numerical algorithm described above may be used to study the finite coupling Schwinger-Dyson equations (3.113) for the $U(n)^2$ model, and assess for example the applicability of the low energy ansatz discussed in Sec. 3.1.4. From previous studies of the complex SYK model [19, 20] we can expect that this should be the case for low enough masses, whereas at large mass the system must be in a gapped perturbative phase (see Sec. 4 for more details).

In Fig. 8 we plot some solutions found for $m = 0$ and $m = 1$ at various tempera-

tures. The $m = 0$ solutions match with the standard SYK model, and we clearly see that the finite-temperature conformal ansatz (3.50) becomes a very good approximation in its range of validity, which in the natural units (3.112) is when $\beta \gg 1$, $\tau \gg 1$ and $\beta - \tau \gg 1$. The $m = 1$ solutions display a distinctly different behaviour, with the two-point function decaying exponentially. An energy gap can be immediately extracted from the log-plot of this type of solutions, see Sec. 4.

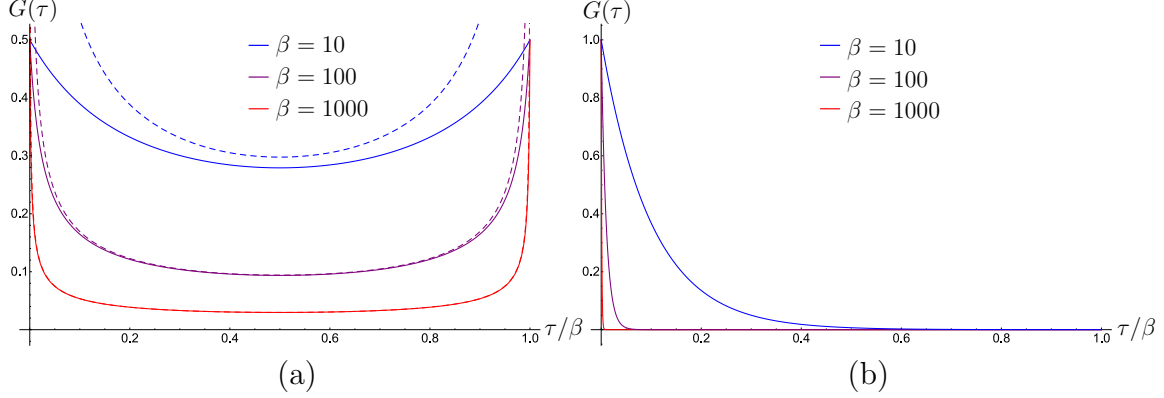


Figure 8: Numerical solution of the Schwinger-Dyson equations for the $U(n)^2$ model at various temperatures. (a) Plots for $m = 0$, with the finite-temperature conformal ansatz (3.50) shown in dashed lines for comparison (for $m = 0$, particle-hole symmetry implies $b_+ = b_-$, so we have set $a = \theta = 0$). (b) Plots for $m = 1$, where the exponential decay of the two-point function is clearly visible. Here and elsewhere in this paper, unless otherwise stated, solutions were computed with $k_{UV} = 2^{16}$, *i.e.* approximately 10^5 frequencies.

For $m \neq 0$ the model loses its particle-hole symmetry and thus $G(\beta - \tau) \neq G(\tau)$. In Fig. 9 we plot solutions for various intermediate values of the mass at $T = 0.01$. The marked change in behaviour around $m \sim 0.3$, from power-law decay at low masses to exponential decay at large masses, will be associated in Sec. 4 to a first order phase transition.

A very rough but simple criterion to distinguish the low and large mass solutions is to test the condition (3.49). Working at finite inverse temperature β the lowest frequency we have access to is $\nu_{1/2} = \pi/\beta$, so that we can use $\text{Re } \Sigma_{1/2}$ as an approximation to $\tilde{\Sigma}(0)$. For the examples illustrated in Fig. 9, we have $\text{Re } \Sigma_{1/2} = 0, -0.07$ and -0.15 for the solutions with $m = 0, 0.1$ and 0.2 respectively, whereas $\Sigma_{1/2} \sim 10^{-4}$ for the solutions with $m = 0.3, 0.4, 0.5$. These values show that the lowest three masses can satisfy the condition (3.49), whereas the three largest ones do not.

There are different ways in which we may quantitatively test the low energy ansatz of Sec. 3.1.4, beyond the naked-eye impression provided by Fig. 8 or the very rough testing of (3.49) through $\text{Re } \Sigma_{1/2}$. One possible approach starting from the zero-temperature expression (3.42) consists in extracting the scaling dimension from a

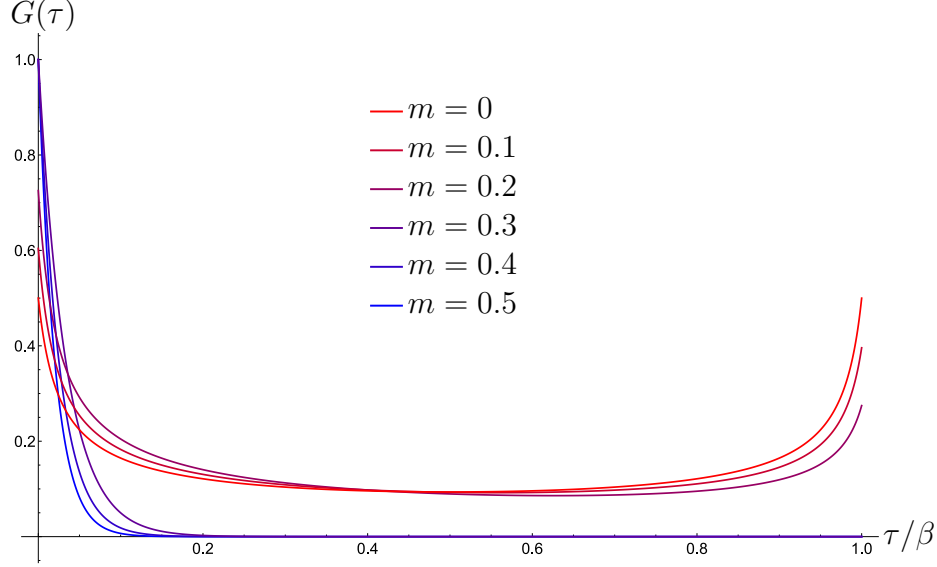


Figure 9: Numerical solution of the Schwinger-Dyson equations for the $U(n)^2$ model at $\beta = 100$ in natural units, for various values of the mass m . The loss of the particle-hole symmetry is signalled by the asymmetric shape of $G(\tau)$ about $\tau = \beta/2$. The endpoint values are related to the charge Q through (3.93), so that $Q = G(\beta^-) = 1 - G(0^+)$. SYK-like solutions similar to Fig. 8a and perturbative solutions similar to Fig. 8b can be immediately distinguished.

numerical solution by fitting

$$\frac{d \ln G(\tau)}{d \ln \tau} \rightarrow -2\Delta \quad \text{as} \quad \tau \rightarrow +\infty. \quad (3.118)$$

However, since the numerical evaluations are always performed at finite temperature we cannot use the above equation directly. Instead, Eq. (3.42) is approximately valid for values of τ satisfying $1 \ll \tau \ll \beta$. The precise interval of time we use in practice may be defined in different ways, but fortunately this choice does not greatly affect the fitted value of the scaling dimension Δ unless obviously unreasonable. Comparing the results obtained with different choices can serve the dual purpose of making sure that the numerical results make sense, as well as providing an estimation of the error or uncertainty associated to the computed values, see App. C for more details.

One important advantage of the procedure just described is that we can repeat it independently for $\tau \rightarrow -\infty$. We then obtain from each numerical solution not one but two values for the scaling dimension, denoted by $\Delta_{\pm}$ for $\tau \rightarrow \pm\infty$ as in (3.55). These values could *a priori* be different, but we find that in the $U(n)^2$ model they are always very similar, and certainly equivalent within the numerical precision we can achieve. This is in sharp contrast with the case of the $U(n)$ model discussed in the next subsection. Applying the same methods, we will clearly see there that

$\Delta_+ \neq \Delta_-$, signalling the fact that the behaviour of the $U(n)$ model indeed follows the non-standard ansatz (3.55).

More precisely, the $\beta \rightarrow \infty$ limit of the finite temperature formula (3.50) predicts

$$\frac{d \log |G(\tau)|}{d \ln |\tau|} \underset{1 \ll |\tau| \ll \beta}{\simeq} -2\Delta - \frac{2\pi a}{\beta} \tau \quad \text{with} \quad \Delta = \frac{1}{4}. \quad (3.119)$$

We thus expect a linear dependence in τ , with a vanishingly small slope in the $\beta \rightarrow \infty$ limit. This is illustrated in Fig. 10 for $T = 10^{-3}$ and $m = 0.1$. We indeed find a linear behaviour within the range of Euclidean times considered, both for $\tau \rightarrow +\infty$ and $\tau \rightarrow -\infty$. The scaling dimensions $\Delta_{\pm}$ are then obtained from the τ -independent part of the corresponding linear fits, and seen to be consistent with the expected value given by (3.46), $\Delta = 1/4$. We can also read the coefficient a from the slope and then derive the angle θ from (3.51). Note that this determination uses a subleading term and thus cannot be very accurate.

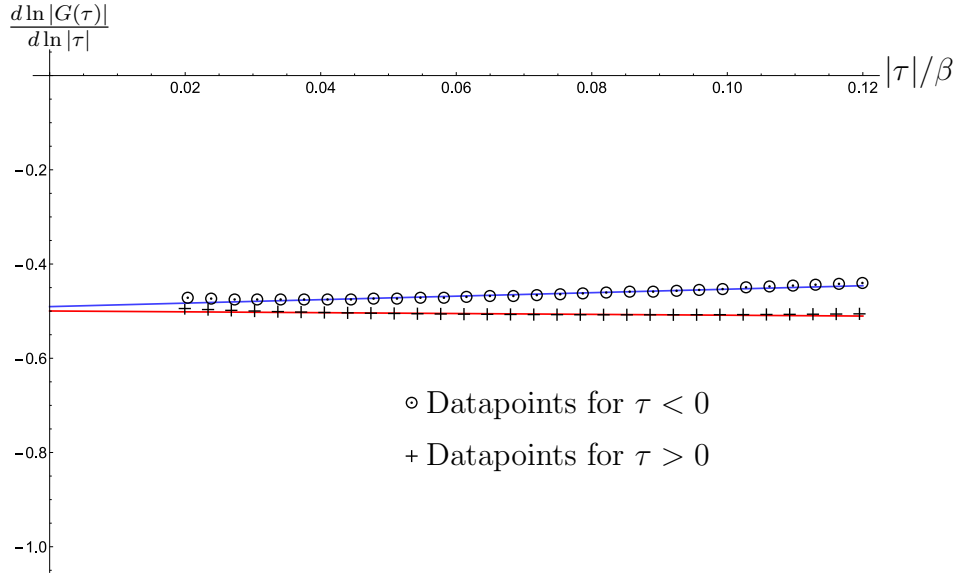


Figure 10: Extraction of the scaling dimension Δ and damping factor a for the $U(n)^2$ model using (3.119), with $m = 0.1$, $\beta = 1000$ and $\lambda = 1$. The linear behaviour of $\frac{d \ln |G(\tau)|}{d \ln |\tau|}$ in the chosen range of Euclidean times is clearly seen. The scaling dimensions $\Delta_{\pm}$ correspond to the y -intercept of the linear fits, in this case $\Delta_+ = 0.250_3$ (red) and $\Delta_- = 0.246_3$ (blue), results consistent with $\Delta_- = \Delta_+ = 1/4$ as predicted by the standard low energy ansatz. From the small slope we get the values $a_+ = 0.02_1$ and $a_- = 0.06_1$ for the parameter a . This result is consistent with $a = a_+ = a_- = 0.04_3$ and through (3.51) yields an angle $\theta = 0.12_9$. A more precise determination of a and θ is obtained in Fig. 11.

We can also use the previously found values for $\Delta_{\pm}$ as input to fit the coefficients $b_{\pm}$ directly in $G(\tau)$, again taking the numerical solution restricted to the finite intervals

considered before. This is illustrated in Fig. 11 for the example with $\beta = 1000$ and $m = 0.1$, and provides in particular a better determination of the parameters a and θ through (3.51). A non-trivial check of the whole procedure is the fact that these satisfy with very good accuracy the relation (3.47).

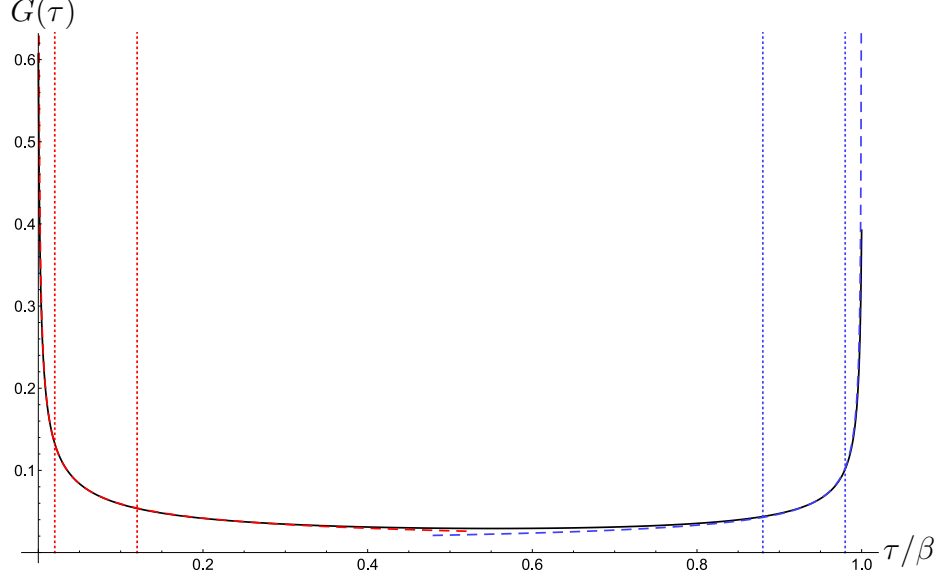


Figure 11: Extraction of the coefficients b_+ (red) and b_- (blue) for the $U(n)^2$ model using the zero-temperature ansatz (3.42), with $m = 0.1$, $\beta = 1000$ and $\lambda = 1$. The intervals used for fitting are delimited by vertical dotted lines; we obtain $b_+ = 0.59_2$ and $b_- = 0.45_2$, which through the parameterization (3.43) result in $C = 1.32_3$, $a = 0.043_6$ and $\theta = 0.13_2$, satisfying (3.47) to within 10%.

Another complementary approach is to fit Δ and $b_{\pm}$, with $a = \frac{1}{2\pi} \ln \frac{b_+}{b_-}$ as before, by using the natural generalization of (3.42) to arbitrary Δ given in App. A, Eq. (A.7). This ansatz can be trusted as soon as $|\tau| \gg 1$, without the need to restrict to $|\tau| \ll \beta$ as before. It is therefore not surprising that the results obtained in this way are even more precise. Some examples are shown in Fig. 12, for solutions with $m = 0.1$ at various temperatures. The fitting interval is delimited by vertical dotted lines, but once more the results are fairly robust with respect to its variations. Repeating this process at a given fixed mass for various temperatures, we observe that the scaling dimension varies linearly with the temperature for low enough temperatures. Such a linear dependence could certainly be explained by looking at subleading corrections to (3.42), but we have not tried to do so. We may consider the appearance of such linear behaviour as an indicator that the IR regime has been properly reached, since it is then immediate to extrapolate the finite temperature results to $T \rightarrow 0$.

The consistency of the results convincingly demonstrates that the solution to the Schwinger-Dyson equations approaches the low energy ansatz discussed in Sec. 3.1.4. Similar results are obtained for other small masses, with the approximation becoming

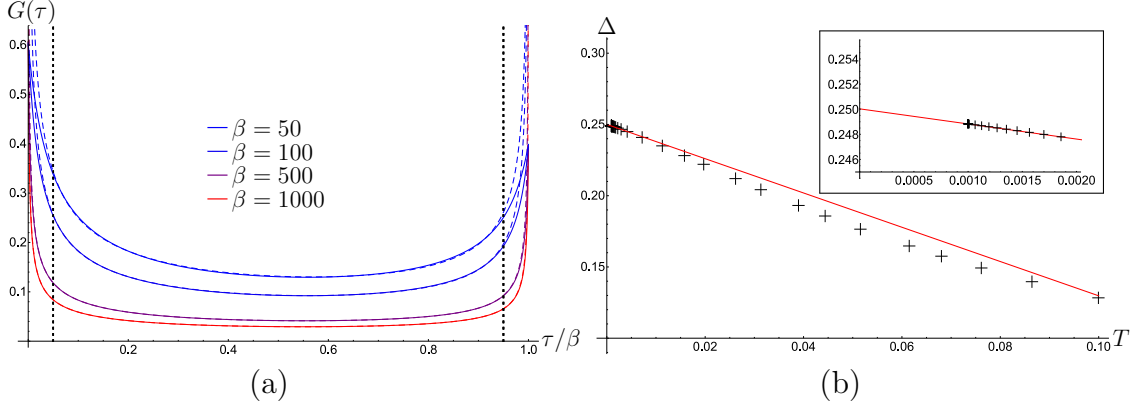


Figure 12: Extraction of the scaling dimension Δ and $b_{\pm}$ coefficients for the $U(n)^2$ model, with $m = 0.1$ at various temperatures, using the finite-temperature ansatz (A.7). (a) Numerical solutions (solid lines) and the fitted finite temperature scaling ansatz (A.7) (dashed lines); the dotted vertical lines delimit the region used for fitting. (b) Dependence of the extracted scaling dimension Δ with the temperature T , showing linear convergence to $\Delta = 1/4$ as $T \rightarrow 0$; the linear fit (red) was performed using ~ 20 points taken with $\beta \in [500, 1000]$ (shown in inset), and results in extrapolated values $\Delta = 0.250_1$, $b_+ = 0.600_1$, $b_- = 0.459_1$, $2\pi a = 0.268_4$, $C = 1.339_3$ and $\theta = 0.132_2$, matching (3.46) and (3.47) to within 2% precision.

better and better as the temperature gets lower. Note that the aim of the detailed analysis of some standard cases, as discussed above, is mainly to test our numerical methods before applying the same ideas to the much more interesting $U(n)$ model. In that case the non-standard ansatz (3.55) will play a crucial role, and we do not have a finite temperature expression analogous to (3.50). We will therefore not have the luxury of added precision. However, the consistency observed in the $U(n)^2$ model between the different methods provides us with a high degree of confidence that the results for the $U(n)$ model that we present in what follows are also very reliable.

3.3.2 Basic numerical results for the $U(n)$ model

We can perform on the $U(n)$ model a numerical analysis similar to the one done in the previous section for the $U(n)^2$ model. Note that for $m = 0$ we can use the particle-hole symmetry to look for solutions satisfying $G(-\tau) = -G(\tau)$, exactly as in the $U(n)^2$ model. The Schwinger-Dyson equations are then equivalent to the standard SYK model [2, 6, 9], the two-point function following at low temperatures the standard ansatz of Sec. 3.1.4.

For $m \neq 0$ the situation is very different already at a qualitative level. To illustrate this, we show in Fig. 13 solutions of the Schwinger-Dyson equations found for various masses at $T = 0.01$ in natural units. Again we expect the system to develop an

energy gap for large enough values of the mass, and indeed we see that the two-point function is exponentially decaying already for $m = 0.5$. However, we observe in this case a gradual, smooth transition into this regime, contrasting with the sharp change of behaviour seen for the $U(n)^2$ model, *e.g.* in Fig. 9.

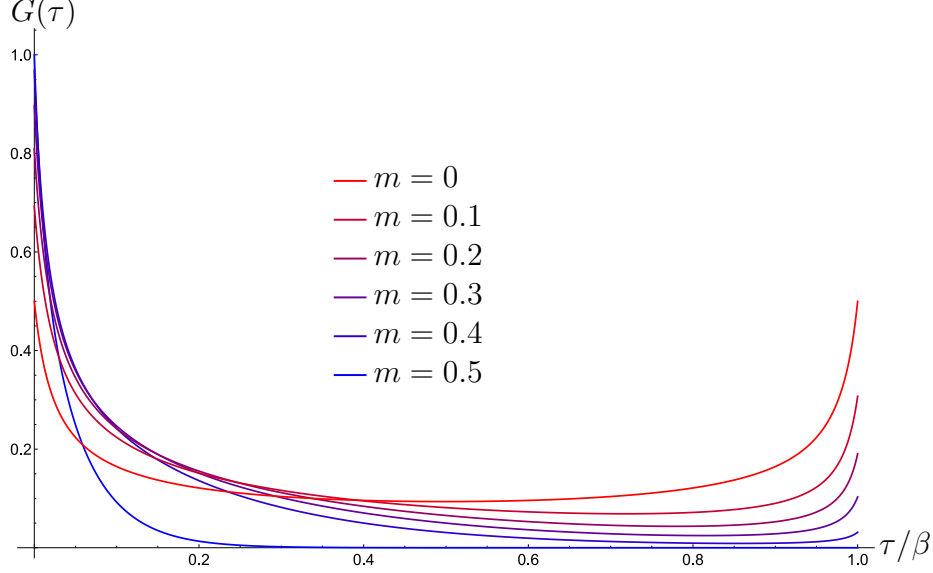


Figure 13: Numerical solution of the Schwinger-Dyson equations for the $U(n)$ model at $\beta = 100$ and $|\xi| = 1$, for various values of the mass m . The loss of the particle-hole symmetry is signalled by the asymmetric shape of $G(\tau)$ about $\tau = \beta/2$. The plot is consistent with a smooth change in the long time behaviour of the two-point function, from power-law to exponential decay, to be contrasted with that of the $U(n)^2$ model shown in Fig. 9.

Further evidence in favor of a smooth transition into the perturbative regime is provided by the constraint in (3.49). Note that this condition is equally valid for the standard and non-standard ansätze discussed in Sec. (3.1.4) and (3.1.5). For the solutions shown in Fig. 13 with $m = 0, 0.1, 0.2, 0.3, 0.4$ and 0.5 , we have $\text{Re } \Sigma_{1/2} = 0, -0.03, -0.11, -0.20, -0.28$ and -0.23 , respectively, so that $\text{Re } \Sigma_{1/2}$ switches from being decreasing to increasing somewhere in the range $0.3 < m < 0.5$. For the moment, we will concentrate on the numerical methods we can use to characterize the small-mass solutions approximately satisfying (3.49), deferring until Sec. 4 a more detailed discussion of this issue.

To elucidate in more details the nature of the gapless regime, we can try to extract scaling dimensions from the numerical solutions at finite temperature, as done before for the $U(n)^2$ model. Having fixed the mass to some small value, a scaling dimension growing unboundedly as the temperature decreases would be a telltale sign that the IR behaviour of the two-point function is trivial. On the other hand, a linear dependence of Δ with the temperature as $T \rightarrow 0$, as seen in the $U(n)^2$ model, would be strong

evidence in favor of some non-trivial scaling in the IR.

A representative case with $m = 0.1$ at $T = 10^{-3}$ in natural units is illustrated in Fig. 14. An exponential decay of the two-point function is clearly ruled out, the results being consistent with a power-law decay both when $\tau \rightarrow +\infty$ and when $\tau \rightarrow -\infty$. However, as already anticipated the exponents $\Delta_{\pm}$ relevant at $\tau \rightarrow \pm\infty$ must now be distinct, in sharp contrast with the situation observed in the $U(n)^2$ model in Sec. 3.3.1. In other words, the general ansatz (3.55) appears to be valid, but the standard ansatz (3.42) is unambiguously violated.

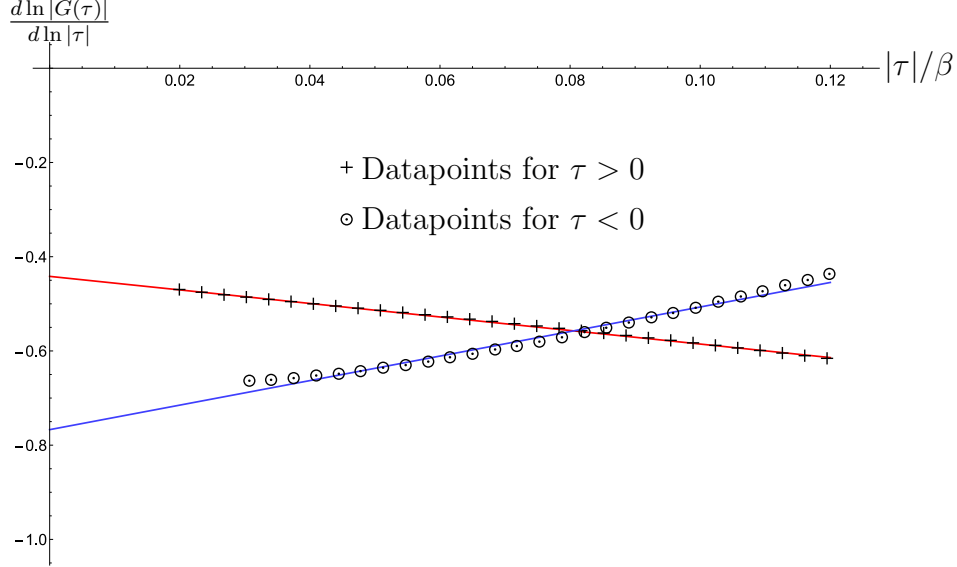


Figure 14: Extraction of the scaling dimensions $\Delta_{\pm}$ for the $U(n)$ model using the zero-temperature ansatz (3.55), with $m = 0.1$, $\beta = 1000$ and $|\xi| = 1$. We find a clear linear behaviour of $\frac{d \ln |G(\tau)|}{d \ln |\tau|}$ in the chosen range of Euclidean times when $\tau > 0$ (crosses), with a scaling dimension $\Delta_+ = 0.22_2$ corresponding to the y -intercept of the linear fit (red line). The behaviour when $\tau < 0$ (circles) is only approximately linear, and does not allow for a very precise determination of Δ_- . The linear fit shown on the plot (blue line) yields $\Delta_- = 0.37_4$, but in fact this may be considered a lower bound for the actual value (see discussion in the main text and App. C). In any case, the data does unambiguously show that $\Delta_- > \Delta_+$.

More precisely, as in Fig. 10 for the $U(n)^2$ model we find an approximately linear behaviour of $\frac{d \log |G(\tau)|}{d \log |\tau|}$ when $1 \ll |\tau| \ll \beta$ and $\tau > 0$. This allows for the determination of an exponent Δ_+ with satisfactory precision, the result being consistent with the expected value of $1/4$, see Eq.(3.57). On the other hand, linearity is not as good when $\tau < 0$, suggesting that the inverse temperature $\beta = 1000$ is not high enough to properly reach the deep IR regime. This yields a significant uncertainty in the determination of Δ_- , but in any case the equality $\Delta_- = \Delta_+$ is ruled out.¹⁷ We

¹⁷Note that we have also tried to fit the data to the finite temperature ansatz (A.7), which is valid

have also proceeded to extract the coefficients $b_{\pm}$ using the zero-temperature ansatz (3.55), as shown for the $U(n)^2$ model in Fig. 11. This yields a rather imprecise determination $b_+ = 0.64_5$ and $b_- = 0.46_7$, again supporting the interpretation that the deep IR regime has not been fully reached.

Further evidence in favour of the generalized ansatz $\Delta_+ \neq \Delta_-$ is found when analyzing the temperature dependence of the various quantities involved. In Fig. 15 we illustrate the typical situation at low enough masses by plotting the results obtained for $m = 0.1$. The scaling dimension Δ_+ extracted for $\tau \rightarrow +\infty$ is seen to develop a very clear linear behaviour at low temperatures. This allows us to confidently extrapolate its value to $T \rightarrow 0$, obtaining $\Delta_+ = 0.251_1$. On the other hand, the linear regime does not seem to be reached for Δ_- . Taking into account the sign of the small non-zero second derivative in the plot, our results should then be interpreted as providing a lower bound on Δ_- , establishing unambiguously that $\Delta_- > \Delta_+$. A similar analysis is made in Fig. 16 for the coefficients $b_{\pm}$. The extrapolation of b_+ , which is observed to be quite reliable, is of particular interest. Indeed, it provides an additional criterion to distinguish the standard ansatz $\Delta_+ = \Delta_-$, for which $b_+ = (1/4\pi)^{1/4} \simeq 0.53$ from Eq. (3.54), from the generalized ansatz $\Delta_+ \neq \Delta_-$, for which $b_+ = (2/\pi)^{1/4} \simeq 0.89$ from Eq. (3.57). We find in this case $b_+ = 0.82_4$, a result clearly in favour of $\Delta_+ \neq \Delta_-$ and consistent with our previous discussion of the scaling dimensions.

In conclusion, we have discovered a new IR behaviour in the $U(n)$ model, inconsistent with the standard ansatz (3.42) discussed in Sec. 3.1.4 and used so far in the literature. It is however in good agreement with the more general, non-standard ansatz of Sec. 3.1.5, at least to the degree in which we can trust the numerical methods involved. More details, in particular a more global picture of the resulting phase diagram, will be provided in Sec. 4, whereas the consequences of this new IR behaviour in the real time physics shall be presented in Sec. 5.

3.4 Euclidean time four-point function

3.4.1 General definitions and properties

Let us now introduce the time-ordered finite temperature Euclidean four-point function,

$$\phi(\tau_1, \tau_2, \tau_3, \tau_4) = \begin{cases} \frac{1}{n^2} \langle \text{Tr} \psi_{\mu E}(\tau_1) \psi_{\mu E}^\dagger(\tau_2) \text{Tr} \psi_{\nu E}(\tau_3) \psi_{\nu E}^\dagger(\tau_4) \rangle_\beta & \text{(matrix models)} \\ \langle \text{Tr} \chi_E^i(\tau_1) \chi_{iE}^\dagger(\tau_2) \chi_E^j(\tau_3) \chi_{jE}^\dagger(\tau_4) \rangle_\beta & \text{(disordered models)} \end{cases}. \quad (3.120)$$

if $\Delta_+ = \Delta_-$. We have found that it is impossible to obtain a good fit, even at very low temperatures, confirming that Δ_+ and Δ_- necessarily differ.

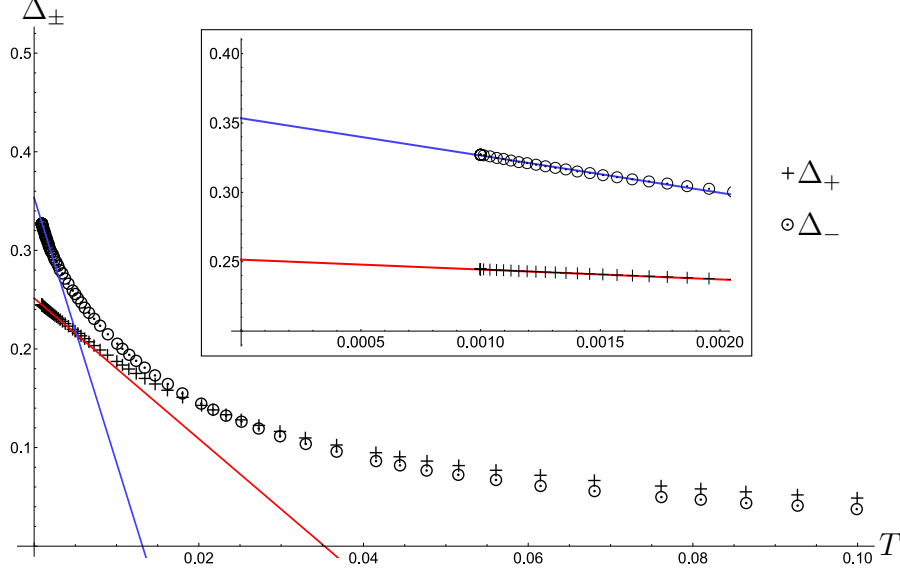


Figure 15: Temperature dependence of the scaling dimensions $\Delta_{\pm}$ obtained for the $U(n)$ model by fitting the zero-temperature ansatz (3.55), for $m = 0.1$ in natural units. Datapoints correspond to scaling dimensions Δ_+ (crosses) and Δ_- (circles), with the $T \rightarrow 0$ linear extrapolations shown with red and blue lines, respectively. The linear fits were performed using ~ 30 points taken with $\beta \in [500, 1000]$ (shown in inset), and result in extrapolated values $\Delta_+ = 0.251_1$ (red) and $\Delta_- = 0.353_2$ (blue). The low-temperature regime of Δ_- is not as linear as for Δ_+ , so that the indicated extrapolated value should conservatively be interpreted as a lower bound for the actual value of Δ_- .

These definitions apply whenever $0 < \tau_i < \beta$ but the KMS condition allows us to extend them to all values of the Euclidean times by antiperiodicity of period β ,

$$\phi(\tau_1 + n_1\beta, \tau_2 + n_2\beta, \tau_3 + n_3\beta, \tau_4 + n_4\beta) = (-1)^{n_1+n_2+n_3+n_4} \phi(\tau_1, \tau_2, \tau_3, \tau_4), \quad n_i \in \mathbb{Z}. \quad (3.121)$$

The four-point function also satisfies the reality and symmetry conditions

$$\phi(\tau_1, \tau_2, \tau_3, \tau_4)^* = \phi(-\tau_4, -\tau_3, -\tau_2, -\tau_1), \quad (3.122)$$

$$\phi(\tau_1, \tau_2, \tau_3, \tau_4) = \phi(\tau_3, \tau_4, \tau_1, \tau_2), \quad (3.123)$$

$$\phi(\tau_1, \tau_2, \tau_3, \tau_4)^* = \phi(\tau_1, \tau_2, \tau_3, \tau_4). \quad (3.124)$$

The last condition above follows from time-reversal invariance, which is always valid in the leading large N melonic limit we are interested in.¹⁸

It is natural to decompose ϕ into a disconnected and a connected piece,

$$\phi(\tau_1, \tau_2, \tau_3, \tau_4) = G(\tau_1, \tau_2)G(\tau_3, \tau_4) + \frac{1}{N} \mathcal{F}(\tau_1, \tau_2, \tau_3, \tau_4). \quad (3.125)$$

¹⁸At finite n and d , Eq. (3.124) is correct only when the couplings are real.

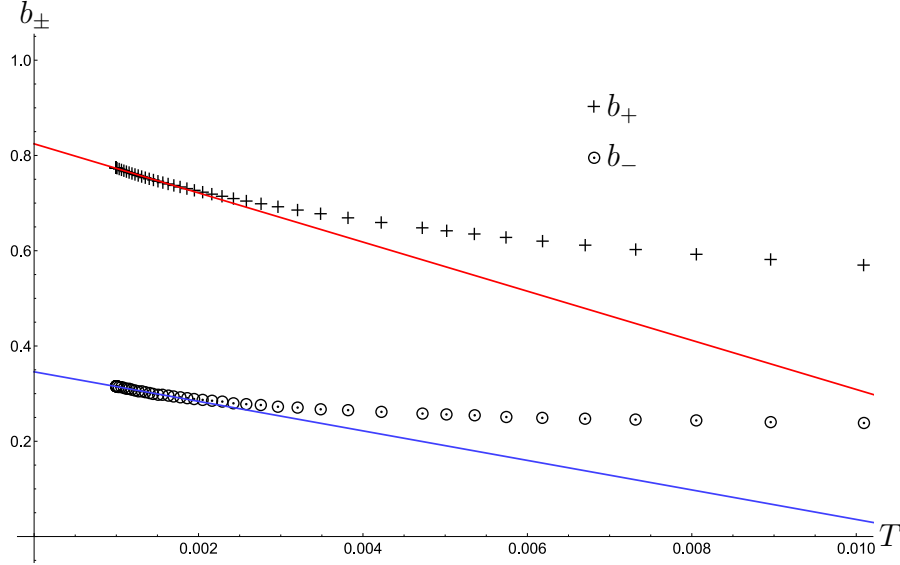


Figure 16: Temperature dependence of the $b_{\pm}$ coefficients obtained for the $U(n)$ model by fitting the zero-temperature ansatz (3.55), for $m = 0.1$ in natural units. Datapoints for the b_+ (crosses) and b_- (circles) coefficients, along with the corresponding $T \rightarrow 0$ extrapolations (red and blue lines, respectively); linear fits were performed using the same points as in Fig. 15, and result in $b_+ = 0.82_4$ and $b_- = 0.35_5$. Once more the linear behaviour is observed to be more accurately followed for $\tau \rightarrow +\infty$ than for $\tau \rightarrow -\infty$, the determination of b_- consequently being more imprecise.

Our goal is to compute the connected piece $\mathcal{F}$ at leading order in the large N limit. We first present a diagrammatic analysis in the next subsection and then a rigorous and very general algebraic proof in Sec. 3.4.3.

3.4.2 Diagrammatic analysis

Graphs contributing to $\mathcal{F}$ at leading order are easy to find. One considers ladders of the form depicted in Fig. 17, with rungs of various types. These ladder graphs can be easily resummed because they form a geometric series. If we introduce a convenient matrix notation for functions of four variables, with a pair of variables forming a “matrix index” and matrix multiplication rule

$$(\mathbb{A} \cdot \mathbb{B})(\tau_1, \tau_2, \tau_3, \tau_4) = \int_{[0, \beta]^2} \mathcal{A}(\tau_1, \tau_2, \tau, \tau') \mathcal{B}(\tau, \tau', \tau_3, \tau_4) d\tau d\tau', \quad (3.126)$$

we can write the geometric series for $\mathcal{F}$ in the form

$$\mathbb{F} = \sum_{n \geq 0} \mathbb{K}^n \cdot \mathbb{F}_0 = (\mathbb{I} - \mathbb{K})^{-1} \cdot \mathbb{F}_0. \quad (3.127)$$

The function

$$\mathcal{F}_0(\tau_1, \tau_2, \tau_3, \tau_4) = -G(\tau_1, \tau_4)G(\tau_3, \tau_2) \quad (3.128)$$

corresponds to the contribution with no rung inserted, whereas the kernel $\mathcal{K}$ is built from all the possible rung structures that can appear in the ladder.

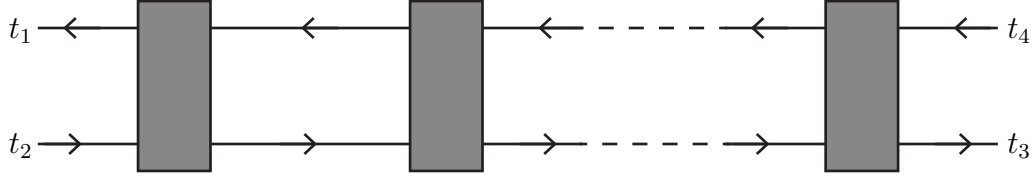


Figure 17: Ladder graph contributing to the connected four-point function $\mathcal{F}$. For better readability, we use a simple oriented line to represent the full two-point function (instead of a line with a shaded disk as in Fig. 5). The shaded rectangles can be any of the rung structures depicted in Fig. 18 for the $U(n)^2$ model, or in Fig. 19 for the $U(n)$ model.

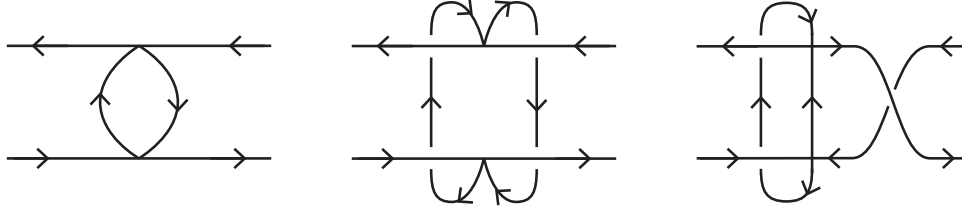


Figure 18: Rungs from which we can build the ladder diagrams contributing to the leading order of the connected four-point function of the $U(n)^2$ model. For better readability, we use a simple oriented line to represent the full two-point function.

In the $U(n)^2$ model, the kernel is given by

$$\mathcal{K}^{U(n)^2}(\tau_1, \tau_2, \tau_3, \tau_4) = -\lambda^2 \left[2G(\tau_1, \tau_3)G(\tau_4, \tau_2)G(\tau_3, \tau_4)G(\tau_4, \tau_3) + G(\tau_1, \tau_4)G(\tau_3, \tau_2)G(\tau_4, \tau_3)^2 \right]. \quad (3.129)$$

The first term on the right-hand side of (3.129) is associated with the left and central rungs in Fig. 18, while the second term corresponds to the rightmost rung depicted in that figure.

In the $U(n)$ model, the kernel is similarly given by

$$\mathcal{K}^{U(n)}(\tau_1, \tau_2, \tau_3, \tau_4) = \frac{|\xi|^2}{4} \left[3(G(\tau_4, \tau_3)^2 + G(\tau_3, \tau_4)^2)G(\tau_1, \tau_3)G(\tau_4, \tau_2) + 6G(\tau_3, \tau_4)G(\tau_4, \tau_3)G(\tau_1, \tau_4)G(\tau_3, \tau_2) \right]. \quad (3.130)$$

The first term on the right-hand side of (3.130) is associated with the three rungs in the first row of Fig. 19; the second term with the three rungs in the second row of Fig. 19; and the third term with the six rungs in the last two rows of Fig. 19.

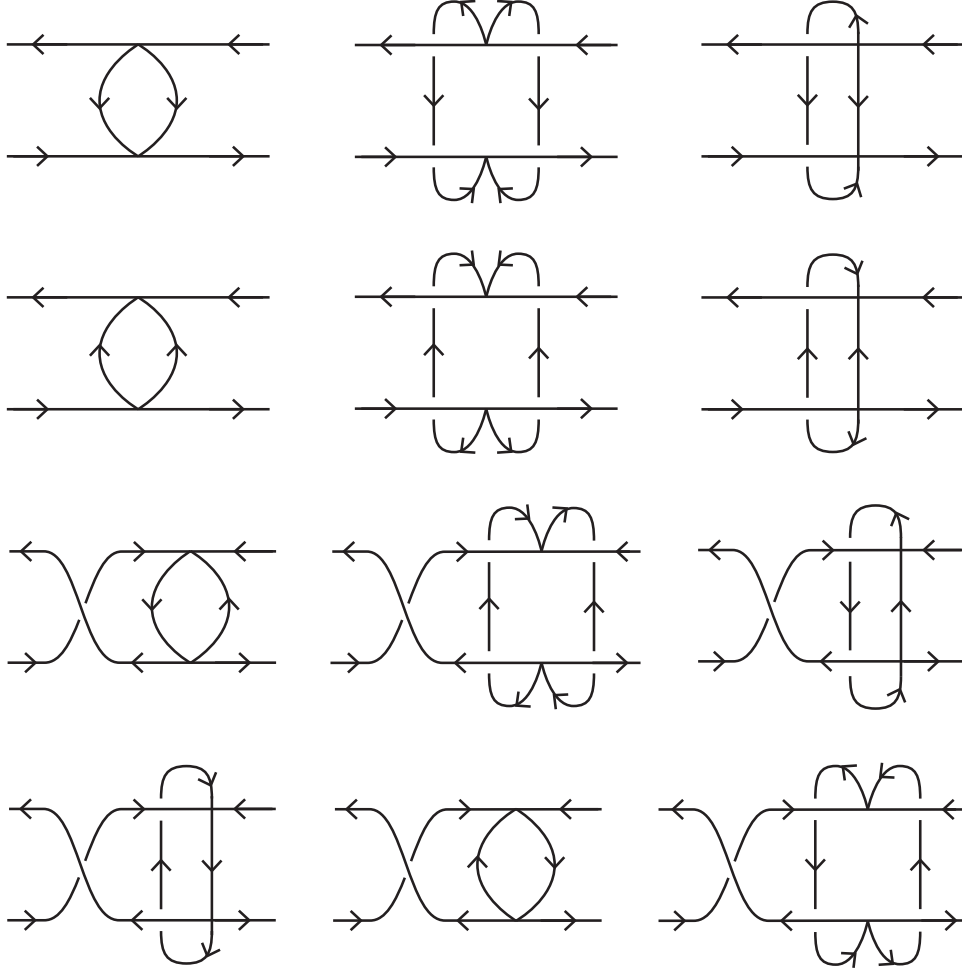


Figure 19: Rungs from which we can build the ladder diagrams contributing to the leading order of the connected four-point function of the $U(n)$ model. For better readability, we use a simple oriented line to represent the full two-point function.

Even though it is completely straightforward to check that the ladder diagrams do contribute at leading order to $\mathcal{F}$, it is less immediate to prove diagrammatically that they are the only diagrams having this property. This part of the proof is often omitted in the literature. In the next subsection, we provide a very general algebraic method that permits the rigorous computation of the connected four-point function kernels of a large class of models. Applied to the $U(n)^2$ and $U(n)$ models this method yields (3.129) and (3.130), respectively, thus proving that the ladder diagrams considered above are indeed the only ones that can contribute at leading large N order.

3.4.3 Algebraic proof

We shall use the language of matrix models to describe the proof of the formulas (3.129) and (3.130) for the four-point function kernels, but the disordered model formulation can be treated in exactly the same way. Only some very general features of the models are needed, namely:

- i) The large N limit is dominated by generalized melonic graphs such that the Schwinger-Dyson equations for the two-point function G and self-energy Σ coincide with the saddle-point equations of an effective action functional S_{eff} , of the general form

$$\begin{aligned} \frac{1}{N} S_{\text{eff}}[\mathbf{G}, \Sigma] = & -\ln(1 + e^{-\beta m}) - \ln \frac{\det \mathcal{O}_\Sigma}{\det \mathcal{O}_0} \\ & + \int_{[0, \beta]^2} \Sigma(\tau_2, \tau_1) \mathbf{G}(\tau_1, \tau_2) d\tau_1 d\tau_2 + s[\mathbf{G}]. \end{aligned} \quad (3.131)$$

The operators $\mathcal{O}_\Sigma$ and $\mathcal{O}_0$ are defined in (3.59) and s is an *a priori* arbitrary functional of $\mathbf{G}$. In particular, G and Σ are the on-shell values of $\mathbf{G}$ and Σ .

- ii) The partition function Z is given at leading order by the on-shell value of the effective action,

$$Z = e^{-S_{\text{eff}}}, \quad (3.132)$$

where $S_{\text{eff}} = S_{\text{eff}}(G, \Sigma)$ as in (3.62).

The Euclidean four-point function (3.120) is then given by (3.125), (3.127) and (3.128), with a kernel defined by the general formula

$$\mathcal{K}(\tau_1, \tau_2, \tau_3, \tau_4) = \int_{[0, \beta]^2} G(\tau_1, \tau) G(\tau', \tau_2) \frac{\delta^2 s}{\delta G(\tau', \tau) \delta G(\tau_3, \tau_4)} d\tau d\tau'. \quad (3.133)$$

To prove this formula, it is convenient to introduce the source formalism. The partition function Z has a standard path integral representation, with matrix Grassmannian integration variables ψ_μ and $\bar{\psi}_\mu$ and an action

$$\begin{aligned} S = & nd \int_0^\beta d\tau \operatorname{tr} (\bar{\psi}_\mu \dot{\psi}_\mu + m \bar{\psi}_\mu \psi_\mu) + S_{\text{int}} \\ = & nd \int_{[0, \beta]^2} d\tau_1 d\tau_2 \operatorname{tr} \bar{\psi}_\mu(\tau_1) \mathcal{O}_0(\tau_1, \tau_2) \psi_\mu(\tau_2) + S_{\text{int}}. \end{aligned} \quad (3.134)$$

The term S_{int} represents the interactions in the Hamiltonian. We can modify this action by adding a source term for the two-point function. The new action is

$$S[J] = nd \int_{[0, \beta]^2} d\tau_1 d\tau_2 \operatorname{tr} \bar{\psi}_\mu(\tau_1) (\mathcal{O}_0(\tau_1, \tau_2) + J(\tau_1, \tau_2)) \psi_\mu(\tau_2) + S_{\text{int}}, \quad (3.135)$$

and the associated path integral yields a source-dependent partition function $Z[J]$ and free energy $F[J] = -T \ln Z[J]$. The two-point function (3.1) and four-point function (3.120) are then given by

$$G(\tau_1, \tau_2) = \frac{1}{N} \frac{1}{Z} \frac{\delta Z}{\delta J(\tau_2, \tau_1)} \Big|_{J=0}, \quad (3.136)$$

$$\phi(\tau_1, \tau_2, \tau_3, \tau_4) = \frac{1}{N^2} \frac{1}{Z} \frac{\delta^2 Z}{\delta J(\tau_2, \tau_1) \delta J(\tau_4, \tau_3)} \Big|_{J=0}. \quad (3.137)$$

Equivalently, we can express G and the connected piece $\mathcal{F}$ of the four-point function defined in (3.125) in terms of the free energy, so that

$$G(t_1, t_2) = -\frac{\beta}{N} \frac{\delta F}{\delta J(\tau_2, \tau_1)} \Big|_{J=0}, \quad (3.138)$$

$$\mathcal{F}(\tau_1, \tau_2, \tau_3, \tau_4) = -\frac{\beta}{N} \frac{\delta^2 F}{\delta J(\tau_2, \tau_1) \delta J(\tau_4, \tau_3)} \Big|_{J=0}. \quad (3.139)$$

Because the source term in (3.135) is quadratic in ψ and $\bar{\psi}$, its addition only modifies the expression of the propagator, *without changing the large n and large d diagrammatics of the models*. The results (3.131) and (3.132), which are *a priori* assumed to be valid without the source term, can thus be immediately generalized to take it into account. This amounts to replacing the mass term $m\delta_\beta(\tau_1 - \tau_2)$ in $\mathcal{O}_\Sigma$ and $\mathcal{O}_0$ defined in (3.59) by $m\delta_\beta(\tau_1 - \tau_2) + J(\tau_1, \tau_2)$. The form (3.131) of the effective action then implies that this modification yields a new, source-dependent effective action $\hat{\mathcal{S}}_{\text{eff}}[\mathbf{G}, \Sigma; J]$ given by

$$\hat{\mathcal{S}}_{\text{eff}}[\mathbf{G}, \Sigma; J] = \mathcal{S}_{\text{eff}}[\mathbf{G}, \Sigma + J] - N \int_{[0, \beta]^2} J(\tau_2, \tau_1) \mathbf{G}(\tau_1, \tau_2) d\tau_1 d\tau_2. \quad (3.140)$$

The solutions to the source-dependent saddle-point equations

$$\frac{\delta \hat{\mathcal{S}}_{\text{eff}}}{\delta \mathbf{G}(\tau_1, \tau_2)} = 0 \quad \text{and} \quad \frac{\delta \hat{\mathcal{S}}_{\text{eff}}}{\delta \Sigma(\tau_1, \tau_2)} = 0 \quad (3.141)$$

are the on-shell values $G[J]$ and $\Sigma[J]$, and the source-dependent partition function and free energy are given by

$$Z[J] = e^{-\beta F[J]} = e^{-\hat{\mathcal{S}}_{\text{eff}}[J]} \quad \text{with} \quad \hat{\mathcal{S}}_{\text{eff}}[J] \equiv \hat{\mathcal{S}}_{\text{eff}}[G[J], \Sigma[J]; J]. \quad (3.142)$$

It is then an exercise in partial differentiation to obtain the formula (3.127) with a kernel given by (3.133), starting from (3.139) and using (3.140) and (3.142). The details are provided in App. B.

For the particular case of the models we are discussing, one then finds the explicit form of the kernels (3.129) and (3.130) by plugging in (3.133) the relevant functionals for $s[\mathbf{G}]$, namely $s_{\mathrm{U}(n)^2}$ and $s_{\mathrm{U}(n)}$ given by

$$s_{\mathrm{U}(n)^2}[\mathbf{G}] = -\frac{\lambda^2}{4} \int_{[0,\beta]^2} \mathbf{G}(\tau_1, \tau_2)^2 \mathbf{G}(\tau_2, \tau_1)^2 d\tau_1 d\tau_2, \quad (3.143)$$

$$s_{\mathrm{U}(n)}[\mathbf{G}] = \frac{|\xi|^2}{4} \int_{[0,\beta]^2} \mathbf{G}(\tau_1, \tau_2)^3 \mathbf{G}(\tau_2, \tau_1) d\tau_1 d\tau_2. \quad (3.144)$$

4 Phase diagrams

In this section, unless explicitly stated otherwise, we continue using for convenience the natural units defined in (3.112). Having set $\lambda = 1$ (for the $\mathrm{U}(n)^2$ model) or $|\xi| = 1$ (for the $\mathrm{U}(n)$ model), we then have two dimensionless parameters, m and T , and we are going to study in detail the properties of the models at large N in the (m, T) -plane.

We can assume $m \geq 0$ without loss of generality. Indeed, the sign of m may be changed by permuting “particles” and “holes,” *i.e.* by exchanging the role of the operators ψ and $\psi^\dagger$ (in the matrix models) or χ and $\chi^\dagger$ (in the disordered models). It is immediate to check that if $G(\tau; m)$ is a solution with mass m of the Schwinger-Dyson equations (3.26)–(3.30) (for the $\mathrm{U}(n)^2$ model) or (3.26)–(3.31) (for the $\mathrm{U}(n)$ model), then $G(\tau; -m) = G(\beta - \tau; m)$ is a solution of the same equations with mass $-m$.

4.1 General discussion

A qualitative picture of the physics to be expected in the (m, T) plane is illustrated in Fig. 20. Our goal in this subsection is to elaborate on this picture and elucidate many important features, without resorting to difficult calculations or numerical analysis. Full details are then provided in the following subsections 4.2 and 4.3.

Zero mass When the mass strictly vanishes, the models have particle-hole symmetry. Assuming that this symmetry is not spontaneously broken, a possibility that we shall not investigate here, we can seek a solution to the Schwinger-Dyson equations such that $G(-\tau) = -G(\tau)$, and this solution matches the standard solution for the original SYK model studied extensively in the literature [6, 9, 62–64]. The only difference with SYK is the doubling of the number of degrees of freedom, because we use complex and not Majorana fermions. This introduces trivial factors of two, for example in the formula for the entropy. The physics at $m = 0$ is thus familiar. The models flow from a trivial perturbative behaviour at high temperature, governed by

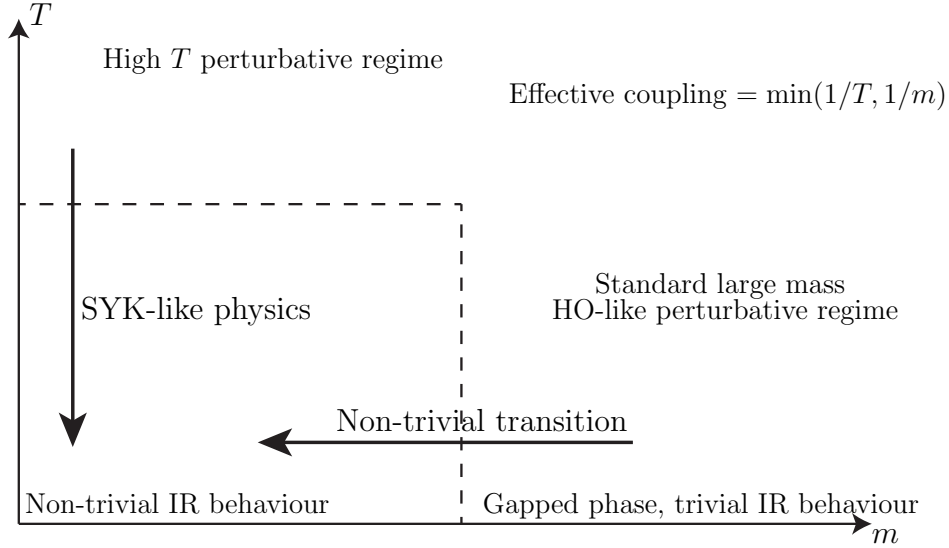


Figure 20: Qualitative picture of the phase diagram in the (m, T) -plane. The physics at small mass is SYK-like, with a flow from a perturbative regime at high T to a strongly coupled conformal phase at low T . The physics at large mass is harmonic-oscillator-like, weakly coupled for all values of the temperature. Our main goal below is to elucidate and study the consequences of the non-trivial transition from the weakly coupled low-temperature region at large mass to the strongly coupled low-temperature conformal region at small mass.

the standard high T perturbation theory for fermionic systems around the maximally entropic state with density matrix $2^{-N}\mathbb{I}$, to a non-trivial, strongly coupled, conformal behaviour at low temperature, reminiscent of extremal black-hole physics. In particular: the zero-temperature Euclidean two-point function G has a power-law decay $|\tau|^{-2\Delta}$ at large times, with $\Delta = 1/4$; there exists a non-vanishing zero-temperature entropy given by (3.107); and the chaos exponent, as read-off from real time out-of-time-ordered four-point functions, is $2\pi T$ at low temperatures, saturating the bound discussed in [33].

Small non-zero mass When the mass is small but non-zero, three qualitatively distinct behaviours could *a priori* be expected.

The first possibility is that a mass gap is immediately created, as would be required if the physics were weakly coupled. This is what happens, for instance, when two copies of the SYK model are coupled quadratically as in [54, 65].

The second possibility is that the physics remains gapless when the mass is turned on. This is what happens in both of the models we consider, over a finite interval of masses $0 \leq m \leq m_*$, where m_* is a strictly positive critical mass beyond which a mass gap is created. The nature of the deformation induced by the mass term is, however, qualitatively different in the $U(n)^2$ and the $U(n)$ models. This was mentioned in

Sec. 3.3.1 and 3.3.2 and will be thoroughly discussed in Sec. 4.2 and 4.3. More precisely, in the $U(n)^2$ model the mass term flows to a marginal deformation in the IR, in the sense that the IR physics smoothly depends on m . In the $U(n)$ model, on the other hand, the mass operator is relevant at strictly $m = 0$, in the sense that the deep IR physics is governed by two disconnected fixed points at $m = 0$ and $m > 0$. It becomes marginal for all $0 < m < m_*$.

The third possibility is that the mass term is irrelevant in the IR. This behaviour would actually be quite natural in the $U(n)$ model, since the standard low energy ansatz (3.42), (3.43) has no free parameter in this case, see Eq. (3.54). However, our numerical results rule out this possibility in favor of the most general behaviour described in Sec. 3.1.5.

Large mass At very large masses the models are weakly coupled for all values of the temperature, and are governed by the standard perturbation theory around a set of N decoupled harmonic oscillators. The physics is then very simple. There is a mass gap $m_{\text{eff}} > 0$ of order m , with the zero-temperature Euclidean two-point function G decaying as $e^{-m_{\text{eff}}\tau}$ at large times; the zero-temperature entropy vanishes, the ground state being the unique Fock vacuum; and the chaos exponent is $o(T)$ at low temperatures.

In the $U(n)^2$ model, the fermion number is conserved and the mass term can then also be interpreted as (minus) a chemical potential for the fermion number. Under these circumstances, it is straightforward to derive, for instance from the spectral decomposition formula (3.12), that in a phase with a mass gap the mass dependence of the zero-temperature two-point function is exactly given by a simple exponential term. More precisely,

$$G(\tau; m_1, T = 0) = G(\tau; m_2, T = 0)e^{-(m_1 - m_2)\tau} \quad (4.1)$$

for m_1 and m_2 any two masses in the gapped regime where the ground state is the Fock vacuum. This very general result is realized in the simplest possible way in the $U(n)^2$ model. Indeed, the tree-level zero-temperature two-point function,

$$G^{(0)}(\tau; m, T = 0) = \Theta(\tau)e^{-m\tau}, \quad (4.2)$$

is actually an exact solution of the zero-temperature Schwinger-Dyson equations for any value of the coupling, simply because the right-hand side of (3.30) involves the product of G at opposite times.

In the $U(n)$ model this simplification of the large mass physics does not occur. The fermion number is not conserved and the mass is not a chemical potential. Even though the theory is still gapped and the ground state is the Fock vacuum, the simple relation (4.1) is not valid. Indeed the zero-temperature two-point function can get non-trivial corrections, due to the presence of the $G(\tau)^3$ term on the right-hand side of (3.31). In particular, the physical mass gap m_{eff} , which is computed in Sec. 4.3, is a non-trivial function of m .

The transition from weak to strong coupling The above discussion implies that, when the mass is decreased from large to small values, the behaviour of the models must drastically change from a gapped perturbative regime in which the ground state is the Fock vacuum to a non-perturbative regime with no gap. Therefore, there exists a critical mass m_* below which the mass gap closes,

$$m_{\text{eff}} > 0 \text{ for } m > m_*, \quad m_{\text{eff}} = 0 \text{ for } m < m_*. \quad (4.3)$$

The zero-temperature transition at $m = m_*$ is reflected in the zero-temperature energy, which vanishes in the Fock vacuum at $m > m_*$ and must be strictly negative at $m < m_*$. The transition also comes along with the creation of a non-vanishing zero-temperature entropy. This follows from (3.110) for the $U(n)^2$ model and will be convincingly demonstrated numerically for the $U(n)$ model in Sec. 4.3.

However, in spite of these similarities between the two models, we can infer from our discussion of the high mass regime that the transition at $m = m_*$ must be qualitatively different in the $U(n)^2$ and $U(n)$ models.

In the $U(n)^2$ model, because the zero-temperature two-point function matches with the tree-level result (4.2) in the gapped regime, the mass gap is exactly $m_{\text{eff}} = m$ for all $m > m_*$. The transition from $m > m_*$ to $m < m_*$ must then be discontinuous, with in particular a discontinuous jump of the mass gap from m_* to zero. Such a first order phase transition was indeed found in [28]. Its existence is made possible by the non-trivial monodromies in the space of solutions of the Schwinger-Dyson equations that we have already mentioned in Sec. 3.1.3. The first order transition is not limited to zero temperature, but extends up to a critical point at finite temperature [28]. More details are given in Sec. 4.2.

In the $U(n)$ model, the zero-temperature two-point function gets non-trivial corrections at $m > m_*$ and the mass gap is a non-trivial function $m_{\text{eff}}(m)$. We shall provide convincing numerical evidence in Sec. 4.3 that m_{eff} smoothly vanishes when $m \rightarrow m_*$ from above. We thus get in this case a large N version of a quantum critical point. In particular, the first order phase transition is replaced by a smooth crossover at finite temperature.

Most of the rest of this paper will focus on providing more details on the transition from weak to strong coupling in both models, using the standard thermodynamical Euclidean point of view in Sec. 4.2 and 4.3 and a more modern and innovative real time point of view in Sec. 5.

The gravitational collapse interpretation There is an interesting qualitative interpretation of the transition from weak to strong coupling when the mass is varied at a fixed (low) temperature in the models we are studying, based on the compelling analogy between black-hole physics and the strongly coupled conformal phase on the one hand and the matrix model formulation of the models on the other hand.

As already mentioned in Sec. 2.3.3, quantum mechanical matrix models can always be interpreted in terms of open strings and D-particles. In this point of view, the operators $\psi_{\mu b}^{\dagger a}$ create open strings whose endpoints are attached to the D-particles labelled by the $U(n)$ indices a and b . Since the mass of a string is proportional to its length, the mass parameter m represents the distance between the D-particles joined by the string.

At large mass, we thus have an assembly of D-particles which are very far away from each other and interact very weakly. Their dynamics is governed by standard matrix model perturbation theory. In the usual terminology, this is the regime where the “brane picture” is reliable. As the mass is decreased, the D-particles are brought closer and closer together and the matrix model becomes more and more strongly coupled. Eventually, we expect the system to undergo gravitational collapse and form a black hole. It then enters a regime in which the “brane picture” is no longer suitable, and should be replaced by a “gravity picture” in which the branes are effectively described by a non-trivial gravitational background containing a black hole.

It is very natural to identify this phenomenon with the expected non-trivial transition at $m = m_*$ discussed above. This is perfectly consistent with the appearance of a macroscopic zero-temperature entropy for $m < m_*$, signalling the formation of a black-hole horizon, as well as with other properties of the strongly coupled regime of both models we consider.

Of course, the matrix models under consideration are still quite far from being truly realistic models of black holes. For instance, they do not contain fundamental bosonic matrix degrees of freedom, which are most naturally associated with transverse dimensions of space-time. Moreover, the strongly coupled phases of the models are “black holes” only in the same sense as the SYK model is one: there is a very suggestive matching of various physical properties at the qualitative level (entropy, quasi-normal behaviour, continuum spectrum, chaos) and some non-trivial quantitative agreement (saturation of the bound on the Lyapunov exponent). However, it is also clear that they can at most represent “string size” black holes for which a purely Einsteinian description cannot be fully adequate. In particular, the models contain an infinite tower of states in the bulk that cannot be decoupled from the low mass modes [9]. Moreover, the strongly coupled phase of the $U(n)$ model is quite exotic and its interpretation as a “black hole” is then even more questionable than in the case of the $U(n)^2$ model (see Sec. 4.3).

Keeping in mind these important caveats, we nevertheless expect that the transition from weak to strong coupling in these models does capture important features of the transition between the “brane description” and the “gravitational description” of a collection of D-branes, *i.e.* of the phenomenon of gravitational collapse. To the best of our knowledge, such a description has never before been given in a fully quantum mechanical framework, even in a very simplified set-up such as the one presented here.

4.2 Phase diagram of the $U(n)^2$ model

The phase diagram of the $U(n)^2$ model was first described in [28], and is depicted in Fig. 21. We will briefly review here its main features to offer a self-contained discussion, and present additional details about its construction, generalizations and most important properties.

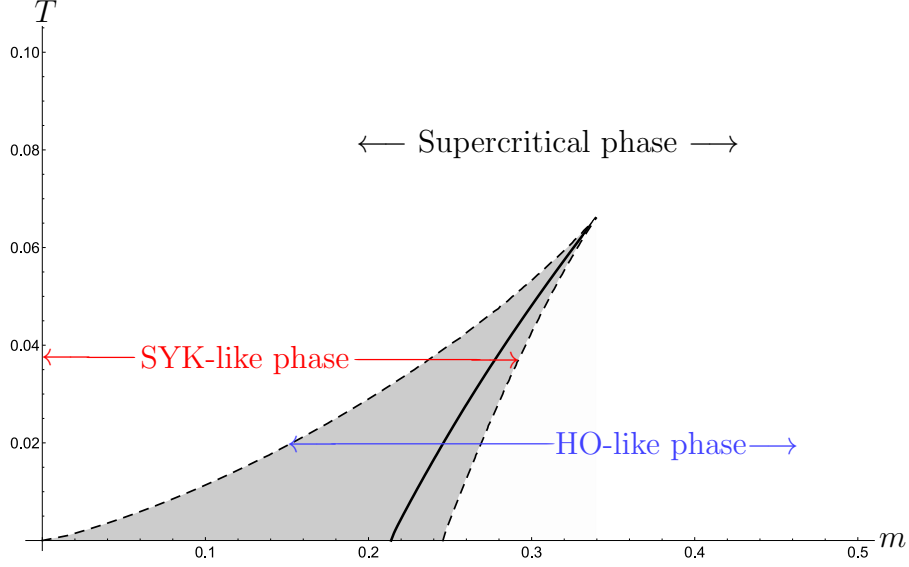


Figure 21: Phase diagram of the $U(n)^2$ model in the (m, T) plane, also in [28]. The dashed lines correspond to $m_-(T)$ (left) and $m_+(T)$ (right), with the triangle-shaped region they delimit (shaded in grey) being where HO-like and SYK-like solutions coexist. The solid line corresponds to $m_t(T)$, the locus where the free energies of the two solutions coincide. It starts at $(T = 0, m = m_* = 0.213_2)$ and terminates at a non-trivial critical point $(T = T_c = 0.067_1, m = m_c = 0.34_1, Q = Q_c = 0.050_1)$. All along this line the system undergoes a first order phase transition.

The phase diagram's structure is governed by the existence of a critical temperature T_c below which two distinct and consistent solutions to the Schwinger-Dyson equations may coexist. The critical temperature itself can be determined numerically using binary search, taking as division criterion the appearance of the hysteresis phenomenon already discussed in Sec. 3.1.3 and exemplified in Fig. 7. The resulting value $T_c = 0.067_1$ is small but finite, and is compatible with theoretical back-of-the-envelope estimates. Having determined the critical temperature, the phase diagram of Fig. 21 was constructed by taking ~ 130 fixed temperatures in the range $T \in [10^{-3}, T_c]$, and sweeping the (m, T) -plane from small to large masses and vice-versa, as explained in Sec. 3.3.

One of the solutions found below the critical temperature T_c is the analytic continuation of the weakly coupled solution at large masses. We shall call this solution

“harmonic oscillator-like”, or simply HO-like solution for short. It can be analytically continued down to a minimum mass $m_-(T)$, but ceases to exist for any mass $m < m_-(T)$. As discussed in Sec. 4.1, at strictly zero temperature this solution matches with the tree-level result, which solves the Schwinger-Dyson equations for all masses, so that $m_-(T=0) = 0$. At finite temperatures, however, we always find $m_-(T) > 0$.

A different solution existing for $T < T_c$ is the SYK solution at $m = 0$, which in turn can be analytically continued to higher masses up to the maximum value $m_+(T)$. This solution is conformal in the IR and retains many of the familiar properties of the SYK model, such as non-vanishing zero-temperature entropy, so we shall call it the SYK-like solution.

The curves $m_-(T)$ and $m_+(T)$ are drawn in Fig. 21 with dashed lines. For any fixed temperature $T < T_c$, there is an interval of masses $m_-(T) < m < m_+(T)$ for which both solutions coexist. Their free energies coincide at a transition mass $m_t(T)$, the HO-like and SYK-like solutions being more stable for $m > m_t(T)$ and $m < m_t(T)$ respectively. The system thus undergoes a phase transition when crossing the line $m_t(T)$, the jump from one solution to the other being discontinuous. The corresponding phase transition is therefore of first order. This is illustrated in Fig. 22, where the thermodynamical potentials are plotted as a function of mass for temperatures $T < T_c$, $T = T_c$ and $T > T_c$.

As we have mentioned at the end of Sec. 4.1, the first order phase transition observed by decreasing m at fixed T , for $T < T_c$, is reminiscent of the phenomenon of gravitational collapse. On the other hand, if we start from very low T in the mass interval $m_* < m < m_c$ and increase T , then we also cross the line of first order phase transitions. In this case, the phenomenon is somehow reminiscent of the Hawking-Page phase transition, where the thermal AdS space becomes thermodynamically unstable compared with the large Schwarzschild-AdS black hole solution [66].

A very important feature of the phase diagram is that the mass interval in which the two solutions coexist shrinks as the temperature is increased, and eventually reduces to the point $m = m_c$ when $T = T_c$. At this point, the two solutions coincide and the transition becomes of second order. Remarkable properties of this critical point and its generalizations will be discussed below in Sec. 4.2.3 and 4.2.4. For $T > T_c$ there is only one solution and the continuation from large to small mass is completely smooth.

In [54], Maldacena and Qi studied a system consisting of two copies of the SYK model coupled via a quadratic “mass” term. The resulting phase diagrams, depicted in Fig. 11 and 18 of Ref. [54], have both similarities and differences with the one described here.

An important similarity is that there are two competing solutions in a region of the (m, T) -plane, with an associated first order phase transition. Moreover, it seems that

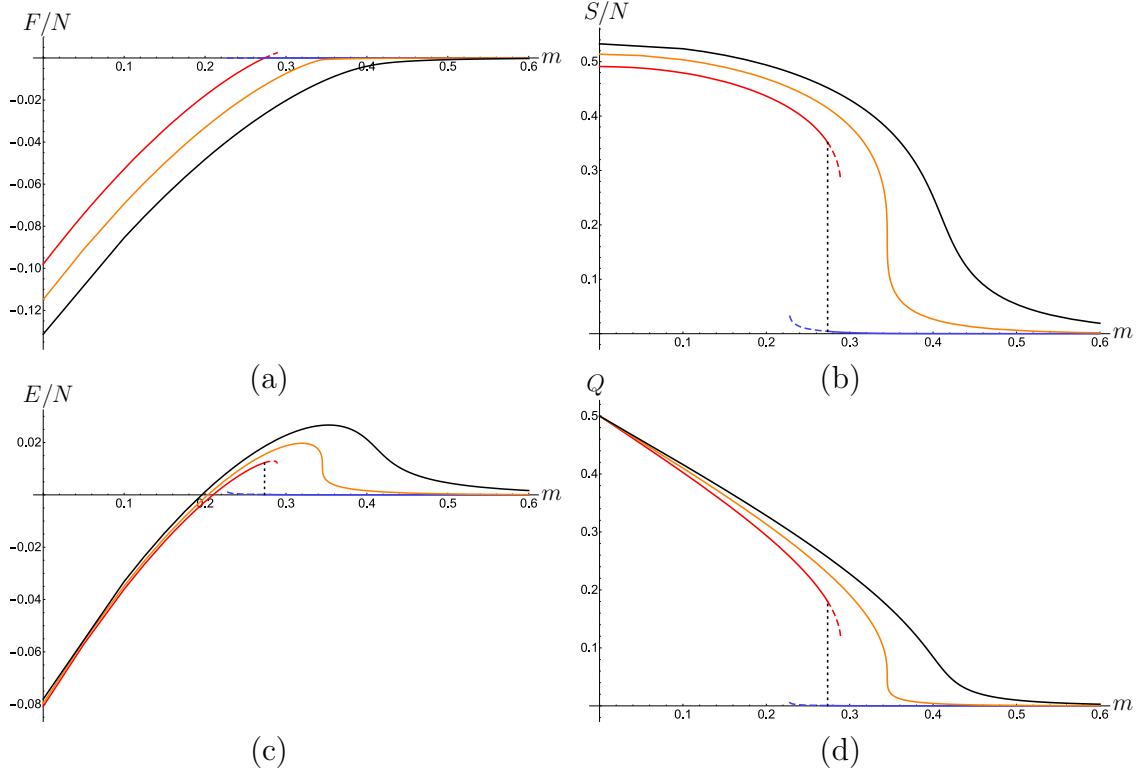


Figure 22: Free energy (a), entropy (b), energy (c) and charge (d) of the $U(n)^2$ model, as functions of the mass, for $\lambda = 1$. Solid lines correspond to the actual physical values associated with the most stable solution, dashed lines to metastable values. For $T < T_c$ (here $T = 0.035$) there are two distinct solutions, one SYK-like (red) and another HO-like (blue), with thermodynamic quantities experiencing finite jumps where their free energies cross (black dotted lines at $m_t = 0.274_1$). For $T = T_c = 0.067_1$ there is a single solution (orange), with a singular behaviour at $m = m_c$. For the supercritical phase at $T > T_c$ (black line for $T = 0.1$ here) all the thermodynamical potentials are smooth.

the line of first order phase transitions terminates at a critical point (see in particular Fig. 18 in [54]), but this critical point is not explicitly mentioned or analyzed in [54]. A detailed discussion of the properties of the critical point found in the $U(n)^2$ model will be provided below in Sec. 4.2.3, 4.2.4 and 5.

On the other hand, a crucial difference is that the mass parameter m in Maldacena-Qi is *relevant*: a mass gap is created as soon as the mass is turned on. All the solutions for the two-point functions in their model are thus gapped when $m \neq 0$, in parallel with our HO-like solutions; in other words, they do not have SYK-like solutions with a non-trivial IR behaviour when $m \neq 0$.

Another interesting aspect of the quadratically coupled SYK model is that it can be fruitfully studied in the large q limit, where q is the order of the interactions. In

this limit a third branch of solutions is found, and allows to interpolate smoothly between the two other types of solutions responsible for the existence of the first order phase transition that we have discussed above. Solutions in this third branch have a negative specific heat. Their existence is irrelevant in the canonical ensemble, but Maldacena and Qi argued that they could smoothen the transition in the micro-canonical ensemble.¹⁹ Note that this novel class of solutions was not found by solving numerically the Schwinger-Dyson equations at finite q , so it might be an artifact of the large q limit, even though this seems unlikely.

The $U(n)^2$ model we are describing here can also be generalized to arbitrary q -body interactions, see App. A and the results in Sec. 4.2.3 and 4.2.4. However, the large q limit, when carried out in a standard way along the lines of [9] and [54], turns out to be rather trivial. In particular, the phase transition is never seen. The reason for this seems to be that the critical masses m_* or m_c , or the critical temperature T_c , become exponentially small with q at large q , instead of being governed by a simple power law. For this reason, we shall not discuss the large q limit any further here. Note also that we have not been able to find numerically any different solution with negative specific heat at finite q either, although we can of course not guarantee it does not exist.

4.2.1 The (m, Q) plane and physics at fixed Q

The phase diagram in the (m, Q) plane is depicted in Fig. 23. One notable feature is the existence of a large forbidden region, shaded in grey. Another notable feature is that for temperatures $T < T_c$ there is an interval of charges that can only be realized if the two phases coexist (vertical dotted lines on the figure).²⁰

In the context of most condensed matter work, it is natural to interpret the mass term as minus a chemical potential, which is adjusted to fix the fermion number Q to any desired value. This is not the point of view adopted in the present paper, where we always consider the mass to be the most natural variable to be fixed. An interesting consequence of our analysis is that for any given temperature $T < T_c$, there is a maximum mass $m_+(T)$, associated with the dashed line on the right in Fig. 21, above which the system cannot exist in the SYK-like phase.

The existence of this maximum mass, and thus of a minimum chemical potential $-m_+(T)$, is a very well established fact. Independently of the numerical results, we can provide the following simple argument. Let us assume that the $m_+(T)$ bound does not exist, so that the SYK-like solution can then be analytically continued

¹⁹A non-trivial issue is that solutions with negative specific heat are thermodynamically unstable in standard set-ups with local interactions and a notion of extensivity; what really happens in the Maldacena-Qi model, which is a disordered model with all-to-all interactions, or in equivalent matrix models, remains an open question.

²⁰Since the model we are considering does not have a notion of space or local interactions, it is unclear how such coexistence might be realized.

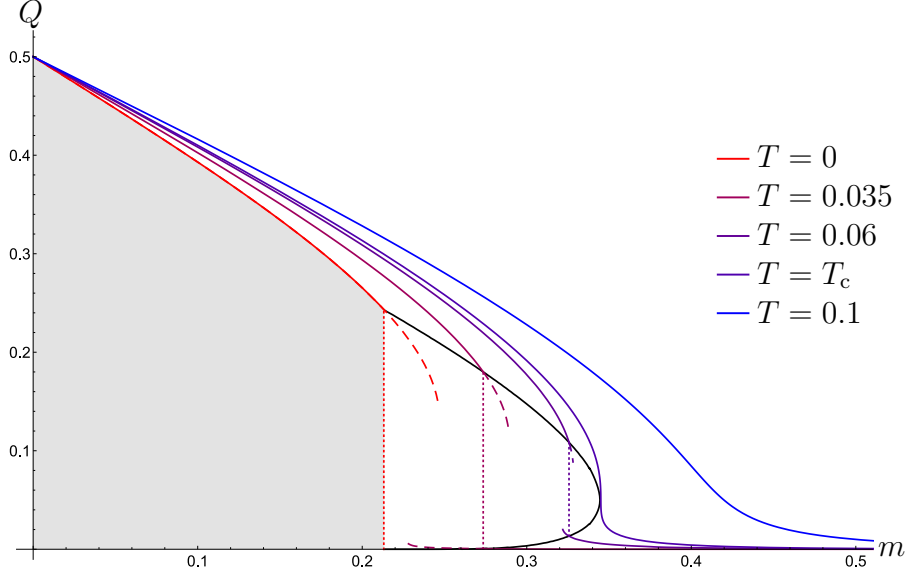


Figure 23: Phase diagram of the $U(n)^2$ model in the (m, Q) plane, also in [28]. We have indicated several isothermal curves, including one for the $T \rightarrow 0$ extrapolation of finite-temperature results (see Sec. 4.2.2 below). Solid lines correspond to thermodynamically stable solutions; dashed lines are used for metastable ones. For $T < T_c$ these curves go through the transition region delimited by the black line, where both phases coexist. The area shaded in grey is a forbidden region.

anywhere in the (m, T) plane. Starting from the grey-shaded region in Fig. 21 where both the SYK and HO-like solutions exist, and performing the analytic continuation of the two solutions to the high-temperature regime along a path that goes around the critical point (m_c, T_c) on its right, we deduce that two distinct solutions must exist in the weakly coupled, high T at fixed m regime. Moreover, these two distinct solutions must have the same UV behaviour $G_k \sim i/\nu_k$ given by (3.18), since this behaviour cannot change under analytic continuation. But this is impossible. At high temperatures, all the Matsubara frequencies are in the UV regime $\nu_k \gg \lambda$ and thus any solution with the standard UV behaviour will necessarily be very near the perturbative solution $G(\tau) \simeq \frac{1}{2} \text{sign}(\tau)$. A unique series expansion for any such solution can be generated from the Schwinger-Dyson equations; this series coincides, of course, with the high-temperature perturbative series. It converges and yields a unique solution, thus proving that the bound $m_+(T)$ necessarily exists.

The numerical results suggest that for the SYK-like phase the lines $Q(m)$ at fixed temperature $T < T_c$ all terminate at a strictly positive charge $Q_{\min}(T) > 0$. This minimum possible charge $Q_{\min}(T)$ corresponds to the endpoints of the upper dashed lines in Fig. 23, and can be computed at any given temperature evaluating $Q(m)$ at $m = m_+(T)$. For instance, we find at zero temperature $Q_{\min}(T = 0) = 0.150_2$. Note that $Q_{\min}$ is of course less than the value of the charge below which the solution

becomes thermodynamically unstable; at $T = 0$, this latter charge corresponds to the upper left endpoint of the black line in Fig. 23 and is found to be approximately $Q_* = Q(T = 0, m = m_*) = 0.243_1$. Our numerical results seem to indicate that $Q_{\min}(T)$ is a decreasing function of T which goes to the value $Q_c = 0.050_1$ at the critical point when $T = T_c$.

However, it is also conceivable that our numerical algorithm may fail to converge for a small interval of masses near $m_+(T = 0)$, for which the slope $|\partial Q/\partial m|$ is large; one may then imagine that the zero-temperature line $Q(m)$ (the red dashed line in Fig. 23) might extend all the way down to zero charge. On the other hand, at non-zero temperature it seems very unnatural to obtain solutions with zero charge and finite chemical potential. Then $Q_{\min}(T)$ should be strictly positive, at least for all $T > 0$, while going to Q_c when $T = T_c$ (since at the critical point the SYK-like solution merges with the HO-like solution).

4.2.2 The physics at zero temperature

It is certainly very interesting to study the phase diagram at zero temperature, but the numerical algorithm used to solve the Schwinger-Dyson equations operates inherently at finite-temperature. Of course, the identification of the two distinct solutions in the $U(n)^2$ model as SYK- and HO-like already constitutes a statement about the low-temperature regime. However, this identification is possible using only very elementary data, such as the value of the entropy (see Fig. 7) or the rough criterion associated to the constraint (3.49). Indeed, as shown in Fig. 24 the value of $\text{Re } \Sigma_{1/2}$ at a small but finite temperature $T = 10^{-3}$ is approximately equal to the mass for the analytic continuation of the $m = 0$ solution, whereas it essentially vanishes for the analytic continuation of the large mass solution.

In order to provide further details about the zero-temperature physics, a careful analysis involving more delicate extrapolations is required. We have already shown in Fig. 12 an example for how such a procedure may be carried out for some small but non-vanishing mass, its results matching the expectations coming from the low-temperature ansatz, (3.46) and (3.47). A more global picture is provided in Fig. 25 by systematically repeating that process for various masses. We clearly see that the agreement is very good all along the range of masses for which the SYK-like phase actually exists.

A further non-trivial check of the extrapolation procedure is made possible by the exact $T = 0$ results for the SYK-like phase reviewed in Sec. 3.2.3. In Fig. 26 we plot the $T \rightarrow 0$ extrapolations of the charge and entropy as functions of the IR parameter θ , along with the exact results (3.105) and (3.111) for comparison. The agreement between numerical and theoretical values is very good for masses up to $m = m_*$ and a bit higher, but starts exhibiting deviations in the entropy for larger masses. We can attribute these discrepancies to numerical artifacts becoming more and more relevant

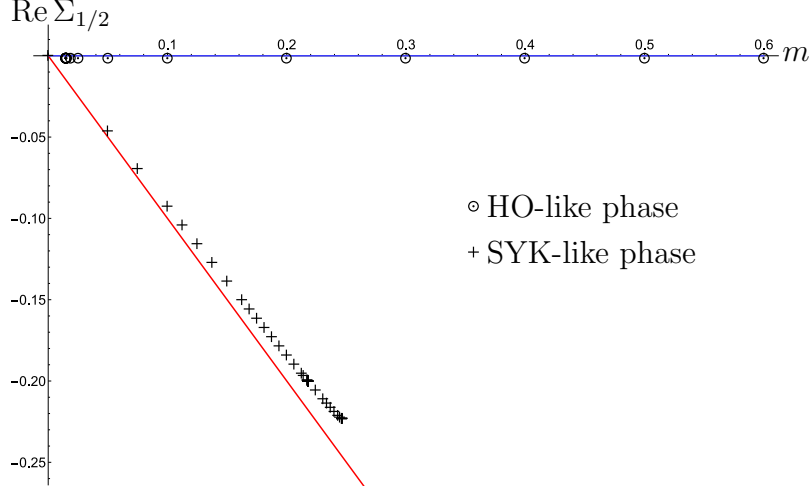


Figure 24: Values of $\text{Re } \Sigma_{1/2} \simeq \tilde{\Sigma}(0)$ as a function of the mass m , for the $U(n)^2$ model at $T = 10^{-3}$ in natural units. Datapoints for the SYK-like phase (crosses) closely follow the constraint (3.49), $\tilde{\Sigma}(0) = m$ (red line), whereas datapoints for the HO-like phase (circles) are compatible with $\tilde{\Sigma}(0) = 0$ (blue line).

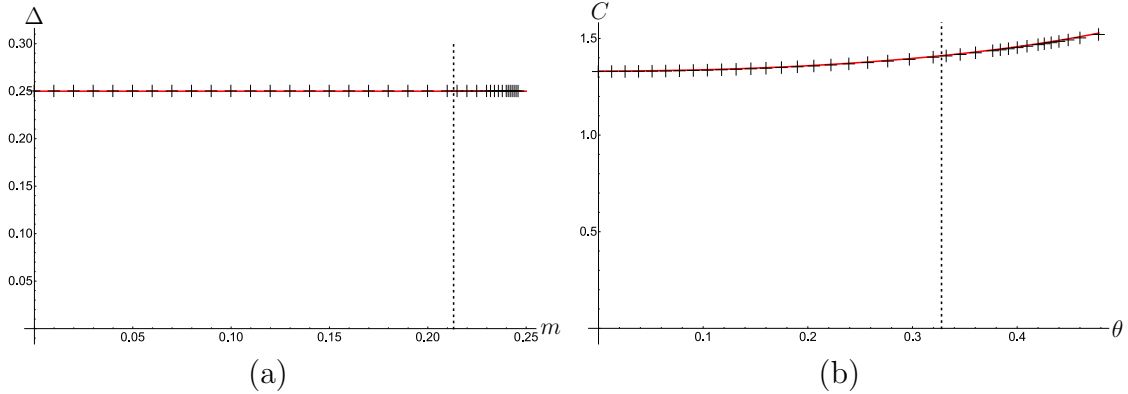


Figure 25: $T \rightarrow 0$ extrapolation of the scaling dimension Δ (a) and the IR parameter C (b), as functions of the mass and the IR parameter θ , respectively. Numerical values were computed using the finite-temperature fitting procedure illustrated in Fig. 12, with linear extrapolations to $T \rightarrow 0$ (represented with crosses) performed using ~ 20 points with $\beta \in [500, 1000]$. The vertical dotted lines mark the zero-temperature phase transition values $m_* = 0.213_2$ and $\theta_* = \theta(m_*) = 0.328_2$, for (a) and (b) respectively. The scaling dimension matches the $\Delta = 1/4$ value of (3.46), whereas the IR parameters C and θ defined by (3.43) satisfy the constraint $C^4 \cos(2\theta) = \pi$ of (3.47) (red lines in each plot).

as the mass increases, the entropy being very sensitive to the truncation implemented by the frequency-cutoff of the numerical algorithm (see App. C). In any case, these demonstrations of the precision of the numerical extrapolation used to implement the

$T \rightarrow 0$ limit give us confidence that our results are very robust at least up to $m \sim m_*$.

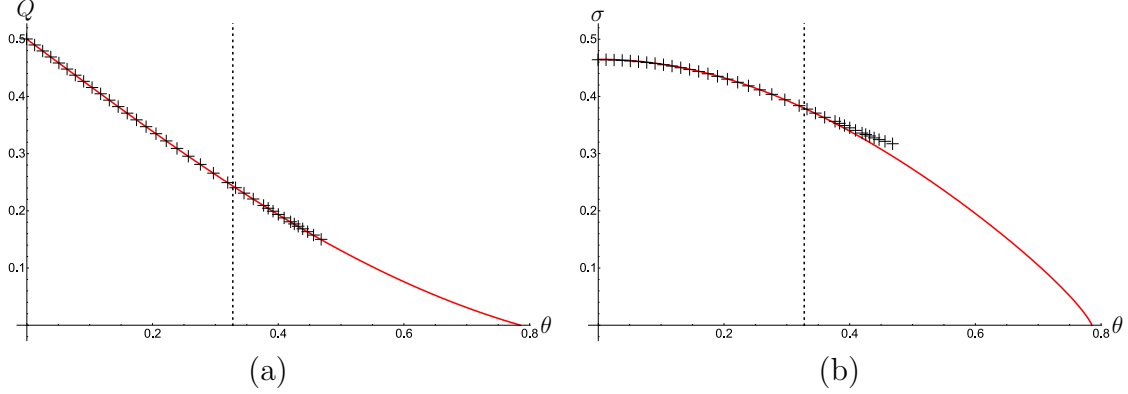


Figure 26: Zero-temperature charge (a) and entropy (b) as functions of the IR parameter θ defined by (3.43). Numerical values are represented with crosses, solid red lines correspond to the formulas (3.105) and (3.111) for (a) and (b), respectively. Extrapolations were performed as explained in Fig. 25, but taking $\beta \in [50, 100]$ for S/N . The vertical dotted line marks the value $\theta_* = \theta(m_*) = 0.3282$.

Moving on, in Fig. 27, we plot the zero-temperature energy, entropy and specific heat divided by T ,

$$\epsilon = \frac{1}{N} \lim_{T \rightarrow 0} E, \quad \sigma = \frac{1}{N} \lim_{T \rightarrow 0} S, \quad c = \frac{1}{N} \lim_{T \rightarrow 0} \frac{C}{T} = \frac{1}{N} \lim_{T \rightarrow 0} \frac{\partial S}{\partial T}, \quad (4.4)$$

as functions of the mass, for the SYK-like solution and the most stable solution. Note that all these quantities are trivially vanishing for the HO-like solution, since it coincides with the tree-level result at zero temperature. The thermodynamical potentials can be expanded in terms of ϵ , σ and c at low temperature as

$$\frac{E}{N} = \epsilon + \frac{1}{2}cT^2 + o(T^2), \quad \frac{S}{N} = \sigma + cT + o(T), \quad \frac{F}{N} = \epsilon - \sigma T - \frac{1}{2}cT^2 + o(T^2). \quad (4.5)$$

The results are consistent with the general discussion in Sec. 4.1. The ground state jumps from the Fock vacuum at $m > m_* = 0.213_2$ to a highly degenerate state at $m < m_*$. The heat capacity c grows from its $c(m=0) = 0.80_1$ value matching the result reported for SYK [9], and experiences a finite jump to zero at $m = m_*$. While the extrapolation for this quantity should certainly experience similar numerical artifacts to those present in the entropy (see Fig. 26b), it seems plausible that c will diverge when $m \rightarrow m_+(T=0)$. We also plot in Fig. 27d the effective mass obtained from the exponential decay of the HO-like solution, extrapolated to zero temperature, finding perfect agreement with the $m_{\text{eff}} = m$ prediction. This result will serve as a baseline reference when studying the $U(n)$ model in the following.

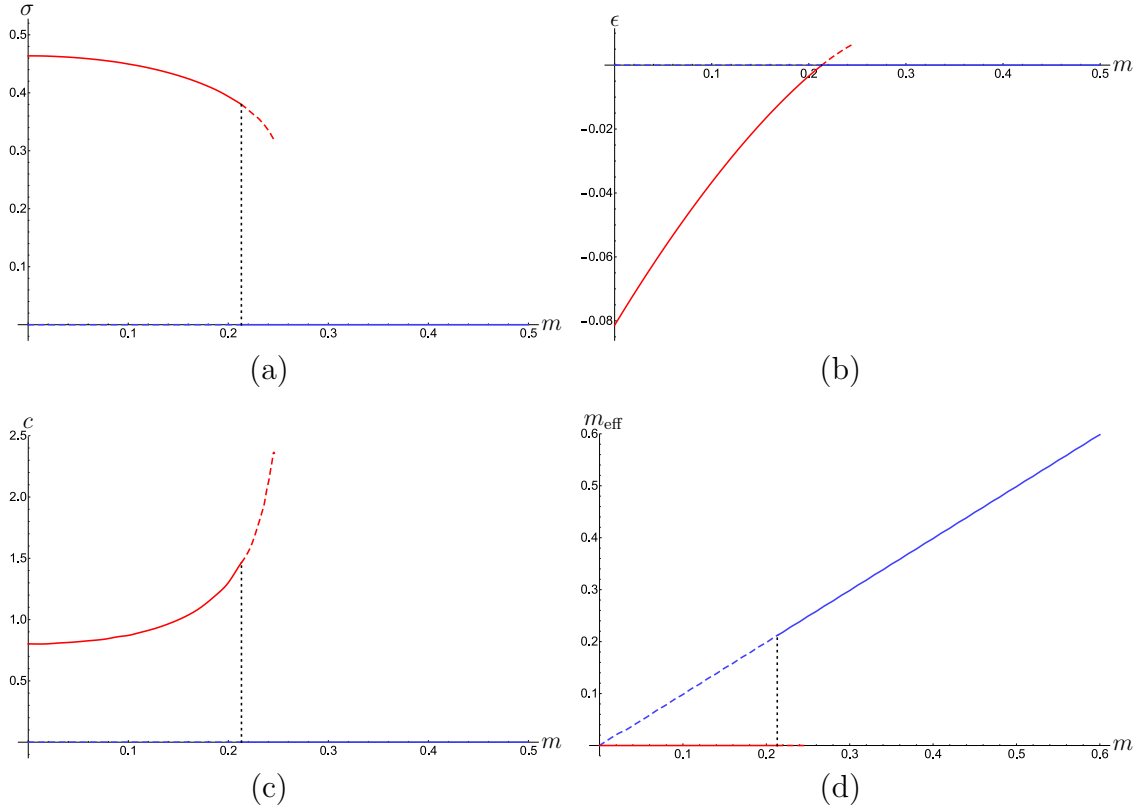


Figure 27: Plots of the zero-temperature entropy σ (a), energy ϵ (b), specific heat coefficient c defined in (4.4) (c) and effective mass m_{eff} (d), as functions of the mass. Solid lines correspond to the thermodynamically stable solution, which is the SYK-like solution for $m < m_*$ (red) and the HO-like solution for $m > m_*$ (blue). Values for metastable solutions (the SYK-like solution for $m > m_*$, the HO-like one for $m < m_*$) are represented with dashed lines. Vertical black dotted lines denote a finite jump in the corresponding functions. All extrapolations are performed as in Fig. 25 and Fig. 26.

Remark: There is no contradiction between the fact that the mass gap m_{eff} jumps discontinuously from $m_{\text{eff}} = m$ for $m > m_*$ to $m_{\text{eff}} = 0$ for $m < m_*$ (Fig. 27d) and the fact that the zero-temperature energy is a continuous function of m at $m = m_*$ (Fig. 27b). This is a large N effect: the eigenstates of the Hamiltonian responsible for the small negative energy ϵ when m is just below m_* do not contribute to the large N two-point function when $m > m_*$.

4.2.3 The q -generalizations and critical exponents

The $U(n)^2$ model admits a straightforward generalization described in App. A, for which the quartic interaction term in the Hamiltonian is replaced by a sum of terms

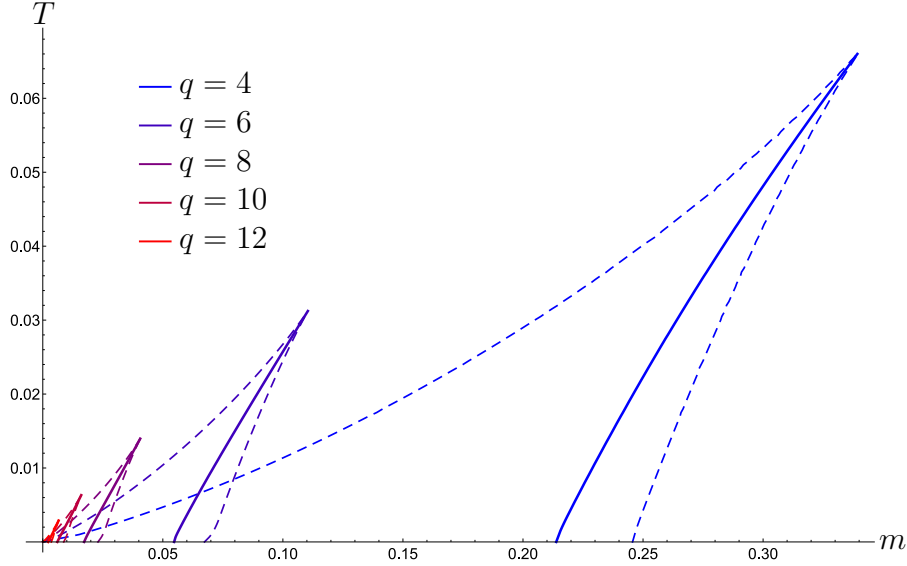


Figure 28: Superimposition of the phase diagram of the q -generalizations of the $U(n)^2$ model in the (m, T) -plane, in the units $\lambda_q = 1$, for $q = 4, 6, 8, 10, 12$. Solid lines correspond to first order phase transitions, dashed lines delimit the regions of coexistence of the SYK-like and HO-like solutions in each model.

of even degrees. In this subsection we restrict our attention to the case where only one such term is included, with degree $q \geq 4$. The qualitative physics of the models is then essentially the same for all values of q , and the discussion presented in the previous subsections for the case $q = 4$ can be straightforwardly generalized. This is illustrated in Fig. 28, in which the phase diagram for the original $q = 4$ model, taken from Fig. 21, is superimposed to those of the $q = 6, 8, 10, 12$ generalizations. In each case, there is a line of first order phase transitions that terminates at a critical point $(m_c^{(q)}, T_c^{(q)})$, with the zero-temperature phase transition occurring at $m = m_*^{(q)}$. The phase diagrams are all very similar, with the transition line shrinking to an ever smaller region in the (m, T) -plane as q is increased (in the natural units $\lambda_q = 1$). Values of $m_c^{(q)}$, $T_c^{(q)}$ and $m_*^{(q)}$ for the first few q 's are presented in Table 1, along with the charge at the critical point $Q_c^{(q)}$.

The q -body interacting generalizations of the SYK model and closely related models have been considered in the literature (see *e.g.* [9, 54, 67]) because the large q limit can be solved analytically and captures interesting aspects of the physics of the models at finite q . However, as we have already mentioned in our comments on the Maldacena-Qi model in Sec. 4.2, the large q limit of the $U(n)^2$ model turns out to be rather trivial when defined in the standard way. In particular it is not able to capture the line of first order phase transitions or the associated critical point, which are the most interesting features of the model. Trying to understand better the large q behaviour in this model, as well as possible q -body generalizations of the $U(n)$ model, could be an interesting direction for future research.

q	$m_*^{(q)}$	$m_c^{(q)}$	$T_c^{(q)}$	$Q_c^{(q)}$
4	0.213 ₂	0.34 ₁	0.067 ₁	0.050 ₁
6	0.055 ₁	0.111 ₂	0.0317 ₅	0.0996 ₁
8	0.017 ₁	0.0412 ₆	0.0143 ₄	0.1343 ₂
10	0.006 ₁	0.0163 ₂	0.0065 ₂	0.1612 ₂
12	0.0022 ₅	0.0068 ₂	0.0030 ₁	0.1837 ₇

Table 1: Values of the phase diagram parameters for the q -generalizations of the $U(n)^2$ model, for the first few values of q , in the units $\lambda_q = 1$. The first order phase transition line starts at $T = 0$ with $m = m_*^{(q)}$ and terminates at the critical point $(m_c^{(q)}, T_c^{(q)})$ with charge $Q_c^{(q)}$.

For our purposes, the main interest of considering the q -generalizations of the $U(n)^2$ model introduced in App. A is that we can study the properties of the critical point for different values of q . There are various natural observables one may consider; we choose to focus on the following quantities:

i) The heat capacity

$$C = T \frac{\partial S}{\partial T}, \quad (4.6)$$

as a function of temperature at $m = m_c^{(q)}$. We shall compute the right and left critical exponents $\alpha_{\pm}^{(q)}$ defined by

$$C(T, m_c^{(q)}) \underset{T \rightarrow T_c^{(q)\pm}}{\sim} \frac{C_{\pm}^{(q)}}{|T - T_c^{(q)}|^{\alpha_{\pm}^{(q)}}}, \quad (4.7)$$

which govern the singular behaviour of C when T approaches the critical value $T_c^{(q)}$ from above ($\alpha_+^{(q)}$ for $T > T_c^{(q)}$) or from below ($\alpha_-^{(q)}$ for $T < T_c^{(q)}$).

ii) The susceptibility

$$\chi = \frac{\partial Q}{\partial m}, \quad (4.8)$$

as a function of temperature at $m = m_c^{(q)}$, with right and left critical exponents $\gamma_{\pm}^{(q)}$ defined by

$$\chi(T, m_c^{(q)}) \underset{T \rightarrow T_c^{(q)\pm}}{\sim} \frac{\chi_{\pm}^{(q)}}{|T - T_c^{(q)}|^{\gamma_{\pm}^{(q)}}}. \quad (4.9)$$

iii) The charge Q as a function of temperature at $m = m_c^{(q)}$, with right and left critical exponents $q_{\pm}^{(q)}$ defined by

$$Q(T, m_c^{(q)}) - Q(T_c^{(q)}, m_c^{(q)}) \underset{T \rightarrow T_c^{(q)\pm}}{\sim} Q_{\pm}^{(q)} |T - T_c^{(q)}|^{q_{\pm}^{(q)}}. \quad (4.10)$$

- iv) The charge difference ΔQ between the two phases along their coexistence line, as a function of temperature. This quantity is of course defined only for $T < T_c^{(q)}$, and the associated critical exponent $\beta^{(q)}$ is defined by

$$\Delta Q(T) \underset{T \rightarrow T_c^{(q)-}}{\sim} \Delta Q_-^{(q)} (T_c^{(q)} - T)^{\beta^{(q)}}. \quad (4.11)$$

Note that we could also consider the same quantity as a function of mass, but since $m_t^{(q)}(T)$ is always smooth the associated critical exponent is also $\beta^{(q)}$, a fact that we have verified numerically.

- v) The entropy S as a function of mass at $T = T_c^{(q)}$, with right and left critical exponents $s_{\pm}^{(q)}$ defined by

$$S(T_c^{(q)}, m) - S(T_c^{(q)}, m_c^{(q)}) \underset{m \rightarrow m_c^{(q)\pm}}{\sim} S_{\pm}^{(q)} |m - m_c^{(q)}|^{s_{\pm}^{(q)}}. \quad (4.12)$$

- vi) The charge Q as a function of mass at $T = T_c^{(q)}$, with right and left critical exponents $\tilde{q}_{\pm}^{(q)}$ defined by

$$Q(T_c^{(q)}, m) - Q(T_c^{(q)}, m_c^{(q)}) \underset{m \rightarrow m_c^{(q)\pm}}{\sim} \tilde{Q}_{\pm}^{(q)} |m - m_c^{(q)}|^{\tilde{q}_{\pm}^{(q)}}. \quad (4.13)$$

To obtain all of these critical exponents we proceed in a way which is very similar to what we did before for the scaling dimension. For example, to extract $q_{\pm}^{(q)}$ we may plot

$$q^{(q)} = \frac{\partial \ln |Q(T, m_c^{(q)}) - Q(T_c^{(q)}, m_c^{(q)})|}{\partial \ln |T - T_c^{(q)}|} \quad (4.14)$$

as a function of $|T - T_c^{(q)}|$, for both $T > T_c^{(q)}$ and $T < T_c^{(q)}$, see Fig. 29a. We observe an approximately linear behaviour of $q^{(q)}$ within a fitting window where T is close, but not too close, to the critical temperature $T_c^{(q)}$. This allows for a reliable extrapolation to $T \rightarrow T_c^{(q)}$, giving $q_+^{(q)}$ and $q_-^{(q)}$ when extrapolating $T > T_c^{(q)}$ and $T < T_c^{(q)}$, respectively. Having fixed the exponents $q_{\pm}^{(q)}$ we can finally fit the $Q_{\pm}^{(q)}$ coefficients in (4.10) directly, again using an appropriate fitting window, as shown in Fig. 29b.

The example of this procedure illustrated in Fig. 29 for $q = 4$ is entirely typical, with other exponents and q -generalizations behaving analogously. The results for $q = 4, 6, 8, 10, 12$ are listed in Table 2.

Let us emphasize three important properties of these results.

First, the critical exponents that we find are not mean-field. In other words, the critical points are non-trivial and strongly coupled. This may appear as a surprise,

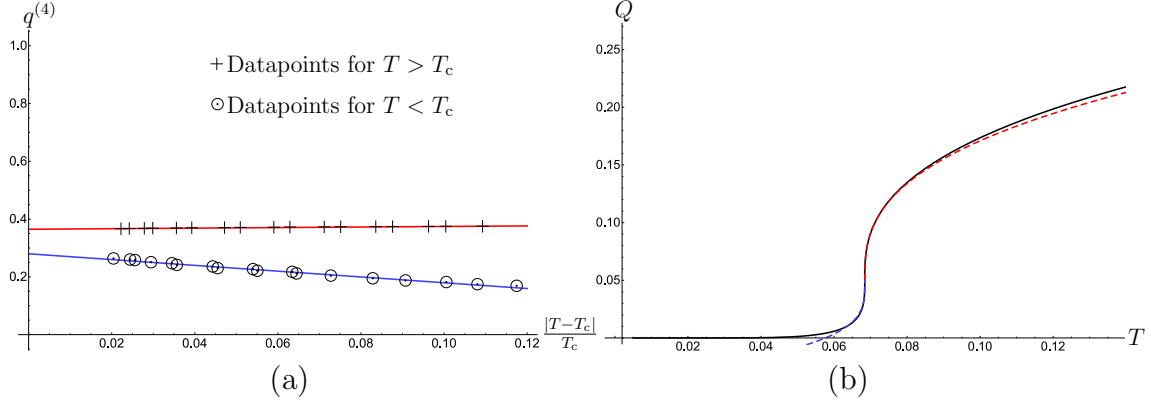


Figure 29: (a) Plot of $q^{(4)} = \partial \ln |Q(T, m_c^{(4)}) - Q(T_c^{(4)}, m_c^{(4)})| / \partial \ln |T - T_c|$ as a function of $|T - T_c|/T_c$, for both $T > T_c$ (crosses) and $T < T_c$ (circles). The critical exponents $q_{\pm}^{(4)}$ are defined by $q_{\pm}^{(4)} = \lim_{T \rightarrow T_c^{\pm}} q^{(4)}(T)$, implemented numerically through linear extrapolations (red and blue lines for $T > T_c$ and $T < T_c$, respectively). (b) Charge Q at $m = m_c^{(4)}$, as a function of temperature, together with the best power-law fits near $T = T_c^{(4)}$ (red dashed line for $T > T_c^{(4)}$ and blue dashed line for $T < T_c^{(4)}$).

	$q = 4$	$q = 6$	$q = 8$	$q = 10$	$q = 12$
$\alpha_+^{(q)}$	0.68 ₂	0.68 ₂	0.67 ₂	0.69 ₂	0.69 ₂
$\alpha_-^{(q)}$	0.66 ₂	0.62 ₁	0.59 ₁	0.60 ₁	0.59 ₁
$\gamma_+^{(q)}$	0.51 ₁	0.59 ₁	0.54 ₁	0.57 ₂	0.57 ₁
$\gamma_-^{(q)}$	0.80 ₁	0.74 ₂	0.86 ₂	0.80 ₁	0.83 ₁
$q_+^{(q)}$	0.36 ₁	0.36 ₁	0.37 ₁	0.36 ₁	0.35 ₁
$q_-^{(q)}$	0.28 ₂	0.32 ₂	0.36 ₁	0.36 ₁	0.36 ₁
$\beta^{(q)}$	0.50 ₂	0.52 ₂	0.54 ₂	0.53 ₂	0.53 ₂
$s_+^{(q)}$	0.28 ₂	0.33 ₁	0.37 ₁	0.37 ₁	0.37 ₁
$s_-^{(q)}$	0.36 ₁	0.34 ₁	0.34 ₁	0.33 ₁	0.32 ₁
$\tilde{q}_+^{(q)}$	0.22 ₂	0.27 ₂	0.30 ₁	0.32 ₂	0.32 ₂
$\tilde{q}_-^{(q)}$	0.43 ₁	0.40 ₁	0.40 ₁	0.38 ₁	0.37 ₂

Table 2: Various critical exponents in the q -generalizations of the $U(n)^2$ model, for $q = 4, 6, 8, 10, 12$. For the purposes of calculating these values, we have taken $(m_c^{(4)}, T_c^{(4)}) = (0.3445, 0.0685)$, $(m_c^{(6)}, T_c^{(6)}) = (0.112, 0.0323)$, $(m_c^{(8)}, T_c^{(8)}) = (0.0417, 0.0147)$, $(m_c^{(10)}, T_c^{(10)}) = (0.0164, 0.0066)$ and $(m_c^{(12)}, T_c^{(12)}) = (0.0068, 0.0030)$. Reported error bars correspond to typical variations when adjusting the fitting window, see App. C. Further variations can be expected when approaching the critical points more precisely, but general properties and trends should however be robust.

since the whole formalism in which the solution of the models is discussed is mean-field-like: all the thermodynamical properties are derived from the classical effective

actions that we extremize, (3.86) and its q -body generalization (A.2). The crucial difference with the standard mean-field theory is that these effective actions depend on *functions* and not simply on a finite number of variables. We see that this functional version of mean-field theory yields a much richer theory of critical phenomena than Landau's, including the possibility for non-trivial critical exponents.

Second, the critical exponents are asymmetric, for instance $\alpha_+^{(q)} \neq \alpha_-^{(q)}$, $\tilde{q}_+^{(q)} \neq \tilde{q}_-^{(q)}$ *etc.* This is an unusual property [68]. In ordinary local quantum field theory, standard renormalization group arguments imply scaling forms near the critical points which ensure that the critical exponents must be symmetric (the constants of proportionality, like $C_\pm^{(q)}$ in (4.6), may however be asymmetric). This general result does not contradict our findings, since the models we are dealing with here have no notion of space or locality.

Third, the critical exponents are q -dependent, especially for lower values of q . We thus obtain an infinite series of inequivalent critical points. We shall argue in the next subsection that these points actually belong to a continuous line of inequivalent critical points.

Clearly, we are here just scratching the surface of a very non-trivial new phenomenon in large N quantum mechanical models. The critical points we are dealing with remain to a large extent very mysterious, and further study is needed to better understand their properties. In particular, we do not know if some version of the renormalization group formalism can be applied in the present context.

4.2.4 Multi- q models, domain walls and a line of critical points

When turning on several couplings λ_q at the same time, one expects to find an even richer structure of critical and possibly multi-critical points in the corresponding phase diagrams. A detailed study is outside the scope of the present work, but we want to display here two interesting features.

RG flow and domain walls One way to visualize the renormalization group flow from the asymptotically free regime $\omega \gg \lambda_q$ to the conformal region $\omega \ll \lambda_q$ at low masses is to plot in log-log scale the imaginary part of the Matsubara coefficients $\text{Im } G_k$ as a function of the Matsubara frequencies ν_k .

This is illustrated in Fig. 30a for the standard case of the $U(n)^2$ model, where only $\lambda_4 = 1$ is turned on and we take $m = 0$ for simplicity. The structure of the RG flow is then clearly visible: we have at large frequencies a -1 slope characteristic of the asymptotic regime (3.18), $G_k \sim i/\nu_k$; and a slope of $2\Delta - 1 = -1/2$ determined by (3.44) in the conformal region at low frequencies. The transition between the two regimes occurs at frequencies of order $\omega \sim \lambda_4 = 1$ and is very visible at the small but finite temperature $T = 10^{-4}$ used here.

Now consider the case shown in Fig. 30b, where an additional coupling, chosen to be $\lambda_8 = 200$ for illustrative purposes, is turned on. There is a deep-UV region $\nu_k \gg \lambda_8$ again governed by asymptotic freedom, as well as a deep-IR region $\nu_k \ll \lambda_4$ governed by the conformal behaviour associated with the most relevant coupling λ_4 . Moreover, between them now lies an intermediate region $\lambda_4 \ll \nu_k \ll \lambda_8$, which is approximately conformal with a slope $2\Delta - 1 \simeq -0.75$, corresponding to the dimension associated with the $q = 8$ interaction term, $\Delta \simeq 1/8$.

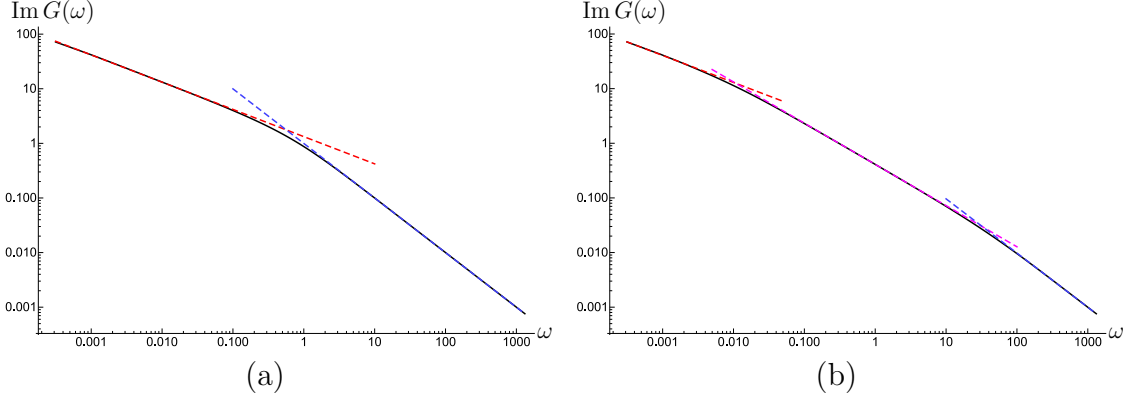


Figure 30: Log-log plot of $\text{Im } G(\omega)$ as a function of the frequency ω at $T = 10^{-4}$ and $m = 0$ in the $U(n)^2$ model (a) and in the model with both $\lambda_4 = 1$ and $\lambda_8 = 200$ turned on (b). The RG flow from the asymptotically free region at high frequencies (-1 slope, blue) to the conformal behaviour dominated by the most relevant quartic coupling at low frequencies ($-1/2$ slope, red) is clearly visible in both cases. In (b) the transition occurs through an intermediate regime dominated by the $q = 8$ interaction term (-0.75 slope, magenta).²²

Note that we are using a model with λ_4 and λ_8 turned on, instead of, for example, λ_4 and λ_6 . This is because the larger difference in the associated exponents Δ makes the intermediate approximately conformal region easier to visualize when one works at small yet finite temperature and ratio λ_4/λ_8 , a limitation inherent to the numerical approach. More generally, one may turn on many couplings $\lambda_{q_1}, \dots, \lambda_{q_r}$, organized hierarchically with $q_{i-1} < q_i$ and $\lambda_{q_{i-1}} \ll \lambda_{q_i}$. In the frequency regime $\lambda_{q_{i-1}} \ll \omega \ll \lambda_{q_i}$ the model is then in an approximate conformal phase dominated by the coupling λ_{q_i} and characterized by an exponent $\Delta = 1/q_i$. In a holographic picture, along the lines of [69], this behaviour should be mapped to interesting domain wall solutions interpolating between different AdS_2 spaces. To the best of our knowledge, these solutions have not yet been discussed in the literature.

²²Both plots in Fig. 30 were made using solutions computed with a weighting factor $\alpha = 0.1$ and $k_{\text{UV}} = 2^{21}$, *i.e.* approximately 4×10^6 frequencies.

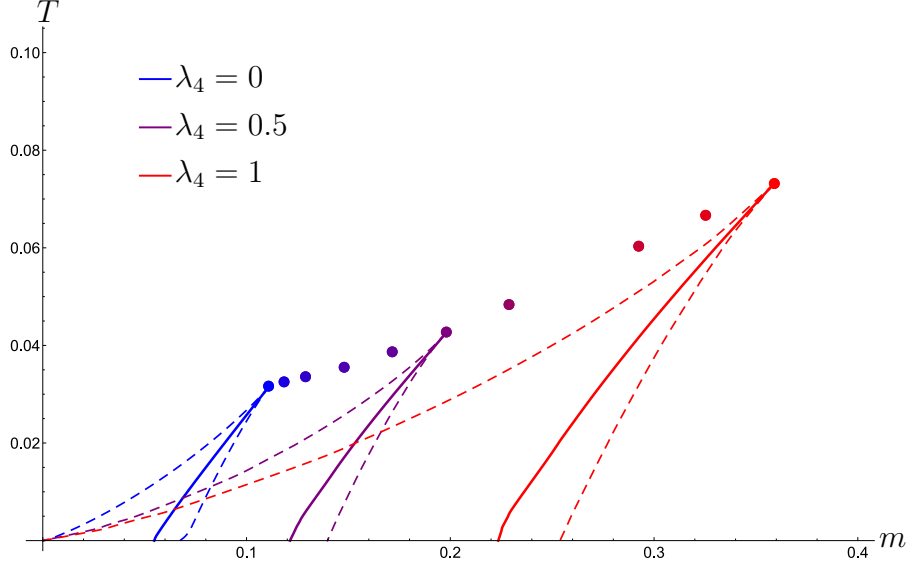


Figure 31: Phase diagram of the model with two couplings λ_4 and $\lambda_6 = 1$ turned on, for three different values of λ_4 . The critical points for $\lambda_4 = 0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1$ are marked with colored circles from blue to red, illustrating the continuous nature of the phase diagram dependence on the λ_4 coupling.

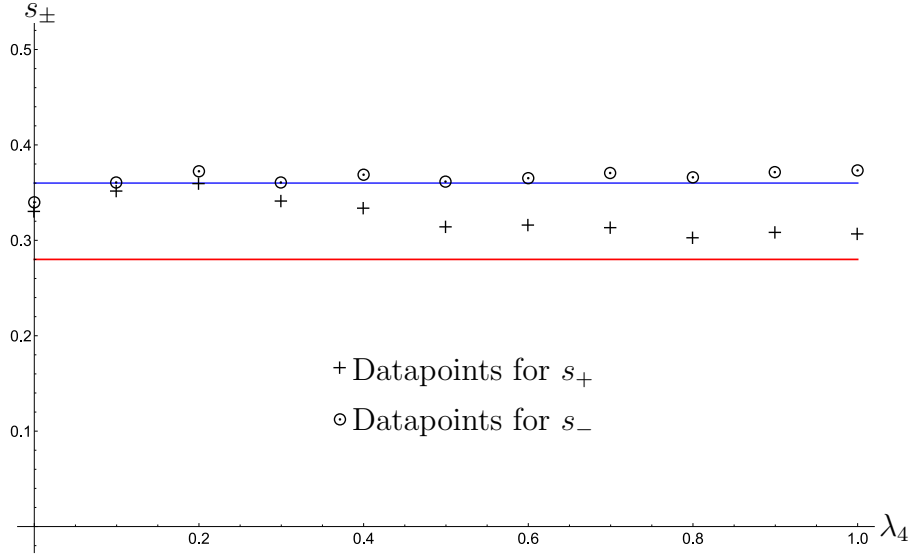


Figure 32: Critical exponents $s_{\pm}$ for the entropy as a function of mass at $T = T_c$, for various values of λ_4 in the model where both λ_4 and $\lambda_6 = 1$ are turned on. The continuous variation of the exponents implies that we have a line of inequivalent critical points with marginal coupling λ_4 . Solid lines mark the values corresponding to $\lambda_4 \gg \lambda_6$ (red for s_+ and blue for s_-).

A line of inequivalent critical points In multi- q models we expect that the phase diagram smoothly depends on the various couplings, so that we have in fact critical lines, surfaces or hyper-surfaces when considering models with two, three or more interaction terms, respectively.

One of the simplest cases to analyze corresponds to the model in which both λ_4 and λ_6 are turned on, with $\lambda_6 = 1$ setting the units. Its phase diagram is depicted in Fig. 31, projecting different values of λ_4 to the (m, T) -plane. We see that there is a critical line interpolating between the critical point of the $q = 6$ model when $\lambda_4 = 0$ and the critical point of the $q = 4$ model when $\lambda_4 \gg \lambda_6$. This is nicely confirmed by our analysis of the critical exponents. For example, in Fig. 32 we show the entropy exponents $s_{\pm}$, computed as explained for $q_{\pm}$ in Fig. 29, as a function of λ_4 . We observe a continuous variation between the values $s_+^{(6)} = 0.33_1$, $s_-^{(6)} = 0.34_1$ and $s_+^{(4)} = 0.28_2$, $s_-^{(4)} = 0.36_1$ (see Table 2), with the gap $s_- - s_+$ smoothly widening as λ_4 becomes larger. This provides convincing evidence that we have a continuous line of inequivalent critical points, for which the coupling λ_4 is exactly marginal.

4.3 Phase diagram of the $U(n)$ model

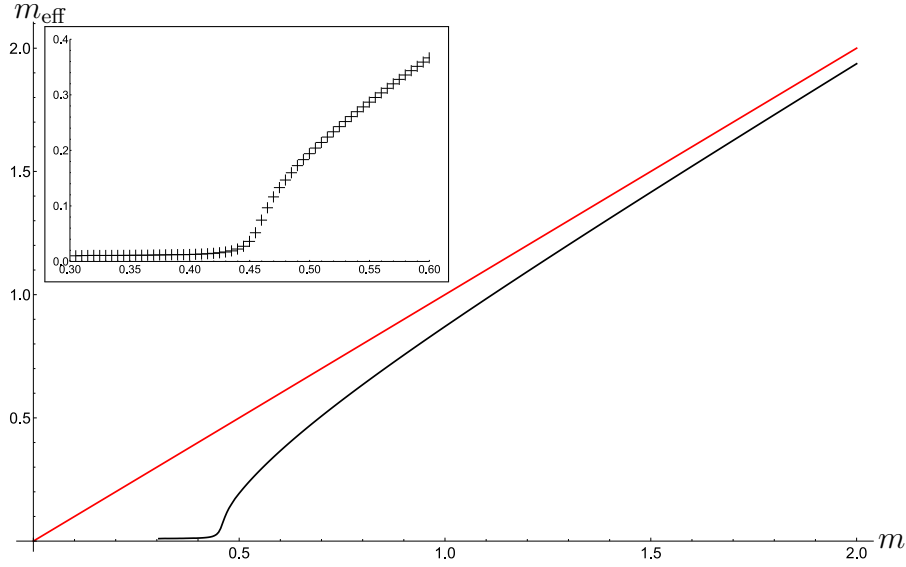


Figure 33: Zero-temperature extrapolation of the mass gap in the $U(n)$ model (black) as a function of the mass parameter m . Numerical values were computed fitting the log-plots of finite-temperature Euclidean two-point functions, linearly extrapolating results for ~ 10 points in the range $\beta \in [50, 100]$. The mass gap decreases from its asymptotic value $m_{\text{eff}} \simeq m$ at large mass (red) to zero for a finite critical value $m = m_* = 0.44_1$ (detail shown in inset, with numerical values represented with crosses).

Unlike the case of the $U(n)^2$ model that we have discussed in the previous Section, the thermodynamics of the $U(n)$ model is governed by a unique solution to the Schwinger-Dyson equations that can be analytically continued for all values of m and T without undergoing monodromies. In particular, we do not find first order phase transitions or non-trivial critical points at finite temperature in the $U(n)$ model. All the interesting physics occurs at zero, or near zero, temperature. As discussed in Sec. 4.1, this important difference with the $U(n)^2$ model is made possible by the fact that the zero-temperature Euclidean two-point function gets non-trivial corrections even at relatively large masses.

More precisely, at large enough mass the system is weakly coupled. The ground-state must then coincide with the Fock vacuum $|0\rangle$ defined by the conditions $\psi_\mu|0\rangle = 0$, with a mass gap m_{eff} of order m . The spectral function (3.9) then has a support only for $\omega \geq m_{\text{eff}} > 0$ at zero temperature, and the spectral decomposition formula (3.12) implies that

$$G(\tau; T=0) = 0 \quad \text{for } \tau < 0, \quad (4.15)$$

$$G(\tau; T=0) \propto e^{-m_{\text{eff}}\tau} \quad \text{for } \tau \rightarrow +\infty. \quad (4.16)$$

In Fig. 33 we plot the mass gap m_{eff} as a function of the mass parameter m , evaluated numerically by matching finite-temperature solutions with the exponential behaviour (4.16) and then extrapolating m_{eff} to $T \rightarrow 0$. The essential feature to note here is that the gap closes for a finite value $m = m_* = 0.44_1$.

A complementary point of view is provided by the constraint (3.49). In Fig. 34 we plot $\text{Re} \Sigma_{1/2}$ as a function of the mass at $T = 0.01$ in natural units. We observe here a clear crossover between the regime at low masses approximately satisfying (3.49), and the regime at large masses for which $\tilde{\Sigma}(0) \simeq 0$. The turning point is at $m = m_* = 0.44_1$, perfectly consistent with the appearance of the effective mass discussed above.

The point $m = m_*$ is a new kind of large N , strongly-coupled quantum critical point. For $m < m_*$ the zero-temperature Euclidean two-point function behaves as (3.55), with asymmetric exponents $\Delta_+ \neq \Delta_-$. This feature was already discussed in detail in Sec. 3.3.2, and is in sharp contrast with the IR behaviour of the two-point function in the $U(n)^2$ model. As pointed out in Sec. 3.1.5, the $SL(2, \mathbb{R})$ subgroup of the low energy reparameterization symmetry (3.38)-(3.39), including the scale transformations $\tau \mapsto \alpha\tau$ for $\alpha > 0$, is now spontaneously broken, even though the phase is gapless. It is thus a phenomenon of spontaneous scale symmetry breaking without the generation of a mass gap.

Plots of the energy, entropy and heat capacity divided by T are provided in Fig. 35. At zero temperature these quantities are all vanishing for $m > m_*$, but become non-zero for $m < m_*$, very much like in the $U(n)^2$ model (compare with Fig. 27). The low-temperature properties of the $U(n)^2$ and $U(n)$ models are thus quite similar as far as the thermodynamical potentials are concerned. For example, they both have

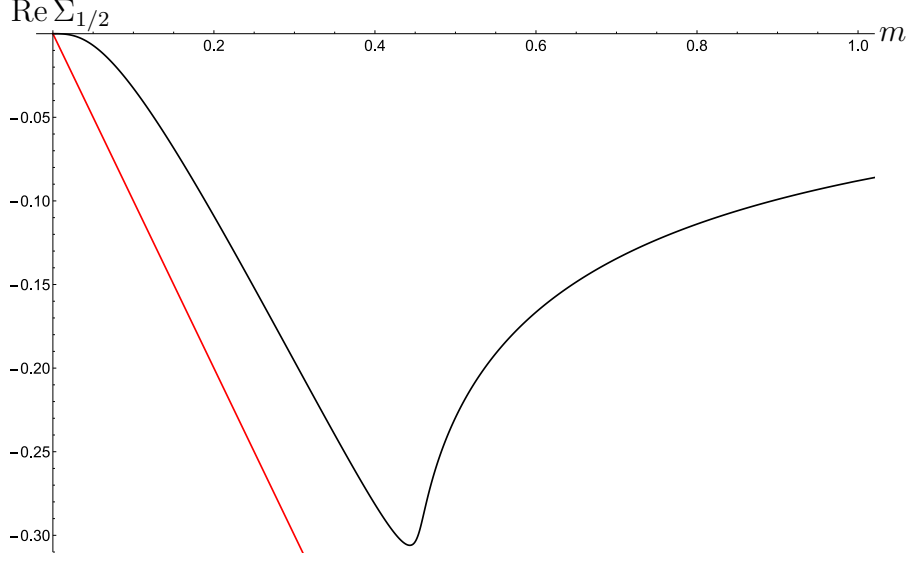


Figure 34: Plot of $\text{Re } \Sigma_{1/2} \simeq \tilde{\Sigma}(0)$ as a function of the mass (black), for the $U(n)$ model at $T = 0.01$ in natural units. We observe a clear crossover from a gapless phase, which must have $\tilde{\Sigma}(0) = m$ at zero temperature (red), as in (3.49), to a gapped phase, for which $\tilde{\Sigma}(0) \simeq 0$ at very large masses. The transition occurs at $m = m_* = 0.44_1$, consistently with the appearance of a non-vanishing effective mass, see Fig. 33.

non-vanishing zero-temperature entropy and a heat capacity vanishing linearly with the temperature in the gapless phase $m < m_*$. A basic difference with the $U(n)^2$ model is that the entropy is now continuous. Its derivative is discontinuous, however, and this translates into a finite jump for C/T at $m = m_*$.

We conclude this Section with a few comments on some subtleties, both conceptual and numerical, associated with the new IR behaviour found in the $U(n)$ model for $0 < m < m_*$.

We have argued that the deep IR behaviour changes discontinuously from the standard SYK behaviour at $m = 0$, with symmetric exponents $\Delta_+ = \Delta_- = 1/4$, to the new asymmetric behaviour with $\Delta_+ \neq \Delta_-$ for any $m > 0$, even if arbitrarily small. In this sense, the mass operator is relevant at $m = 0$ (and marginal for all $0 < m < m_*$), even though no mass gap is created: as soon as it is turned on, the IR fixed point of the model jumps from one gapless solution to another very different gapless solution. The fact that the jump must be discontinuous can be understood *e.g.* as a consequence of the formulas (3.54) and (3.57) for b_+^4 , which show that this coefficient is discontinuous when one goes from $\Delta_+ = \Delta_-$ to $\Delta_+ < \Delta_-$, even if the difference between Δ_+ and Δ_- is very small.

This discontinuity at $m = 0$ contrasts with the fact that all the thermodynamical quantities are found to be perfectly continuous when $m \rightarrow 0$, including at $T = 0$, as shown in Fig. 35. At $m = 0$, they match with the results for the $U(n)^2$ model, which

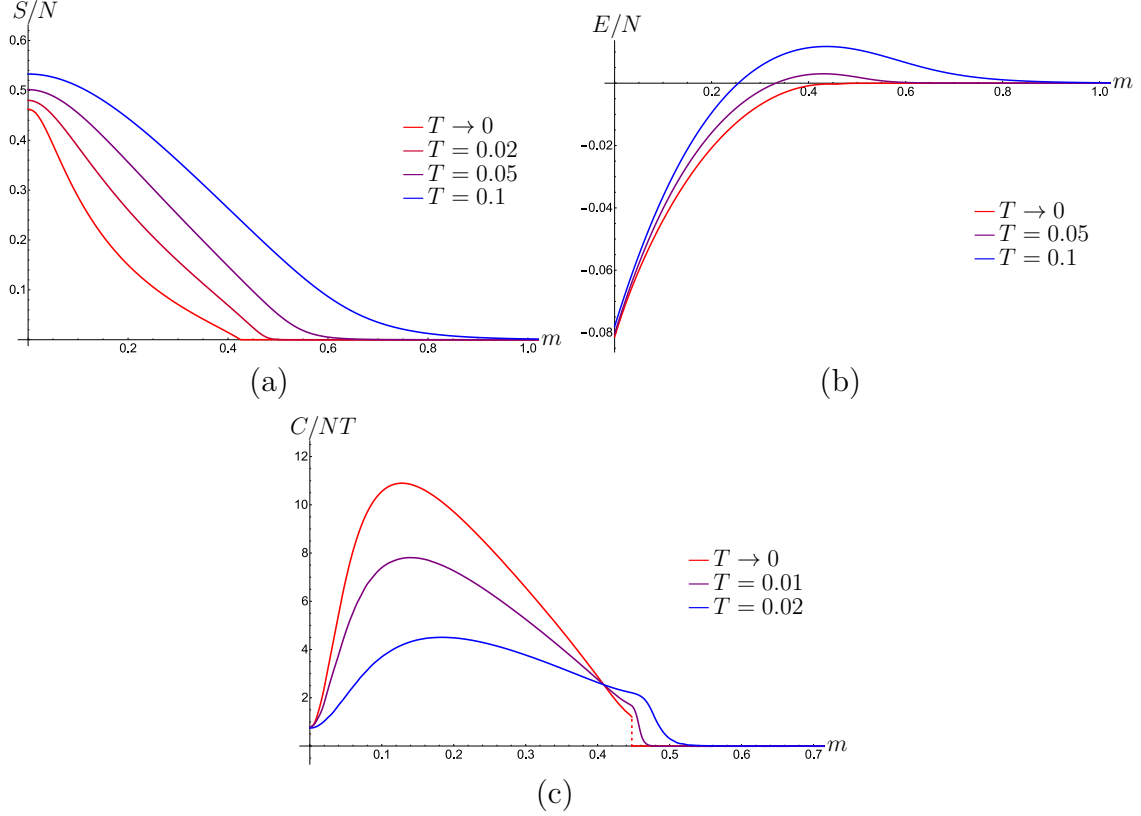


Figure 35: Plots of the entropy S/N (a), energy E/N (b) and heat capacity coefficient $\frac{C}{NT} = \frac{1}{N} \frac{\partial S}{\partial T}$ (c), as functions of the mass in the $U(n)$ model. $T \rightarrow 0$ extrapolations were performed using linear fits of ~ 10 finite-temperature solutions with $\beta \in [50, 100]$. At zero temperature all quantities become non-trivial below the quantum critical point $m = m_* = 0.44_1$, where the heat capacity coefficient experiences a finite jump (red dotted line).

is essentially just SYK. But there is no contradiction here; the discontinuous change in the IR behaviour occurs only for a range of frequencies $|\omega| \ll m$ that shrinks to zero size when $m \rightarrow 0$. At very small $m \ll |\xi|$, the model thus has an interesting RG flow: from the asymptotically free regime at frequencies $\omega \gg |\xi|$, it flows to an approximate conformal phase *à la* SYK with $\Delta = 1/4$ in the regime $m \ll \omega \ll |\xi|$ before eventually flowing to the non-standard deep IR behaviour with $\Delta_+ < \Delta_-$ when $\omega \ll m$.

The fact that the correct IR behaviour can emerge only for $\omega \ll m$ implies that the numerical characterization of the deep IR regime at small mass is particularly difficult. To do this, one should probe temperatures satisfying $\beta m \gg 1$ as well as $\beta|\xi| \gg 1$, but the former condition is much stronger than the latter in the regime $m \ll |\xi|$.

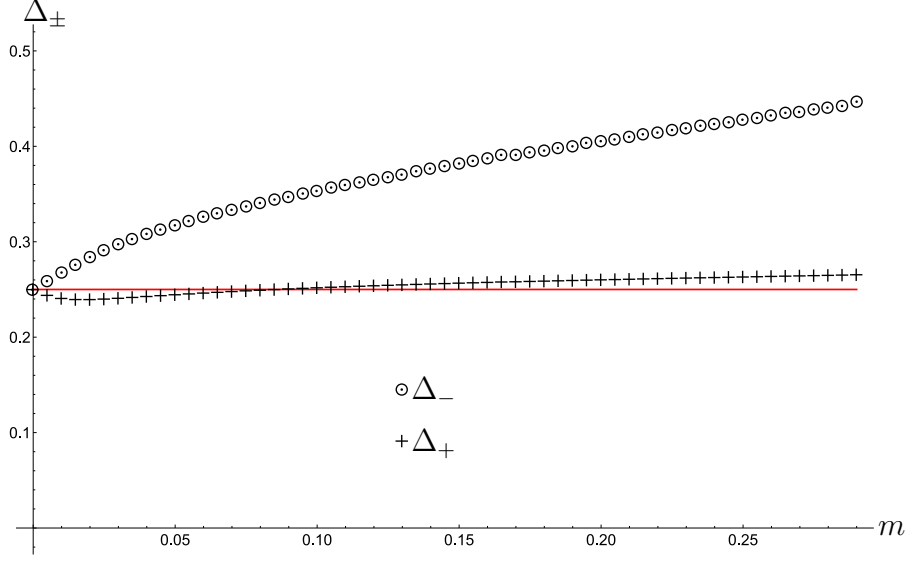


Figure 36: Numerical scaling dimensions $\Delta_{\pm}$ of the $U(n)$ model as functions of the mass, for small and intermediate values of m . Each of the datapoints for Δ_+ (crosses) and Δ_- (circles) was computed by linearly extrapolating results obtained from fitting finite-temperature two-point functions, as exemplified in Fig. 15. While Δ_+ remains close to the value $1/4$ (red line), as required by the generalized ansatz (3.57), the numerical evaluation of Δ_- exhibits mass dependence and quickly departs from this value, in clear violation of the standard ansatz (3.54). See the main text for a detailed discussion.

To illustrate this difficulty, we plot in Fig. 36 the exponents Δ_+ and Δ_- for various values of m , evaluated numerically by following exactly the same procedure as in Fig. 15 of Sec. 3.3.2. We see that Δ_+ stays reasonably close to the value $\Delta_+ = 1/4$, as required by (3.57), with small variations that may be attributed to numerical artifacts surviving the $T \rightarrow 0$ extrapolation. This is not surprising, since the value $\Delta_+ = 1/4$ coincides for both the standard SYK conformal phase and the new non-standard phase with $\Delta_+ < \Delta_-$. A precise determination of Δ_+ is thus expected as soon as $\beta|\xi| \gg 1$, even if $m \ll |\xi|$. The situation is very different for Δ_- . The numerical values we find exhibit a marked mass dependence, starting from $\Delta_- = 0.250_1$ at $m = 0$ and increasing up to $\Delta_- = 0.446_2$ at $m = 0.29$.

The fact that the numerical result for Δ_- varies smoothly from its $m = 0$ value is an automatic consequence of the discussion above: at extremely small masses the lowest temperature we can study numerically will always be in the near-SYK regime $1/|\xi| \ll \beta \ll 1/m$ and the genuine deep IR regime $\beta \gg 1/m$ is not probed. It is therefore not possible to decide, based on the numerical analysis alone, whether the actual value of Δ_- in the phase $\Delta_+ \neq \Delta_-$ does depend continuously on m or not. Two possibilities thus remain open:

- i) The deep IR value of Δ_- depends smoothly on m and our numerical analysis can actually provide a reasonable approximation to the actual value, possibly even for small masses.
- ii) The deep IR value of $\Delta_- > \Delta_+$ is fixed and does not depend on m , but our numerical analysis cannot reliably determine it for small m . The mass dependence of Δ_- in Fig. 36 may then be interpreted as mimicking a RG flow, since the deep IR is probed better and better when m increases (as explained below, this is no longer true when the mass is too close to m_*).

Note that, in either case, the fact that $\Delta_+ < \Delta_-$ when $m > 0$ is unambiguously established. As will be explained elsewhere [70], the possibility ii) seems to be strongly favoured.

Finally, let us mention that obtaining numerical solutions at very low temperatures becomes increasingly harder as the mass approaches the critical value $m_* = 0.44_1$, which explains why the highest mass appearing in Fig. 36 is $m = 0.29$. Indeed, the solutions near the quantum critical point are quite subtle. They must be very near the low entropy, perturbative-like solutions above some effective mass scale that goes to zero when $m \rightarrow m_*$ from below, but at the same time should reproduce the non-standard IR behaviour with $\Delta_+ < \Delta_-$ below this small effective mass scale.

5 Real time physics

5.1 Real time two-point functions

The goal of this subsection is to explain in detail how to compute real time two-point functions. After recalling standard definitions and results in Sec. 5.1.1, we present a simple derivation of the real time version of the Schwinger-Dyson equations in Sec. 5.1.2. The argument can be used in any melonic quantum mechanics. The resulting equations are the real time counterparts of the Euclidean equations (3.26), (3.30), (3.31). They can be solved numerically using a method outlined in Sec. 5.1.3. Note that our results generalize and are consistent with known formulas, presented for example in [9] for the SYK model; the only true novelty here may be the rigorous justification of the derivations, especially in Sec. 5.1.2, and the pedagogical presentation.

Real time two-point functions can be used to compute quasi-normal frequencies. The results are presented in Sec. 5.1.4. We find a very interesting behaviour near the finite temperature critical point in the $U(n)^2$ model. Real time two-point functions will also serve as a crucial building block in the analysis of the four-point functions in Sec. 5.2, allowing the calculation of Lyapunov exponents in the following subsections.

5.1.1 General definitions and properties

For any operator $\mathcal{O}$ and complex “time” u we define

$$\mathcal{O}(u) = e^{Hu} \mathcal{O} e^{-Hu}, \quad (5.1)$$

denoting

$$\tau_u = \text{Re } u, \quad t_u = \text{Im } u. \quad (5.2)$$

If $u = \tau_u$ is real (5.1) is equivalent to the Euclidean time evolution (3.2). If $u = it_u$ is purely imaginary, on the other hand, (5.1) yields the standard real time Heisenberg evolution. We can then define a generalized two-point function for complex time arguments u and v satisfying $0 < |\tau_u - \tau_v| < \beta$,

$$G(u, v) = \begin{cases} \frac{1}{n} \langle \text{T tr } \psi_\mu(u) \psi_\mu^\dagger(v) \rangle_\beta & \text{(matrix models)} \\ \langle \text{T } \chi^i(u) \chi_i^\dagger(v) \rangle_\beta & \text{(disordered models)}, \end{cases} \quad (5.3)$$

where T indicates an ordering with respect to the real parts of u and v , *i.e.* with respect to the Euclidean time. This is an important point: it is the Euclidean time that always dictates the ordering of operators in a correlator. Clearly $G(u, v)$ depends on $u - v$ only, and we shall thus also note $G(u, v) = G(u - v)$. The KMS condition

$$G(u + \beta) = -G(u) \quad (5.4)$$

can be used to extend $G(u)$ to all values of τ_u . It is sometimes convenient to decompose the generalized two-point function in terms of Wightman correlators,

$$G(u) = \Theta(\tau_u) G_+(u) + \Theta(-\tau_u) G_-(u), \quad (5.5)$$

which are defined by²³

$$G_+(u) = \frac{1}{n} \langle \text{tr } \psi_\mu(u) \psi_\mu^\dagger \rangle_\beta \quad \text{and} \quad G_-(u) = -\frac{1}{n} \langle \text{tr } \psi_\mu^\dagger \psi_\mu(u) \rangle_\beta. \quad (5.6)$$

At tree level,

$$G^{(0)}(u; m) = \begin{cases} n_F(m) e^{m(\beta-u)} & \text{if } 0 < \tau_u < \beta \\ -n_F(m) e^{-mu} & \text{if } -\beta < \tau_u < 0, \end{cases} \quad (5.7)$$

generalizing (3.7). The Euclidean spectral decomposition (3.12) is generalized to

$$G(u) = \int_{-\infty}^{+\infty} G^{(0)}(u; \omega) \rho(\omega) d\omega, \quad (5.8)$$

²³We stick to the matrix model notation from now on.

where the spectral function ρ is defined in (3.9). We can also use the Fourier decomposition

$$G(\tau + it) = \frac{1}{2\pi\beta} \sum_{k \in \mathbb{Z} + \frac{1}{2}} \int_{-\infty}^{+\infty} \tilde{G}_k(\omega) e^{-i\omega t - i\nu_k t} d\omega, \quad (5.9)$$

with

$$\tilde{G}_k(\omega) = \frac{2\pi\rho(\omega)}{\omega - i\nu_k}. \quad (5.10)$$

It is important to note that the correct analytic continuation from Euclidean time to complex time cannot be obtained by replacing $\tau \mapsto \tau + it$ in the Euclidean Fourier expansion (3.4). The analytic continuation to complex time is simple in correlators for which the operator ordering is fixed, because the complex time operator evolution (5.1) is a simple analytic continuation of the Euclidean time evolution (3.2), but is non-trivial in time-ordered correlators.

Purely real time correlators are obtained from the complex time correlators by letting τ go to zero, the ordering of operators being fixed by the way the limit is taken, *i.e.* $\tau \rightarrow 0^+$ or $\tau \rightarrow 0^-$. Eq. (5.8) implies that all the real time two-point functions, for any ordering of the operators, can be expressed in terms of the spectral function ρ . For instance,

$$S_+(t) = \frac{1}{2n} \langle \text{tr}[\psi_\mu(it), \psi_\mu^\dagger] \rangle_\beta = \frac{1}{2} (G_+(it) + G_-(it)) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \tilde{S}_+(\omega) e^{-i\omega t} d\omega, \quad (5.11)$$

$$S_-(t) = \frac{1}{2n} \langle \text{tr}\{\psi_\mu(it), \psi_\mu^\dagger\} \rangle_\beta = \frac{1}{2} (G_+(it) - G_-(it)) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \tilde{S}_-(\omega) e^{-i\omega t} d\omega \quad (5.12)$$

are given by

$$\tilde{S}_+(\omega) = \pi \frac{e^{\beta\omega} - 1}{e^{\beta\omega} + 1} \rho(\omega) \quad \text{and} \quad \tilde{S}_-(\omega) = \pi \rho(\omega). \quad (5.13)$$

It is also convenient to introduce the retarded and advanced two-point functions

$$G_r(t) = 2i\Theta(t)S_-(t) = i\Theta(t)(G_+(it) - G_-(it)) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \chi_r(\omega) e^{-i\omega t} d\omega, \quad (5.14)$$

$$G_a(t) = -2i\Theta(-t)S_-(t) = -i\Theta(-t)(G_+(it) - G_-(it)) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \chi_a(\omega) e^{-i\omega t} d\omega, \quad (5.15)$$

for which

$$\chi_r(\omega) = - \int_{-\infty}^{+\infty} \frac{\rho(\omega')}{\omega - \omega' + i0^+} d\omega', \quad \chi_a(\omega) = - \int_{-\infty}^{+\infty} \frac{\rho(\omega')}{\omega - \omega' - i0^+} d\omega'. \quad (5.16)$$

From the definition of the resolvent in (3.22), we see that

$$\chi_r(\omega) = -R(\omega + i0^+) \quad \text{and} \quad \chi_a(\omega) = -R(\omega - i0^+). \quad (5.17)$$

The retarded and advanced Fourier transforms χ_r and χ_a can thus be analytically continued to holomorphic functions on the upper and lower half-planes respectively, an important and well-known property. Of course, they are not independent since the reality of the spectral function implies that

$$G_a(t)^* = G_r(-t), \quad \chi_a(z)^* = \chi_r(z^*). \quad (5.18)$$

Real time two-point functions are in this way completely encoded in the spectral function ρ , whereas the Euclidean two-point function is encoded in the Fourier-Matsubara coefficients G_k . The real time and Euclidean time data thus look very different: frequencies in real time are always continuous, whereas their Euclidean counterparts are discrete. There is a powerful theorem due to Carlson [71] that shows that these two different-looking sets of data are nevertheless strictly equivalent: the real time data determines uniquely the Euclidean time data and vice versa. That the real time data determines the Euclidean time data is easy to understand: from the spectral function ρ one can get the G_k 's from (3.22) and (3.23). The converse is much more subtle. The knowledge of the Matsubara-Fourier coefficients fixes the values (3.23) of the resolvent on a discrete set of points $z = i\nu_k$ on the imaginary axis. Moreover, we know that the resolvent is a holomorphic function for $\text{Im } z > 0$ and $\text{Im } z < 0$ and has the simple asymptotic behaviour (3.24). Carlson's theorem then implies that $R(z)$ is determined uniquely for all complex values of z .

The theorem itself does not provide an explicit procedure to reconstruct ρ from the G_k 's, but such a procedure was discovered in [72] and generalized in [73]. The explicit formulas in [73] actually allow to reconstruct the resolvent from the knowledge of the Euclidean data above an arbitrary cut-off k_0 in frequency, *i.e.* the set of Matsubara-Fourier coefficients $\{G_k, k \geq k_0 > 0\}$. A startling consequence of this fact is the existence in any quantum mechanical theory of a rigorous version of the Renormalization Group, as a consequence solely of analyticity properties [73].²⁴ The explicit reconstruction formulas are thus very interesting conceptually, but in practice their use is limited by the fact that a reliable reconstruction requires the knowledge of the G_k 's to an enormous precision. They are therefore impractical except in simple cases where ρ is known to be very smooth, without sharp features. Since we do not want to make any such *a priori* assumptions, we prefer to base our analysis on the real time version of the Schwinger-Dyson equations derived in the next subsection. Interestingly, the derivation of these equations that we now present is going to use Carlson's theorem in a central way.

²⁴Ref. [73] focuses on the bosonic case, but the fermionic case can be treated in exactly the same way.

5.1.2 Melonic real time two-point functions

The standard general approach to real time correlation functions is based on the Schwinger-Keldysh formalism. This formalism is a direct consequence of the path integral formulation, suitably adapted to cases where operators are inserted at arbitrary complex times, as in (5.1). One introduces an oriented closed Keldysh contour γ on the complex time cylinder where u and $u + \beta$ are identified. The contour starts at $u = 0$, ends at $u = \beta$ and goes through all the operator insertion points, say $\mathcal{O}_1(u_1), \dots, \mathcal{O}_r(u_r)$. The correlator associated with such a contour is

$$G_\gamma(u_1, \dots, u_r) = \langle T_\gamma \mathcal{O}_1(u_1) \cdots \mathcal{O}_r(u_r) \rangle_\beta, \quad (5.19)$$

where the symbol T_γ orders the operators according to their position along the contour. For instance, to compute the two-point function $G(u_0)$, Eq. (5.3), we can use the contour depicted in Fig. 37, with an insertion of $\psi_\mu^\dagger$ at $u = 0$ and an insertion of ψ_μ at $u = u_0$. The standard Euclidean formalism corresponds to the special case where all the insertions are on the Euclidean time axis $\text{Im } u = 0$, so that we can then use the contour $\text{Im } u = 0$. In practice, it is convenient to use contours along which $\text{Re } u$ is monotonically increasing, so that the contour ordering coincides with the usual Euclidean time ordering. This is why the real time piece of γ in Fig. 37 was chosen to have a triangular shape.

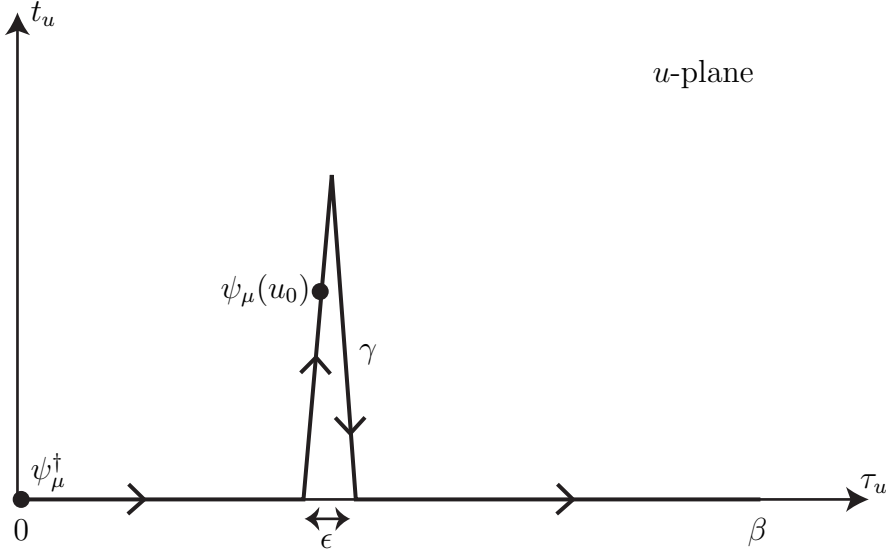


Figure 37: Keldysh contour γ (thick line) used to compute the two-point functions at an arbitrary complex time u_0 . The insertions of the operators ψ_μ at $u = 0$ and $\psi_\mu^\dagger$ at $u = u_0$ are represented by bullets along the contour. The width ϵ of the triangle-shaped piece of the contour may be taken to zero. Note that the contour may be chosen such that the insertion point at $u = u_0$ sits at the top of the triangle, but this is not required.

The advantage of this approach is that the Feynman rules take a familiar form. The only differences with the Euclidean case is that integrals over τ are replaced by contour integrals along γ , and propagators are associated with the complex time tree-level two-point function (5.7). The arguments leading to the Euclidean Schwinger-Dyson equations (3.30) and (3.31), in their path integral or diagrammatic versions for the disordered and matrix models respectively, can be repeated without change in the present, more general context. The self-energy for the contour of Fig. 37 is defined by the equation

$$\left(\frac{d}{du} + m\right)G(u - u') + \int_{\gamma} \Sigma(u - v)G(v - u') dv = \delta_{\gamma}(u - u'), \quad (5.20)$$

generalizing (3.27). Eqs. (3.30) and (3.31) are now replaced by

$$\Sigma(u) = \lambda^2 G(u)^2 G(-u) \quad \text{for the } U(n)^2 \text{ model}, \quad (5.21)$$

$$\Sigma(u) = -\frac{|\xi|^2}{4} (G(u)^3 + 3G(u)G(-u)^2) \quad \text{for the } U(n) \text{ model}. \quad (5.22)$$

Traditionally, one proceeds by using the decomposition (5.5) to rewrite (5.20) in the form of the so-called Kadanov-Baym equations; see *e.g.* [74] for a pedagogical introduction. The Kadanov-Baym equations can then be integrated numerically using (5.21) or (5.22).²⁵

We are not going to use this rather tedious route here. Instead, we present a simple argument based on Carlson's theorem, leading to a very convenient expression of the solution which entirely avoids having to deal with the details of the Kadanov-Baym equations.

Let us start by defining

$$R_{\Sigma}(z) = \frac{1}{R(z)} + m - z, \quad (5.23)$$

so that Eqs. (3.23) and (3.26) imply

$$R_{\Sigma}(i\nu_k) = -\Sigma_k. \quad (5.24)$$

Using (3.22) and denoting $x = \text{Re } z$ and $y = \text{Im } z$, we get

$$\text{Im } R(x + iy) = -y \int_{-\infty}^{+\infty} \frac{\rho(\omega)}{(x - \omega)^2 + y^2} d\omega. \quad (5.25)$$

From the positivity of the spectral function ρ , which is manifest from its definition (3.9), we deduce that $\text{Im } R(z) \neq 0$ away from the real axis, and thus also $R(z) \neq 0$

²⁵For a recent interesting use of the Kadanov-Baym equations in the context of the SYK model, see *e.g.* [75].

there. Since R is holomorphic and has no zeros for $\text{Im } z > 0$ or $\text{Im } z < 0$, Eq. (5.25) implies that R_Σ is also holomorphic for $\text{Im } z > 0$ and $\text{Im } z < 0$. Moreover, the expansion (3.24) shows that

$$R_\Sigma(z) \underset{z \rightarrow \infty}{=} \mathcal{O}(1/z). \quad (5.26)$$

All the conditions for the applicability of Carlson's theorem to the function R_Σ are thus met: if we are able to find some function $\tilde{R}_\Sigma(z)$, holomorphic for $\text{Im } z > 0$ and $\text{Im } z < 0$, behaving as $\mathcal{O}(1/z)$ when $z \rightarrow \infty$ and such that $\tilde{R}_\Sigma$ and R_Σ match for all $z = i\nu_k$, then the theorem states that $\tilde{R}_\Sigma = R_\Sigma$ for all z :

$$\tilde{R}_\Sigma(i\nu_k) = R_\Sigma(i\nu_k) \quad \forall k \in \mathbb{Z} + \frac{1}{2} \implies R_\Sigma(z) = \tilde{R}_\Sigma(z) \quad \forall z \in \mathbb{C}. \quad (5.27)$$

We can combine the equations (5.21) or (5.22) with the spectral decomposition (5.8) to guess a formula for $R_\Sigma(z)$. We shall then check that the above criteria are satisfied.

In the case of the $U(n)^2$ model, we consider

$$\tilde{R}_\Sigma(z) = -\lambda^2 \int_{-\infty}^{+\infty} \prod_{j=1}^3 [d\omega_j \rho(\omega_j) n_F(\omega_j)] \frac{e^{\beta(\omega_1 + \omega_2)}}{n_F(\omega_3 - \omega_1 - \omega_2)} \frac{1}{z + \omega_3 - \omega_1 - \omega_2}. \quad (5.28)$$

The convergence of the integral is trivial to check. The function $\tilde{R}_\Sigma$ so defined is manifestly holomorphic for $\text{Im } z > 0$ and $\text{Im } z < 0$ and is $\mathcal{O}(1/z)$ at large z . Moreover, if we evaluate Σ_k by taking the Fourier transform of (5.21), in which we use the spectral decomposition formula (5.8) for G and the explicit form (5.7) for $G^{(0)}$, we immediately find that $\Sigma_k = -\tilde{R}_\Sigma(i\nu_k)$ is satisfied. From (5.24) we conclude that $\tilde{R}_\Sigma(i\nu_k) = R_\Sigma(i\nu_k)$ and thus from (5.27) $\tilde{R}_\Sigma = R_\Sigma$.

Exactly the same reasoning can be applied to the $U(n)$ model starting from

$$\begin{aligned} \tilde{R}_\Sigma(z) = & -\frac{|\xi^2|}{4} \int_{-\infty}^{+\infty} \prod_{j=1}^3 [d\omega_j \rho(\omega_j) n_F(\omega_j)] \\ & \times \left[\frac{1}{n_F(\omega_1 + \omega_2 + \omega_3)} \frac{1}{z - \omega_1 - \omega_2 - \omega_3} + \frac{3e^{\beta(\omega_2 + \omega_3)}}{n_F(\omega_1 - \omega_2 - \omega_3)} \frac{1}{z + \omega_2 + \omega_3 - \omega_1} \right], \end{aligned} \quad (5.29)$$

and then showing that $\tilde{R}_\Sigma(i\nu_k) = R_\Sigma(i\nu_k)$ and thus $\tilde{R}_\Sigma = R_\Sigma$.

We should stress that it is crucial to understand that the property $\tilde{R}_\Sigma(i\nu_k) = R_\Sigma(i\nu_k) = -\Sigma_k$ is *not* sufficient in itself to conclude $\tilde{R}_\Sigma(z) = R_\Sigma(z)$. For example, a naive way to do the analytic continuation from Euclidean time to real time in the $U(n)^2$ model is to start from the Euclidean equation (3.30) written in terms of the Matsubara-Fourier coefficients,

$$\Sigma_k = \frac{\lambda^2}{\beta^2} \sum_{k_1, k_2 \in \mathbb{Z} + \frac{1}{2}} G_{k_1} G_{k_2} G_{k_1 + k_2 - k} = -\frac{\lambda^2}{\beta^2} \sum_{k_1, k_2 \in \mathbb{Z} + \frac{1}{2}} G_{k_1} G_{k_2} R(i\nu_{k_1 + k_2} - i\nu_k). \quad (5.30)$$

This might induce us to propose

$$R_\Sigma(z) \stackrel{?}{=} \frac{\lambda^2}{\beta^2} \sum_{k_1, k_2 \in \mathbb{Z} + \frac{1}{2}} G_{k_1} G_{k_2} R(i\nu_{k_1+k_2} - z). \quad (5.31)$$

Although this expression satisfies by construction $R_\Sigma(i\nu_k) = -\Sigma_k$, it does not have the required analytic properties; indeed, the right-hand side of (5.31) is holomorphic only within bands $ni < \text{Im } z < (n+1)i$ where $n \in \mathbb{Z}$, but not on the whole half-planes $\text{Im } z > 0$ and $\text{Im } z < 0$ as required.

For the purpose of improving the efficiency of the numerical algorithm presented in the next subsection, it is useful to perform explicitly one of the integrals over the frequencies in (5.28) and (5.29). For example, in the $U(n)^2$ model we write

$$R_\Sigma(z) = -\lambda^2 \int_{-\infty}^{+\infty} d\omega_1 d\omega_2 \rho(\omega_1) n_F(\omega_1) \rho(\omega_2) n_F(\omega_2) e^{\beta(\omega_1+\omega_2)} \mathcal{I}(\omega_1 + \omega_2; z), \quad (5.32)$$

where

$$\mathcal{I}(\Omega; z) = \int_{-\infty}^{+\infty} d\omega \rho(\omega) \frac{n_F(\omega)}{n_F(\omega - \Omega)} \frac{1}{z + \omega - \Omega}. \quad (5.33)$$

We can then use (3.25) and the well-behaved properties of the integrand at infinity to get

$$\mathcal{I}(\Omega; z) = \frac{i}{2\pi} \int_{C_+ \cup C_-} \frac{n_F(w)}{n_F(w - \Omega)} \frac{R(w)}{z + w - \Omega} dw, \quad (5.34)$$

where C_+ and C_- are semi-infinite rectangles on the upper and lower half complex w -plane respectively, see Fig. 38. Picking up the poles at $w = i\nu_k$ for all $k \in \mathbb{Z} + \frac{1}{2}$, as well as at $w = \Omega - z$, yields

$$\mathcal{I}(\Omega; z) = \frac{e^{-\beta\Omega} - 1}{\beta} \sum_{k \in \mathbb{Z} + \frac{1}{2}} \frac{G_k}{z + i\nu_k - \Omega} - (1 + e^{-\beta z}) \frac{R(\Omega - z)}{e^{\beta(\Omega - z)} + 1}. \quad (5.35)$$

Overall, we thus obtain for the $U(n)^2$ model

$$\begin{aligned} R_\Sigma(z) = & -\lambda^2 \int_{-\infty}^{+\infty} d\omega_1 d\omega_2 \rho(\omega_1) n_F(\omega_1) \rho(\omega_2) n_F(\omega_2) \\ & \times \left[\frac{1 - e^{\beta(\omega_1+\omega_2)}}{\beta} \sum_{k \in \mathbb{Z} + \frac{1}{2}} \frac{G_k}{z + i\nu_k - \omega_1 - \omega_2} - (1 + e^{\beta z}) \frac{R(\omega_1 + \omega_2 - z)}{e^{\beta(z - \omega_1 - \omega_2)} + 1} \right]. \end{aligned} \quad (5.36)$$

In the $U(n)$ model, similar manipulations yield

$$\begin{aligned}
R_\Sigma(z) = & -\frac{|\xi|^2}{4} \int_{-\infty}^{+\infty} d\omega_1 d\omega_2 \rho(\omega_1) n_F(\omega_1) \rho(\omega_2) n_F(\omega_2) \\
& \times \left[\frac{2(e^{\beta(\omega_1+\omega_2)} - 1)}{\beta} \sum_{k \in \mathbb{Z} + \frac{1}{2}} \frac{i\nu_k + 2(\omega_1 + \omega_2) - z}{(z - i\nu_k)^2 - (\omega_1 + \omega_2)^2} G_k \right. \\
& \left. + (e^{\beta z} + 1) \left(\frac{R(z - \omega_1 - \omega_2)}{e^{\beta(z - \omega_1 - \omega_2)} + 1} + 3e^{-\beta z} \frac{R(z + \omega_1 + \omega_2)}{e^{-\beta(z + \omega_1 + \omega_2)} + 1} \right) \right]. \tag{5.37}
\end{aligned}$$

Equations (5.23) and (5.36) (for the $U(n)^2$ model) or (5.36) (for the $U(n)$ model) are the precise real time counterparts of the Euclidean equations (3.26) and (3.30) or (3.31), respectively, and completely determine the spectral function ρ .

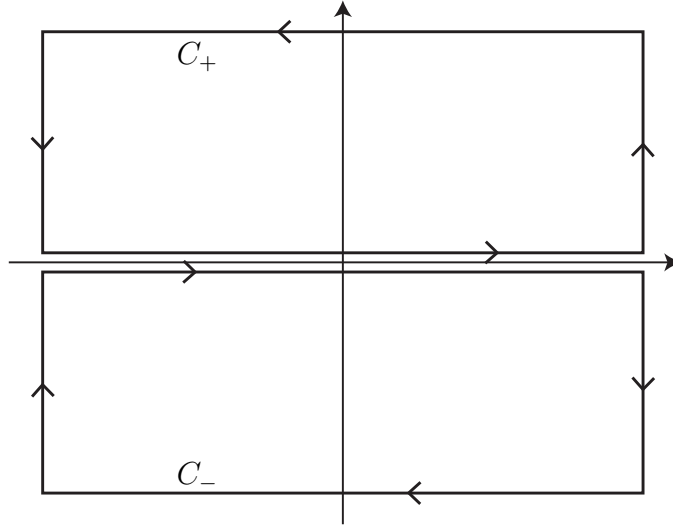


Figure 38: Contours C_+ and C_- in the upper and lower w -plane used in the main text for computing (5.33) and similar integrals. The limit for which the rectangles have an infinite size is taken.

Let us conclude this derivation with the following remark. The Kadanov-Baym equations obtained from (5.20) must of course be satisfied by our solution, and we have checked this explicitly. This is an interesting albeit somewhat tedious exercise in contour integration. One must consider several cases depending on the position of u and u' along the Keldysh contour γ . It is convenient to use double Fourier decompositions of the form (5.9) for both $G(u)$ and $\Sigma(u)$. All the required identities follow from the analytic properties of R and R_Σ discussed above.

5.1.3 Numerical solution of real time equations

As already mentioned, all the real time two-point functions can be straightforwardly derived from the spectral function $\rho(\omega)$. The basic formula relating ρ to the resolvent R is (3.25), and noting that for any real ϵ

$$R(\omega + i\epsilon)^* = R(\omega - i\epsilon), \quad (5.38)$$

we can rewrite this equation as

$$\rho(\omega) = -\frac{1}{\pi} \text{Im} R(\omega + i0^+). \quad (5.39)$$

We will therefore focus on the computation of the resolvent through equations (5.23) and (5.36) or (5.37). In practice, we evaluate $R(\omega + i\epsilon)$ for a small positive ϵ , and assume throughout

$$\rho(\omega) \simeq \rho_\epsilon(\omega) = -\frac{1}{\pi} \text{Im} R(\omega + i\epsilon). \quad (5.40)$$

The real time Schwinger-Dyson equations can be solved numerically using an iterative algorithm that is very similar to the one applied in the Euclidean case, described in detail in Sec. 3.3. There are however a few new subtleties associated with the real time case.

The algorithm starts from an approximate solution $R^{[0]}(\omega + i\epsilon)$, and progressively refines it in successive iterations. Denoting $R^{[i]}(\omega + i\epsilon)$ the approximate solution at step i , and $\rho_\epsilon^{[i]}(\omega)$ the spectral function associated to it through (5.40), we first compute $R_\Sigma^{[i+1]}(\omega + i\epsilon)$ by plugging these in the right-hand side of (5.36) or (5.37), according to the model under study. To complete the iteration, we define the resolvent at step $i + 1$ to be

$$R^{[i+1]}(\omega + i\epsilon) = (1 - \alpha)R^{[i]}(\omega + i\epsilon) + \frac{\alpha}{\omega + i\epsilon - m + R_\Sigma^{[i+1]}(\omega + i\epsilon)}, \quad (5.41)$$

where α is an *a priori* arbitrary real weighting factor.

Because equations (5.36) and (5.37) involve integrals over ω , in practice we need to choose a finite number of discrete frequencies over which the resolvent is evaluated,

$$\omega_p = \delta\omega p \quad \text{for} \quad p = p_i, p_i + 1, \dots, p_f - 1, p_f. \quad (5.42)$$

This defines a unidimensional grid on a finite interval $[\omega_{\min}, \omega_{\max}] = [\delta\omega p_i, \delta\omega p_f]$, outside which we use asymptotic freedom to identify the solution with the tree-level result (5.44), thus transforming all integrals into finite sums. The discrete step $\delta\omega$ is chosen consistently with the parameter $\epsilon > 0$, and sets an intrinsic upper limit on the resolution with which ρ can be reconstructed from (5.40).

Note that the Euclidean two-point function also appears on the right-hand side of equations (5.36) and (5.37) through its Matsubara-Fourier coefficients G_k . We will therefore assume that a solution for these is already known, possibly by having run the Euclidean algorithm beforehand, so that they are simply constant coefficients in this context.

The iteration of the algorithm is repeated until the equations are solved up to a prescribed tolerance ε , *i.e.*

$$\frac{1}{P} \sum_{p=p_i}^{p_f} \left| R^{[i]}(\omega_p + i\epsilon) - \frac{1}{\omega + i\epsilon - m + R_{\Sigma}^{[i]}(\omega_p + i\epsilon)} \right| < \varepsilon \quad \text{with} \quad P = p_f - p_i + 1. \quad (5.43)$$

As in the Euclidean case, we will take $\varepsilon = 10^{-12}$ (in natural units $\lambda = |\xi| = 1$), ensuring that the solution is found with a very good accuracy while staying safely above machine-precision issues.

Before moving on, let us make a few further comments on the algorithm described above. First, the choice of initial approximation is very important for its convergence. We may use the tree-level result corresponding to $R_{\Sigma}^{(0)}(\omega + i\epsilon) = 0$,

$$R^{(0)}(\omega + i\epsilon) = \frac{1}{\omega + i\epsilon - m}, \quad (5.44)$$

or any previously found solution deemed to be close to the actual solution we are seeking. As in the Euclidean case, this allows us to map the parameter space by taking small steps in either the m or T directions to “trace” solutions.

Another common element with the Euclidean algorithm is the presence of an arbitrary weighting factor α . This comes about for the same reasons discussed in Sec. 3.3, having a similar impact on convergence and performance as was seen there, and again we find that $\alpha = 0.75$ works well for most cases.

A new ingredient in the real time algorithm is the arbitrary frequency-domain discretization (5.42), which was naturally given by the Matsubara frequencies in the Euclidean setting. While as already mentioned $\delta\omega$ is related to the resolution of the resulting spectral function, its support is determined by p_i and p_f . Taking Eq. (3.21) into account, these parameters need to be chosen such that the interval $[\omega_{\min}, \omega_{\max}]$ is approximately centered around $\omega = m$, and extends to either side at least up to the regime where asymptotic freedom takes over.

More generally, the choice (5.42) used here has the advantage that it is additive and symmetric about the origin, so that it satisfies the convenient properties

$$\omega_{p_1} + \omega_{p_2} = \omega_{p_1+p_2} \quad \text{and} \quad -\omega_p = \omega_{-p}. \quad (5.45)$$

These properties simplify the practical implementation of the algorithm and improve its performance by completely avoiding costly interpolations. Furthermore, they allow

for the efficient pre-calculation of the Matsubara-Fourier coefficient sums in (5.36) and (5.37), which after truncation of the Euclidean data can be done in $\mathcal{O}(P)$ memory and $\mathcal{O}(P \times k_{UV})$ time, instead of the naive $\mathcal{O}(P^3)$ and $\mathcal{O}(P^3 \times k_{UV})$, respectively. It is for this reason that we chose to use the formulas (5.36) and (5.37), with two integrals over frequencies, instead of (5.28) and (5.29), with three integrals over frequencies.²⁶

There is no guarantee that the algorithm described here terminates, but if it does so in $I_\alpha(\varepsilon)$ steps then the total running time is $\mathcal{O}(I_\alpha(\varepsilon) \times P^3 + P \times k_{UV})$. Additional pre-processing in $\mathcal{O}(P^2)$ time and memory is useful in practice to reduce constant factors and the incidence of numerical errors, but this is not relevant asymptotically. For our C++ implementation running on a desktop computer one may then take $P = \mathcal{O}(10^3)$, which makes the pre-processing subleading for typical values of $k_{UV} = \mathcal{O}(10^5)$. Depending on the model and values of the mass, this is enough to reach temperatures of order $T = \mathcal{O}(10^{-2})$ in natural units, but quickly becomes insufficient below that point.

To illustrate the results obtained with this algorithm, we show in Fig. 39 examples of the spectral function computed in the $U(n)^2$ model, for various masses and temperatures. At $m = 0$ it is an even function as a consequence of the particle-hole symmetry, and coincides with the spectral function of the $U(n)$ model. These properties are lost when $m > 0$. At small masses and temperatures, $\rho(\omega)$ approaches the conformal form (3.53), whereas for large masses it converges to the perturbative result $\delta(\omega - m)$.

In order to check the precision of our numerical methods, we show in Fig. 40a the values of the first two moments of the spectral function, ρ_0 and ρ_1 defined in Eq. (3.17), which are to be compared with the exact results (3.20) and (3.21), *i.e.* $\rho_0 = 1$ and $\rho_1 = m$. Another check is provided in Fig. 40b, where we plot the energy computed from the real time data via Eq. (3.103), and compare with the results obtained from the Euclidean formula (3.102). Similar results are obtained for the $U(n)^2$ model at temperatures other than $T = T_c$, used here for illustrative purposes, as well as for other thermodynamic quantities such as the fermion number (computed from Euclidean or real time data through (3.96) and (3.97), respectively).

While Fig. 39 and 40 correspond to the $U(n)^2$ model, very similar results are also obtained for the $U(n)$ model. The main difference in this case resides in the fact that the spectral function has a more involved structure for $m > 0$, presenting multiple peaks instead of just one, see Fig. 41. While this is not intrinsically problematic, these peaks become sharper as the temperature is lowered, thus increasing our resolution requirements and therefore raising the lower bound on the temperatures we can probe numerically in a reasonable amount of time.

²⁶One should note that precision is also improved by this choice, since Euclidean data is generally of better quality than real time one.

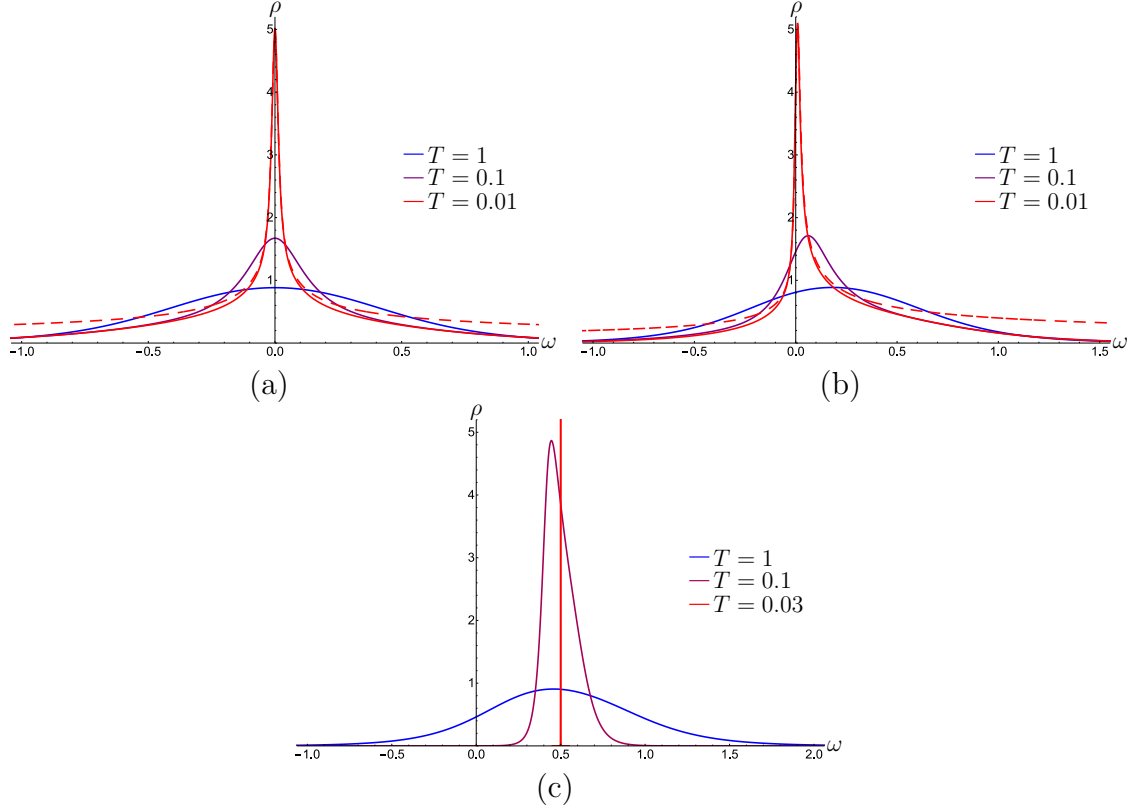


Figure 39: The spectral function $\rho(\omega)$ in the $U(n)^2$ model, for $m = 0$ (a), $m = 0.2$ (b) and $m = 0.5$ (c), at various temperatures in natural units. Plots (a) and (b) correspond to the SYK-like phase, in which the spectral function approaches the conformal form at low temperatures for $|\omega| \ll \lambda = 1$ (red dashed lines show Eq. (3.53) with parameters θ and a fitted as in Fig. 12 for the solution at $T = 0.01$). Plot (c) corresponds to the HO-like phase, where the low-temperature limit $\rho(\omega) \rightarrow \delta(\omega - m)$ is already evident for $T = 0.03$ (vertical axis is truncated for clarity). All solutions were obtained with $P = 1001$, $\Delta\omega = \delta\omega(p_f - p_i) = 10$ and $\epsilon = 10^{-5}$, except for the lowest temperature in (c) for which $\Delta\omega = 0.1$ and $\epsilon = 10^{-7}$.

5.1.4 Application: quasi-normal frequencies

Quasi-normal frequencies $\nu_{\text{qn}}^{(p)}$, $p = 0, 1, 2, \dots$, govern the way a system responds to a small perturbation. Linear response theory relates these frequencies to the long time behaviour of the retarded two-point function G_r , with frequency $\nu_{\text{qn}}^{(p)}$ contributing a term proportional to $\exp(-i\nu_{\text{qn}}^{(p)}t)$. Equivalently, they correspond to poles of χ_r on the lower-half complex frequency plane.²⁷ A non-zero negative imaginary part in $\nu_{\text{qn}}^{(p)}$

²⁷Or to poles of χ_a on the upper-half plane, see (5.18).

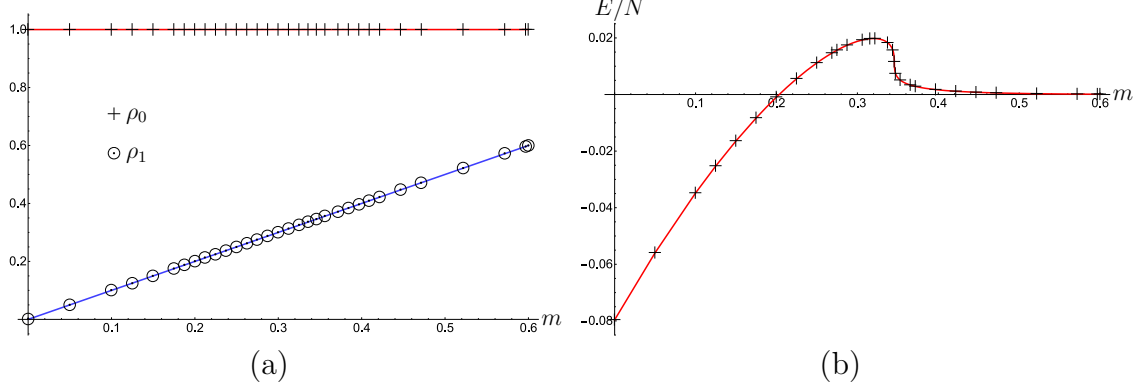


Figure 40: (a) First and second moments ρ_0 (crosses) and ρ_1 (circles) of the spectral function of the $U(n)^2$ model, for various masses at $T = T_c$; solid lines correspond to the exact values $\rho_0 = 1$ (red) and $\rho_1 = m$ (blue). (b) Energy of the $U(n)^2$ model, as calculated from the spectral function through Eq. (3.103), for various masses at $T = T_c$ (datapoints shown with crosses); the solid red line corresponds to the numerical values computed independently from purely Euclidean data, through the formula (3.102). Consistency between the real time numerics and the analytic and Euclidean results is excellent. Note that $T = T_c$ is used here for illustrative purposes and similar results are obtained for other temperatures.

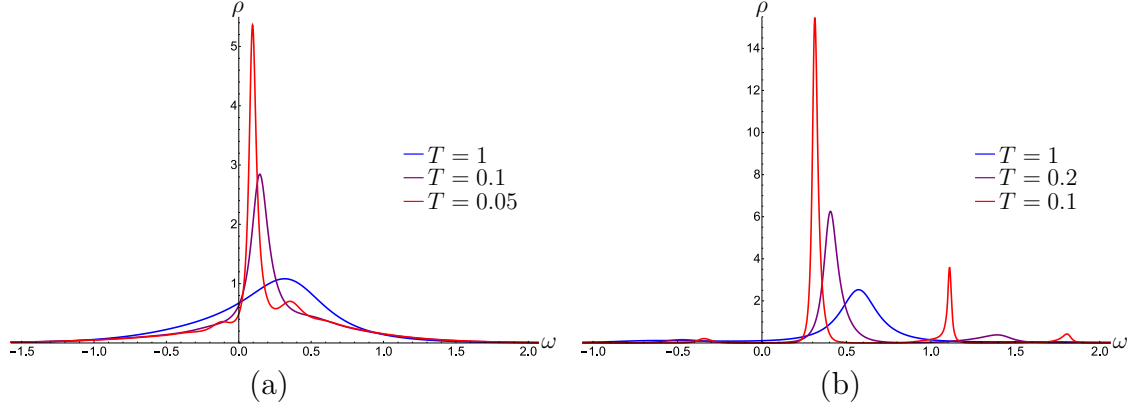


Figure 41: The spectral function $\rho(\omega)$ in the $U(n)$ model, for $m = 0.2$ (a) and $m = 0.5$ (b), at various temperatures in natural units. All solutions were obtained with $P = 1001$, $\Delta\omega = \delta\omega(p_f - p_i) = 10$ and $\epsilon = 10^{-5}$.

indicates an exponential decay at long times of the associated mode, with time scale

$$t_{\text{qn}}^{(p)} = -\frac{1}{\text{Im } \nu_{\text{qn}}^{(p)}}. \quad (5.46)$$

In the context of black holes, having $\text{Im } \nu_{\text{qn}}^{(p)} < 0$ is a landmark feature of horizon physics, signalling the absorption of any small external perturbation by the black

hole and its subsequent return to equilibrium.

For example, for the $U(n)^2$ model at low temperatures in the SYK-like phase, we can use (3.50) and (3.43) to write down the retarded function in the regime $\beta\lambda \gg 1$, $t \gg 1/\lambda$,

$$G_r(t) = i\Theta(t) \frac{C}{\sqrt{\beta}} \frac{e^{-i(\theta+2\pi at/\beta)}}{\sqrt{\sinh(\pi t/\beta)}}. \quad (5.47)$$

As expected, the large t asymptotic expansion of this function is a sum of decreasing exponentials of the form $\exp(-i\nu_{\text{qn}}^{(p)}t)$, with

$$\nu_{\text{qn}}^{(p)} \underset{\beta \rightarrow \infty}{=} \frac{2\pi}{\beta} \left(a - \frac{i}{4}(1+4p) \right), \quad p = 0, 1, 2, \dots \quad (5.48)$$

Amongst all the quasi-normal frequencies $\nu_{\text{qn}}^{(q)}$, the one having the largest time scale (5.46) dominates at long times. It is called the leading quasi-normal frequency and will be denoted simply by ν_{qn} ,

$$G_r(t) \underset{t \rightarrow +\infty}{\propto} e^{-i\nu_{\text{qn}}t}. \quad (5.49)$$

The associated time scale

$$t_{\text{qn}} = -\frac{1}{\text{Im } \nu_{\text{qn}}} \quad (5.50)$$

is often referred to in the literature as the thermalization or dissipation time scale.

In Fig. 42, we plot $-\text{Im } \nu_{\text{qn}}$ and $\text{Re } \nu_{\text{qn}}$ as functions of the temperature, for various values of the mass in the $U(n)^2$ model. As explained in Sec. 4, for masses $m < m_* = 0.213_2$ the IR physics is dominated by the SYK-like solution; our numerical results are then consistent with the asymptotic values (5.48) for $p = 0$. For masses $m > m_*$, on the other hand, we find that the behaviour of $-\text{Im } \nu_{\text{qn}}$ is $o(T)$.

The situation in the $U(n)$ model is illustrated in Fig. 43. For low enough masses, in the gapless phase, the low-temperature behaviour appears to be linear for both the real and imaginary parts of ν_{qn} , but the slope is seemingly dependent of the mass. Note that we have no analytical results to contrast with this observation. For larger masses, in the gapped phase, numerics do not allow us to reliably probe the very low temperature regime.

Just as with the thermodynamic quantities studied in Sec. 4.2.3, the behaviour of the leading quasi-normal frequency near the critical point of the $U(n)^2$ model turns out to be very interesting. In Fig. 44 we plot $-\text{Im } \nu_{\text{qn}}$ as a function of the mass at $T = T_c$ and as a function of the temperature at $m = m_c$. The results are consistent with a critical power-law behaviour,

$$\frac{\partial \text{Im } \nu_{\text{qn}}}{\partial m}(m, T_c) \underset{m \rightarrow m_c^\pm}{\sim} \frac{\nu_\pm}{|m - m_c|^{f_\pm}}, \quad \frac{\partial \text{Im } \nu_{\text{qn}}}{\partial T}(m_c, T) \underset{T \rightarrow T_c^\pm}{\sim} \frac{\tilde{\nu}_\pm}{|T - T_c|^{\tilde{f}_\pm}}, \quad (5.51)$$

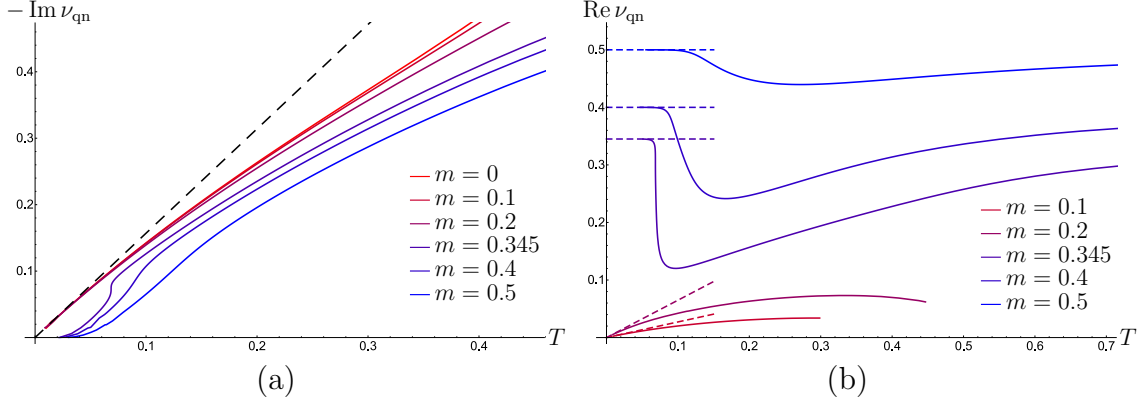


Figure 42: The leading quasi-normal frequency ν_{qn} in the $U(n)^2$ model, as a function of the temperature and for a few fixed masses. The low-temperature behaviour for the SYK-like phase with $m < m_* = 0.213_2$ is given by Eq. (5.48). (a) Imaginary part $-\text{Im } \nu_{\text{qn}}$, matching $-\text{Im } \nu_{\text{qn}} = \pi T/2$ for the SYK-like phase at small T (black dashed line). (b) Real part $\text{Re } \nu_{\text{qn}}$, matching $\text{Re } \nu_{\text{qn}} = 2\pi a T$ for the SYK-like phase and $\text{Re } \nu_{\text{qn}} = m$ for the HO-like phase at small T (dashed lines, with the parameter a extrapolated from the finite-temperature fits of ~ 10 solutions with $\beta \in [50, 100]$).

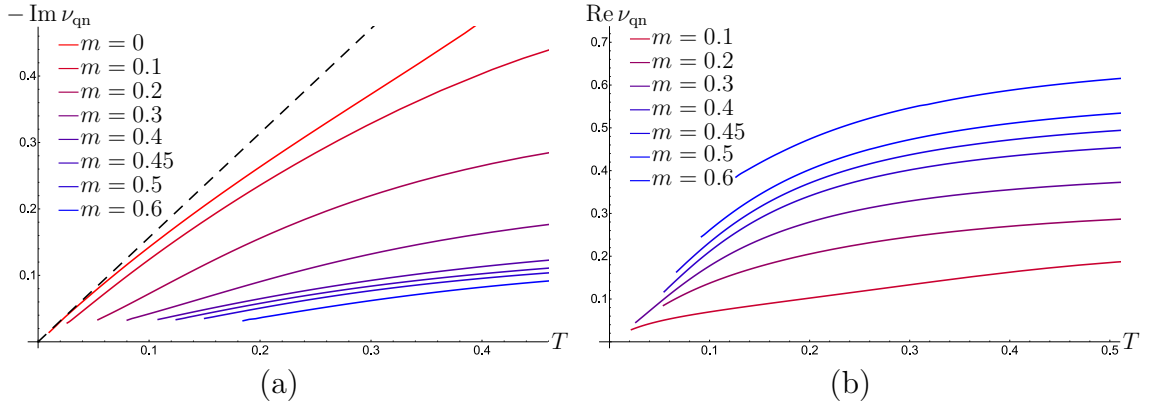


Figure 43: The leading quasi-normal frequency ν_{qn} in the $U(n)$ model, as a function of the temperature and for a few fixed masses. (a) Imaginary part $-\text{Im } \nu_{\text{qn}}$; (b) Real part $\text{Re } \nu_{\text{qn}}$. In the gapless phase, the low-temperature behaviour is linear with seemingly mass-dependent slopes, matching $-\text{Im } \nu_{\text{qn}} = \pi T/2$ only for $m = 0$ (black dashed line).

with critical exponents found to be

$$f_+ = 0.64_1, \quad f_- = 0.67_1, \quad \tilde{f}_+ = 0.64_2, \quad \tilde{f}_- = 0.57_2. \quad (5.52)$$

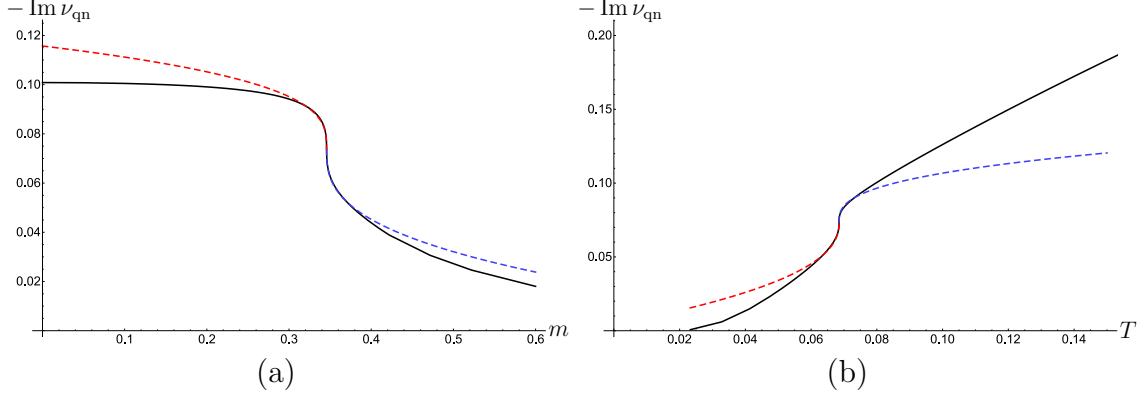


Figure 44: Leading quasi-normal frequency $-\text{Im } \nu_{\text{qn}}$ as a function of mass at $T = T_c$ (a) and as a function of temperature at $m = m_c$ (b), together with the best power-law fits near the critical point (m_c, T_c) (dashed lines). Power-law fits were performed as in Sec. 4.2.3, see *e.g.* Fig. 29.

5.2 Real time four-point functions

This subsection is devoted to the computation of real time four-point functions, in particular in the out-of-time-ordered (OTOC) case which is relevant to probe chaos [6, 33, 76, 77]. We review basic properties in Sec. 5.2.1 and the Schwinger-Keldysh set-up in Sec. 5.2.2. We introduce the Lyapunov exponent η_L in Sec. 5.2.3 and explain in Sec. 5.2.4 how the exact formulas derived in Sec. 5.2.2 can be simplified if one wishes to compute η_L only, instead of the detailed time evolution of the four-point function. The simple numerical algorithms presented in Sec. 5.2.5 are then used in Sec. 5.2.6 to compute η_L in various cases. We uncover in the process new and interesting features. In particular, we find that the chaos bound of [33] is saturated in the SYK-like phase of the $U(n)^2$ model, but find evidence that this is not so in the $U(n)$ model except at strictly $m = 0$. We also find that the Lyapunov exponent has a singular behaviour near the critical point of the $U(n)^2$ model, governed by non-trivial critical exponents.

5.2.1 General definitions and properties

A four-point function for complex time arguments u_i satisfying $0 < \tau_{u_i} < \beta$, generalizing (3.120), can be defined by

$$\phi(u_1, u_2, u_3, u_4) = \begin{cases} \frac{1}{n^2} \langle \text{T tr } \psi_\mu(u_1) \psi_\mu^\dagger(u_2) \text{tr } \psi_\nu(u_3) \psi_\nu^\dagger(u_4) \rangle_\beta & \text{(matrix models)} \\ \langle \text{T } \chi^i(u_1) \chi_i^\dagger(u_2) \chi^j(u_3) \chi_j^\dagger(u_4) \rangle_\beta & \text{(disordered models)}, \end{cases} \quad (5.53)$$

where as usual T indicates the ordering with respect to the Euclidean times $\tau_{u_i} = \text{Re } u_i$. This definition can be extended to all values of τ_{u_i} using the KMS condition

$$\phi(u_1 + n_1\beta, u_2 + n_2\beta, u_3 + n_3\beta, u_4 + n_4\beta) = (-1)^{n_1+n_2+n_3+n_4} \phi(u_1, u_2, u_3, u_4), \quad n_i \in \mathbb{Z}. \quad (5.54)$$

This four-point function satisfies the properties

$$\phi(u_1, u_2, u_3, u_4)^* = \phi(-u_4^*, -u_3^*, -u_2^*, -u_1^*), \quad (5.55)$$

$$\phi(u_1, u_2, u_3, u_4) = \phi(u_3, u_4, u_1, u_2), \quad (5.56)$$

$$\phi(u_1, u_2, u_3, u_4)^* = \phi(u_1^*, u_2^*, u_3^*, u_4^*), \quad (5.57)$$

generalizing the similar conditions (3.122), (3.123) and (3.124) in the Euclidean setting. One may decompose

$$\phi(u_1, u_2, u_3, u_4) = \sum_{\sigma \in \mathfrak{S}_4} \Theta(\tau_{u_{\sigma(1)}}, \tau_{u_{\sigma(2)}}, \tau_{u_{\sigma(3)}}, \tau_{u_{\sigma(4)}}) \phi_\sigma(u_1, u_2, u_3, u_4), \quad (5.58)$$

where the sum is over all the elements of the group $\mathfrak{S}_4$ of permutations of four elements,

$$\Theta(\tau_1, \tau_2, \tau_3, \tau_4) = \Theta(\tau_1 - \tau_2) \Theta(\tau_2 - \tau_3) \Theta(\tau_3 - \tau_4) \quad (5.59)$$

orders the Euclidean times and ϕ_σ is one of the $4! = 24$ possible four-point correlators with a fixed given ordering of the operators. Note that, by construction, the Euclidean times appearing as the arguments of ϕ_σ are always ordered according to the permutation σ , with $\tau_{u_{\sigma(1)}} > \tau_{u_{\sigma(2)}} > \tau_{u_{\sigma(3)}} > \tau_{u_{\sigma(4)}}$ and $0 < \tau_{u_{\sigma(1)}} - \tau_{u_{\sigma(4)}} < \beta$. The real times (imaginary parts of the u_i), on the other hand, are unconstrained. Any purely real time four-point function can be obtained by taking the Euclidean times to zero while keeping their ordering fixed. Spectral decomposition formulas read

$$\phi_\sigma(u_1, u_2, u_3, u_4) = \int_{-\infty}^{+\infty} \prod_{j=1}^4 [d\omega_j] \delta\left(\sum_{j=1}^4 \omega_j\right) \mathfrak{r}_\sigma(\omega_1, \omega_2, \omega_3, \omega_4) e^{-\sum_{j=1}^4 \omega_j u_j}, \quad (5.60)$$

with spectral functions that can be expressed in terms of sums over the eigenstates $|p\rangle$ and eigenvalues E_p of the Hamiltonian,

$$\begin{aligned} \delta\left(\sum_{j=1}^4 \omega_j\right) \mathfrak{r}_\sigma(\omega_1, \omega_2, \omega_3, \omega_4) &= \delta\left(\sum_{j=1}^4 \omega_j\right) \frac{1}{Z} \sum_{p,q,r,s} e^{-\beta E_p} \alpha_{pqrs}^\sigma \\ &\delta(\omega_{\sigma(1)} + E_p - E_q) \delta(\omega_{\sigma(2)} + E_q - E_r) \delta(\omega_{\sigma(3)} + E_r - E_s). \end{aligned} \quad (5.61)$$

The coefficients α_{pqrs}^σ depend on the matrix elements of the operators and on the permutation σ in a natural way. If we use the explicit notation

$$\sigma = [\sigma(1)\sigma(2)\sigma(3)\sigma(4)], \quad (5.62)$$

we have for example in the matrix model notation

$$\alpha_{pqrs}^{[1234]} = \langle p | \psi_{\mu b}^a | q \rangle \langle q | \psi_{\mu a}^{\dagger b} | r \rangle \langle r | \psi_{\nu d}^c | s \rangle \langle s | \psi_{\nu c}^{\dagger d} | p \rangle, \quad (5.63)$$

$$\alpha_{pqrs}^{[2134]} = -\langle p | \psi_{\mu a}^{\dagger b} | q \rangle \langle q | \psi_{\mu b}^a | r \rangle \langle r | \psi_{\nu d}^c | s \rangle \langle s | \psi_{\nu c}^{\dagger d} | p \rangle, \text{ etc.} \quad (5.64)$$

Correlators with the same cyclic ordering of the operators are related by the KMS conditions (5.54). These conditions follow, for instance, from the cyclicity property of the coefficients α_{pqrs}^σ ,

$$\alpha_{pqrs}^{[\sigma(1)\sigma(2)\sigma(3)\sigma(4)]} = -\alpha_{spqr}^{[\sigma(4)\sigma(1)\sigma(2)\sigma(3)]}. \quad (5.65)$$

In terms of the spectral functions (5.61), they read

$$\mathfrak{r}_{[\sigma(4)\sigma(1)\sigma(2)\sigma(3)]}(\omega_1, \omega_2, \omega_3, \omega_4) = -e^{\beta\omega_{\sigma(4)}} \mathfrak{r}_{[\sigma(1)\sigma(2)\sigma(3)\sigma(4)]}(\omega_1, \omega_2, \omega_3, \omega_4). \quad (5.66)$$

The KMS conditions thus reduce from $4! = 24$ to $3! = 6$ the number of independent spectral functions or correlators. Moreover, in our particular case one may check that the additional symmetry (5.56) leaves only four independent spectral functions or correlators,

$$\phi_{[1234]}(u_1, u_2, u_3, u_4) = \langle \psi_{\mu b}^a(u_1) \psi_{\mu a}^{\dagger b}(u_2) \psi_{\nu d}^c(u_3) \psi_{\nu c}^{\dagger d}(u_4) \rangle_\beta, \quad (5.67)$$

$$\phi_{[2134]}(u_1, u_2, u_3, u_4) = -\langle \psi_{\mu a}^{\dagger b}(u_2) \psi_{\mu b}^a(u_1) \psi_{\nu d}^c(u_3) \psi_{\nu c}^{\dagger d}(u_4) \rangle_\beta, \quad (5.68)$$

$$\phi_{[1324]}(u_1, u_2, u_3, u_4) = -\langle \psi_{\mu b}^a(u_1) \psi_{\nu d}^c(u_3) \psi_{\mu a}^{\dagger b}(u_2) \psi_{\nu c}^{\dagger d}(u_4) \rangle_\beta, \quad (5.69)$$

$$\phi_{[3214]}(u_1, u_2, u_3, u_4) = -\langle \psi_{\nu d}^c(u_3) \psi_{\mu a}^{\dagger b}(u_2) \psi_{\mu b}^a(u_1) \psi_{\nu c}^{\dagger d}(u_4) \rangle_\beta. \quad (5.70)$$

The other two *a priori* independent correlators are given by

$$\phi_{[2314]}(u_1, u_2, u_3, u_4) = -\phi_{[1324]}(u_3, u_4, u_1, u_2 - \beta), \quad (5.71)$$

$$\phi_{[3124]}(u_1, u_2, u_3, u_4) = -\phi_{[2134]}(u_3, u_4 + \beta, u_1, u_2). \quad (5.72)$$

The above extends the discussion of Sec. 5.1.1 for the two-point functions, for which there is only one independent spectral function. It can be straightforwardly generalized to arbitrary p -point functions, with generically $(p-1)!$ independent spectral functions, see *e.g.* [78] for a thorough discussion.

5.2.2 Melonic real time four-point functions

By time translation invariance, we can restrict our attention to $\phi(u_1, u_2, u_3, u_4 = 0)$. For the purpose of studying chaos, along the lines of [6, 33], we can further limit ourselves to a real

$$u_3 = \tau_3. \quad (5.73)$$

We also fix the real parts of u_1 and u_2 to be τ_+ and τ_- (or τ_- and τ_+) such that

$$0 < \tau_- < \tau_3 < \tau_+ < \beta, \quad (5.74)$$

and we note

$$t_1 = \text{Im } u_1, \quad t_2 = \text{Im } u_2. \quad (5.75)$$

The associated Keldysh contour Γ for these choices of insertion points has two triangle-shaped pieces and is depicted in Fig. 45. The height of the triangles must be at least $\max(t_1, t_2)$ but may conveniently be taken to be infinite.

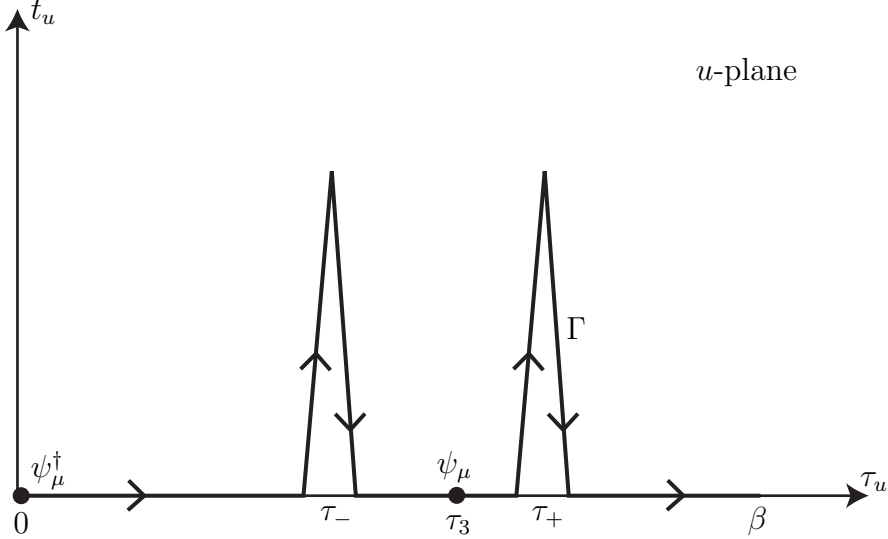


Figure 45: Keldysh contour Γ (thick line) used to compute the four-point functions with a choice of triplet (τ_-, τ_3, τ_+) for the Euclidean part of their time arguments. The triangle-shaped pieces of the contour are centered on the Euclidean times τ_- and τ_+ . Their width ϵ is taken to zero and their height to infinity. Fixed insertions at $u = 0$ and $u = \tau_3$ are indicated.

As already emphasized in Sec. 5.1.2, the advantage of the Keldysh contour formalism is that the Feynman rules and diagrammatics take the familiar form, with only simple and obvious modifications with respect to the standard Euclidean case. In particular, the arguments presented in Sec. 3.4 go through. The leading Feynman graphs are still the ladders depicted in Fig. 17, 18 and 19, but now the positions of

the vertices have to be integrated over the Keldysh contour Γ instead of the Euclidean interval $[0, \beta]$. If we decompose

$$\phi(u_1, u_2, u_3, u_4) = G(u_1, u_2)G(u_3, u_4) + \frac{1}{N}\mathcal{F}(u_1, u_2, u_3, u_4), \quad (5.76)$$

as in (3.125), $\mathcal{F}$ is determined by the equation

$$\mathcal{F}(u_1, u_2, \tau_3, 0) = \mathcal{F}_0(u_1, u_2, \tau_3, 0) + \int_{\Gamma \times \Gamma} \mathcal{K}(u_1, u_2, u, u') \mathcal{F}(u, u', \tau_3, 0) du du'. \quad (5.77)$$

Generalizing the matrix multiplication rule (3.126) to

$$(\mathbb{A} \cdot \mathbb{B})(u_1, u_2, u_3, u_4) = \int_{\Gamma \times \Gamma} \mathcal{A}(u_1, u_2, u, u') \mathcal{B}(u, u', u_3, u_4) du du', \quad (5.78)$$

with integrals taken along Γ , (5.77) is solved by resumming a geometric series or equivalently inverting the matrix $\mathbb{I} - \mathbb{K}$,

$$\mathbb{F} = \sum_{n \geq 0} \mathbb{K}^n \cdot \mathbb{F}_0 = (\mathbb{I} - \mathbb{K})^{-1} \cdot \mathbb{F}_0, \quad (5.79)$$

exactly as in (3.127). The function $\mathcal{F}_0$ and the kernel $\mathcal{K}$ are given by the same formulas as in the Euclidean case, (3.128) and (3.129) (in the $U(n)^2$ model) or (3.130) (in the $U(n)$ model), with arguments $(\tau_1, \tau_2, \tau_3, \tau_4)$ replaced by (u_1, u_2, u_3, u_4) and the Euclidean two-point function $G(\tau_1, \tau_2)$ replaced by the more general complex time two-point function $G(u_1, u_2)$ defined in (5.3).

The points u_1 and u_2 in (5.77) may be chosen to be anywhere on the contour Γ . The real time OTOCs correspond to insertions on the two triangles, *i.e.*

$$u_1 = \tau_+ + it_1, \quad u_2 = \tau_- + it_2 \quad \text{or} \quad u_1 = \tau_- + it_1, \quad u_2 = \tau_+ + it_2, \quad (5.80)$$

which yield

$$\begin{aligned} \mathcal{C}_1(t_1, t_2) &= -\phi_{[1324]}(\tau_+ + it_1, \tau_- + it_2, \tau_3, 0) \\ &= \langle \psi_{\mu b}^a(\tau_+ + it_1) \psi_{\nu d}^c(\tau_3) \psi_{\mu a}^{\dagger b}(\tau_- + it_2) \psi_{\mu c}^{\dagger d} \rangle_{\beta}, \end{aligned} \quad (5.81)$$

$$\begin{aligned} \mathcal{C}_2(t_1, t_2) &= \phi_{[2314]}(\tau_- + it_1, \tau_+ + it_2, \tau_3, 0) \\ &= \langle \psi_{\mu a}^{\dagger b}(\tau_+ + it_2) \psi_{\nu d}^c(\tau_3) \psi_{\mu b}^a(\tau_- + it_1) \psi_{\mu c}^{\dagger d} \rangle_{\beta}. \end{aligned} \quad (5.82)$$

These correlators are decomposed, following (5.76), as

$$\mathcal{C}_1(t_1, t_2) = -G(\tau_+ - \tau_- + i(t_1 - t_2))G(\tau_3) + \frac{1}{N}C_1(t_1, t_2), \quad (5.83)$$

$$\mathcal{C}_2(t_1, t_2) = G(\tau_- - \tau_+ + i(t_1 - t_2))G(\tau_3) + \frac{1}{N}C_2(t_1, t_2), \quad (5.84)$$

with

$$C_1(t_1, t_2) = -\mathcal{F}(\tau_+ + it_1, \tau_- + it_2, \tau_3, 0), \quad (5.85)$$

$$C_2(t_1, t_2) = \mathcal{F}(\tau_- + it_1, \tau_+ + it_2, \tau_3, 0). \quad (5.86)$$

The two connected OTOCs (5.85) and (5.86) can be inserted on the left-hand side of Eq. (5.77) by setting u_1 and u_2 to the appropriate values (5.80). However, it is important to realize that several other types of correlators then necessarily appear on the right-hand side of (5.77) when one performs the integrals over Γ . This includes, for example, correlators like

$$\langle \psi_{\mu b}^a(\tau_+ + it_1) \psi_{\mu a}^{\dagger b}(\tau_+ + it_2) \psi_{\nu d}^c(\tau_3) \psi_{\mu c}^{\dagger d} \rangle_{\beta}, \quad \langle \psi_{\mu a}^{\dagger b}(\tau_+ + it_2) \psi_{\mu b}^a(\tau_+ + it_1) \psi_{\nu d}^c(\tau_3) \psi_{\mu c}^{\dagger d} \rangle_{\beta}, \quad (5.87)$$

when u and u' are both on the triangle at $\text{Re } u = \text{Re } u' = \tau_+$, similar correlators when u and u' are both on the triangle at $\text{Re } u = \text{Re } u' = \tau_-$, and also correlators with only one real time turned on or purely Euclidean correlators when both u and u' are on the Euclidean segments of Γ .

It is thus impossible to write an exact closed-form system of integral equations similar to (5.77) involving the OTOCs alone.²⁸ However, we shall explain in Sec. 5.2.4 that such an *approximate* system can be used if we are only concerned with evaluating the asymptotic, long time behaviour. It is this behaviour which is relevant to compute the chaos Lyapunov exponents, which we discuss in the following.

5.2.3 On Lyapunov exponents

Let us start by summarizing the standard lore on the late real time behaviour of four-point functions such as $\phi(u_1, u_2, \tau_3, 0)$ in models like the ones we are studying. Recall that t_1 and t_2 are defined in (5.75); we want to discuss the regime where $t = t_1 = t_2$ is very large.

We assume that the number N of degrees of freedom is large but finite. This is important, because time scales of the order of $\ln N$ play a role in the discussion. Much larger time scales, like the Poincaré recurrence time scale, which are exponentially large with N , are irrelevant for our purposes and can be considered to be infinite.

Depending on the ordering of the operators in the correlators, *i.e.* on the positions of u_1 and u_2 on the Keldysh contour of Fig. 45, two qualitatively distinct behaviours are expected:

- i) If u_1 and u_2 are on the same triangle-shaped piece of Γ , then the correlator will factorize and approach a constant value as soon as $t = t_1 = t_2$ is much larger than the dissipation time scale defined in (5.50),

$$\phi(u_1, u_2, \tau_3, 0) \Big|_{t=t_1=t_2 \gg t_{\text{qn}}} \simeq G_{\pm}(0)G(\tau_3). \quad (5.88)$$

²⁸This point seems to be overlooked in part of the existing literature.

The function G_+ or G_- appears on the right-hand side of (5.88) depending on the relative position of u_1 and u_2 along Γ . Note that the result (5.88) is consistent with the large N expansion, since the right-hand side of this equation matches with the first term on the right-hand side of (5.76). In other words, the long time and large N limits commute for this class of correlators.²⁹

- ii) If u_1 and u_2 are on distinct triangle-shaped pieces of Γ , which is the case corresponding to the OTOCs (5.81) and (5.82), it is expected that the correlator will go to zero when $t = t_1 = t_2 \rightarrow \infty$,

$$\lim_{t=t_1=t_2 \rightarrow +\infty} \phi(u_1, u_2, \tau_3, 0) = 0. \quad (5.89)$$

Intuitively, this is because the insertions at u_1 and u_2 are infinitely far apart on the Keldysh contour Γ in the limit $t \rightarrow \infty$. The result (5.89) can actually be used as a good criterion to detect chaos in a strongly coupled quantum system [33]. Indeed, (5.89) is inconsistent with the first term on the right-hand side of (5.76), which predicts (5.88) at $N \rightarrow \infty$. In other words, the long time and large N limits do not commute for the OTOCs. The $\mathcal{F}/N$ term, which naively is a small correction at large N , must eventually become of order one to cancel out the constant GG term. More precisely, it is expected that

$$\mathcal{F}(u_1, u_2, \tau_3, 0) \propto e^{\eta_L t} \quad (5.90)$$

as soon as $t = t_1 = t_2 \gg t_{\text{qn}}$, the exponential growth of the “small” correction $\mathcal{F}/N$ signalling the onset of chaos. The constant η_L is called the Lyapunov exponent. Eventually, the OTOCs should vanish on a time scale, called the scrambling time scale, that can be estimated by setting $\mathcal{F}/N \sim 1$ in (5.90),

$$t_{\text{scr}} \approx \frac{\ln N}{\eta_L}. \quad (5.91)$$

A few clarifying remarks may be made at this stage. First, in the strict $N \rightarrow \infty$ limit studied in the present paper the scrambling time $t_{\text{scr}} \sim \ln N$ is not seen and the vanishing of the OTOCs cannot be rigorously derived. This is a non-perturbative effect in the large N expansion.

On the other hand, the behaviour (5.90) can in principle be derived by solving the exact equation (5.77). In practice, this can be done only in the low-temperature limit and in the conformal phases of the models. The method, presented in [9] for SYK (see also [30, 79–81]), does not use (5.77) explicitly. Instead, one first works in the Euclidean setting and exploits the conformal symmetry to explicitly invert the operator $\mathbb{I} - \mathbb{K}$ in (3.127), and then performs the appropriate analytic continuations to real time on the exact solution. The result is the famous formula

$$\lim_{T \rightarrow 0} \frac{\eta_L}{T} = 2\pi \quad (5.92)$$

²⁹As indicated above, we do not consider time scales larger than $\ln N$ in our discussion.

for the Lyapunov exponent in the zero-temperature limit, which matches with the expectations from black hole physics and holography [77]. The universality of the result (5.92) can be understood as being the consequence of the small explicit breaking of the low energy reparameterization invariance (3.38)-(3.39) by UV effects, which yields a universal Schwarzian action for the reparameterization zero modes from which the leading behaviour (5.90) can be obtained.

At non-zero temperature one can in principle derive the behaviour (5.90) by solving (5.77) numerically. We first compute the kernel by using the solution for the two-point functions presented in Sec. 5.1, and then invert the operator $\mathbb{I} - \mathbb{K}$ to get $\mathcal{F}$ from (5.79). This procedure is in principle straightforward and allows the computation of the OTOCs from the very short time scales $0 \leq t_1, t_2 \ll t_{\text{qn}}$ to the chaos regime $t_1, t_2 \gg t_{\text{qn}}$. However, the size of the matrix $\mathbb{I} - \mathbb{K}$ that one needs to invert to obtain reliable results when $t_1, t_2 \gg t_{\text{qn}}$ is very large. To the best of our knowledge, this has never been done and we let this interesting challenge for the future.

Now, if the detailed time evolution of the OTOCs is not required, but one is only interested in the regime $t \gg t_{\text{qn}}$, the exact equation (5.77) can be simplified. This is sufficient to extract the Lyapunov exponents, and is explained in the next subsection. The drawback of this approach is that it relies on the standard expectations concerning the behaviour of the real time correlators reviewed above, discarding in a somewhat heuristic way many terms on the right-hand side of (5.77). It is therefore not a first principles derivation strictly proving that this standard lore is indeed correct. However, the results obtained in this way are fully consistent with the original hypothesis and we believe that they are very reliable, see Sec. 5.2.5 for examples.

An important and well-motivated conjecture [33] about the Lyapunov exponents η_{L} is the famous upper bound

$$\eta_{\text{L}} \leq 2\pi T. \quad (5.93)$$

This bound is supposed to be valid in a large class of models for which a large hierarchy exists between the dissipation time scale and the scrambling time scale, $t_{\text{scr}} \gg t_{\text{d}}$ [33]. It is saturated by black holes and, as Eq. (5.92) shows, by the SYK model and many generalizations thereof, including the $U(n)^2$ model in the standard conformal phase (as emphasized in Sec. 4, the standard conformal phase is not realized in the $U(n)$ model; this will be further discussed in Sec. 5.2.5).

The OTOCs we are considering are labelled by an arbitrary triplet (τ_-, τ_3, τ_+) constrained by (5.74). The most natural choice, which yields physical, real time finite temperature correlation functions, corresponds to taking the limits

$$\tau_+ \rightarrow 0, \quad \tau_3 \rightarrow 0, \quad \tau_- \rightarrow 0, \quad (5.94)$$

keeping the ordering (5.74) fixed. We shall call the associated OTOCs the “physical OTOCs.” The existing literature has instead mainly focused on a different choice, which is singled out in [33] and corresponds to

$$\tau_- = \beta/4, \quad \tau_3 = \beta/2, \quad \tau_+ = 3\beta/4. \quad (5.95)$$

We will refer to these as the “standard” correlators, which can be interpreted as particular expectation values in a model for which we take two copies of the original Hilbert space and in a state which is a deformed version of the standard thermofield double state [33].

It is not obvious whether the bound (5.93) is valid independently of the choice of triplet (τ_-, τ_3, τ_+) , even though we believe it is very likely to be so in melonic models. We shall compute Lyapunov exponents for different choices of triplets (τ_-, τ_3, τ_+) in Sec. 5.2.5 and our results support this claim. However, it is important to stress that the proof presented in [33] does not apply to the general case of arbitrary (τ_-, τ_3, τ_+) but only to the special choice $\tau_- = \beta/4$, $\tau_3 = \beta/2$, $\tau_+ = 3\beta/4$.³⁰ Indeed, the reasoning crucially depends on the possibility to analytically continue the OTOC from the real time $t = t_1 = t_2$ to a strip $-\beta/4 < \text{Im } t < \beta/4$ in the complex t -plane. This is indeed possible when $\tau_- = \beta/4$, $\tau_3 = \beta/2$, $\tau_+ = 3\beta/4$, but, unfortunately, no such analytic continuation exists in the physical case. In general, an analytic continuation is possible in a strip $-\min(\tau_3 - \tau_-, \beta - \tau_+) < \text{Im } t < \min(\tau_-, \tau_+ - \tau_3)$, whose width can never exceed, but is in general strictly less than, $\beta/2$; the argument of [33] may then be used to derive an upper bound which is less strict than (5.93).

5.2.4 Computing Lyapunov exponents

Lyapunov exponents are determined by analyzing the long time exponential growth of the connected OTOCs (5.85) and (5.86), and comparing with (5.90). We start by writing (5.77) for the relevant values (5.80) of u_1 and u_2 . Using the complex time version of (3.128), we get

$$C_1(t_1, t_2) = G_+(\tau_+ + it_1)G_+(\tau_3 - \tau_- - it_2) - \int_{\Gamma \times \Gamma} \mathcal{K}(\tau_+ + it_1, \tau_- + it_2, u, u') \mathcal{F}(u, u', \tau_3, 0) du du', \quad (5.96)$$

$$C_2(t_1, t_2) = -G_+(\tau_- + it_1)G_-(\tau_3 - \tau_+ - it_2) + \int_{\Gamma \times \Gamma} \mathcal{K}(\tau_- + it_1, \tau_+ + it_2, u, u') \mathcal{F}(u, u', \tau_3, 0) du du'. \quad (5.97)$$

The idea is that we can simplify the right-hand sides of these equations, because we are interested in the regime in which the OTOCs are exponentially large. We thus keep only the terms that can develop such an exponential growth in the integrals and neglect all the others. From the discussion in Sec. 5.2.3, we expect that these terms correspond to the integration regions $u = \tau_+ + it$ and $u' = \tau_- + it'$ or $u = \tau_- + it$ and $u' = \tau_+ + it'$, for which $\mathcal{F}(u, u', \tau_3, 0)$ coincides with the OTOCs (5.85), (5.86). We are left with integrals running along the two sides of the triangles, which are non-trivial because the kernel function $\mathcal{K}$ can undergo discontinuities when u or u' goes through

³⁰We would like to thank Douglas Stanford for a discussion of this point.

an insertion point at u_1 or u_2 . To take into account these discontinuities, it is useful to decompose the kernel along the lines of (5.58),

$$\mathcal{K}(u_1, u_2, u_3, u_4) = \sum_{\sigma \in \mathfrak{S}_4} \Theta(\tau_{u_{\sigma(1)}}, \tau_{u_{\sigma(2)}}, \tau_{u_{\sigma(3)}}, \tau_{u_{\sigma(4)}}) \mathcal{K}_\sigma(u_1, u_2, u_3, u_4). \quad (5.98)$$

We get in this way a closed system of equations involving C_1 and C_2 only,

$$C_1(t_1, t_2) \approx G_+(\tau_+ + it_1)G_+(\tau_3 - \tau_- - it_2) + \int_0^{+\infty} \left[\mathcal{K}_{1,1}(t_1, t_2, t, t')C_1(t, t') + \mathcal{K}_{1,2}(t_1, t_2, t, t')C_2(t, t') \right] dt dt', \quad (5.99)$$

$$C_2(t_1, t_2) \approx -G_+(\tau_- + it_1)G_-(\tau_3 - \tau_+ - it_2) + \int_0^{+\infty} \left[\mathcal{K}_{2,1}(t_1, t_2, t, t')C_1(t, t') + \mathcal{K}_{2,2}(t_1, t_2, t, t')C_2(t, t') \right] dt dt', \quad (5.100)$$

with

$$\begin{aligned} \mathcal{K}_{1,1}(t_1, t_2, t_3, t_4) &= -\Theta(t_1 - t_3)\Theta(t_2 - t_4) \\ &\quad \left[\mathcal{K}_{[1324]} - \mathcal{K}_{[1342]} - \mathcal{K}_{[3124]} + \mathcal{K}_{[3142]} \right] (\tau_+ + it_1, \tau_- + it_2, \tau_+ + it_3, \tau_- + it_4), \\ \mathcal{K}_{1,2}(t_1, t_2, t_3, t_4) &= \Theta(t_2 - t_3)\Theta(t_1 - t_4) \\ &\quad \left[\mathcal{K}_{[1423]} - \mathcal{K}_{[4123]} - \mathcal{K}_{[1432]} + \mathcal{K}_{[4132]} \right] (\tau_+ + it_1, \tau_- + it_2, \tau_- + it_3, \tau_+ + it_4), \\ \mathcal{K}_{2,1}(t_1, t_2, t_3, t_4) &= \Theta(t_2 - t_3)\Theta(t_1 - t_4) \\ &\quad \left[\mathcal{K}_{[2314]} - \mathcal{K}_{[2341]} - \mathcal{K}_{[3214]} + \mathcal{K}_{[3241]} \right] (\tau_- + it_1, \tau_+ + it_2, \tau_+ + it_3, \tau_- + it_4), \\ \mathcal{K}_{2,2}(t_1, t_2, t_3, t_4) &= -\Theta(t_1 - t_3)\Theta(t_2 - t_4) \\ &\quad \left[\mathcal{K}_{[2413]} - \mathcal{K}_{[4213]} - \mathcal{K}_{[2431]} + \mathcal{K}_{[4231]} \right] (\tau_- + it_1, \tau_+ + it_2, \tau_- + it_3, \tau_+ + it_4). \end{aligned} \quad (5.101)$$

The $\approx$ signs in (5.99) and (5.100) emphasize the fact that this is a truncated version of the exact equations, to be used only to evaluate the late time behaviour of the OTOCs. These equations can be presented in a convenient matrix form

$$\begin{pmatrix} \mathbb{C}_1 \\ \mathbb{C}_2 \end{pmatrix} \approx \begin{pmatrix} \mathbb{C}_1^{(0)} \\ \mathbb{C}_2^{(0)} \end{pmatrix} + \begin{pmatrix} \mathbb{K}_{1,1} & \mathbb{K}_{1,2} \\ \mathbb{K}_{2,1} & \mathbb{K}_{2,2} \end{pmatrix} \star \begin{pmatrix} \mathbb{C}_1 \\ \mathbb{C}_2 \end{pmatrix}, \quad (5.102)$$

where

$$C_1^{(0)}(t_1, t_2) = G_+(\tau_+ + it_1)G_+(\tau_3 - \tau_- - it_2), \quad (5.103)$$

$$C_2^{(0)}(t_1, t_2) = -G_+(\tau_- + it_1)G_-(\tau_3 - \tau_+ - it_2), \quad (5.104)$$

and we define the $\star$ -product between components in (5.102) as

$$(\mathbb{K} \star \mathbb{C})(t_1, t_2) = \int_0^{+\infty} \mathcal{K}(t_1, t_2, t, t')C(t, t') dt dt'. \quad (5.105)$$

The above formulas can be applied in any melonic model with ladder kernel $\mathcal{K}$. For the special cases of the $U(n)^2$ and $U(n)$ models (3.129) and (3.130), the right-hand sides of (5.101) can be conveniently expressed in terms of the retarded and advanced functions defined in (5.14) and (5.15), as well as the particular Wightman two-point correlators

$$G_{W+}(t) = G_+(\tau_+ - \tau_- + it), \quad G_{W-}(t) = G_-(\tau_- - \tau_+ + it). \quad (5.106)$$

We find

$$\begin{aligned} \mathcal{K}_{1,1}^{U(n)^2}(t_1, t_2, t_3, t_4) &= -2\lambda^2 G_r(t_1 - t_3) G_a(t_4 - t_2) G_{W+}(t_3 - t_4) G_{W-}(t_4 - t_3), \\ \mathcal{K}_{1,2}^{U(n)^2}(t_1, t_2, t_3, t_4) &= \lambda^2 G_r(t_1 - t_4) G_a(t_3 - t_2) G_{W+}(t_4 - t_3)^2, \\ \mathcal{K}_{2,1}^{U(n)^2}(t_1, t_2, t_3, t_4) &= \lambda^2 G_r(t_1 - t_4) G_a(t_3 - t_2) G_{W-}(t_4 - t_3)^2, \\ \mathcal{K}_{2,2}^{U(n)^2}(t_1, t_2, t_3, t_4) &= -2\lambda^2 G_r(t_1 - t_3) G_a(t_4 - t_2) G_{W+}(t_4 - t_3) G_{W-}(t_3 - t_4), \end{aligned} \quad (5.107)$$

and

$$\begin{aligned} \mathcal{K}_{1,1}^{U(n)}(t_1, t_2, t_3, t_4) &= \frac{3}{4} |\xi|^2 G_r(t_1 - t_3) G_a(t_4 - t_2) [G_{W+}(t_3 - t_4)^2 + G_{W-}(t_4 - t_3)^2], \\ \mathcal{K}_{1,2}^{U(n)}(t_1, t_2, t_3, t_4) &= -\frac{3}{2} |\xi|^2 G_r(t_1 - t_4) G_a(t_3 - t_2) G_{W+}(t_4 - t_3) G_{W-}(t_3 - t_4), \\ \mathcal{K}_{2,1}^{U(n)}(t_1, t_2, t_3, t_4) &= -\frac{3}{2} |\xi|^2 G_r(t_1 - t_4) G_a(t_3 - t_2) G_{W+}(t_3 - t_4) G_{W-}(t_4 - t_3), \\ \mathcal{K}_{2,2}^{U(n)}(t_1, t_2, t_3, t_4) &= \frac{3}{4} |\xi|^2 G_r(t_1 - t_3) G_a(t_4 - t_2) [G_{W+}(t_4 - t_3)^2 + G_{W-}(t_3 - t_4)^2]. \end{aligned} \quad (5.108)$$

5.2.5 Numerical implementation

The equations (5.99) and (5.100), or equivalently (5.102), can be used to compute $C_1(t_1, t_2)$ and $C_2(t_1, t_2)$ numerically by introducing a cutoff $t_{\max}$, discretizing the real time variables $(t_1, t_2) \in [0, t_{\max}]^2$, and inverting the resulting finite size matrix. Note that the integrals effectively extend from zero to t_1 or t_2 due to the presence of the Θ functions in (5.101), so that this procedure allows us to obtain $C_1(t_1, t_2)$ and $C_2(t_1, t_2)$ for all $t_1, t_2 < t_{\max}$ without further truncations. The choice of the cutoff is also related to the support of the kernels (5.107) or (5.108), which can be estimated at low temperatures using the conformal forms derived from (5.8) with $\rho(\omega)$ given by (3.53), or alternatively determined beforehand from the explicit numerical solution for the real time two-point functions.

In order to see the exponential growth of the OTOCs we need to reliably evaluate C_1 and C_2 up to long times. To accurately approximate the integrals and achieve an appropriate resolution we may therefore need a significant number of points P

in the $[0, t_{\max}]$ grid. The runtime of the algorithm is dominated by the $P^2 \times P^2$ matrix inversion, which for the generic matrices we are dealing with here can be done in $\mathcal{O}(P^{2\delta})$ for some $2 < \delta < 3$. This is the main reason why the truncated system (5.99)-(5.100) is much more convenient than the exact equation (5.77), since the latter involves more correlators and integrals thus enlarging the final size of the matrix to be inverted.

In practice, our Mathematica implementation allows for $P \sim 10^2$ points in the interval $[0, t_{\max}]$, resulting in the inversion of $\sim 10^4 \times 10^4$ matrices performed in a few seconds with a desktop computer.³¹ The typical time scale beyond which the exponential behaviour of C_1 and C_2 is seen is of the order t_{qn} , and we have shown in Figs. 42a and 43a that $t_{\text{qn}}^{-1} = -\text{Im} \nu_{\text{qn}}$ behaves linearly with T at low-temperatures in the SYK-like phase of the $U(n)^2$ and $U(n)$ models. We are thus required to take $t_{\max} \sim \beta$, so that the resolution requirements then impose a lower bound on the temperatures we can reach. We find that this bound is of the same order of that imposed by the real time algorithm used for the computation of the two-point functions themselves, so it is only occasionally a limitation.

Finally, we should note that while the physical correlators defined by (5.94) are better motivated, there is a practical advantage associated to the standard choice corresponding to (5.95). Indeed, in the latter case the system (5.102) decouples when taking the combinations $C_1 \pm C_2$, further reducing the size of the problem. Moreover, $C_1 + C_2$ is real and displays exponential growth, whereas we find that $|C_1 - C_2|$ is bounded and need not be computed if only Lyapunov exponents are sought. We show in Fig. 46 the combination $C_1 + C_2$, obtained from (5.102) as explained above, for typical values of the parameters m and T .³² The exponential behaviour is clearly seen at long times and yields a precise determination of the associated Lyapunov exponents through linear fitting in log-scale. Note that the values at short times are not reliable, since we use the truncated system (5.102) and not the exact equations (5.77).

We can also use an alternative method to obtain the Lyapunov exponents, due to Kitaev, that can be justified in the following way. We start from the truncated system (5.99)-(5.100) in which we plug the following ansatz,

$$C_j(t_1 + t_0, t_2 + t_0) = e^{\frac{\eta_L}{2}(t_1+t_2)+\eta_L t_0} D_j\left(\frac{t_1 - t_2}{2}\right) \quad \text{for } j = 1, 2, \quad (5.109)$$

for some *a priori* unknown functions D_1 and D_2 . In the limit of very large t_0 , the first term on the right-hand sides of (5.99) and (5.100) drops. It is convenient to use

³¹Much improvement can be expected from a GPU implementation, which might enable the numerical solution of the full set of equations (5.77). We found however that this was not necessary for our purposes here.

³²The truncated system (5.102) can be used only to compute the exponentially growing contributions to the correlators and thus does not yield a reliable evaluation of the combination $C_1 - C_2$. All we can say in this case is that $|C_1 - C_2|$ is not exponentially growing.

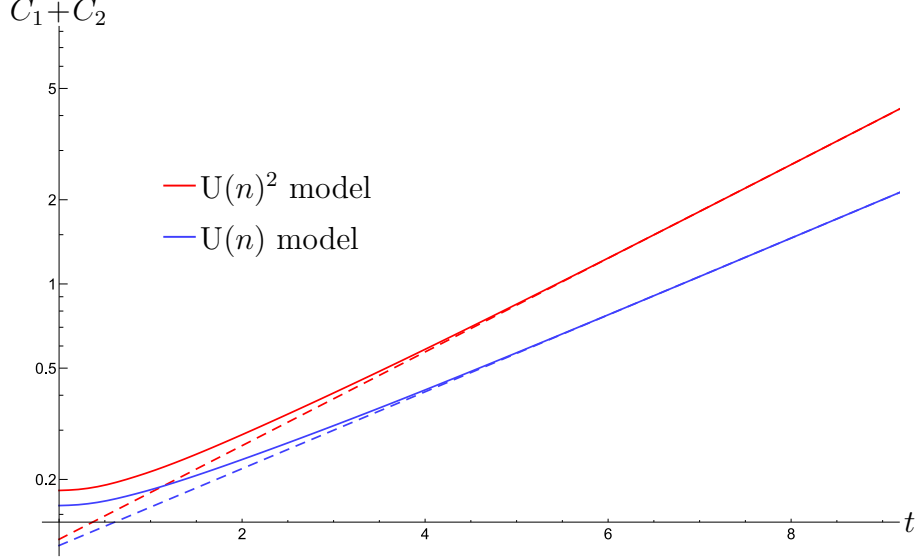


Figure 46: Log-plot of the numerical evaluation of $C_1(t, t) + C_2(t, t)$, as obtained from (5.102), for the standard choice of correlators (5.95) at $m = 0.2$ and $T = 0.1$ in natural units ($U(n)^2$ model in red, $U(n)$ model in blue). The exponential growth is clearly seen at long times, with associated Lyapunov exponents $\frac{\eta_L}{2\pi T} = 0.614_5$ and $\frac{\eta_L}{2\pi T} = 0.503_5$ obtained for the $U(n)^2$ and $U(n)$ models, respectively, through linear fits in log-scale (shown with dashed lines). These results were calculated using $P = 100$ points and a cutoff of $t_{\max} = 10.5$.

the new variables

$$T = \frac{t_1 + t_2}{2}, \quad X = \frac{t_1 - t_2}{2}, \quad T' = \frac{t + t'}{2}, \quad X' = \frac{t - t'}{2}, \quad (5.110)$$

to get

$$e^{\eta_L T} D_j(X) = 2 \int_{-\infty}^{+\infty} dT' dX' e^{\eta_L T'} \left[\mathcal{K}_{j,1}(T + X, T - X, T' + X', T' - X') D_1(X') \right. \\ \left. + \mathcal{K}_{j,2}(T + X, T - X, T' + X', T' - X') D_2(X') \right], \quad j = 1, 2. \quad (5.111)$$

The integrals over T' can be performed and we are left with a new linear system of the form

$$\begin{pmatrix} \mathbb{D}_1 \\ \mathbb{D}_2 \end{pmatrix} = \begin{pmatrix} \mathbb{M}_{1,1} & \mathbb{M}_{1,2} \\ \mathbb{M}_{2,1} & \mathbb{M}_{2,2} \end{pmatrix} * \begin{pmatrix} \mathbb{D}_1 \\ \mathbb{D}_2 \end{pmatrix}, \quad (5.112)$$

where the $*$ -product between components now corresponds to

$$(\mathbb{M} * \mathbb{D})(X) = \int_{-\infty}^{+\infty} \mathcal{M}(X, X') D(X') dX'. \quad (5.113)$$

More explicitly, if

$$a(x) = \int_{|x|}^{+\infty} G_r(x+T)G_a(u-T)e^{-\eta_L T} dT, \quad (5.114)$$

we find for the $U(n)^2$ and $U(n)$ models respectively,

$$\begin{aligned} \mathcal{M}_{1,1}^{U(n)^2}(X, X') &= -4\lambda^2 G_{W+}(2X')G_{W-}(-2X')a(X - X'), \\ \mathcal{M}_{1,2}^{U(n)^2}(X, X') &= 2\lambda^2 G_{W+}(-2X')^2 a(X + X'), \\ \mathcal{M}_{2,1}^{U(n)^2}(X, X') &= 2\lambda^2 G_{W-}(-2X')^2 a(X + X'), \\ \mathcal{M}_{2,2}^{U(n)^2}(X, X') &= -4\lambda^2 G_{W+}(-2X')G_{W-}(2X')a(X - X'), \end{aligned} \quad (5.115)$$

and

$$\begin{aligned} \mathcal{M}_{1,1}^{U(n)}(X, X') &= \frac{3}{2}|\xi|^2 [G_{W+}(2X')^2 + G_{W-}(-2X')^2] a(X - X'), \\ \mathcal{M}_{1,2}^{U(n)}(X, X') &= -3|\xi|^2 G_{W-}(2X')G_{W+}(-2X')a(X + X'), \\ \mathcal{M}_{2,1}^{U(n)}(X, X') &= -3|\xi|^2 G_{W-}(-2X')G_{W+}(2X')a(X + X'), \\ \mathcal{M}_{2,2}^{U(n)}(X, X') &= \frac{3}{2}|\xi|^2 [G_{W+}(-2X')^2 + G_{W-}(2X')^2] a(X - X'). \end{aligned} \quad (5.116)$$

Equation (5.112) shows that the ansatz (5.109) is consistent as long as the matrix $\begin{pmatrix} \mathbb{M}_{1,1} & \mathbb{M}_{1,2} \\ \mathbb{M}_{2,1} & \mathbb{M}_{2,2} \end{pmatrix}$ has an eigenvalue one. The Lyapunov exponent will correspond to the largest η_L for which the matrix does have this eigenvalue, since such a value will obviously dominate at long times. Note that this largest η_L always exists; indeed, (5.114) shows that the matrix elements go to zero when $\eta_L \rightarrow \infty$. We thus get η_L by shooting for the largest eigenvalue of $\begin{pmatrix} \mathbb{M}_{1,1} & \mathbb{M}_{1,2} \\ \mathbb{M}_{2,1} & \mathbb{M}_{2,2} \end{pmatrix}$ to be one. Of course, the drawback is that this approach uses the exponential ansatz (5.109) *a priori*, whereas the exponential growth of the correlators emerges on its own from the solution of (5.102).

The method described above can be efficiently implemented numerically. The dimensions of the matrix $\begin{pmatrix} \mathbb{M}_{1,1} & \mathbb{M}_{1,2} \\ \mathbb{M}_{2,1} & \mathbb{M}_{2,2} \end{pmatrix}$ are linear in the number P of discretization points used to perform the integrals, instead of quadratic as in (5.102). This allows for greater values of $t_{\max}$, which then results in higher matrix sparsity. Specialized algorithms such as Arnoldi iteration can then be used to compute the largest eigenvalue of the matrix faster than its inverse. At this point, it also becomes important to avoid the $\mathcal{O}(P^2)$ naive computation of the one-dimensional kernel $a(x)$. Instead, one may rewrite (5.114) as a convolution to be computed in Fourier space in $\mathcal{O}(P \log P)$. At the end of the day, taking into account the binary search used to shoot η_L for the maximal eigenvalue equal to one, the algorithm runs in $\mathcal{O}(\log \frac{1}{\varepsilon} \times (P \log P + P^{\delta'}))$, where ε is the desired precision of η_L and δ' characterizes the diagonalization method

used. In practice, we find $P \sim 10^3$ and $\varepsilon = 10^{-5}$ to be feasible on a desktop computer within a few seconds.

We used both methods in all our evaluations of the Lyapunov exponents below, and the results turn out to be perfectly consistent.

Remark: Kitaev's method is sometimes justified by taking the large time $t = t_1 = t_2$ limit directly in (5.99) and (5.100), or (5.102), and using the quasi-normal behaviour of the two point functions to neglect the first term on the right-hand sides of these equations. This is not correct. In particular, the matrix $\begin{pmatrix} \mathbb{K}_{1,1} & \mathbb{K}_{1,2} \\ \mathbb{K}_{2,1} & \mathbb{K}_{2,2} \end{pmatrix}$ does *not* have an eigenvalue one. If it had one, the system (5.102) would not be invertible! Note also that the integrals in (5.99), (5.100) or (5.105) run from zero to infinity, unlike the integrals in (5.111) or (5.113) for which one integrates from minus infinity to plus infinity.

5.2.6 Applications

Zero-temperature limit and chaos bound In Fig. 47 and 48, we plot $\eta_L/(2\pi T)$ as a function of inverse temperature β , for various values of the mass parameter in the $U(n)^2$ and $U(n)$ models. These results were computed for the standard case (5.95), but complete agreement is seen with the Lyapunov exponents of the physical one (5.94), as exemplified for $m = 0.2$ in Fig. 49.

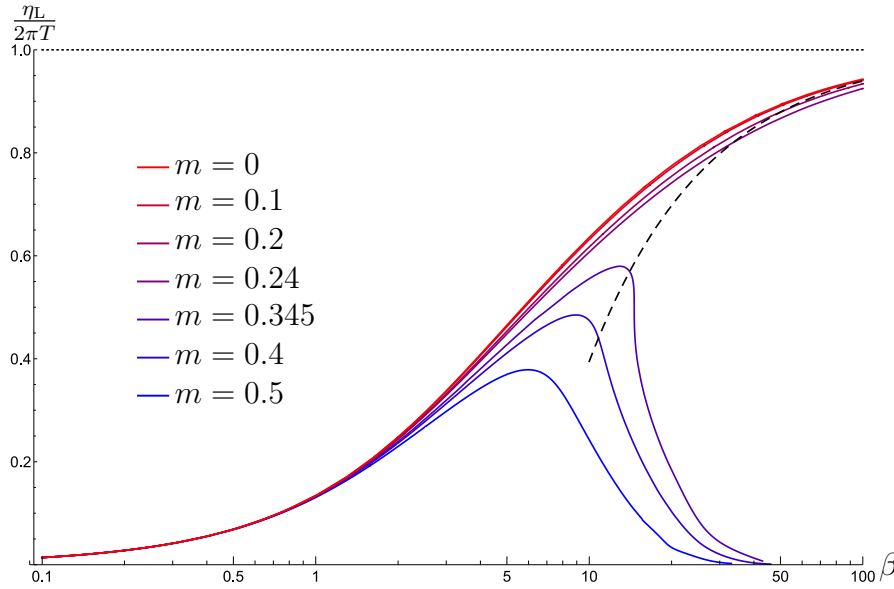


Figure 47: Normalized Lyapunov exponent $\eta_L/(2\pi T)$ of the $U(n)^2$ model, as a function of the inverse temperature β (in logarithmic scale) and for various masses. The bound $\eta_L/(2\pi T) = 1$ is marked by a horizontal dotted line, with the dashed black curve corresponding to the asymptotic result $\eta_L \simeq \frac{2\pi}{\beta}(1 - \frac{6.05}{\beta})$ computed for SYK in [9]. The bound is saturated in the SYK-like phase in the zero-temperature limit.

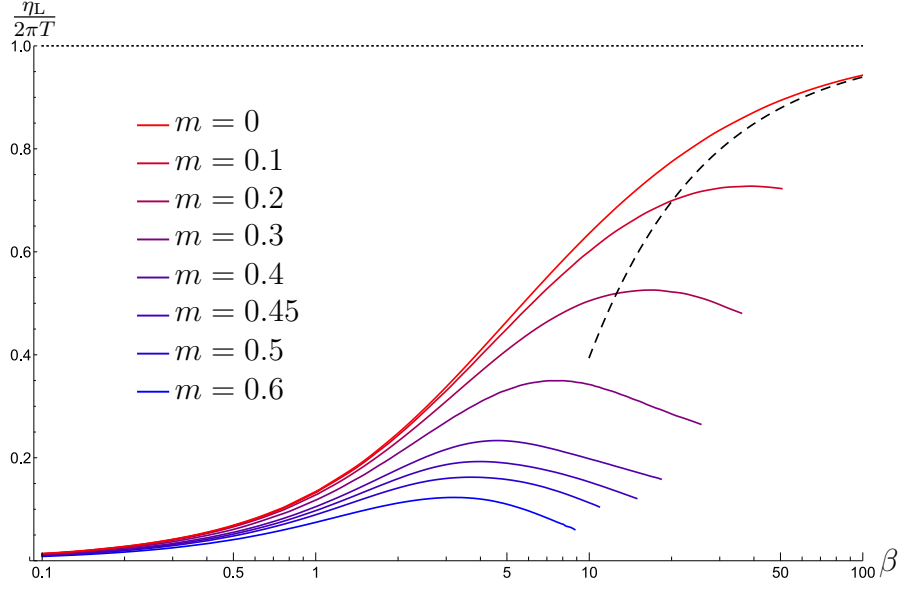


Figure 48: Normalized Lyapunov exponent $\eta_L/(2\pi T)$ of the $U(n)$ model, as a function of the inverse temperature β (in logarithmic scale) and for various masses. The bound $\eta_L/(2\pi T) = 1$ is marked by a horizontal dotted line, with the dashed black curve corresponding to the asymptotic result $\eta_L \simeq \frac{2\pi}{\beta}(1 - \frac{6.05}{\beta})$ computed for SYK in [9]. Saturation of the bound is only clearly seen for $m = 0$ and seems to be violated for all $m > 0$, including in the gapless phase $0 < m < m_*$.

In the case of the $U(n)^2$ model, the results are as expected. We find that the bound $\eta_L \leq 2\pi T$ of (5.93) is saturated for the SYK-like phase in the zero-temperature limit, including for $m > m_*$ where this solution is metastable (*e.g.* for $m = 0.24$ in Fig. 47).

The situation seems to be very different in the case of the $U(n)$ model. As soon as $m > 0$, we find that the curve for $\eta_L/(2\pi T)$ reaches a maximum and then starts to decrease as the temperature is lowered. Our numerics do not allow us to probe the deep IR regime, so we must remain cautious about the conclusions we may draw for the strict zero-temperature limit. However, the behaviour is clearly qualitatively different from the $U(n)^2$ model and suggests that the bound (5.93) is not saturated already in the phase where the non-standard ansatz discussed in Sec. 3.1.5 is valid. The fact that the $SL(2, \mathbb{R})$ subgroup of the reparameterization symmetry (3.38)-(3.39) is broken, including the scale invariance, is certainly crucial to understand this property. This seemingly very different behaviour of the Lyapunov exponent is contrasting with other properties that we have discussed previously, like the existence of a non-vanishing zero-temperature entropy, that are similar to the SYK-like phase of the $U(n)^2$ model. Clearly, this raises interesting questions with respect to the gravitational interpretation of the $U(n)$ model, in particular in relation with a possible description in terms of a suitable generalization of the Schwarzian action, and deserves further study.

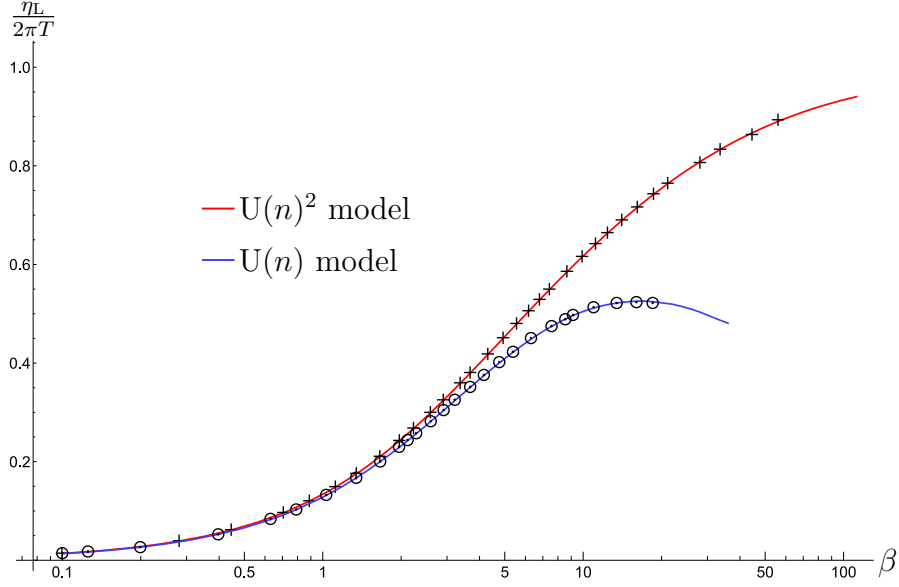


Figure 49: Lyapunov exponents of the $U(n)^2$ and $U(n)$ models, as functions of inverse temperature β (in logarithmic scale), for $m = 0.2$ and in natural units. Red and blue solid lines correspond to the standard correlators (5.95), with values for the physical case (5.94) represented with crosses and circles, for the $U(n)^2$ and $U(n)$ model respectively. All results were computed using Kitaev’s method with $P = 1001$ and $\varepsilon = 10^{-5}$. As discussed in Sec. 5.2.4 equations (5.102) can be decoupled in the standard case, leading to a faster and more precise solution allowing for slightly lower temperatures. An equally good agreement is observed between Lyapunov exponents for standard and physical correlators for other masses, both in SYK-like and HO-like phases.

Remark: The non-saturation of the chaos bound discussed above in the gapless phase of the $U(n)$ model must not be confused with the usual and expected non-saturation of the bound found in the gapped phase, for $m > m_*$. The gapped phases of both the $U(n)^2$ and $U(n)$ models are similar to what is found in the SYK model with a random mass term or in the Maldacena-Qi model, where non-saturation of the chaos bound was already observed [38, 54, 65].

Critical properties of the Lyapunov exponent We close our analysis with a discussion of the very interesting behaviour of the Lyapunov exponent near the critical point (m_c, T_c) of the $U(n)^2$ model. In Fig. 50a, $\eta_L/(2\pi T)$ is plotted as a function of m for $T = T_c$. We find a singular behaviour near $m = m_c$, consistent with

$$\frac{\partial \eta_L}{\partial m}(m, T_c) \underset{m \rightarrow m_c^\pm}{\sim} \frac{\eta_L^\pm}{|m - m_c|^{g_\pm}}, \quad (5.117)$$

critical exponents being

$$g_+ = 0.69_1 \quad , \quad g_- = 0.60_1 . \quad (5.118)$$

We have also studied $\eta_L/(2\pi T)$ as a function of T at $m = m_c$, see Fig. 50b. It is very plausible that critical exponents can be defined in this case as well, but our data is not precise enough to extract them in a reliable way.

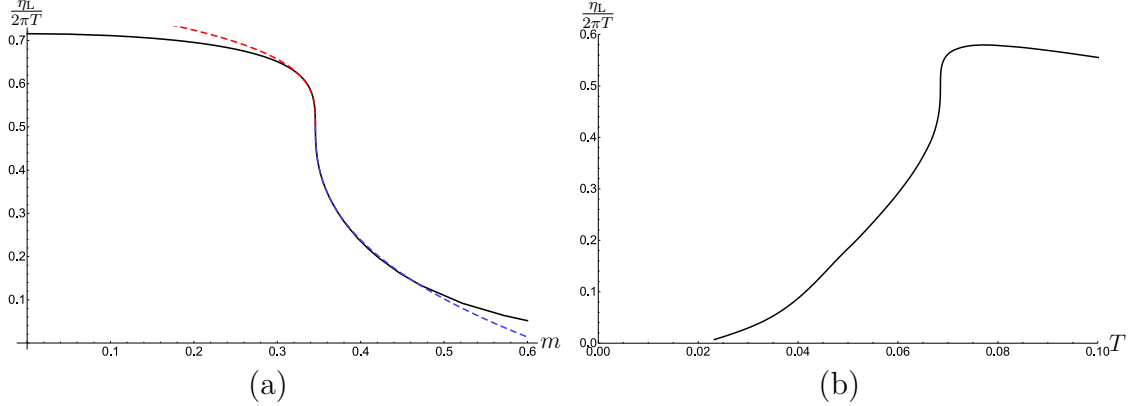


Figure 50: Lyapunov exponent for the $U(n)^2$ model, as a function of mass at fixed temperature $T = T_c$ (a) and as a function of temperature at fixed mass $m = m_c$ (b). For (a) critical exponents were obtained as in Sec. 4.2.3, see *e.g.* Fig. 29. In (b) a singular behaviour is clearly seen when $T \rightarrow T_c$, but numerical data is not precise enough to extract critical exponents in a reliable way.

Acknowledgements

We would like to thank Igor Klebanov, Juan Maldacena, Subir Sachdev, Douglas Stanford and Herman Verlinde for useful discussions and comments. F. S. M. acknowledges affiliation to the IBS Center for Theoretical Physics of the Universe (Seoul, South Korea) during the initial stages of this project. This research is supported in part by the Belgian Fonds National de la Recherche Scientifique FNRS (convention IISN 4.4503.15) and the Fédération Wallonie-Bruxelles (Advanced ARC project “Holography, Gauge Theories and Quantum Gravity”). This work was performed in part at the Aspen Center for Physics, which is supported by the National Science Foundation grant PHY-1607611, and was partially supported by a grant from the Simons Foundation.

A q -body interacting Hamiltonians

One can consider the following extension of the $U(n)^2$ model, which has interesting features discussed in Sec. 4.2. We generalize the random Hamiltonian (2.17) to

$$H_{\mathbf{q}} = Nm\chi_i^\dagger\chi^i + \sqrt{N} \sum_{q \in \mathbf{q}} \lambda_{j_1 \dots j_{q/2}}^{i_1 \dots i_{q/2}} \chi_{i_1}^\dagger \dots \chi_{i_{q/2}}^\dagger \chi^{j_1} \dots \chi^{j_{q/2}}, \quad (\text{A.1})$$

where the sum is over even integers $q \in \mathbf{q} = \{q_1, \dots, q_p\}$, $4 \leq q_1 < q_2 < \dots < q_p$, and the $\lambda_{j_1 \dots j_{q/2}}^{i_1 \dots i_{q/2}}$ are random Gaussian couplings satisfying constraints similar to the λ_{ij}^{kl} in (2.14) and (2.15). One can also provide a matrix model version of the model, using the complete interaction vertex studied in [4]. The effective action (3.86) is generalized to

$$\begin{aligned} \frac{1}{n^2 d} \mathcal{S}_{\text{eff}}^{U(n)^2}[\mathbf{G}, \Sigma] = & -\ln(1 + e^{-\beta m}) - \ln \frac{\det \mathcal{O}_\Sigma}{\det \mathcal{O}_0} + \int_{[0, \beta]^2} \left[\Sigma(\tau_2, \tau_1) \mathbf{G}(\tau_1, \tau_2) \right. \\ & \left. - \sum_{q \in \mathbf{q}} \frac{(-1)^{q/2}}{q} \lambda_q^2 \mathbf{G}(\tau_1, \tau_2)^{q/2} \mathbf{G}(\tau_2, \tau_1)^{q/2} \right] d\tau_1 d\tau_2, \quad (\text{A.2}) \end{aligned}$$

and the Schwinger-Dyson equation (3.30) is replaced by

$$\Sigma(\tau) = \sum_{q \in \mathbf{q}} (-1)^{q/2} \lambda_q^2 G(\tau)^{q/2} G(-\tau)^{q/2-1}. \quad (\text{A.3})$$

At low energy and when the mass parameter is below a critical value (see Sec. 4), the models flow to a conformal phase. In this regime, only the q_1 -body interaction term contributing in (A.3) is relevant. The ansatz (3.42) and Eq. (3.44), (3.45) are still valid. The generalizations of (3.43), (3.46), (3.47), (3.50) and (3.51) read

$$b_+ = \frac{C}{\Gamma(1-2\Delta)} \frac{\sin(\pi\Delta + \theta)}{\sin(2\pi\Delta)}, \quad b_- = \frac{C}{\Gamma(1-2\Delta)} \frac{\sin(\pi\Delta - \theta)}{\sin(2\pi\Delta)}, \quad (\text{A.4})$$

$$\Delta = \frac{1}{q_1}, \quad (\text{A.5})$$

$$C^{q_1} = \frac{(1-2\Delta)(\Gamma(1-2\Delta))^{q_1}}{\pi \lambda_{q_1}^2} \frac{(\sin(2\pi\Delta))^{q_1-1}}{(\sin(\pi\Delta + \theta) \sin(\pi\Delta - \theta))^{\frac{q_1}{2}-1}}, \quad (\text{A.6})$$

$$G(\tau) = \left(\frac{\pi}{\beta}\right)^{2\Delta} \frac{e^{-2\pi a \tau / \beta}}{(\sin(\pi|\tau|/\beta))^{2\Delta}} \left(b_+ \Theta(\tau) - b_- \Theta(-\tau) \right), \quad (\text{A.7})$$

$$a = \frac{1}{2\pi} \ln \frac{b_+}{b_-} = \frac{1}{2\pi} \ln \frac{\sin(\pi\Delta + \theta)}{\sin(\pi\Delta - \theta)}. \quad (\text{A.8})$$

In Sec. 4.2 we study the thermodynamics of the models. The formulas (3.95) or (3.96) for the fermion number are still valid, but the formulas for the free energy

(3.98) and the energy (3.101) or (3.102) are generalized to

$$\begin{aligned} \frac{F}{N} = m - T \ln(1 + e^{\beta m}) + T \sum_{k \in \mathbb{N} + \frac{1}{2}} \ln \left| \frac{G_k}{G_k^{(0)}} \right|^2 \\ + \sum_{q \in \mathbf{q}} \frac{q-1}{q} (-1)^{q/2} \lambda_q^2 \int_0^\beta G(\tau)^{q/2} G(-\tau)^{q/2} d\tau, \quad (\text{A.9}) \end{aligned}$$

$$\begin{aligned} \frac{E}{N} = m n_F(m) - \frac{2m}{\beta} \sum_{k \in \mathbb{N} + \frac{1}{2}} \text{Re}(G_k - G_k^0) \\ - 2 \sum_{q \in \mathbf{q}} \frac{(-1)^q}{q} \lambda_q^2 \int_0^\beta G(\tau)^{q/2} G(-\tau)^{q/2} d\tau. \quad (\text{A.10}) \end{aligned}$$

If only one coupling λ_q is turned on, these formulas simplify to

$$\frac{F}{N} = m - T \ln(1 + e^{\beta m}) + T \sum_{k \in \mathbb{N} + \frac{1}{2}} \left[\ln \left| \frac{G_k}{G_k^{(0)}} \right|^2 + \frac{2(q-1)}{q} \left(1 - \text{Re} \frac{G_k}{G_k^{(0)}} \right) \right], \quad (\text{A.11})$$

$$\frac{E}{N} = m n_F(m) - \frac{1}{\beta} \sum_{k \in \mathbb{N} + \frac{1}{2}} \left[2m \text{Re}(G_k - G_k^0) + \frac{4}{q} \left(1 - \text{Re} \frac{G_k}{G_k^{(0)}} \right) \right]. \quad (\text{A.12})$$

B Algebraic derivation of the kernel formula

The goal of this Appendix is to provide all the elementary details of the algebraic derivation of the formulas (3.127) and (3.133) given in Sec. 3.4 for the connected four-point function.

We start from the equations (3.139), (3.140) and (3.142). In order to simplify the notation we write functions or operators like $G(t_1, t_2)$ or $\mathcal{O}_\Sigma$ as G_{ij} or $(\mathcal{O}_\Sigma)_{ij}$, integrals over t are replaced by sums and functional derivatives by partial derivatives. In these notations, and using (3.142) to express the free energy in terms of the on-shell effective action, the four-point function (3.139) is given by

$$\mathcal{F}_{ijkl} = -\frac{1}{N} \frac{\partial^2 \hat{S}_{\text{eff}}}{\partial J_{ji} \partial J_{lk}} \Big|_{J=0}, \quad (\text{B.1})$$

and the equations (3.127) and (3.133) that we want to prove are equivalent to

$$\mathcal{F}_{ijkl} = -G_{il} G_{kj} + \sum_{p,q} \mathcal{K}_{ijpq} \mathcal{F}_{pqkl}, \quad (\text{B.2})$$

$$\mathcal{K}_{ijkl} = \sum_{p,q} G_{ip} G_{qj} \frac{\partial^2 s}{\partial G_{qp} \partial G_{kl}}. \quad (\text{B.3})$$

From the explicit expressions (3.131) and (3.140), we get

$$\frac{1}{N} \frac{\partial \hat{\mathbf{S}}_{\text{eff}}}{\partial \Sigma_{ji}} = -(\mathcal{O}_{\Sigma+J}^{-1})_{ij} + \mathbf{G}_{ij}, \quad (\text{B.4})$$

$$\frac{1}{N} \frac{\partial \hat{\mathbf{S}}_{\text{eff}}}{\partial \mathbf{G}_{ji}} = \Sigma_{ij} + \frac{\partial s}{\partial \mathbf{G}_{ji}}[\mathbf{G}], \quad (\text{B.5})$$

$$\frac{1}{N} \frac{\partial \hat{\mathbf{S}}_{\text{eff}}}{\partial J_{ji}} = -(\mathcal{O}_{\Sigma+J}^{-1})_{ij}. \quad (\text{B.6})$$

In particular, the on-shell values $G[J]$ and $\Sigma[J]$ satisfy the saddle-point equations

$$G_{ij} = (\mathcal{O}_{\Sigma+J}^{-1})_{ij}, \quad (\text{B.7})$$

$$\Sigma_{ij} = -\frac{\partial s}{\partial G_{ji}}[G]. \quad (\text{B.8})$$

Taking the derivative with respect to J of the definition of the on-shell action,

$$\hat{S}_{\text{eff}}[J] = \hat{\mathbf{S}}_{\text{eff}}[G[J], \Sigma[J], J], \quad (\text{B.9})$$

and using the saddle-point equations together with (B.6), we get

$$\frac{1}{N} \frac{\partial \hat{S}_{\text{eff}}}{\partial J_{ji}} = \frac{1}{N} \frac{\partial \hat{\mathbf{S}}_{\text{eff}}}{\partial J_{ji}}[G[J], \Sigma[J], J] = -G_{ij}. \quad (\text{B.10})$$

The definition (B.1) thus implies that

$$\mathcal{F}_{ijkl} = \frac{\partial G_{ij}}{\partial J_{lk}}. \quad (\text{B.11})$$

Let us now differentiate the saddle-point equations (B.7) and (B.8) with respect to J . We get

$$\frac{\partial G_{ij}}{\partial J_{lk}} = -G_{il}G_{kj} - \sum_{r,s} G_{ir}G_{sj} \frac{\partial \Sigma_{rs}}{\partial J_{lk}}, \quad (\text{B.12})$$

$$\frac{\partial \Sigma_{rs}}{\partial J_{lk}} = -\sum_{p,q} \frac{\partial^2 s}{\partial G_{sr} \partial G_{pq}} \frac{\partial G_{pq}}{\partial J_{lk}}, \quad (\text{B.13})$$

so that plugging (B.13) into (B.12) we find that (B.11) yields

$$\mathcal{F}_{ijkl} = -G_{il}G_{kj} + \sum_{p,q,r,s} G_{ir}G_{sj} \frac{\partial^2 s}{\partial G_{sr} \partial G_{pq}} \mathcal{F}_{pqkl}. \quad (\text{B.14})$$

This is precisely the equation (B.2) we wanted to prove, with the definition (B.3) for the kernel.

For the models we are studying, the relevant $s[\mathbf{G}]$ functionals are (3.143) and (3.144), or in the notation of this Appendix

$$s_{\mathrm{U}(n)^2}[\mathbf{G}] = -\frac{\lambda^2}{4} \sum_{i,j} \mathbf{G}_{ij}^2 \mathbf{G}_{ji}^2, \quad (\text{B.15})$$

$$s_{\mathrm{U}(n)}[\mathbf{G}] = \frac{|\xi|^2}{4} \sum_{i,j} \mathbf{G}_{ij}^3 \mathbf{G}_{ji}. \quad (\text{B.16})$$

Then

$$(\mathcal{K}^{\mathrm{U}(n)^2})_{ijkl} = -\lambda^2 \left[G_{il} G_{kj} G_{lk}^2 + 2G_{ik} G_{lj} G_{lk} G_{kl} \right], \quad (\text{B.17})$$

$$(\mathcal{K}^{\mathrm{U}(n)})_{ijkl} = \frac{3}{4} |\xi|^2 \left[(G_{lk}^2 + G_{kl}^2) G_{ik} G_{lj} + 2G_{kl} G_{lk} G_{il} G_{kj} \right]. \quad (\text{B.18})$$

C On the numerical errors

Many of the results presented in this paper are numerical in nature, consisting either of specific values for various relevant quantities (critical temperatures, masses, exponents, *etc.*) or observations concerning the solutions of Schwinger-Dyson equations, such as *e.g.* the existence of monodromies in the space of parameters of the $\mathrm{U}(n)^2$ model. It is therefore important to understand the limitations of the numerical methods employed, which we briefly discuss in this Appendix.

Lacking statistical fluctuations in our results given the exclusive use of deterministic algorithms, there are two main classes of numerical errors or uncertainties to consider:

- i) Systematic errors are incurred in whenever imposing a cutoff to truncate an infinite set of equations or realize a limiting procedure, such as in (3.113) for the Euclidean setting and (5.40) for the real time one; or introducing a discretization to approximate continuous quantities, as in (5.42) for the real time case. This is essentially unavoidable, and completely trumps the numerical errors inherent to the machine-precision floating-point arithmetic used during all computations.³³
- ii) Methodological uncertainties are present whenever fitting numerical results to analytical models, either for parameter extraction (scaling dimensions, critical exponents, *etc.*) or for zero-temperature extrapolations. They also naturally appear whenever using binary search to determine quantities such as critical temperatures and masses, or Lyapunov exponents as explained in Sec. 5.2.5.

³³Proper care needs to be taken to ensure numerical stability in the real time case, given the potentially large exponential factors in the coefficients of (5.36) and (5.37).

As already mentioned in Sec. 3.2.3 and 3.3, systematic errors translate into any and all quantities computed using the numerical solutions we obtain. While we have not rigorously bounded them in terms of the various parameters involved, they can be roughly estimated by comparing with exact analytical results whenever possible. For example, to estimate the precision of the thermodynamic functions computed for the $U(n)^2$ model at $T = 10^{-3}$ in natural units $\lambda = 1$, we may choose a large enough mass to make sure that the only solution to the Schwinger-Dyson equations is the perturbative one, very close to (3.7) at this temperature. In this case taking $m = 1$ is certainly enough, so that using a cutoff of $k_{UV} = 2^{16}$ we find $F/N \simeq 6 \times 10^{-5}$ instead of $F/N = 0$ as expected for (3.7) at $T = 0$. We may then assume that a systematic error of $\sim 10^{-4}$ is present when computing free energies at $T = 10^{-3}$ with a cutoff $k_{UV} = 2^{16}$, also for other large masses in the $U(n)^2$ model. An analogous procedure can be carried out for the entropy at small masses comparing with the exact formula (3.111), yielding an estimate of ~ 0.1 for the systematic error in the numerical results when using a cutoff $k_{UV} = 2^{16}$ at $T = 10^{-3}$. This is consistent with the $\beta = 10^3$ factor relating the entropy and free energy in (3.104), so that the systematic errors are observed to be approximately mass-independent in this case, at least for large enough cutoffs and low enough temperatures.

The methodological uncertainties, on the other hand, originate in most cases both in the fitting models and techniques we use, as well as in the actual data we choose to analyze. The effect of the latter is generally the most relevant one, with the statistical error of the fitting procedures being considerably smaller than the typical fluctuations observed when modifying the underlying data. While we have not attempted here a rigorous treatment of these issues, we can again roughly estimate their effects. For example, when computing scaling dimensions at small temperatures by fitting the zero-temperature ansätze (3.42) or (3.55), as shown in Fig. 10 and 14, we need to choose a fitting interval satisfying $1 \ll \tau \ll \beta$ in natural units $\lambda = |\xi| = 1$. One simple choice is to take a 10% section of the full Euclidean time interval, say $\tau/\beta \in [0.2, 0.12]$ for some large β , and keep track of the fluctuations of the fitted scaling dimensions when considering 10% fluctuations of its endpoints, *i.e.* when fitting the ranges $\tau/\beta \in [0.1, 0.11]$ and $\tau/\beta \in [0.3, 0.13]$. Similar considerations apply to the calculation of critical exponents, where there is an additional uncertainty in the critical temperatures and masses to be used, so that fluctuations in these parameters need to be accounted for in an analogous manner.

Other error estimation methods are of course possible, but the important observation here is that results show certain universality: for all reasonable choices, even if arbitrary, the numerical fluctuations we encounter are consistent. The methodological error bar obtained in this way is then fairly robust. Moreover, when the method above is applied systematically it serves the additional purpose of providing an indicator for relative certainty in these determinations. It is in this sense that we may begin to quantify the naked-eye impression that the determination of Δ_- in Fig. 14 is less precise than that of Δ_+ .

In practice, both types of numerical errors and uncertainties described above are intertwined. For example, the systematic error in numerically computed thermodynamic functions prevents us from distinguishing branches of solutions to the Schwinger-Dyson equations with arbitrary precision, thereby forcing us to prematurely interrupt binary searches for critical temperatures (thus prompting a methodological uncertainty). Alternatively, we may use extrapolation procedures as shown in Sec. 4.2.2 to account for systematic errors, obtaining numerical results that have better precision than one could naively expect, see *e.g.* Fig. 25.

The error bars we report in the various quantities discussed in this work account for the methodological uncertainties, estimated as explained above, and for systematic errors only as far as they are encompassed by the former. They are therefore not exhaustive, and subject to vary at least slightly for different data and methodologies. It is for these reasons we do not present them in figures, but provide them only for the most important values explicitly referenced in the main text. In any case, the general features and conclusions we derive from these values never depend on the fine-grained details. Indeed, wherever this could be an issue, such as in the detailed characterization of the low-temperature conformal phase of the $U(n)$ model, we opt for the more conservative route and only provide definite assertions if the numerical results are conclusive beyond reasonable doubt.

References

- [1] R. Gurau, “The $1/N$ expansion of colored tensor models,” *Annales Henri Poincaré* **12** (2011) 829 [arXiv:1011.2726],
R. Gurau and V. Rivasseau, “The $1/N$ expansion of colored tensor models in arbitrary dimension,” *Europhys. Lett.* **95** (2011) 50004 [arXiv:1101.4182],
R. Gurau, “The complete $1/N$ expansion of colored tensor models in arbitrary dimension,” *Annales Henri Poincaré* **13** (2012) 399 [arXiv:1102.5759].
- [2] S. Sachdev and J. Ye, “Gapless Spin-Fluid Ground State in a Random Quantum Heisenberg Magnet,” *Phys. Rev. Lett.* **70** (1993) 3339 [cond-mat/9212030].
- [3] F. Ferrari, “The Large D Limit of Planar Diagrams,” to appear in *Ann. I. H. Poincaré D*, arXiv:1701.01171 [hep-th].
- [4] F. Ferrari, V. Rivasseau and G. Valette, “A New Large N Expansion for General Matrix-Tensor Models,” arXiv:1709.07366 [hep-th].
- [5] T. Azeyanagi, F. Ferrari, P. Gregori, L. Leduc and G. Valette, “More on the New Large D Limit of Matrix Models,” *Ann. Phys.* **393**, 308 (2018) [arXiv:1710.07263].

- [6] A. Kitaev, “A Simple Model of Quantum Holography,” KITP Program *Entanglement in Strongly-Correlated Quantum Matter*, unpublished; see <http://online.kitp.ucsb.edu/online/entangled15/>.
A. Kitaev and S. J. Suh, “The soft mode in the Sachdev-Ye-Kitaev model and its gravity dual,” *JHEP* **1805** (2018) 183 [arXiv:1711.08467].
- [7] W. Fu, D. Gaiotto, J. Maldacena and S. Sachdev, “Supersymmetric Sachdev-Ye-Kitaev models,” *Phys. Rev. D* **95** (2017) no.2, 026009 Addendum: [*Phys. Rev. D* **95** (2017) no.6, 069904] [arXiv:1610.08917],
J. Yoon, “Supersymmetric SYK Model: Bi-local Collective Superfield/Supermatrix Formulation,” *JHEP* **1710** (2017) 172 [arXiv:1706.05914],
K. Bulycheva, “ $\mathcal{N} = 2$ SYK model in the superspace formalism,” *JHEP* **1804** (2018) 036 [arXiv:1801.09006],
C. Peng, “ $\mathcal{N} = (0, 2)$ SYK, Chaos and Higher-Spins,” *JHEP* **1812** (2018) 065 [arXiv:1805.09325],
P. Narayan and J. Yoon, “Supersymmetric SYK Model with Global Symmetry,” *JHEP* **1808** (2018) 159 [arXiv:1712.02647].
- [8] D. J. Gross and V. Rosenhaus, “A Generalization of Sachdev-Ye-Kitaev,” *JHEP* **1702** (2017) 093 [arXiv:1610.01569],
V. Bonzom, L. Lionni and A. Tanasa, “Diagrammatics of a colored SYK model and of an SYK-like tensor model, leading and next-to-leading orders,” *J. Math. Phys.* **58** (2017) no.5, 052301 [arXiv:1702.06944].
- [9] J. Maldacena and D. Stanford, “Remarks on the Sachdev-Ye-Kitaev model,” *Phys. Rev. D* **94** (2016) no.10, 106002 [arXiv:1604.07818].
- [10] V. Rosenhaus, “An introduction to the SYK model,” arXiv:1807.03334 [hep-th].
- [11] O. Parcollet and A. Georges, “Non-Fermi liquid regime of a doped Mott insulator,” *Phys. Rev. B* **59** (1999) 5341 [cond-mat/9806119].
- [12] A. Georges, O. Parcollet and S. Sachdev, “Mean-Field Theory of a Quantum Heisenberg Spin Glass,” *Phys. Rev. Lett.* **85** (2000) 840 [cond-mat/9909239].
- [13] O. Parcollet and A. Georges, “Transition from overscreening to underscreening in the multichannel Kondo model: exact solution at large N ,” *Phys. Rev. Lett.* **79** (1997) 4665 [cond-mat/9707337],
O. Parcollet, A. Georges, G. Kotliar and A. Sengupta, “Overscreened multi-channel $SU(N)$ Kondo model: large- N solution and Conformal Field Theory,” *Phys. Rev. B* **58** (1998) 3794 [cond-mat/9711192],
A. Georges, O. Parcollet and S. Sachdev, “Quantum Fluctuations of a Nearly Critical Heisenberg Spin Glass,” *Phys. Rev. B* **63** (2001) 134406 [cond-mat/0009388].

- [14] S. Sachdev, “Holographic metals and the fractionalized Fermi liquid,” *Phys. Rev. Lett.* **105** (2010) 151602 [arXiv:1006.3794].
- [15] E. Witten, “An SYK-Like Model Without Disorder,” arXiv:1610.09758 [hep-th].
- [16] C. Peng, M. Spradlin and A. Volovich, “A Supersymmetric SYK-like Tensor Model,” *JHEP* **1705** (2017) 062 [arXiv:1612.03851],
 S. Carrozza and V. Pozsgay, “SYK-like tensor quantum mechanics with $Sp(N)$ symmetry,” *Nucl. Phys. B* **941** (2019) 28 [arXiv:1809.07753],
 S. Prakash and R. Sinha, “A Complex Fermionic Tensor Model in d Dimensions,” *JHEP* **1802** (2018) 086 [arXiv:1710.09357],
 D. Benedetti, S. Carrozza, R. Gurau and A. Sfondrini, “Tensorial Gross-Neveu models,” *JHEP* **1801** (2018) 003 [arXiv:1710.10253],
 C. Krishnan, S. Sanyal and P. N. Bala Subramanian, “Quantum Chaos and Holographic Tensor Models,” *JHEP* **1703** (2017) 056 [arXiv:1612.06330],
 J. Kim, I. R. Klebanov, G. Tarnopolsky and W. Zhao, “Symmetry Breaking in Coupled SYK or Tensor Models,” arXiv:1902.02287 [hep-th].
- [17] J. S. Cotler *et al.*, “Black Holes and Random Matrices,” *JHEP* **1705** (2017) 118, Erratum: *JHEP* **1809** (2018) 002 [arXiv:1611.04650].
- [18] P. Diaz, “Tensor and Matrix models: a one-night stand or a lifetime romance?,” *JHEP* **1806** (2018) 140 [arXiv:1803.04471],
 T. Li, J. Liu, Y. Xin and Y. Zhou, “Supersymmetric SYK model and random matrix theory,” *JHEP* **1706** (2017) 111 [arXiv:1702.01738],
 C. Krishnan, K. V. P. Kumar and S. Sanyal, “Random Matrices and Holographic Tensor Models,” *JHEP* **1706** (2017) 036 [arXiv:1703.08155],
 D. Anninos and G. A. Silva, “Solvable Quantum Grassmann Matrices,” *J. Stat. Mech.* **1704** (2017) no.4, 043102 [arXiv:1612.03795],
 P. Saad, S. H. Shenker and D. Stanford, “A semiclassical ramp in SYK and in gravity,” arXiv:1806.06840 [hep-th].
- [19] S. Sachdev, “Bekenstein-Hawking Entropy and Strange Metals,” *Phys. Rev. X* **5** (2015) no.4, 041025 [arXiv:1506.05111].
- [20] R. A. Davison, W. Fu, A. Georges, Y. Gu, K. Jensen and S. Sachdev, “Thermoelectric transport in disordered metals without quasiparticles: The Sachdev-Ye-Kitaev models and holography,” *Phys. Rev. B* **95** (2017) no.15, 155131 [arXiv:1612.00849].
- [21] S. Sachdev, “Universal low temperature theory of charged black holes with AdS_2 horizons,” [arXiv:1902.04078].
- [22] D. Stanford and E. Witten, “Fermionic Localization of the Schwarzian Theory,” *JHEP* **1710** (2017) 008 [arXiv:1703.04612],

- A. Jevicki, K. Suzuki and J. Yoon, “Bi-Local Holography in the SYK Model,” JHEP **1607** (2016) 007 [arXiv:1603.06246],
- D. J. Gross and V. Rosenhaus, “The Bulk Dual of SYK: Cubic Couplings,” JHEP **1705** (2017) 092 [arXiv:1702.08016],
- S. R. Das, A. Jevicki and K. Suzuki, “Three Dimensional View of the SYK/AdS Duality,” JHEP **1709** (2017) 017 [arXiv:1704.07208],
- S. R. Das, A. Ghosh, A. Jevicki and K. Suzuki, “Space-Time in the SYK Model,” JHEP **1807** (2018) 184 [arXiv:1712.02725],
- S. Frste, J. Kames-King and M. Wiesner, “Towards the Holographic Dual of $N = 2$ SYK,” JHEP **1803** (2018) 028 [arXiv:1712.07398],
- H. A. Gonzalez, D. Grumiller and J. Salzer, “Towards a bulk description of higher spin SYK,” JHEP **1805** (2018) 083 [arXiv:1802.01562],
- J. Maldacena, D. Stanford and Z. Yang, “Diving into traversable wormholes,” Fortsch. Phys. **65** (2017) no.5, 1700034 [arXiv:1704.05333 [hep-th]].
- [23] C. Krishnan, K. V. Pavan Kumar and D. Rosa, “Contrasting SYK-like Models,” JHEP **1801** (2018) 064 [arXiv:1709.06498],
- Y. Chen, H. Zhai and P. Zhang, “Tunable Quantum Chaos in the Sachdev-Ye-Kitaev Model Coupled to a Thermal Bath,” JHEP **1707** (2017) 150 [arXiv:1705.09818],
- Y. Gu, A. Lucas and X. L. Qi, “Energy diffusion and the butterfly effect in inhomogeneous Sachdev-Ye-Kitaev chains,” SciPost Phys. **2** (2017) no.3, 018 [arXiv:1702.08462],
- A. M. Garcia-Garca, Y. Jia and J. J. M. Verbaarschot, “Universality and Thouless energy in the supersymmetric Sachdev-Ye-Kitaev Model,” Phys. Rev. D **97** (2018) no.10, 106003 [arXiv:1801.01071],
- S. Banerjee and E. Altman, “Solvable model for a dynamical quantum phase transition from fast to slow scrambling,” Phys. Rev. B **95** (2017) no.13, 134302 [arXiv:1610.04619],
- Y. Gu, X. L. Qi and D. Stanford, “Local criticality, diffusion and chaos in generalized Sachdev-Ye-Kitaev models,” JHEP **1705** (2017) 125 [arXiv:1609.07832],
- P. Zhang, “Dispersive Sachdev-Ye-Kitaev model: Band structure and quantum chaos,” Phys. Rev. B **96** (2017) no.20, 205138 [arXiv:1707.09589],
- D. S. Ageev and I. Y. Aref’eva, “When things stop falling, chaos is suppressed,” JHEP **1901** (2019) 100 [arXiv:1806.05574].
- [24] A. Tanasa, “Multi-orientable Group Field Theory,” J. Phys. A **45** (2012) 165401 [arXiv:1109.0694],
- S. Dartois, V. Rivasseau and A. Tanasa, “The $1/N$ expansion of multi-orientable random tensor models,” Annales Henri Poincare **15** (2014) 965 [arXiv:1301.1535],
- E. Fusy and A. Tanasa, “Asymptotic expansion of the multi-orientable random tensor model,” The electronic journal of combinatorics **22**(1) (2015), #P1.52

- [arXiv:1408.5725],
S. Carrozza and A. Tanasa, “ $O(N)$ Random Tensor Models,” *Lett. Math. Phys.* **106** (2016) no.11, 1531 [arXiv:1512.06718].
- [25] I. R. Klebanov and G. Tarnopolsky, “Uncolored random tensors, melon diagrams, and the Sachdev-Ye-Kitaev models,” *Phys. Rev. D* **95**, no. 4, 046004 (2017) [arXiv:1611.08915].
- [26] S. S. Gubser, C. Jepsen, Z. Ji and B. Trundy, “Higher melonic theories,” *JHEP* **1809** (2018) 049 [arXiv:1806.04800],
S. Choudhury, A. Dey, I. Halder, L. Janagal, S. Minwalla and R. Poojary, “Notes on melonic $O(N)^{q-1}$ tensor models,” *JHEP* **1806** (2018) 094 [arXiv:1707.09352],
C. M. Chang, S. Colin-Ellerin and M. Rangamani, “On Melonic Supertensor Models,” *JHEP* **1810** (2018) 157 [arXiv:1806.09903].
- [27] S. Dartois, O. Evnin, L. Lionni, V. Rivasseau and G. Valette, “Melonic Turbulence,” arXiv:1810.01848 [math-ph].
- [28] T. Azeyanagi, F. Ferrari and F. I. Schaposnik Massolo, “Phase Diagram of Planar Matrix Quantum Mechanics, Tensor, and Sachdev-Ye-Kitaev Models,” *Phys. Rev. Lett.* **120** (2018) no.6, 061602 [arXiv:1707.03431].
- [29] W. Cai, X. H. Ge and G. H. Yang, “Diffusion in higher dimensional SYK model with complex fermions,” *JHEP* **1801** (2018) 076 [arXiv:1711.07903].
- [30] K. Bulycheva, “A note on the SYK model with complex fermions,” *JHEP* **1712** (2017) 069 [arXiv:1706.07411].
- [31] R. Bhattacharya, D. P. Jatkar and A. Kundu, “Chaotic Correlation Functions with Complex Fermions,” arXiv:1810.13217 [hep-th].
- [32] V. Balasubramanian, B. Craps, B. Czech and G. Srofi, “Echoes of chaos from string theory black holes,” *JHEP* **1703** (2017) 154 [arXiv:1612.04334].
- [33] J. Maldacena, S. H. Shenker and D. Stanford, “A bound on chaos,” *JHEP* **1608** (2016) 106 [arXiv:1503.01409].
- [34] J. Zinn-Justin, “Vector models in the large N limit: a few applications,” hep-th/9810198,
M. Moshe and J. Zinn-Justin, “Quantum Field Theory in the Large N Limit: a review,” *Phys. Rep.* **385** (2003) 69 [hep-th/0306133].
- [35] A. J. Bray and M. A. Moore, “Replica theory of quantum spin glasses,” *J. Phys. C: Solid State Phys.* (1980) **13** L655.
- [36] R. Gurau, “Quenched equals annealed at leading order in the colored SYK model,” *EPL* **119** (2017) no.3, 30003 [arXiv:1702.04228].

- [37] G. 't Hooft, “A Planar Diagram Theory for Strong Interactions,” *Nucl. Phys.* **B 72** (1974) 461.
- [38] A. M. Garca-Garca, B. Loureiro, A. Romero-Bermdez and M. Tezuka, “Chaotic-Integrable Transition in the Sachdev-Ye-Kitaev Model,” *Phys. Rev. Lett.* **120** (2018) no.24, 241603 [arXiv:1707.02197].
- [39] T. Nosaka, D. Rosa and J. Yoon, “The Thouless time for mass-deformed SYK,” *JHEP* **1809** (2018) 041 [arXiv:1804.09934].
- [40] R. Gurau, “Random Tensors,” Oxford University Press (2016),
SIGMA special issue, “Tensor Models, Formalism and Applications” (2016).
- [41] I. R. Klebanov, F. Popov and G. Tarnopolsky, “TASI Lectures on Large N Tensor Models,” *PoS TASI 2017* (2018) 004 [arXiv:1808.09434].
- [42] N. Delporte and V. Rivasseau, “The Tensor Track V: Holographic Tensors,” arXiv:1804.11101 [hep-th].
- [43] R. Gurau, “Colored Group Field Theory,” *Comm. Math. Phys.* **304** (2011) 69 [arXiv:0907.2582],
R. Gurau and J. P. Ryan, “Colored Tensor Models - a review,” *SIGMA* **8**, 020 (2012) [arXiv:1109.4812].
- [44] V. Bonzom and F. Combes, “The calculation of expectation values in Gaussian random tensor theory via meanders,” *AIHP-D*, 443-485 (2014) [arXiv:1310.3606],
V. Bonzom and L. Lionni, “Counting gluings of octahedra,” *Elec. J. Combin.* **24** (2017) #P3.36 [arXiv:1608.00347],
L. Lionni and J. Thürigen, “Multi-critical behaviour of 4-dimensional tensor models up to order 6,” arXiv:1707.08931,
V. Bonzom, L. Lionni and V. Rivasseau, “Colored triangulations of arbitrary dimensions are stuffed Walsh maps,” *The Electronic Journal of Combinatorics*, **24**(1)(2017), #P1.56 [arXiv:1508.03805],
V. Bonzom, “Large N Limits in Tensor Models: Towards More Universality Classes of Colored Triangulations in Dimension $d \geq 2$,” *SIGMA* **12**, 073 (2016), [arXiv:1603.03570],
L. Lionni, “Colored Discrete Spaces: Higher dimensional combinatorial maps and quantum gravity,” PhD thesis, U. Paris-Saclay (2017) [arXiv:1710.03663].
- [45] V. Bonzom, R. Gurau and V. Rivasseau, “Random tensor models in the large N limit: Uncoloring the colored tensor models,” *Phys. Rev.* **D 85** (2012) 084037 [arXiv:1202.3637].

- [46] V. Bonzom, R. Gurau, A. Riello and V. Rivasseau, “Critical behaviour of colored tensor models in the large N limit,” *Nucl. Phys. B* **853** (2011) 174 [arXiv:1105.3122].
- [47] R. Gurau and G. Schaeffer, “Regular colored graphs of positive degree,” *Ann. Inst. Henri Poincare Comb. Phys. Interact.*, 3:257–320 (2016) [arXiv:1307.5279].
- [48] S. Carrozza and A. Tanasa, “ $O(N)$ Random Tensor Models,” *Lett. Math. Phys.* **106** (2016) 1531 [arXiv:1512.06718].
- [49] I. R. Klebanov and G. Tarnopolsky, “On Large N Limit of Symmetric Traceless Tensor Models,” *JHEP* **1710** (2017) 037 [arXiv:1706.00839].
- [50] D. Benedetti, S. Carrozza, R. Gurau and M. Kolanowski, “The $1/N$ expansion of the symmetric traceless and the antisymmetric tensor models in rank three,” arXiv:1712.00249 [hep-th],
S. Carrozza, “Large N limit of irreducible tensor models: $O(N)$ rank-3 tensors with mixed permutation symmetry,” *JHEP* **1806**, 039 (2018) [arXiv:1803.02496].
- [51] R. Gurau, “The $1/N$ expansion of tensor models with two symmetric tensors,” *Commun. Math. Phys.* **360**, no. 3, 985 (2018) [arXiv:1706.05328].
- [52] R. Emparan, R. Suzuki and K. Tanabe, “The large D limit of General Relativity,” *J. High Energy Phys.* **06** (2013) 009 [arXiv:1302.6382].
- [53] R. de Mello Koch, R. Mello Koch, D. Gossman and L. Tribelhorn, “Gauge Invariants, Correlators and Holography in Bosonic and Fermionic Tensor Models,” *JHEP* **1709**, 011 (2017) [arXiv:1707.01455],
K. Bulycheva, I. R. Klebanov, A. Milekhin and G. Tarnopolsky, “Spectra of Operators in Large N Tensor Models,” *Phys. Rev. D* **97**, no. 2, 026016 (2018) [arXiv:1707.09347],
K. Pakrouski, I. R. Klebanov, F. Popov and G. Tarnopolsky, “Spectrum of Majorana Quantum Mechanics with $O(4)^3$ Symmetry,” *Phys. Rev. Lett.* **122** (2019) no.1, 011601 [arXiv:1808.07455],
I. R. Klebanov, A. Milekhin, F. Popov and G. Tarnopolsky, “Spectra of eigenstates in fermionic tensor quantum mechanics,” *Phys. Rev. D* **97** (2018) no.10, 106023 [arXiv:1802.10263],
C. Krishnan and K. V. Pavan Kumar, “Complete Solution of a Gauged Tensor Model,” arXiv:1804.10103 [hep-th].
- [54] J. Maldacena and X. L. Qi, “Eternal traversable wormhole,” arXiv:1804.00491 [hep-th].
- [55] S. R. Das, A. Ghosh, A. Jevicki and K. Suzuki, “Duality in the Sachdev-Ye-Kitaev Model,” *Springer Proc. Math. Stat.* **255** (2017) 43.

- [56] F. Ferrari, “The Analytic Renormalization Group,” Nucl. Phys. B **909**, 880 (2016) [arXiv:1602.07355].
- [57] N. Seiberg and E. Witten, “Electric - magnetic duality, monopole condensation, and confinement in N=2 supersymmetric Yang-Mills theory,” Nucl. Phys. B **426**, 19 (1994) Erratum: [Nucl. Phys. B **430**, 485 (1994)] [hep-th/9407087].
- [58] S. Giombi, I. R. Klebanov and G. Tarnopolsky, “Bosonic tensor models at large N and small ϵ ,” Phys. Rev. D **96** (2017) no.10, 106014 [arXiv:1707.03866].
- [59] S. Giombi, I. R. Klebanov, F. Popov, S. Prakash and G. Tarnopolsky, “Prismatic Large N Models for Bosonic Tensors,” Phys. Rev. D **98** (2018) no.10, 105005 [arXiv:1808.04344].
- [60] D. Benedetti and R. Gurau, “2PI effective action for the SYK model and tensor field theories,” JHEP **1805** (2018) 156 [arXiv:1802.05500].
- [61] J. W. Cooley and J. W. Tukey, “An Algorithm for the Machine Calculation of Complex Fourier Series,” Math. Comput. **19** (1965) 297.
- [62] J. Polchinski and V. Rosenhaus, “The Spectrum in the Sachdev-Ye-Kitaev Model,” JHEP **1604** (2016) 001 [arXiv:1601.06768].
- [63] A. M. Garca-Garca and J. J. M. Verbaarschot, “Spectral and thermodynamic properties of the Sachdev-Ye-Kitaev model,” Phys. Rev. D **94** (2016) no.12, 126010 [arXiv:1610.03816].
- [64] D. J. Gross and V. Rosenhaus, “All point correlation functions in SYK,” JHEP **1712** (2017) 148 [arXiv:1710.08113].
- [65] A. M. Garca-Garca, T. Nosaka, D. Rosa and J. J. M. Verbaarschot, “Quantum chaos transition in a two-site SYK model dual to an eternal traversable wormhole,” arXiv:1901.06031 [hep-th].
- [66] S. W. Hawking and D. N. Page, “Thermodynamics of Black Holes in anti-De Sitter Space,” Commun. Math. Phys. **87** (1983) 577.
- [67] S. R. Das, A. Ghosh, A. Jevicki and K. Suzuki, “Three Dimensional View of Arbitrary q SYK models,” JHEP **1802** (2018) 162 [arXiv:1711.09839],
G. Tarnopolsky, “Large q expansion in the Sachdev-Ye-Kitaev model,” Phys. Rev. D **99** (2019) no.2, 026010 [arXiv:1801.06871],
M. Berkooz, M. Isachenkov, V. Narovlansky and G. Torrents, “Towards a full solution of the large N double-scaled SYK model,” arXiv:1811.02584 [hep-th].
- [68] F. Lonard and B. Delamotte, “Critical Exponents Can Be Different on the Two Sides of a Transition: A Generic Mechanism,” Phys. Rev. Lett. **115** (2015) no.20, 200601 [arXiv:1508.07852].

- [69] A. Almheiri and J. Polchinski, “Models of AdS₂ backreaction and holography,” *JHEP* **1511**, 014 (2015) [arXiv:1402.6334],
J. Maldacena, D. Stanford and Z. Yang, “Conformal symmetry and its breaking in two dimensional Nearly Anti-de-Sitter space,” *PTEP* **2016**, no. 12, 12C104 (2016) [arXiv:1606.01857].
- [70] F. Ferrari, to appear.
- [71] L.A. Rubel, “Necessary and sufficient conditions for Carlsens theorem on entire functions,” *Trans. Amer. Math. Soc.* **83** (1956) 417.
- [72] G. Cuniberti, E. De Micheli, G.A. Viano, “Reconstructing the thermal Green functions at real times from those at imaginary times,” *Comm. Math. Phys.* **216** (2001) 59 [cond-mat/0109175].
- [73] F. Ferrari, “The Analytic Renormalization Group,” *Nucl. Phys. B* **909** (2016) 880 [arXiv:1602.07355].
- [74] R. van Leeuwen, N. E. Dahlen, G. Stefanucci, C-O. Almbladh and U. von Barth, “Introduction to the Keldysh formalism and applications to time-dependent density-functional theory,” *Lecture Notes in Physics* **706** (2006) 33 [cond-mat/0506130].
- [75] A. Eberlein, V. Kasper, S. Sachdev and J. Steinberg, “Quantum quench of the Sachdev-Ye-Kitaev Model,” *Phys. Rev. B* **96**, no. 20, 205123 (2017) [arXiv:1706.07803].
- [76] A.I. Larkin and Y.N. Ovchinnikov, “Quasiclassical Method in the Theory of Superconductivity,” *JETP* **28**, No. 6 (1969) 1200.
- [77] S. H. Shenker and D. Stanford, “Black holes and the butterfly effect,” *JHEP* **1403**, 067 (2014) [arXiv:1306.0622],
S. H. Shenker and D. Stanford, “Multiple Shocks,” *JHEP* **1412**, 046 (2014) [arXiv:1312.3296],
S. H. Shenker and D. Stanford, “Stringy effects in scrambling,” *JHEP* **1505**, 132 (2015) [arXiv:1412.6087].
- [78] F. M. Haehl, R. Loganayagam, P. Narayan, A. A. Nizami and M. Rangamani, “Thermal out-of-time-order correlators, KMS relations, and spectral functions,” *JHEP* **1712**, 154 (2017) [arXiv:1706.08956].
- [79] J. Murugan, D. Stanford and E. Witten, “More on Supersymmetric and 2d Analogs of the SYK Model,” *JHEP* **1708** (2017) 146 [arXiv:1706.05362].
- [80] C. Peng, M. Spradlin and A. Volovich, “Correlators in the $\mathcal{N} = 2$ Supersymmetric SYK Model,” *JHEP* **1710** (2017) 202 [arXiv:1706.06078].

- [81] J. Yoon, “SYK Models and SYK-like Tensor Models with Global Symmetry,” JHEP **1710** (2017) 183 [arXiv:1707.01740].