Out-of-time-order correlators and Lyapunov exponents in sparse SYK

Elena Cáceres,[a] Tyler Guglielmo,[a,b] Brian Kent[a] and Anderson Misobuchi[a]

[a]*Physics Department, University of Texas, Austin,*
Austin, TX 78712, U.S.A.

[b]*Lawrence Livermore National Laboratory,*
Livermore, CA 94550, U.S.A.

E-mail: elenac@utexas.edu, guglielmo2@llnl.gov, brian_kent@utexas.edu,
anderson.misobuchi@utexas.edu

ABSTRACT: We use a combination of analytical and numerical methods to study out-of-time order correlators (OTOCs) in the sparse Sachdev-Ye-Kitaev (SYK) model. We find that at a given order of N, the standard result for the q-local, all-to-all SYK, obtained through the sum over ladder diagrams, is corrected by a series in the sparsity parameter, k. We present an algorithm to sum the diagrams at any given order of $1/(kq)^n$. We also study OTOCs numerically as a function of the sparsity parameter and determine the Lyapunov exponent. We find that numerical stability when extracting the Lyapunov exponent requires averaging over a massive number of realizations. This trade-off between the efficiency of the sparse model and consistent behavior at finite N becomes more significant for larger values of N.

Contents

1 Introduction

Understanding the underlying mechanisms of the spread of information in strongly interacting quantum systems is a challenging problem that connects quantum information, condensed matter theory and quantum gravity. Out-of-time order correlators (OTOCs) provide a useful measure of how fast a local perturbation spreads across a quantum many-body system under time evolution [1–14]. Originally introduced as a semiclassical approach to superconductivity [15], OTOCs experienced a renaissance with the discovery of a deep connection between quantum chaos and black holes, where black holes have been shown to be fast scramblers [1, 16] and a bound on the rate of growth has been conjectured [3].

The Sachdev-Ye-Kitaev (SYK) model [7, 17–19], a system of N Majorana fermions with q-body interactions, i.e. q-local, has been extensively explored as a solvable model of quantum chaos and holography. At low temperatures, the system displays an approximate

conformal symmetry and, as in generic black holes, it scrambles information at the fastest possible rate. Among its many variants, the sparse-SYK model [20, 21] stands out as a simplified version with fewer interactions which allows one to probe finite N corrections via numerical simulations. Remarkably, the sparse-SYK model with sparseness parameter k and only kN interaction terms in the Hamiltonian, is still a fast scrambler [20, 21] and it can be shown to approximately reproduce properties of the original SYK model while allowing to uncover other features as in the spectral form factor studied in [22]. Additionally, the sparse-SYK model can serve as a replacement for the original SYK model when modeling traversable wormholes [23, 24] including, remarkably, its simulation on a quantum processor [24].

In this work, we study out-of-time order correlators in the sparse-SYK model, both analytically and numerically. Even though the sparse Hamiltonian contains fewer terms than the all-to-all SYK, there are more diagrams contributing to the OTOCs at order $\mathcal{O}(\frac{1}{N})$. We investigate ladder diagrams contributing to the OTOCs and present an algorithm to sum the diagrams at any given order of $1/(kq)^n$. However, an analytic expression for the sum of the diagrams to all orders of $1/(kq)^n$ is out of reach because the methods used in the all-to-all SYK are not applicable here. Another of our main results is the numerical study of the OTOCs. We investigate them as a function of the sparsity parameter k, with the goal of determining the Lyapunov exponent. We also explore to what extent the symmetry describing the large N behavior of the OTOCs in SYK is present in the sparse model.

2 Preliminaries

2.1 Chaos and OTOCs

Scrambling in quantum systems can be quantified via out-of-time order correlators (OTOCs). In a chaotic system, OTOCs display an exponential growth at intermediate times characterized by a Lyapunov exponent λ_L, and the fast scrambling property translates into a Lyapunov exponent that saturates the chaos bound $\lambda_L \leq 2\pi/\beta$ [3]. The definition of OTOCs can be motivated as the quantum analog of the classical characterization of chaos using Poisson brackets, where the brackets are promoted to commutators (or anticommutators for fermionic operators)

$$C(t) = -\langle [W(t), V(0)]^2 \rangle_\beta, \tag{2.1}$$

where V and W are generic local Hermitian operators. The operators evolve according to the Hamiltonian H of the system via $W(t) = e^{iHt}W(0)e^{-iHt}$. The brackets denote the thermal expectation value $\langle \cdot \rangle_\beta = Z^{-1}\mathrm{tr}[. \, e^{-\beta H}]$ at inverse temperature $\beta = 1/T$ and $Z = \mathrm{tr}[e^{-\beta H}]$ is the partition function. Intuitively, this correlator measures how much an early perturbation V affects the later measurement of W. After the expansion of the commutator squared, we see the appearance of a four-point correlator that is out-of-time ordered

$$F^{(u)}(t) = \langle W(t)V(0)W(t)V(0)\rho \rangle_\beta, \tag{2.2}$$

where $\rho = e^{-\beta H}$. In the SYK model, we typically consider W and V as two distinct Majorana operators. The above expression is the standard OTOC that is also referred to

in the literature as 'unregularized' OTOC. In the study of quantum chaos at finite N, a slightly modification of this correlator have been considered, named 'regularized' OTOC

$$F^{(r)}(t) = \langle W(t)\rho^{1/4}V(0)\rho^{1/4}W(t)\rho^{1/4}V(0)\rho^{1/4}\rangle_\beta, \qquad (2.3)$$

in which we split ρ evenly between the operators. One of the reasons to work with the regularized version is because this version is less sensitive to finite-size effects, and it can be shown that the extracted Lyapunov exponent is the same as in the unregularized version [25].

From the OTOC, one could derive the Lyapunov exponent by fitting an exponential function $a + b\, e^{\lambda_{\text{fit}}t}$ to the region of exponential growth. However, this method is known to lead to incorrect results in the SYK model for several reasons. The main subtlety is that, due to strong finite-size effects, the OTOC does not display a well-defined exponential growth window at intermediate times and finite size N. An alternative method was proposed [25], where it is assumed that the OTOC admits the following large N expansion based upon analytical arguments [5, 11]

$$F(t) = C_0 + C_1\left(\frac{e^{\lambda_L t}}{N}\right) + C_2\left(\frac{e^{\lambda_L t}}{N}\right)^2 + \dots \qquad (2.4)$$

for $t \lesssim \frac{1}{\lambda_L}\log N$, which means that $F(t)$ obeys the rescaling symmetry

$$N \to c\,N, \qquad t \to t + \frac{1}{\lambda_L}\log c. \qquad (2.5)$$

This symmetry is expected to hold for any many-body chaotic model governed by ladder diagrams close to the semiclassical limit.

2.2 Sparse SYK

The Sachdev-Ye-Kitaev (SYK) model [7, 17–19] is a toy model of lower dimensional quantum black holes consisting of a system of Majorana fermions with all-to-all interactions. The sparse-SYK model [20, 21] is a variant of the original model, which we will refer to as all-to-all SYK, with the advantage of allowing for more efficient computer simulation while preserving important black hole physics behavior. The Hamiltonian of the sparse-SYK model with N Majorana fermions χ_j, $j = 1, \dots, N$, and q-body interactions is defined as

$$H = i^{q/2}\sum_{j_1 < \dots < j_q} J_{j_1 \dots j_q} x_{j_1 \dots j_q} \chi_{j_1} \cdots \chi_{j_q}. \qquad (2.6)$$

The parameters $x_{j_1 \dots j_q}$ are either 0 or 1 and they can be defined in different ways leading to different sparse models. The couplings $J_{j_1 \dots j_q}$ are drawn from a Gaussian distribution with zero mean and variance given by[1]

$$\langle (J_{j_1 \dots j_q})^2\rangle = \frac{(q-1)!J^2}{pN^{q-1}}, \qquad (2.7)$$

where p is the fraction of terms in the Hamiltonian, i.e., the number of the terms such that $x_{j_1 \dots j_q} = 1$ divided by $\binom{N}{q}$, the number of terms in the all-to-all version. The parameter J

---

[1]Sparse models with non-Gaussian couplings have also been considered [26].

has dimensions of energy and sets the energy scale of the theory. It will be convenient to define the parameter

$$k \equiv \frac{p}{N} \binom{N}{q},$$

(2.8)

which is such that the Hamiltonian is a sum of kN independent terms. Due to the random nature of the couplings, physical observables are obtained after performing an average over different disorder realizations of the system. If we choose all $x_{j_1 \ldots j_q}$ to be 1, we recover the original all-to-all SYK Hamiltonian

$$H_{\text{all-to-all}} = i^{q/2} \sum_{1 \leq j_1 < \ldots < j_q \leq N} J_{j_1 \ldots j_q} \chi_{j_1} \cdots \chi_{j_q}.$$

(2.9)

In this case we have $p = 1$ and we also recover the variance of the all-to-all SYK model

$$\langle (J_{j_1 \ldots j_q})^2 \rangle_{\text{all-to-all}} = \frac{(q-1)! J^2}{N^{q-1}}.$$

(2.10)

One possible implementation of sparseness in the SYK model consists of taking $x_{ijkl} = 1$ with probability p and $x_{ijkl} = 0$ with probability $1 - p$. The system becomes more sparse as $p \to 0$ and we recover the all-to-all SYK when $p = 1$. This procedure, known as *random pruning*, is perhaps the simplest to implement. However, it makes necessary to take an additional average since x_{ijkl} are treated as random variables and can potentially lead to disconnected clusters of Majorana fermions that do not interact with the rest of the system [21].

A different approach to characterize the sparseness is to use *regular hypergraphs*, in which the Majorana fermions are identified as vertices of a hypergraph whose hyperedges contain q vertices. The regularity condition imposes that each fermion appears exactly the same number of times in the list of hyperedges. The hypergraph is also uniform, meaning that all hyperedges contain the same number, q, of vertices. An important property of random-uniform-regular hypergraphs is that they are expected to be *expanders* [27]. That is, they are expected to have good connectivity even though they are sparse. Note that the q-uniformity condition is related to the Hamiltonian being q-local but the regularity condition is imposed for convenience. A simple counting argument shows that the hypergraph should be kq-regular. Measures of connectivity of this type of hypergraph in the sparse-SYK context were studied in [23]. An example of a regular hypergraph is given in figure 1.

3 OTOCs in sparse SYK

In [7, 11], it was shown that, in the all-to-all SYK, to obtain the leading order in N and late time contributions to the OTOC, we should calculate ladder diagrams of dressed propagators. This is also true for the sparse model. In this section, we will present an analytic calculation of ladder diagrams in the sparse-SYK. We find that the standard contribution obtained through the sum over ladder diagrams is corrected, in the sparse-SYK, by higher order terms of order $1/(kq)^n$. In the all-to-all SYK the ladder diagrams can be summed exactly, giving an expression for the OTOC from which a Lyapunov exponent
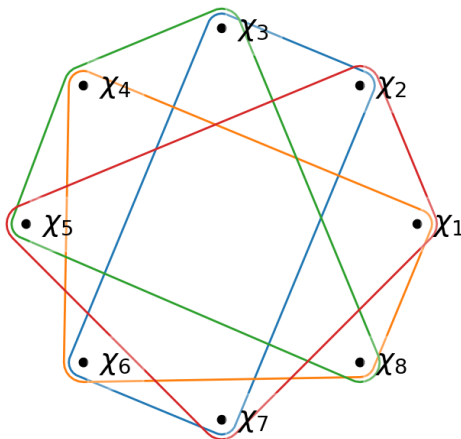
Figure 1. A concrete regular hypergraph representation of the sparse-SYK Hamiltonian with $N = 8$, $q = 4$, and $k = 1/2$ with hyperedges $= \{(2, 3, 6, 7), (1, 4, 6, 8), (3, 4, 5, 8), (1, 2, 5, 7)\}$. Each vertex is contained in exactly $kq = 2$ hyperedges.

can be computed [7]. The same ressumation does not go through in the case of the sparse model and an analytic expression for the ladder diagrams to all orders of $1/(kq)^n$ cannot be obtained. However, we found an algorithm to sum the diagrams up to any given order of $1/(kq)^n$ that we present in this section. The intricacies of the procedure are in appendix A.

3.1 Ladder diagrams

As a first step, we will analyze the combinatorics related to ladder diagrams in a sparse-SYK model with random pruning [20]. In this model, each interaction term in the Hamiltonian comes with a factor, x_{abcd}, which is either 0 or 1 with probability p. This is telling us that each term in the Hamiltonian is deleted with probability $1 - p$. In this model, we normalize the couplings to have a p dependent variance,

$$\left\langle J_{abcd}^2 \right\rangle = \frac{(q-1)!J^2}{pN^{q-1}}, \tag{3.1}$$

where after averaging, the multiplication of two couplings is only non-zero if their indices are equal (up to permutations of one another). As in the standard SYK model, the ladder diagrams will be the diagrams that contribute to the OTOC [5, 7, 11],

$$\mathcal{F}(t_1, t_2, t_3, t_4) = \frac{1}{N^2} \text{Tr}\left(\rho^{1/4} \chi_i(t_1) \rho^{1/4} \chi_j(t_2) \rho^{1/4} \chi_i(t_3) \rho^{1/4} \chi_j(t_4) \right), \tag{3.2}$$

at leading order in late time t. We will now proceed by showing how the ladder diagrams can be summed at zero temperature — although the OTOC is clearly a finite temperature quantity, the zero temperature calculation will be sufficient for showing the general procedure of ladder diagram summation. The zero temperature calculation can be extended to finite temperatures in a fairly straightforward manner by writing retarded propagators along the real time portion of a Schwinger-Keldish contour and Wightman functions between the two real parts of the contour [5, 7, 12].

We find, as in [20], a series in both $1/N$ and $1/(kq)$. This means that at each order in $1/N$ we will find a series expression for a prefactor that is a proportionality constant for each diagram in $1/(kq)$. This series terminates at an order corresponding to the order of the ladder diagram being considered. At zeroth order in $1/(kq)$, we have the familiar diagrams of the dense model. But at higher orders, there are corrections organized in a $1/(kq)$ expansion. To leading order in N, we find that a ladder diagram of $\mathcal{O}(J^{2n})$ (think a ladder with n rungs) will contain a series in $1/(kq)$ up to order $1/(kq)^{n-1}$, i.e.

$$\overline{\underset{}{\left(\!\bigcirc\!\right)\cdots\times n}} \equiv \mathcal{F}_n \sim \frac{(-1)^n(q-1)^n J^{2n}}{N}\left(1 + \frac{c_1}{kq} + \frac{c_2}{(kq)^2} + \cdots + \frac{c_{n-1}}{(kq)^{n-1}}\right). \quad (3.3)$$

The dashed line implies that the diagram is disorder averaged. For the rest of this section, dashed lines will be dropped from diagrams, and should be assumed.

Let us now write down the first few terms of the ladder diagram sum

$$\quad (3.4)$$

The first two diagrams are given by

$$\mathcal{F}_0 = \frac{1}{N}\left[-G(\tau_{13})G(\tau_{24}) + G(\tau_{14})G(\tau_{23})\right] \quad (3.5)$$

$$\mathcal{F}_1 = \frac{(q-1)J^2}{N}\int d\tau d\tau'\left[G(\tau_1-\tau)G(\tau_2-\tau')G^{q-2}(\tau-\tau')G(\tau-\tau_3)G(\tau'-\tau_4) - (\tau_3 \leftrightarrow \tau_4)\right], \quad (3.6)$$

where $G(\tau)$ is the Euclidean-time propagator, i.e. the time-ordered product

$$G(\tau) = \langle \mathcal{T}(\chi(\tau)\chi(0))\rangle. \quad (3.7)$$

Let us now focus on the combinatoric factor that appears in front of the integral, and is given generally by (3.3). First, let us look at the overall factor outside of the parenthesis in (3.3). The $(-1)^n$ comes from the addition of a fermion loop at each successive order. The $(q-1)^n$ is a combinatoric factor coming from the number of ways Wick contractions can be done on the interaction vertices. Inside the parenthesis, we can find each term by considering the ways $\langle J_{ia_1b_1c_1}J_{ia_1b_1c_2}\cdots\rangle$ can be contracted at leading order in N. Focusing on $\mathcal{F}_1$, we have $\langle J_{iabj}J_{iabj}\rangle$ which gives us back (3.1) a priori, except now there is a sum over i,j,a,b which will give a factor of N^q. We will also obtain a factor of p since, $x_{iabj}x_{iabj} = x_{iabj}$ and $\langle x_{iabj}\rangle = p$. Ignoring the Wick contraction combinatorics, which amounts to the $(q-1)$ factor above, we have

$$\frac{1}{N^2}\frac{J^2}{pN^{q-1}}N^q p = \frac{J^2}{N}. \quad (3.8)$$

This is the 1 inside the parenthesis in (3.3). Every diagram in the sum will contribute the same 1, however the terms of order $1/(kq)$ and up will change based on the order of perturbation theory we are working in.

The two rung diagram will give

$$\langle J_{iabc}J_{iabc'}J_{ja'b'c}J_{ja'b'c'}\rangle = \langle J_{iabc}J_{iabc'}\rangle \langle J_{ja'b'c}J_{ja'b'c'}\rangle + \langle J_{iabc}J_{ja'b'c}\rangle \langle J_{iabc'}J_{ja'b'c'}\rangle$$
$$+ \langle J_{iabc}J_{ja'b'c'}\rangle \langle J_{iabc'}J_{ja'b'c}\rangle. \tag{3.9}$$

There are now two possibilities to consider. The first is when $ja'b'$ is not a permutation of iab. In this case, the second two terms in the cumulant expansion vanish while the first will give us a factor of 1 as above. When $ja'b'$ is a permutation of iab, then the combinatoric pre-factor will be

$$3\frac{1}{N^2}\frac{J^4}{p^2 N^{2q-2}}N^q(q-1)! \ p = 3\frac{J^4(q-1)!}{NN^{q-1}p} = \frac{J^4}{N}\frac{3}{kq}, \tag{3.10}$$

where the 3 is from the three different terms in the sum. The last equality can be seen from the identity $p = kq!/N^{q-1}$. So, we see that the one-rung diagram has two terms that contribute at leading order in N. One that is k independent and one that is of order $1/(kq)$. We can proceed in this way for higher order rung diagrams order by order and find a series in $1/(kq)$ at leading order in N. At each order, care must be taken to consider which indices are permutations of the others and in how many ways they can be made permutations of each other. The procedure is quite tedious; we have worked out the first six pre-factors and present them below for the reader.

$$\mathcal{F}_1 \sim 1 \tag{3.11}$$

$$\mathcal{F}_2 \sim 1 + \frac{3}{kq} \tag{3.12}$$

$$\mathcal{F}_3 \sim 1 + \frac{9}{kq} + \frac{15}{(kq)^2} \tag{3.13}$$

$$\mathcal{F}_4 \sim 1 + \frac{18}{kq} + \frac{90}{(kq)^2} + \frac{105}{(kq)^3} \tag{3.14}$$

$$\mathcal{F}_5 \sim 1 + \frac{30}{kq} + \frac{300}{(kq)^2} + \frac{1035}{(kq)^3} + \frac{945}{(kq)^4} \tag{3.15}$$

$$\mathcal{F}_6 \sim 1 + \frac{45}{kq} + \frac{750}{(kq)^2} + \frac{5205}{(kq)^3} + \frac{13675}{(kq)^4} + \frac{10395}{(kq)^5}. \tag{3.16}$$

A general expression for the first two terms and the last term in the series is shown below

$$\mathcal{F}_n \sim \frac{(q-1)^n J^{2n}}{N}\left(1 + \frac{3}{2}n(n-1)\frac{1}{kq} + \frac{5}{4}n(n-1)^2(n-2)\frac{1}{(kq)^2} + \cdots + \frac{(2n)!}{2^n n!}\frac{1}{(kq)^{n-1}}\right). \tag{3.17}$$

In appendix A, we present a straightforward and general algorithm for calculating a given coefficient to arbitrary order.

We would now like to make an attempt at summing the ladder diagrams. Since we do not have an analytic expression for the ladder diagram to all orders in J^2 and $1/(kq)$, we

choose to work at order $1/(kq)^2$. We can move between successive ladders (adding rungs to the zero-order diagram) through multiplication of a kernel $\mathcal{K}_n$,

$$\mathcal{F}_n(\tau_1, \tau_2, \tau_3, \tau_4) = \int d\tau d\tau' \, \mathcal{K}_n(\tau_1, \tau_2; \tau, \tau') \mathcal{F}_0(\tau, \tau'; \tau_3, \tau_4) \tag{3.18}$$

$$= \int d\tau d\tau' \, K^n \left(1 + \frac{3}{2} n(n-1) \frac{1}{kq} + \frac{5}{4} n(n-1)^2 (n-2) \frac{1}{(kq)^2}\right) \mathcal{F}_0, \tag{3.19}$$

where

$$K(\tau_1, \tau_2; \tau_3, \tau_4) = -(q-1) J^2 G(\tau_{13}) \, G(\tau_{24}) \, G^{q-2}(\tau_{34}). \tag{3.20}$$

We can now attempt to perform the sum over all ladders by viewing the integral operation in (3.18) as matrix multiplication. We obtain

$$\mathcal{F} = \sum_{n=0}^{\infty} \mathcal{F}_n = \sum_{n=0}^{\infty} K^n \left(1 + \frac{3}{2} n(n-1) \frac{1}{kq} + \frac{5}{4} n(n-1)^2(n-2) \frac{1}{(kq)^2}\right) \mathcal{F}_0$$

$$= \frac{1}{1-K} \left(1 + \frac{3K^2}{kq(1-K)^2} + \frac{15K^3(1+K)}{(kq)^2(1-K)^4}\right) \mathcal{F}_0. \tag{3.21}$$

In all-to-all SYK, the sum of all ladder diagrams at early times can be shown to scale as $\mathcal{F} \sim e^{\lambda_L t}/N$. For large kq the sum of ladder diagrams is the same as the all-to-all case, as expected. However, it is unclear how the corrections of the form $1/kq$ will contribute to the exponential behavior of $\mathcal{F}$.

3.2 Diagram counting intuition and hypergraphs

The importance of ladder diagrams in all-to-all SYK is because they contain all $\mathcal{O}(\frac{1}{N})$ contributions to the four-point function. As we showed in the previous section, in the sparse model there are additional contributions (compared to all-to-all) of the form $\frac{1}{kq}$. This implies that *sparse-SYK has more diagrams at* $\mathcal{O}(\frac{1}{N})$. This may seem counterintuitive at first since sparse-SYK *decreases* the number of couplings, however, an analysis of how the couplings themselves change elucidates this problem. Note that we have redefined the variance in the pruning model to be

$$\left\langle J_{ijkl}^2 \right\rangle = \frac{(q-1)! J^2}{p N^{q-1}} \approx \frac{J^2}{kq}, \tag{3.22}$$

where we have used $k = p/N\binom{N}{q} \approx p N^{q-1}/q!$. Therefore, the contributions of $\mathcal{O}(\frac{1}{N})$ no longer depend on the number of couplings J_{ijkl}, rather the $\mathcal{O}(\frac{1}{N})$ counting moves to the pruning coupling x_{ijkl}

$$\langle x_{ijkl} \rangle = p \approx \frac{kq!}{N^{q-1}} = \left\langle x_{ijkl}^n \right\rangle, \tag{3.23}$$

where n is any positive integer. Intuitively, the right equality follows because if $x_{ijkl} = 1$, then the couplings with labels $\{i, j, k, l\}$ exist in the Hamiltonian. Hence, a diagram vertex representing this coupling is non-zero, so we can insert these vertices an arbitrary number of times and the diagram will still exist in the theory. In other words, if couplings with indices

$\{i, j, k, l\}$ exist with probability p, then the probability of the same couplings contributing twice is not less likely (such as p^2), because said couplings already exist in the theory. Another version of this statement is that since x_{ijkl} can only be 1 or 0, $1^n = 1$ and $0^n = 0$. Additionally, any non-zero term will have the same $1/kq$ contribution from the J's, specifically the nth order ladder diagram will have $1/(kq)^n$ *from the J's*. Therefore, the order of any specific term in $\mathcal{O}(1/kq)$ is determined by how many p's (from the pruning couplings) contribute in that term. Hence, the $1/kq$ counting for any given term is also determined by the sparsity couplings x_{ijkl}.

In summary, this means that the $\mathcal{O}(\frac{1}{N})$ counting moves to the x_{ijkl} couplings, and since these behave differently from the J_{ijkl} couplings, the contributions of $\mathcal{O}(\frac{1}{N})$ should change in the sparse pruning model.

An open question remains: we have performed this analysis using the random pruning method, yet the preferable numerical choice appears to be the hypergraph method, do we expect the ladder diagrams in the hypergraph method to yield the same expansion in $1/kq$? While we cannot comment on the expected behavior for a fixed hypergraph configuration as any given diagram would be dependent on which couplings are present, we also find in the subsequent section that a fixed hypergraph is insufficient to extract the Lyapunov exponent using the known numerical methods for the SYK model. It is only upon averaging over hypergraphs that the Lyapunov exponent can be extracted accurately, it is important to note that averaging over random pruning and averaging over hypergraphs (for sufficiently large averages) will lead to the same behavior. This is because when averaging over hypergraphs, the hypergraph configuration is chosen randomly, in which case (upon sufficient averaging) each coupling will occur will probability p by construction. This is analogous to the random pruning model and hence we expect the same behavior, albeit the hypergraph should require less averaging due to its higher connectivity.

4   Lyapunov numerics

Numerical methods are needed to study the Lyapunov exponent at finite q and finite coupling in the all-to-all SYK because no closed form is currently known in this regime.[2] The need for numerical methods to study the growth of OTOCs is even more pressing in the sparse-SYK model, where the general analytic relation between the Lyapunov exponent, λ_L, and the sparsity parameter k is unknown. As we showed in the previous section, the early time behavior of the OTOC contains $1/kq$ contributions and it is unclear how to account for these contributions in models for extracting analytically the Lyapunov exponent.

In this section we numerically simulate the behavior of the regularized OTOCs (3.2) at finite N in the sparse-SYK model. We employ the *regular hypergraph* method of generating sparseness, as it ensures greater connectivity for any given realization of the Hamiltonian. The qualitative behavior from varying k is shown in figure 2, for which we see a decrease in the rate of exponential growth as k decreases. This behavior is physically expected, as the sparse model reduces the number of connections between fermions, and hence it decreases the sensitivity to operator perturbations. At this stage, the computational benefit from

---

[2]Perturbative expressions in $1/\beta J$ and in the $q \to \infty$ limit [7] do exist.
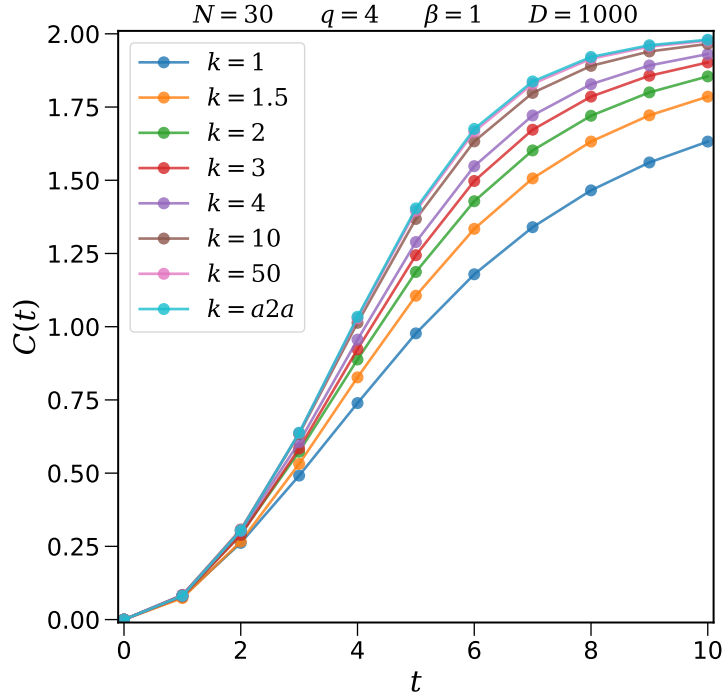
Figure 2. A comparison between the double commutators $C(t) = 2 - 2\mathcal{F}(t)$ of the all-to-all model (denoted 'a2a') and various sparsity parameters k for the sparse model. All curves use the same number of fermions ($N = 30$), number of fermions per interactions ($q = 4$), temperature ($\beta = 1$), and number of disorder realizations ($D = 1000$). For each k, hypergraphs are averaged over, a choice justified in the main body of the paper. Higher k's closely resemble the curves of the all-to-all case, however as $k \to \mathcal{O}(1)$, i.e. the limit still expected to retain chaos, the exponential growth starkly decreases.

sparse-SYK is undeniable. Table 1 demonstrates that even for relatively low N, sparse simulations reap massive efficiency gains, an effect that further widens the higher N one chooses to probe.

If one wishes to capitalize on these efficiency benefits to quantify quantum chaos by means of a Lyapunov exponent, one needs a model for use in extraction. In the all-to-all model, the best known method for accurate extraction of the Lyapunov exponent [25] utilizes a symmetry of the OTOC expansion

$$\text{OTOC} \approx C_0 + C_1 \frac{e^{\lambda_L t}}{N} + C_2 \frac{e^{2\lambda_L t}}{N^2} + \cdots, \tag{4.1}$$

which is invariant under

$$t \to t + \frac{1}{\lambda_L} \log c, \quad N \to c\,N. \tag{4.2}$$

We know the sparse model will approach the all-to-all model in the large k limit, and hence this symmetry is expected to be recovered in said limit. Our aim is to determine whether the aforementioned symmetry still holds in the sparse-SYK model, even if only approximately, in such a way that it allows a quantifiable extraction of the Lyapunov exponent for a given k.

Time Comparisons Average, $q = 4$ $\beta = 1$ (112 Samples)				
Size (N)	All-to-All	$k = 50$	$k = 10$	$k = 4$
26	00:02:14	00:00:33	00:00:10	00:00:04
28	00:06:22	00:01:19	00:00:23	00:00:10
30	00:17:44	00:03:09	00:00:52	00:00:22
32	00:47:42	00:07:15	00:01:55	00:00:48
34	02:06:27	00:16:26	00:04:14	00:01:48
36	05:27:22	00:37:26	00:09:28	00:04:02

Table 1. Time comparisons between all-to-all SYK and various sparseness (k's) in sparse-SYK using the hypergraph method over a range of N's from $N = 26$ to $N = 36$. All simulations were run on a single processor with $\beta = 1$ and $q = 4$, averaging over 112 samples to account for time differences in individual runs. Times are rounded to the nearest second.

4.1 Extraction methods

Typical numerical methods for extracting the Lyapunov exponent via the OTOC involve performing an exponential fit in the exponential regime of the OTOC [28–31]. However, this method was found to be insufficient at finite N and $q = 4$ for the all-to-all model, where the OTOC at finite-size could only be matched to a simple exponential for a very short time interval. It has been shown [25] that a numerical derivative method exploiting the symmetry (4.2) is most effective in extracting the Lyapunov exponent. For each value of N, a fitted Lyapunov exponent $\lambda_{\text{fit}}(N)$ was found as follows. Since a shift under the symmetry (4.2) leaves the OTOC invariant, we can pick a constant value of the OTOC, say $\mathcal{F}^*$, between all N's simulated and use interpolation to find the time t^* associated to each $\mathcal{F}^*$. It follows that the fitted Lyapunov exponent between two curves (assuming the symmetry holds) is

$$\frac{1}{\lambda_{\text{fit}}} = \frac{\text{Log}(N_2) - \text{Log}(N_1)}{t_2^* - t_1^*}, \tag{4.3}$$

which could then be plotted vs $1/N$ to see the effect as the system approached large N. Taking a second order fit in $1/N$ for λ_{fit}, λ_L is defined as the value when $1/N \to 0$

$$\lambda_{\text{fit}}(N) \approx \lambda_L + \frac{\lambda_1}{N} + \frac{\lambda_2}{N^2}. \tag{4.4}$$

In our simulations, we use $\mathcal{F}^* = 0.4$. See figure 3 for a concrete example of this procedure. We test the robustness of this choice by using different values in appendix C. Due to the success of the numerical derivative method exploiting the symmetry (4.2), we choose to closely match the numerical tools employed by [25] in their extraction of the Lyapunov exponent. This includes utilizing *dynamite* [32], a Python package allowing fast simulation of quantum dynamics with Krylov subspace implementation [33]. As stated in section 2.1, a *regularized* OTOC

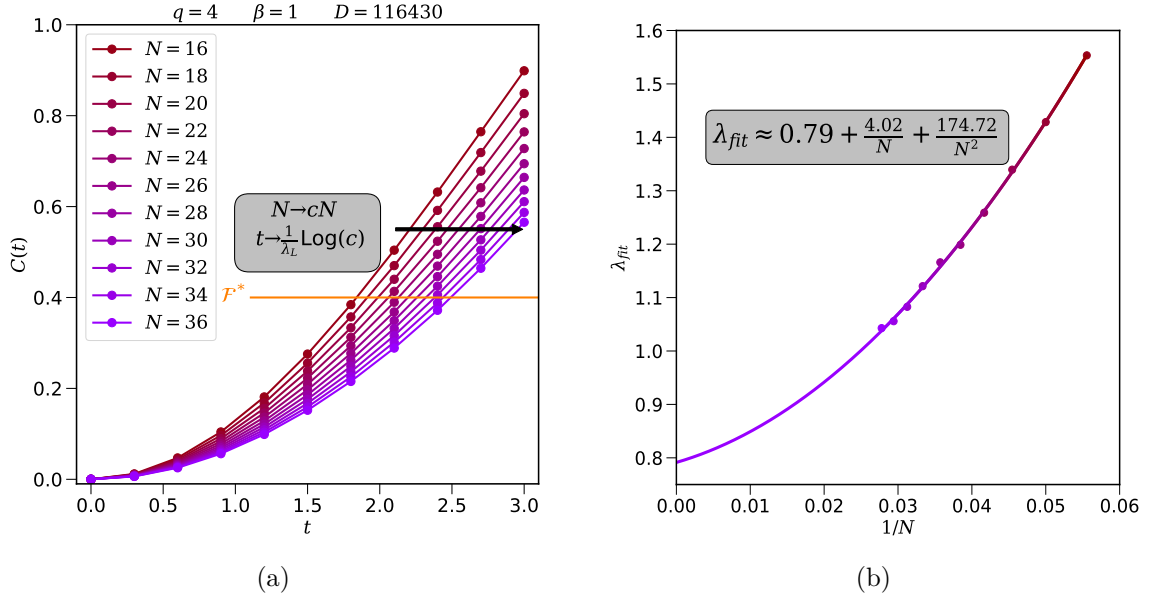$$\mathcal{F}(t) = \langle W(t)\rho^{1/4}V(0)\rho^{1/4}W(t)\rho^{1/4}V(0)\rho^{1/4}\rangle_\beta \tag{4.5}$$

Figure 3. Plot visualizing the extraction of λ_{fit} in all-to-all SYK ($q = 4, \beta = 1$, 116430 disorder realizations) from the double commutator $C(t)$ via a numerical derivative. (a) Plots of the $C(t)$ versus time for $N = 16$ to $N = 36$. To calculate the numerical derivative, a horizontal cross section $\mathcal{F}^*$ must be chosen, here $\mathcal{F}^* = 0.4$. The symmetry (labelled within the grey box) describes at what times subsequent curves should intersect $\mathcal{F}^*$. (b) After the numerical derivative between subsequent curves is calculated via (4.3), λ_{fit} is plotted versus $1/N$ and fitted to a 2nd order polynomial (in grey box) to extract the large N Lyapunov exponent, here $\lambda_L = 0.79$.

has been shown to reduce finite-size effects, and here V and W are chosen as distinct single Majorana operators. Additionally, it is known for large N that the thermal expectation value involving the trace can be closely approximated by calculating the expectation value via Haar random states [25, 34, 35], i.e. for a Haar random state $|\psi\rangle$

$$\text{tr}\left[e^{-\beta H}\hat{O}\right] \approx \overline{\langle\psi|e^{-\beta H}\hat{O}|\psi\rangle}. \tag{4.6}$$

The finite N error in this approximation can be reduced through sufficient averaging over different initial random states, represented by the overline. It has been shown that averaging these states separately or simultaneously with the couplings J_{ijkl} in the all-to-all case had no significant difference in error. Hence, in this work, we also employ simultaneous averaging of the couplings and initial random states.

We find another source of averaging that is important to the extraction of the Lyapunov exponent not present in the all-to-all model, namely averaging over many hypergraphs (by generating regular hypergraphs using different seeds). As we will explain below, the fixed hypergraph method does not produce the expected symmetry. Only after averaging over many hypergraphs does the desired symmetry becomes apparent (see appendix C). Thus, only for an aggregate of hypergraphs the symmetry-based method could be used to extract the Lyapunov exponent.

In order to quantify the error of λ_L, we simulate a large number of realizations $\gtrsim 100,000$ and subdivide these realizations into subgroups, from which we extract the Lyapunov exponent for each subgroup. By performing statistics on many of these subgroups, we can derive a mean Lyapunov exponent $\bar{\lambda}$, a standard deviation σ, and a coefficient of determination R^2. We explain this procedure in more detail in what follows.

4.2 Averaging over hypergraphs

The ultimate goal of this section is to glean whether the symmetry as described by (4.2) holds in the sparse-SYK model, and if so to extract Lyapunov exponents and determine their dependence on the sparsity parameter k. We can get a qualitative idea of whether the symmetry holds from figure 4, where we plotted the Double Commutators of the sparse model in comparison to the all-to-all model over $D = 1000$ realizations. These results employ simultaneous averaging over hypergraphs, couplings, and random initial states. At this level of averaging, the inconsistent behavior in spacing between subsequent curves suggest that the symmetry breaks down as sparsity increases. This behavior admits at least three possible explanations. Either the symmetry itself does not hold in the sparse model, the sparse model is more susceptible to finite-sized effects, or the averaging over hypergraphs produces random fluctuations requiring a higher degree of averaging.

A quantitative approach must be employed to separate these possibilities. To test the impact of random fluctuations, we note that an average over a sufficiently large number of realizations must converge to a certain value, and therefore repeating the experiment using the same number of realizations should not change that value. We measure the dependence on the number of realizations by repeating the extraction of the Lyapunov exponent many times, where each extraction utilizes D_g realizations. From these repetitions, we obtain a mean value for the Lyapunov exponent $\bar{\lambda}$ and a standard deviation σ for each D_g. By varying D_g, we can analyze how the error on $\bar{\lambda}$ is modified. The choice of how many repetitions to average over is completely arbitrary, we find that the choice yielding the smoothest results for Lyapunov extraction is to average the "maximum" number of times, i.e. if there are D_{tot} total number of realizations, then we can make D_{tot}/D_g unique groupings of D_g realizations and hence we average over all of these groups. When $D_{\text{tot}}/D_g < 10$ we stop increasing D_g, so as to allow sufficient samples to be averaged over. Another possible choice is to choose a set number of samples to average over regardless of D_g, we do this with 10 and 20 samples in appendix C as a test of robustness.

The other quantitative tool utilized is the coefficient of determination R^2, which represents in our case how well the numerical derivative data of a given group fits (4.4). It has the formula

$$R^2 = 1 - \frac{SS_{\text{res}}}{SS_{\text{tot}}}. \tag{4.7}$$

The term SS_{res} refers to the sum of squares of residuals, $SS_{\text{res}} = \sum_i (x_i - \hat{x}_i)^2$, x_i refers to our numerical data and $\hat{x}_i$ refers to the prediction based upon our model from (4.4). This represents the total sum of variances between our model's prediction and the numerical data. SS_{tot} on the other hand is the sum of the variance of the data, $SS_{\text{tot}} = \sum_i (x_i - \overline{x})^2$.
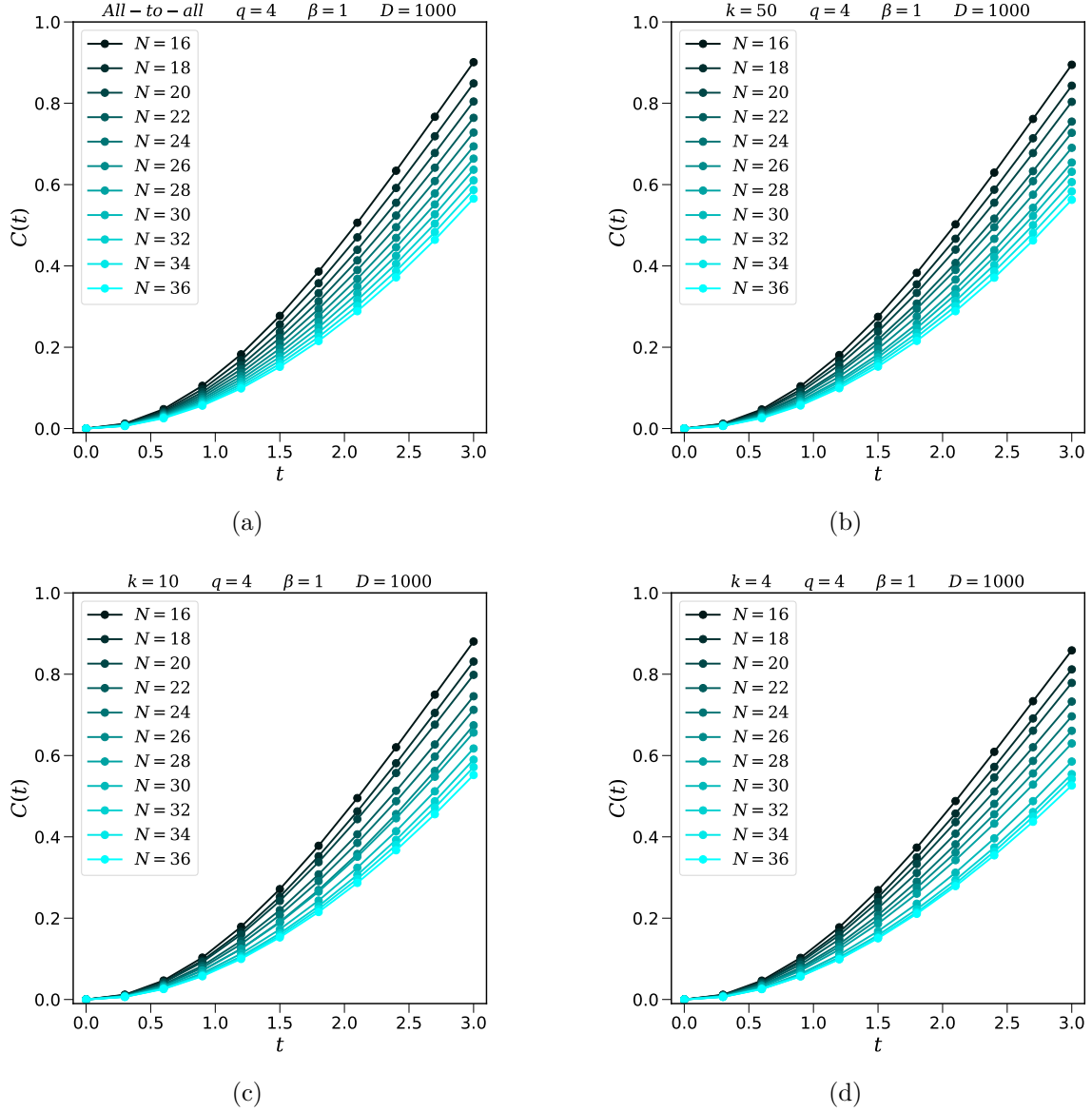
Figure 4. Comparison of $C(t) = 2 - 2\mathcal{F}(t)$ between All-to-all and various Sparse models with $D = 1000$ disorder realizations. Presented data is averaged over many different hypergraphs. (a) In the All-to-all model, the symmetry visually holds as N increases; subsequent curves get successively closer together. (b, c, d) Plots for $k = 50, 10, 4$ respectively. The symmetry visually appears to break down in the Sparse regime at this level of realizations, i.e. the spacing between subsequent curves is not consistent with the symmetry.

The ratio $SS_{\text{res}}/SS_{\text{tot}}$, also called the fraction of variance unexplained, represents the ratio between the "unexplained" variance (from the discrepancy between model and data) and the total variance. If a set of data implements the symmetry given by (4.2), λ_{fit} will closely match the data, implying $SS_{\text{res}}/SS_{\text{tot}} \to 0$ and therefore $R^2 \approx 1$. Hence, by analyzing how R^2 changes with D_g, we can see whether simply greater averaging is sufficient to recover the symmetry.

Due to the outputs of the mean Lyapunov exponent, associated standard deviation, and R^2 all roughly lying in the range $[0,1]$, we can plot these values together in the same plot, which can be seen in figure 5. These plots indicate that random fluctuations, caused in large part because hypergraphs must be averaged over, have a much larger effect for sparse models. It is not until far larger numbers of disorder realizations in the Sparse case that similar values for error and R^2 as compared to the all-to-all model are achieved. At this level of disorder groupings, the all-to-all model demonstrates convergence in the mean, standard deviation, and R^2 values with groups of roughly $D_g \sim 10,000$, with $R^2 \to 1$ suggesting the symmetry based model closely matches the data. As k decreases, the stability of this convergence slightly decreases, with somewhat higher fluctuations. However, given a large enough D_g, such as $D_g \sim 175,000$ for $k = 4$, we can also recover $R^2 \to 1$ and some convergence in error and mean Lyapunov exponent. We illustrate this point in figure 6.

Overall, we find that as k decreases, the number of disorder realizations goes up to (and exceeds) an order of magnitude greater than in the all-to-all case. This suggests the following interpretation: as k decreases, the structure/connections of the sparse Hamiltonian further and further diverge from that of the all-to-all model. To compensate for a given hypergraph being a poor representation of the all-to-model from which it is constructed, hypergraphs (and hence the connections/structure) must be averaged over if the symmetry is to be recovered. The subsequent section provides further evidence for this interpretation.

One important observation from figure 5 is the appearance of the $k = 50$ average Lyapunov exponent being higher than that of the all-to-all model, this is surprising as increasing sparseness should decrease the chaotic properties. This is however an artifact of numerical error, specifically that the Lyapunov exponent approaches its asymptotic value from above. At our level of disorder realizations, the standard deviation and R^2 (for all k) make it clear that full convergence has not yet been achieved, the $k = 10$ and $k = 4$ cases demonstrate that as one approaches convergence by increasing D_g, the average Lyapunov exponent continues to decrease. They also demonstrate that for lower k, convergence is not achieved until higher levels of R^2 are achieved as compared to the all-to-all model. We expect that given enough disorder realizations, the extracted Lyapunov exponent for $k = 50$ would drop below that of the all-to-all model. A reasonable cause for why the standard deviation may asymptote above zero is due to finite N effects, and it appears for the $k = 50$ case this asymptote may be above that of the all-to-all case. It is also possible that some disagreement with symmetry-based models is due to finite N behavior acting more strongly in the sparse model. Hence, our next goal is to investigate whether this is the case, and if simulating larger N can lead to more accurate extraction of the Lyapunov exponent.

4.3 Is accurate extraction possible at larger N?

It is natural to ask if going to larger N will allow us to extract the Lyapunov exponent more accurately, as it does for the all-to-all model. Unfortunately, we find that at fixed k, the answer seems to be no. In this section, we probe higher N's in the sparse model to test if finite-sized effects can explain some of the disagreement with symmetry-based models. Our methods match those used in the previous section, except rather than comparing different k's we compare different ranges of N's for the $k = 4$ model, which approaches the minimum
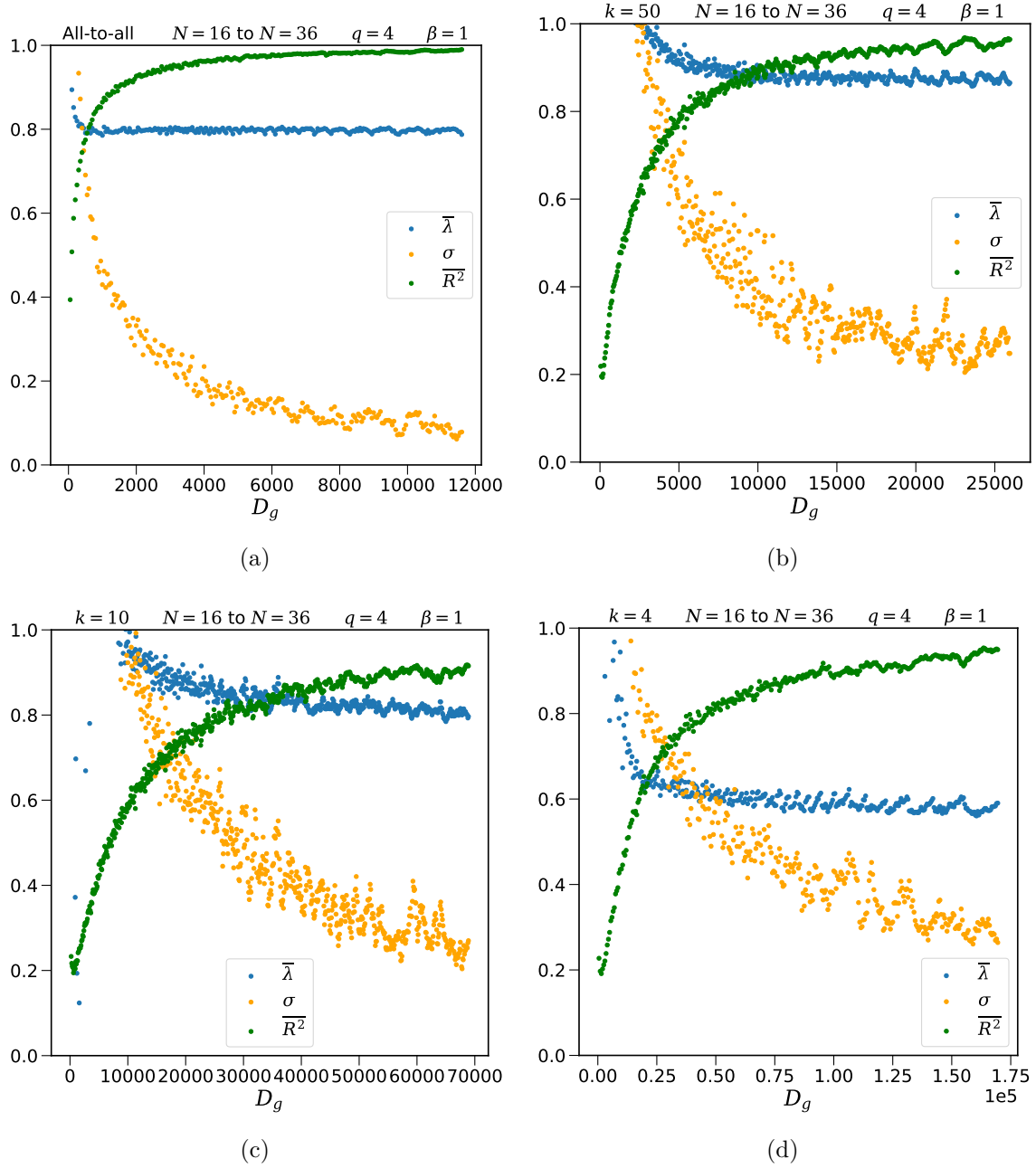
Figure 5. Many realizations are subdivided into groups sized D_g, from which a Lyapunov exponent and R^2 is extracted. This process is repeated D_{tot}/D_g times, the average values of λ_L (*blue*) and R^2 (*green*) are plotted, as well as the standard deviation of λ_L, σ (*orange*). Resolution of plots all have $\Delta D_g = 50$. **(a)** all-to-all model. Exhibits convergence for plotted quantities; $R^2 \to 1$ shows agreement with the symmetry. **(b,c,d)** $k = 50, 10, 4$ respectively. For similar values of R^2, display similar values of error (even compared to (a)), however fluctuations increase as compared to all-to-all.
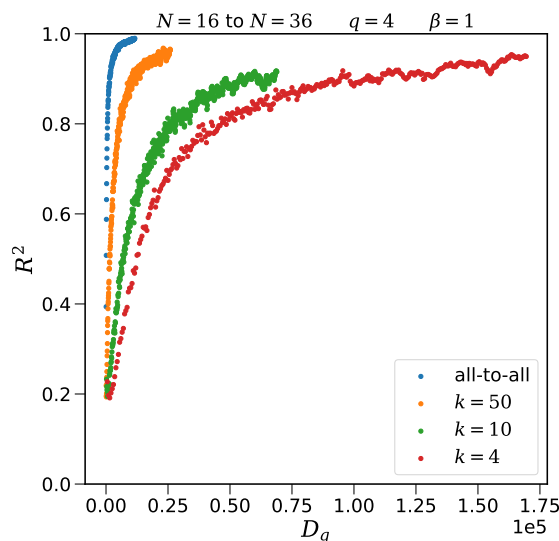
Figure 6. R^2 for the plots in figure 5. As the model becomes more sparse, the number of realizations has to increase considerably to obtain reliable results. As in previous plots, $N = 16$ to 36, $q = 4$, $\beta = 1$.

limit of sparsity while still retaining chaos, and is correspondingly the fastest of our choices of k to simulate. We chose to retain the lower bound of $N = 16$ when extracting the Lyapunov exponent for all higher N's as it led to better agreement with symmetry methods. In appendix C we have also included an alternative analysis with a constant range of N's (e.g. for $N = 42$ the range would be $N = 22$ to $N = 42$).

The corresponding plots for these simulations can be found in figure 7. We find that not only do random fluctuations actually increase in size with higher N, but also the R^2 value has no better agreement as N increases. There is a likely explanation for these trends. Recall that the number of interactions scales as $\sim N^q$ for the all-to-all model and as $\sim kN$ in the sparse models. Therefore, as N increases, the ratio of interaction terms between the all-to-all and sparse models grow as N^{q-1}. Our results in section 4.2 show that the further one strays from the all-to-all configuration, the more averaging one must do to recover the symmetry. Thus, as N increases and the difference in interaction terms grows larger, the sparse model requires higher levels of averaging to recover the symmetry.

Hence we have demonstrated that for low k, there does not appear to be a simple method to consistently extract the Lyapunov exponent at finite N using currently known methods without averaging over massive number of realizations. Another approach one could employ is varying β, however it is unlikely this would lead to better behavior, certainly for $\beta > 1$ as fluctuations rise in this case, in part due to fewer values of N being suitable for extraction since the OTOC expansion being valid requires $N \gtrsim \beta J$.

4.4 Efficiency and Lyapunov exponents

The previous sections have made evident that higher degrees of sparsity (lower k) require greater averaging (and hence greater computation times) in order to accurately extract
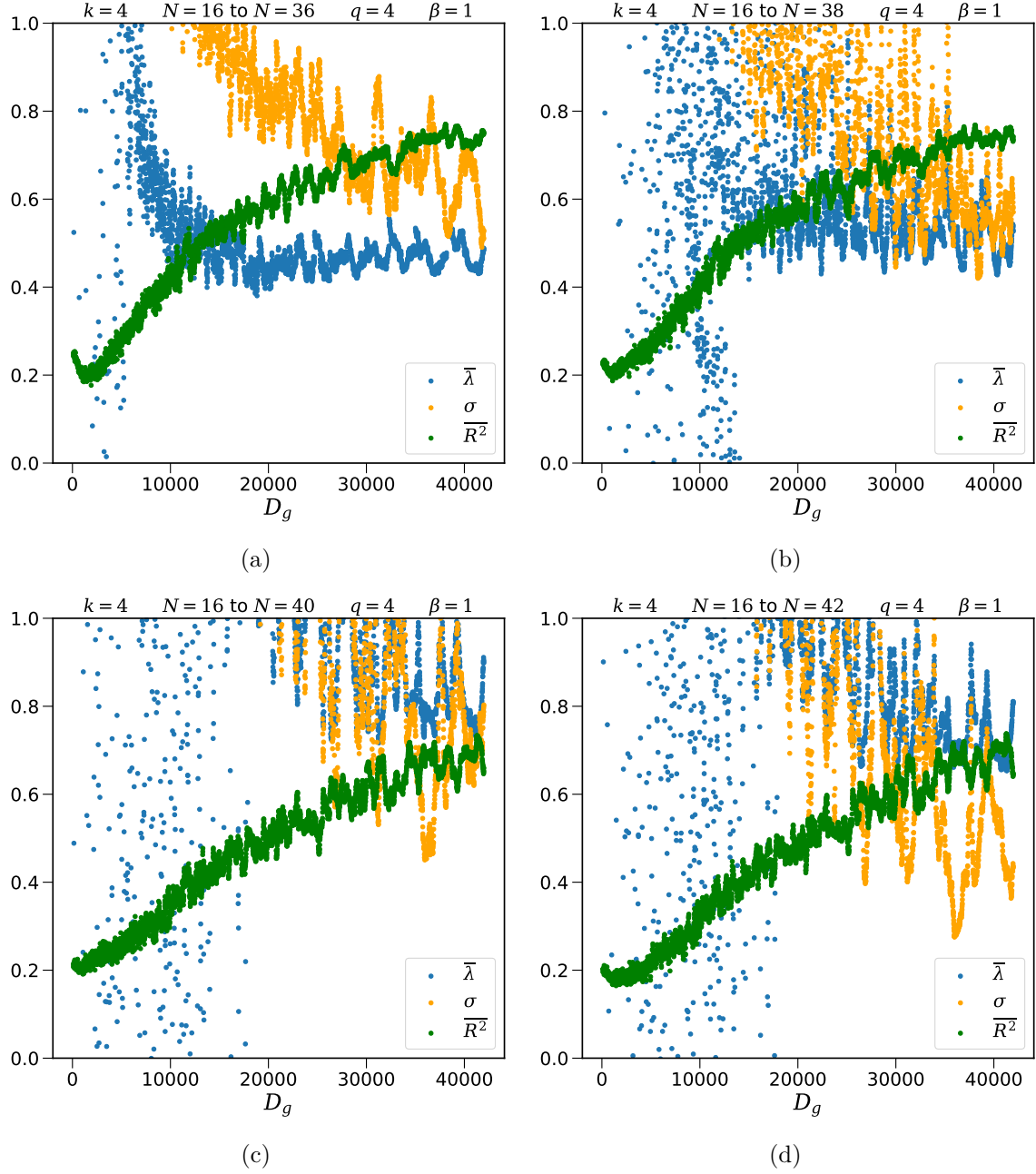
Figure 7. Many realizations are subdivided into groups sized D_g, from a Lyapunov exponent and R^2 is extracted. This process is repeated D_{tot}/D_g times, the average values of the λ_L (*blue*) and R^2 (*green*) are plotted, as well as the standard deviation of λ_L (*orange*). Resolution of plots all have $\Delta D_g = 10$. All plots are of k=4 sparse model, and have the same max D_g. **(a)** Max value $N = 36$, scaled down version of figure 5(d), baseline of comparison. **(b,c,d)** Max value $N = 38, 40, 42$ respectively. All exhibit greater fluctuations and no significant change to R^2 for this range of D_g.

the Lyapunov exponent. We now seek to determine whether the computational benefit inherit in sparse-SYK is retained after the large amounts of disorder averaging that must be performed. In this section, we present a rudimentary method to determine the relative efficiencies of sparse models as compared to all-to-all, granting that more sophisticated methodology could be employed to perhaps gain greater insight as to how they compare. Based upon our methodology, we find that the computational trade-off between inherent speed increase in sparse models and time increase due to greater disorder averaging roughly cancel out, meaning accurate Lyapunov extraction is no more efficient for low k models as it is for all-to-all models. We also present the numerically calculated Lyapunov exponents for our given range of k's.

Our rudimentary choice of calculating efficiency starts with choosing what qualifies as "accurate" Lyapunov extraction. Given our ranges of simulated D_g, a goodness of fit of $R^2 = 0.9$ was our choice of where to begin considering the extraction as accurate enough. Since we do not have exact curves corresponding to how R^2 varies with D_g, we instead chose to analyze the corresponding values of D_g where the numerically calculated values of R^2 were within 0.005 of $R^2 = 0.9$. This gives us a range of D_g values, interpreted to mean that on average, samples extracted with those values of D_g will produce curves for λ_{fit} with a goodness of fit of $R^2 \approx 0.9$. For data that passes through $R^2 = 0.9$ multiple times, the largest continuous cluster is chosen.

We also need to quantify the time advantage sparse-models have over all-to-all models, which can be straightforwardly done by finding the time (on average) it takes to simulate a single disorder realization for our range of N's ($N = 16$ to $N = 36$). The full list of times per N is provided in appendix B, we present in this section just the total time for all values of N. By dividing the total time (in seconds) by the total time of the all-to-all model, one achieves a "time factor", representing the ratio of how much faster a given sparse model is per realization as compared to the all-to-all model. Multiplying the previously described ranges of D_g by these time factors will yield the relative number of disorder realizations needed to achieve accurate (as deemed by $R^2 = 0.9$) Lyapunov extraction. For the calculated values of the Lyapunov exponent and associated standard deviation, we average over the data with $R^2 > 0.9$ (rounded within 0.01), which all achieve convergence within those ranges. Results of this analysis can be found in table 2. For discussion regarding the appearance of the $k = 50$ extracted Lyapunov exponent being higher than that of the all-to-all model, refer to the end of section 4.2.

This analysis demonstrates that the necessity of greater disorder averaging required for accurate Lyapunov extraction within sparse-SYK roughly negates the computational benefit provided per realization by sparse models, since the relative efficiency is roughly comparable between all models.

5 Conclusions

We studied the quantum chaotic behavior in the sparse-SYK model as diagnosed by out-of-time ordered correlators (OTOCs). We approached the problem through both analytical and numerical methods. In the large N limit, where the OTOCs are given by a sum of

Efficiency Table (112 Samples for Time)				
	All-to-All	$k = 50$	$k = 10$	$k = 4$
Total Time ($N = 16$ to 36)	08:49:51	01:06:44	00:17:11	00:07:18
Time Factor	1	0.126	0.0324	0.0138
D_g Range, $R^2 \approx 0.9$	1400–1500	9700–12250	53500–68000	90000–111000
Relative Efficiency	1400–1500	1222–1543	1733–2203	1242–1531
Lyapunov Exponent	0.80	0.88	0.81	0.58
Standard Deviation	0.16	0.30	0.28	0.34

Table 2. Efficiency comparisons between all-to-all SYK and various sparse models. The total time is the average total time (over 112 samples) to simulate a single disorder realization. The time factor is the fraction of time (compared to all-to-all) spend per disorder realization. The D_g range is the range of disorder groupings which yield a value of $R^2 = 0.9$ within 0.005. The relative efficiency is the range of D_g multiplied by the corresponding time factor. The Lyapunov exponent and standard deviation is the average of the corresponding quantities over the range of D_g's with $R^2 > 0.9$ from figure 5.

ladder diagrams, we found an algorithm for calculating the $1/kq$ corrections to the ladder diagrams. This allows us to obtain analytic expressions for ladder diagrams of arbitrary order expanded to arbitrary orders in $1/kq$ — see appendix A for details. Importantly, the kernel in (3.18) obtained in this procedure is dependent on the ladder diagram order, a feature not shared by the all-to-all SYK. Thus, the summation over ladder diagrams needs to be truncated at some order in $1/kq$, and the summation to extract the corrections to the early time exponential behavior of the OTOC cannot be performed with the same tools as in the all-to-all model.

The numerical approach employed in this work followed the symmetry-based method in [25], previously applied to the all-to-all SYK, which is the only known accurate method for extraction of the Lyapunov exponent from finite N simulations. We found that sparse models with $k \lesssim 50$ require far larger averaging of disorder realizations for accurate extraction as compared to the all-to-all model. In a fixed regular hypergraph sparse model, the extraction of the Lyapunov exponent led to inconsistent results, i.e., two fixed regular hypergraphs with the same level of sparseness resulted in different Lyapunov exponents even when we significantly increase the number of disorder realizations. This indicates that the symmetry (4.2) does not hold for a fixed hypergraph configuration. In order to recover the symmetry, we performed an additional average over regular hypergraphs with the same sparseness parameter k. We find that even with the extra average, the fluctuations in the Lyapunov exponent are still larger compared to the all-to-all SYK, and these fluctuations increase even more as k decreases.

As shown in section 4, a reliable extraction of the Lyapunov exponent is still possible for low values of k but a very large number of realizations is needed. Therefore, there is a trade-off between efficiency and variance in the extracted Lyapunov exponent in the sparse-SYK. One could argue that these fluctuations can be reduced through probing larger values of N. However, as we argued in section 4.3, the opposite effect takes place for

Lyapunov extraction in the sparse-SYK model. The reduction in computation time of a single realization is offset by the larger number of simulations needed to be performed. In our numerical tests, the offset of the two effects is roughly comparable, i.e., the computational resources required to achieve similar levels of accuracy for Lyapunov extraction are roughly the same between all-to-all and sparse models.

We remark that our numerical results are consistent with the previous statement in [21] that quantum chaos is present in the sparse-SYK model for $k \sim \mathcal{O}(1)$, derived from level statistics [36]. Interestingly, the analysis of [21] also revealed the existence of emergent symmetries in the sparse-SYK model. It is possible that these emergent symmetries are correlated with the large fluctuations observed in the Lyapunov exponent.

Our results suggest several directions to explore in the future.

- In [37] the authors calculate the Lyapunov exponent in a variant of the SYK model by solving the Schwinger-Dyson equations numerically. It would be interesting to investigate if this method can be an alternative way to extract the Lyapunov exponent of the sparse-SYK model.

- Our numerical results indicate that the early time behavior of the OTOCs is affected by the sparseness of the model. It would be interesting to further explore the summation of the ladder diagrams to determine the leading corrections to the early time exponential ansatz $\mathcal{F} \sim e^{\lambda t}/N$ in the sparse-SYK model.

- A better understanding of the emergent symmetries found in [21], especially their effect at larger N, could also provide insights on how the sparseness affect the behavior of OTOC expansion (2.4).
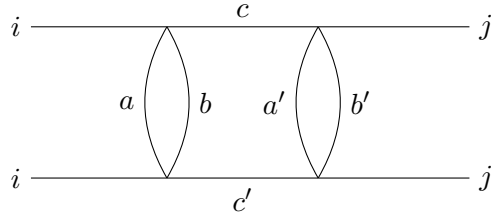
Acknowledgments

A Sparse SYK ladder diagram general algorithm

A.1 Simplifying diagram calculations

As seen from the calculations in section 3, general ladder diagram calculations will heavily involve combinatorics, especially as their order increases. Additionally, since the kernel is dependent on the order of the diagram, finding a given ladder diagram to arbitrary order is a non-trivial procedure. However, as we will see from calculating the first several ladder diagrams, a pattern emerges. In this section, we will develop the notation and motivation behind this procedure.

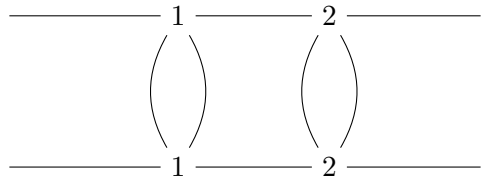Let us return to the example of the two rung ladder diagram

$$\text{(A.1)}$$

General ladder diagrams require $c = c'$ to be $\mathcal{O}(\frac{1}{N})$. In this case, we notice the Gaussian moment expansion (3.9) simplifies to

$$\langle J_{iabc} J_{iabc} J_{ja'b'c} J_{ja'b'c} \rangle = \langle J_{iabc} J_{iabc} \rangle \langle J_{ja'b'c} J_{ja'b'c} \rangle + 2 \langle J_{iabc} J_{ja'b'c} \rangle \langle J_{iabc} J_{ja'b'c} \rangle . \quad \text{(A.2)}$$

Recall from (3.1) that only if two couplings' indices are equal (up to permutation) may the multiplication of the two be non-zero after averaging. Let us adopt a convention of labelling vertices from left to right in the ladder diagram. At the order of $1/N$ the vertices connected by rungs have the same indices, so we can label them by the same vertex number. Our convention will involve replacing couplings by their corresponding vertex numbers, for instance $J_{iabc} \rightarrow 1$. In this new labeling scheme, (A.1) corresponds to

$$\text{(A.3)}$$

Similarly, we can replace (3.9) with

$$\langle 1122 \rangle_J = \langle 11 \rangle_J \langle 22 \rangle_J + 2 \langle 12 \rangle_J \langle 12 \rangle_J . \quad \text{(A.4)}$$

In our new convention, the term $\langle 12 \rangle_J$ should be interpreted to mean that after averaging, $\langle 12 \rangle_J = \frac{J^2}{kq}$, only if vertices 1 and 2 have indices which are equal up to permutations of each other. Recall that the two relevant cases here are if vertex 1 and 2 are or are not permutations of each other. In the latter case (not permutations), $\langle 12 \rangle_J = 0$ (not permutations), so only the term $\langle 11 \rangle_J \langle 22 \rangle_J$ contributes. Let us represent the sparsity couplings $\langle x_{abcd} x_{abcd} \rangle = \langle x_{abcd} \rangle$ as $\langle 11 \rangle_x = \langle 1 \rangle_x$, we found the pruning couplings contributed a factor of $\langle 1 \rangle_x \langle 2 \rangle_x = p^2$.

In the other case, vertices 1 and 2 do have indices equal up to permutations. However since all vertices are equal, the averaging over x reduces to $\langle 1122 \rangle_x = \langle 1 \rangle_x = p$, so overall only a single factor of p is contributed from the x's. Recall that the total expression (in our new notation) is $\langle 1122 \rangle_J \langle 1122 \rangle_J$, so the factors from averaging over J's must also be considered. Overall, our goal is to find the numerical/combinatorial terms multiplying these p's since these are what determine the order in $1/kq$, which are embedded in averaging over the J's. By using (3.9) and dropping the factors of J^2 (these can be pulled out across the entire expression), let us create a table which counts all relevant contributions:

$\langle 1122 \rangle_J =$	$\langle 11 \rangle_J \langle 22 \rangle_J + 2 \langle 12 \rangle_J \langle 12 \rangle_J$			Total
p^2	1			1
p	1	$+$	2	3

Let us repeat this process for the ladder diagram with three rungs in full detail. For this case we have

$$\langle 112233 \rangle_J = \langle 11 \rangle_J \langle 2233 \rangle_J + 2 \langle 12 \rangle_J \langle 1233 \rangle_J + 2 \langle 13 \rangle_J \langle 1223 \rangle_J \tag{A.5}$$

$$= \langle 11 \rangle_J \langle 22 \rangle_J \langle 33 \rangle_J + 2 \langle 11 \rangle_J \langle 23 \rangle_J \langle 23 \rangle_J + 2 \langle 12 \rangle_J \langle 12 \rangle_J \langle 33 \rangle_J +$$
$$+ 2 \langle 13 \rangle_J \langle 13 \rangle_J \langle 22 \rangle_J + 8 \langle 13 \rangle_J \langle 12 \rangle_J \langle 23 \rangle_J. \tag{A.6}$$

Let us organize again by order in p, the case p^3 is given by no vertices being equal up to permutations. In that case, as was the case for $n = 2$, only the first term contributes. For order p^2, the counting gets trickier. There are three cases which contribute, if vertices 1 and 2 are equal, vertices 1 and 3 equal, and vertices 2 and 3 are equal. For 1 and 2 equal, we have $\langle 13 \rangle_J = 0 = \langle 23 \rangle_J$, leaving a coefficient of $1 + 2 = 3$. Each other case of order p^2 has this same numerical coefficient, making the total numerical coefficient 9. For order p, all terms contribute giving a numerical constant of 15.

Since all we care about is the order of p, not the specific case of which vertex is a permutation of another, we notice that every term with coefficient 2 in (A.6) is the same up to relabelling of vertices, therefore they contribute at the same order of p. We therefore wish to label these terms by a factor which suggests that they all contribute equally at the same order of p. If the highest order of p a term contributes is p^k, we will relabel these terms by the factor $\mathcal{P}^k$. In this way may rewrite (A.6) as

$$\langle 112233 \rangle_J \sim \mathcal{P}^3 + 6\mathcal{P}^2 + 8\mathcal{P}. \tag{A.7}$$

Our current notation is still burdensome, as the expansion contains terms with both $\langle \cdot \rangle_J$ and $\langle \cdot \rangle_x$, our goal will be to combine them. We can do this by notationally decomposing our average over x into groups of 2 vertices (like that which is naturally done from the Gaussian moment expansion for the J's). To see this concretely, let $\langle \cdot \rangle_d$ be a term in the decomposed average, such that $\langle aabb \ldots \rangle_x \equiv \langle aa \rangle_d \langle bb \rangle_d \langle \ldots \rangle_d$. Note that these decomposed terms are *not* a true average nor should they be interpreted as one, they can recombine/decompose freely (to conserve the relevant structure) but only after a total recombination of the $\langle \cdot \rangle_d$'s do they represent something physically meaningful, namely it recovers the original average

over x. Their purpose is simply for notational convenience. The use of this becomes apparent when we let $\langle ab \rangle \equiv \langle ab \rangle_J \langle ab \rangle_d$, meaning we have the following identities

$$\langle aa \rangle_J \langle aa \rangle_x = \langle aa \rangle \sim \mathcal{P} \tag{A.8}$$

$$\langle ab \rangle_J \langle ab \rangle_J \langle abab \rangle_x = \Big(\langle ab \rangle_J \langle ab \rangle_d \Big) \Big(\langle ab \rangle_J \langle ab \rangle_d \Big) = \langle ab \rangle \langle ab \rangle \sim \mathcal{P} \tag{A.9}$$

$$\langle ab \rangle_J \langle ac \rangle_J \langle abac \rangle_x = \langle ab \rangle \langle ac \rangle \sim \langle bc \rangle. \tag{A.10}$$

The second identity is true because the term is only non-zero when $\langle ab \rangle_J \neq 0$, meaning vertices a and b are permutations of one another. In that case, $\langle abab \rangle_x = \langle a \rangle_x = p$, meaning the highest order the term can contribute is p^1, hence the term corresponding to $\mathcal{P}$. The third identity can not be further simplified (ie, we cant set it equal to $\mathcal{P}$) because if $\langle bc \rangle \sim \mathcal{P}$, that would imply $\langle bc \rangle \langle bc \rangle \sim \mathcal{P}^2$, which contradicts the second identity. The origin of this comes from equivalent vertices appearing in pairs in the ladder diagram.

Another way to see why we make this replacement is to note as stated earlier that a relabelling of indices will not change the way in which the counting at each order of p is done, therefore we can always choose to relabel terms such that they combine into the form (A.7), for example

$$\begin{aligned}
\langle 112233 \rangle &= \langle 11 \rangle \langle 2233 \rangle + 2 \langle 12 \rangle \langle 1233 \rangle + 2 \langle 13 \rangle \langle 1223 \rangle \\
&= \langle 11 \rangle \langle 2233 \rangle + 4 \langle 12 \rangle \langle 1233 \rangle \\
&= \langle 11 \rangle \langle 22 \rangle \langle 33 \rangle + 2 \langle 11 \rangle \langle 23 \rangle \langle 23 \rangle + 4 \langle 12 \rangle \langle 12 \rangle \langle 33 \rangle + 8 \langle 13 \rangle \langle 12 \rangle \langle 23 \rangle \\
&= \langle 11 \rangle \langle 22 \rangle \langle 33 \rangle + 6 \langle 11 \rangle \langle 23 \rangle \langle 23 \rangle + 8 \langle 13 \rangle \langle 12 \rangle \langle 23 \rangle \\
&\sim \mathcal{P}^3 + 6\mathcal{P}^2 + 8\mathcal{P}. \tag{A.11}
\end{aligned}$$

Note that the process of finding the coefficient at $\mathcal{O}(p^k)$ is not as simple as just reading off the coefficient of $\mathcal{P}^k$ as written above, recall for example the term $\langle 11 \rangle \langle 22 \rangle \langle 33 \rangle \sim \mathcal{P}^3$ contributed for every possible case of order p^2. Therefore we need to multiply each coefficient by all the cases where it contributes. Let us define two vertices as "connected" (forming a "connection") if their indices are equal up to permutations. Additionally, we will define the number of "free" vertices as the number of vertices whose indices may take on independent values while still ensuring the term is non-zero. For instance, $\langle aa \rangle \langle bb \rangle$ has two free vertices, since vertices a and b can both take on independent indices. On the other hand, $\langle ab \rangle \langle ab \rangle$ is only non-zero if a and b are connected, meaning there is only one free vertex (say a, since b must have its vertices equal up to permutation of a).

For the third-order ladder diagram, there must therefore be one additional connection (involving two vertices) for the leading term to contribute at order p^2. The term $\langle 11 \rangle \langle 22 \rangle \langle 33 \rangle$ has three free vertices (1,2, and 3), of which we choose two to be equal at order p^2, hence its contribution will be multiplied by $\binom{3}{2}$. This is not the case for the term $\langle 11 \rangle \langle 23 \rangle \langle 23 \rangle$; vertices 2 and 3 already must be equal for the term to be non-zero.

The general counting procedure will therefore be as follows: create a table with the left column denoting the order of p, and the top row corresponding to the expansion of the average of the couplings in terms of $\mathcal{P}$. For order p^k, take the coefficient of $\mathcal{P}^k$ and all higher powers of $\mathcal{P}$ and multiply them by the number of combinations required to form

enough connections to be of order p^k. For instance at order p, $\langle 11 \rangle \langle 22 \rangle \langle 33 \rangle$ must undergo 2 connections involving 3 free indices, so that means out of 3 free vertices choose 3 to be connected, which is $\binom{3}{3} = 1$. Similarly for $\langle 11 \rangle \langle 23 \rangle \langle 23 \rangle$ there are 2 free vertices (2 or 3 must be equal to the other to be non-zero, removing a free vertex) with one connection involving 2 vertices, implying a factor of $\binom{2}{2} = 1$. In general this pattern will hold. Another reason to see why their combinatoric factors should be 1 is that at order p, all vertices are equal so there is only one way to combine them all. So our table for the third order ladder diagram becomes

$\langle 112233 \rangle =$	$\mathcal{P}^3 + 6\mathcal{P}^2 + 8\mathcal{P}$	Total
p^3	1	1
p^2	$1 \binom{3}{2} + 6$	9
p	$1 \ + \ 6 \ + \ 8$	15

We will continue to 4th order to further demonstrate how these combinatoric factors work. At order p^2, the factor $\mathcal{P}^4$ undergoes two connections, however this time there are two cases. With four free vertices, three vertices could be all connected (making 2 total connections) or two sets of two connections could be formed. The latter of which has $\binom{4}{2}$ for the first choice and $\binom{2}{2} = 1$ for the second choice. Therefore the table becomes

$\langle 112233 \rangle =$	$\mathcal{P}^4 + 12\mathcal{P}^3 + 44\mathcal{P}^2 + 48\mathcal{P}$	Total
p^4	1	1
p^3	$\binom{4}{2} + 12$	18
p^2	$\binom{4}{3} + \binom{4}{2}\binom{2}{2} + 12\binom{3}{2} + 44$	90
p	$1 \ + \ 12 \ + \ 44 \ + \ 48$	105

For this procedure to work for arbitrary order of ladder diagram, we must find a closed form for the expansion of averaged couplings into $\mathcal{P}$'s as done above, and find a general method for calculating the combinatoric factors. These will both be addressed in the next section.

A.2 General Procedure

Let us first address the combinatoric factors. Recall based on our definition of free vertices that the monomial $\mathcal{P}^k$ has k free vertices. At order p^k no connections need be formed, so the term does not gain a combinatoric factor (although still has a coefficient multiplying $\mathcal{P}^k$). The term $\mathcal{P}^{k+1}$ has $k + 1$ free vertices, and must undergo 1 connection to be of order p^k. Out of the $k + 1$ free vertices, two must be chosen to form a connection, hence the combinatoric factor becomes $\binom{k+1}{2}$. For $\mathcal{P}^{k+2}$, there must be two connections between free vertices at order p^k. However we must choose how to partition these connections, they can either all be grouped together, meaning three free vertices are all permutations of one another (2 connections) or be separated into two different connections (1+1 connections). For any grouping of connections, there is always one more vertex than connection, which is what the combinatorics are performed on. So for the two connections grouped together,

the combinatoric factor will be $\binom{k+2}{2+1} = \binom{k+2}{3}$. For two separate connections, we do the combinatorics out of all the free vertices with one connection $\binom{k+2}{1+1}$ and then do combinatorics on the remaining free vertices for the next connection $\binom{k}{1+1}$ making a total combinatoric factor of $\binom{k+2}{2}\binom{k}{2}$.

As an example, let us calculate the order p^2 contribution at $n = 6$ order in ladder diagrams. Additionally, we will denote the nth order diagram by $a_n(\mathcal{P})$. It turns out at $n = 6$ this becomes

$$a_6(\mathcal{P}) \equiv \langle 112233445566 \rangle \sim \mathcal{P}^6 + 30\mathcal{P}^5 + 340\mathcal{P}^4 + 1800\mathcal{P}^3 + 4384\mathcal{P}^2 + 3840\mathcal{P}. \quad \text{(A.12)}$$

So at order p^2, we start with the combinatoric factors of $\mathcal{P}^6$. $6 - 2 = 4$ connections must be made. These connections can be partitioned into the following sets: $\{4\}$, $\{3, 1\}$, $\{2, 2\}$, $\{2, 1, 1\}$, $\{1, 1, 1, 1\}$. We add one to each of these sets to represent the possible combinatorics on the free vertices, so $\{5\}$, $\{4, 2\}$, $\{3, 3\}$, $\{3, 2, 2\}$, $\{2, 2, 2, 2\}$. We then need to find the combinatoric factors by taking the $j = 6$ free vertices and doing successive choose functions on each member of the set, this will give a factor of the form

$$\mathcal{P}^6 \text{ at order } p^2 : \quad \binom{6}{5} + \binom{6}{4}\binom{2}{2} + \binom{6}{3}\binom{3}{3} + \binom{6}{3}\binom{3}{2}\binom{1}{2} + \binom{6}{2}\binom{4}{2}\binom{2}{2}\binom{0}{2} \tag{A.13}$$

Notice that terms like $\binom{0}{2} = 0 = \binom{1}{2}$ make those coefficients vanish, so those terms do not contribute. These factors effectively mean it is not possible to construct the sufficient connections with that partition set. Repeating this process for all $\mathcal{P}$'s along with using the coefficients from (A.12) yields

$$\binom{6}{5} + \binom{6}{4}\binom{2}{2} + \binom{6}{3} + 30\left(\binom{5}{4} + \binom{5}{3}\right) + 340\left(\binom{4}{3} + \binom{4}{2}\right) + 1800\binom{3}{2} + 4384 = 13675 \tag{A.14}$$

Since $\langle 112233445566 \rangle$ contains 6 powers of $\langle J_{abcd}^2 \rangle \sim 1/p$, at order p^2 to total contribution from p's is $1/p^4 \sim 1/(kq)^4$, which matches our coefficient in the expansion from (3.11).

The general procedure for combinatoric factors should become clear from this. At order p^k and ladder diagram order n, the combinatoric factor for $\mathcal{P}^j$ is as follows: if $k > j$, then the factor is zero. If $k = j$, the factor is 1 (effectively zero choices of free vertices need to be made). If $j > k$, then $j - k \equiv m$ number of connections need to be made. The number of distinct combinatoric factors is equal to the number of partition numbers of m, since m different connections must be partitioned into groups. For any given group of connections, the number of vertices is one more than the number of connections in said group. Create a list corresponding to each possible partition of m (so the number of different lists will be the number of partitions of m), and add 1 to each member of the list to represent the number of free vertices. For a given list m_l, place each member $m_{ls} \in m_l$ into the lower value of a choose function, $\binom{\cdot}{m_{ls}}$, and multiply all of these choose functions together. The upper value will start with j (coming from the free vertices available from $\mathcal{P}^j$), and then decrease by the number of free vertices connected m_{l1}, which will be of the form $\binom{j}{m_{l1}}\binom{j-m_{l1}}{m_{l2}}\binom{j-m_{l1}-m_{l2}}{m_{l3}}\cdots$.

This process will repeat for all sets of partitions m_l, and then for all values j from $\mathcal{P}^j$ such that $j > k$. After multiplication by the corresponding coefficient of $\mathcal{P}^j$ and summing all terms together, this will form the total factor contributing to p^k.

The last step of our procedure is to find the expansion of $a_n(\mathcal{P})$ in terms of $\mathcal{P}$'s for arbitrary n. The first step is to form a recursion relation for $\langle 11 \cdots nn \rangle$ in terms of $\mathcal{P}$'s. There are 2 ways for vertices 1 and 2 to be paired, and $n-1$ different vertices that may be paired with vertex 1. Since we have the ability to relabel vertices without impacting the expansion of $\mathcal{P}$'s, we arrive at

$$a_n(\mathcal{P}) = \langle 11 \cdots nn \rangle = \langle 11 \rangle \underbrace{\langle 2233 \cdots nn \rangle}_{n-1} + 2(n-1)\langle 12 \rangle \langle 12 \underbrace{33 \cdots nn}_{n-2} \rangle, \qquad (A.15)$$

where the underbrace refers to the number of unique vertices (before considering permutations). Since $\langle 22 \cdots nn \rangle$ involves $n-1$ unique vertices which we are free to relabel, it is therefore equivalent to $a_{n-1}(\mathcal{P})$. Additionally $\langle 11 \rangle = \mathcal{P}$, which leaves the second term. We can choose to define the second term as $b_k(\mathcal{P})$ and find its recursion relation by again expanding and relabelling

$$b_k(\mathcal{P}) \equiv \langle ab \rangle \langle ab \underbrace{11 \cdots kk}_{k} \rangle = \langle ab \rangle \langle ab \rangle \langle 11 \cdots kk \rangle + 2k \langle ab \rangle \langle a1 \rangle \langle b1 \underbrace{22 \cdots kk}_{k-1} \rangle. \qquad (A.16)$$

Using (A.9) and our above expressions we can rewrite our recursion equations as a set of two coupled recursion relations

$$a_n(\mathcal{P}) = \mathcal{P} a_{n-1}(\mathcal{P}) + 2(n-1)b_{n-2} \qquad (A.17)$$

$$b_k(\mathcal{P}) = \mathcal{P} a_k(\mathcal{P}) + 2k b_{k-1}(\mathcal{P}). \qquad (A.18)$$

These sets of equations can be solved to find

$$a_n(\mathcal{P}) = 2^{n-1} \mathcal{P} \frac{\Gamma(\frac{\mathcal{P}}{2} + n)}{\Gamma(1 + \frac{\mathcal{P}}{2})}. \qquad (A.19)$$

While this solution may at first glance seem complicated, an expansion for the first few n's makes the pattern apparent:

$$a_1(\mathcal{P}) = \langle 11 \rangle = \mathcal{P} \qquad (A.20)$$

$$a_2(\mathcal{P}) = \langle 1122 \rangle = \mathcal{P}(\mathcal{P} + 2) \qquad (A.21)$$

$$a_3(\mathcal{P}) = \langle 112233 \rangle = \mathcal{P}(\mathcal{P} + 2)(\mathcal{P} + 4) \qquad (A.22)$$

$$a_4(\mathcal{P}) = \langle 11223344 \rangle = \mathcal{P}(\mathcal{P} + 2)(\mathcal{P} + 4)(\mathcal{P} + 6). \qquad (A.23)$$

Our goal is therefore to find the coefficient of $\mathcal{P}^k$ for every $1 < k < n$ for the nth order ladder diagram. This is a solved problem for the polynomial $x(x+1)(x+2)\cdots(x+n-1)$, where the coefficient of x^k is given by the unsigned *Stirling numbers of the first kind*, and are denoted by $\left[{n \atop k}\right]$ [38]. Therefore the solution of $a_n(\mathcal{P})$ is given by

$$a_n(\mathcal{P}) = \sum_{k=1}^{n} c_k \mathcal{P}^k; \qquad c_k = 2^{n-k} \left[{n \atop k}\right]. \qquad (A.24)$$

Putting this all together, we arrive at the algorithm for calculating the nth order ladder diagram:

1. Find the expansion of $a_n(\mathcal{P}) = \sum_{k=1}^{n} c_k \mathcal{P}^k$ by using $c_k = 2^{n-k} \left[{n \atop k}\right]$

2. To find the coefficient of p^k (which corresponds to $\frac{1}{(kq)^{n-k}}$):

   (a) For every $\mathcal{P}^j$ such that $j \geq k$:

      i. Find the number of connections needed to be of order p^k, which is $j - k \equiv m$

      ii. Calculate the sets of partitions for m, and then add 1 to every element to represent the free vertices

      iii. For each set, place the elements into the lower values of multiplied choose functions, and fill the upper values with $j - \sum m_l$, where m_l correspond to the lower values of any choose function multiplying to the left.

      iv. Sum all these choose functions together, and then multiply by the coefficient c_j, call this number h_{jk}

   (b) Sum together all coefficients corresponding to p^k, which will be $\sum_{j \geq k}^{n} h_{jk}$, this is the coefficient of p^k, call this coefficient g_k

3. The nth order ladder diagram will have an expansion in terms of p's as $\sum_{z=1}^{n} g_k p^k$.

As a demonstration, we will calculate the first 4 coefficients of an nth order ladder diagram using these techniques:

$\langle 11 \cdots nn \rangle =$	$\mathcal{P}^n + 2\left[{n \atop n-1}\right]\mathcal{P}^{n-1} + 4\left[{n \atop n-2}\right]\mathcal{P}^{n-2} + 8\left[{n \atop n-3}\right]\mathcal{P}^{n-3} + \cdots$	Total
$p^n \sim 1$	1	1
$p^{n-1} \sim \frac{1}{(kq)}$	$\binom{n}{2} + \left[{n \atop n-1}\right]2$	$\frac{3}{2}n(n-1)$
$p^{n-2} \sim \frac{1}{(kq)^2}$	$\binom{n}{3} + \binom{n}{2}\binom{n-2}{2} + 2\left[{n \atop n-1}\right]\binom{n-1}{2} + 4\left[{n \atop n-2}\right]$	$\frac{5}{4}n(n-1)^2(n-2)$
$p^{n-3} \sim \frac{1}{(kq)^3}$	$\binom{n}{4} + \binom{n}{3}\binom{n-3}{2} + \binom{n}{2}\binom{n-2}{2}\binom{n-4}{2}$ $+2\left[{n \atop n-1}\right]\left(\binom{n-1}{3} + \binom{n-1}{2}\binom{n-3}{2}\right) + 4\left[{n \atop n-2}\right]\binom{n-2}{2} + 8\left[{n \atop n-3}\right]$	$\frac{1}{24}n(n-1)(n-2)(n-3)$ $\times(77 + n(19n - 69))$

B Computational specifications

All numerical simulations presented in this paper were done through the Texas Advanced Computing Center (TACC) on the Frontera supercluster [39]. The supercluster is comprised of 8,386 nodes (of which a maximum of 200 were ever used per simulation), with each node contains 56 cores. For ease of access, compute node specifics are transcribed here:

Intel Xeon Platinum 8280 ("Cascade Lake") Compute Node	
Total cores per CLX node:	56 cores on two sockets (28 cores/socket)
Hardware threads per core:	1
Clock rate:	2.7GHz nominal
RAM:	192GB (2933 MT/s) DDR4
Cache:	32KB L1 data cache per core;
	1MB L2 per core;
	38.5 MB L3 per socket.
	Each socket can cache up to 66.5 MB
Local storage:	144GB /tmp partition on a 240GB SSD.

Time Comparisons Average, $q = 4$ $\beta = 1$ (112 Samples)				
Size (N)	All-to-All	k=50	k=10	k=4
16	00:00:02	00:00:02	00:00:00	00:00:00
18	00:00:05	00:00:03	00:00:01	00:00:00
20	00:00:16	00:00:05	00:00:01	00:00:01
22	00:00:48	00:00:11	00:00:03	00:00:01
24	00:00:48	00:00:14	00:00:04	00:00:01
26	00:02:14	00:00:33	00:00:10	00:00:04
28	00:06:22	00:01:19	00:00:23	00:00:10
30	00:17:44	00:03:09	00:00:52	00:00:22
32	00:47:42	00:07:15	00:01:55	00:00:48
34	02:06:27	00:16:26	00:04:14	00:01:48
36	05:27:22	00:37:26	00:09:28	00:04:02

Total (16 to 36)	08:49:51	01:06:44	00:17:11	00:07:18
Time Factor	1	0.126	0.0324	0.0138

C Fixed hypergraph configuration & tests of robustness

The analytical tools used in section 4 can be applied to a fixed hypergraph configuration, results of which can be seen in figure 8. For this set of figures the mean, standard deviation, and R^2 must be separated as their outputs lie in vastly different ranges. The lack of convergence and small value of R^2 demonstrate that a given fixed hypergraph exhibits complete disagreement with the symmetry based methods. These results remain true for various choices of hypergraph seeds, k, and D_g.

As a check that the results of this paper are robust to the various choices of analysis used in the main body of this work, we provide here alternative methodology to demonstrate our conclusions remain consistent. The first encountered choice which could make a significant change is which value $\mathcal{F}^*$ is chosen in the analysis. The main body of the paper utilized $\mathcal{F}^* = 0.4$, however in figures 9, 10 we demonstrate that the behavior exhibited in figure 5 is consistent with $\mathcal{F}^* = 0.3$ and $\mathcal{F}^* = 0.5$ respectively.

The next natural choice of variation is the size of the groupings statistically analyzed for each D_g, as it may be argued the proper choice would be to have a consistent number of samples regardless of D_g. Hence, we provide plots for two different choices, the first being 10 samples (figure 11) and the second being 20 samples (figure 12). These plots are compared to figure 5 from the main body, they exhibit a higher degree of fluctuation for smaller values of D_g. This result is expected, as lower values of D_g (meaning a given Lyapunov exponent is extracted with low number of realizations) will inherently have higher error than for higher values of D_g. This inherent instability makes Lyapunov extraction more inconsistent, hence the choice of extraction used in the main body. However even given these choices, our main findings (namely $R^2 \to 1$ can be achieved at low k with sufficiently high D_g) still hold.

Lastly, we analyze whether including the same lower bound $N = 16$ regardless of the maximum simulated N is a good choice. Instead, we could keep a set number of N in a given range (in this case, 11 different N's) leading to a change in the lower bounds. Hence the ranges compared are $N = 16$ to $N = 36$, $N = 18$ to $N = 38$, $N = 20$ to $N = 40$, and $N = 22$ to $N = 42$, which can be found in figure 13. As compared to figure 7, these simulations actually exhibit worse correspondence with the symmetry based model, and exhibit higher amounts of random fluctuations. This behavior is consistent with our arguments from section 4.3; by disregarding data points from lower N values, the remaining data (higher N's) of the sparse models have higher discrepancy with the all-to-all model, hence requiring more averaging.
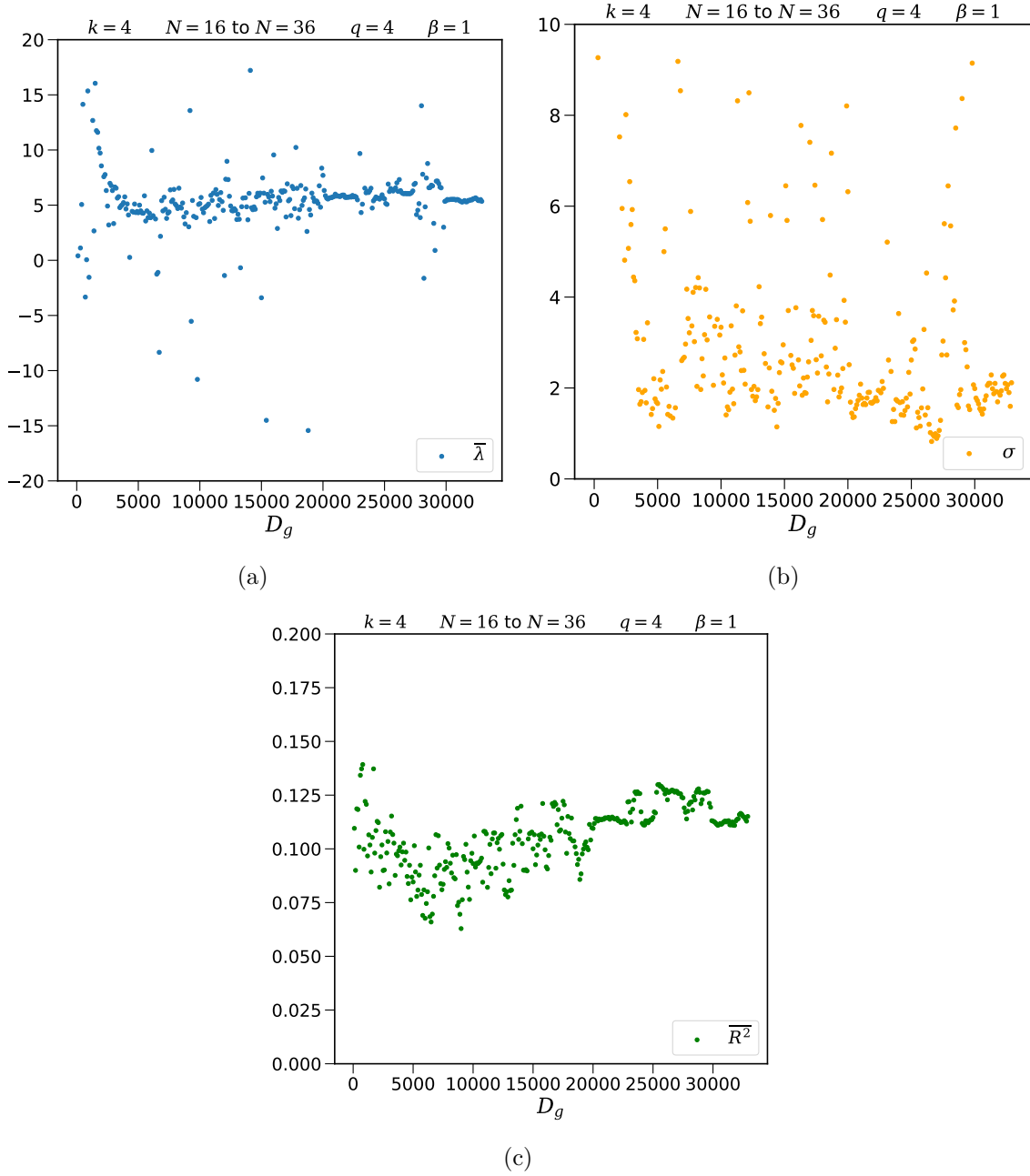
(a)

(b)

(c)

Figure 8. Data analysis of Lyapunov extraction for a fixed hypergraph seed (meaning for each N, the Hamiltonian interaction configuration is fixed for each disorder realization). Random variations persist through random states used in each disorder realization. Extraction was performed on a sparse model with $k = 4$ between $N = 16$ to $N = 36$ with $\beta = 1$. All plots demonstrate high variance and lack of convergence, suggesting the numerical derivative method of extraction is unsuitable for a fixed hypergraph seed. Each D_g is averaged over 10 times. (a) The mean fitted Lyapunov exponent $\overline{\lambda_L}$ versus D_g (b) The standard deviation of the fitted λ_L's, σ, versus D_g (c) The mean R^2 value over each set of λ_L's. The very low values of R^2 suggest a very poor fit to the data.
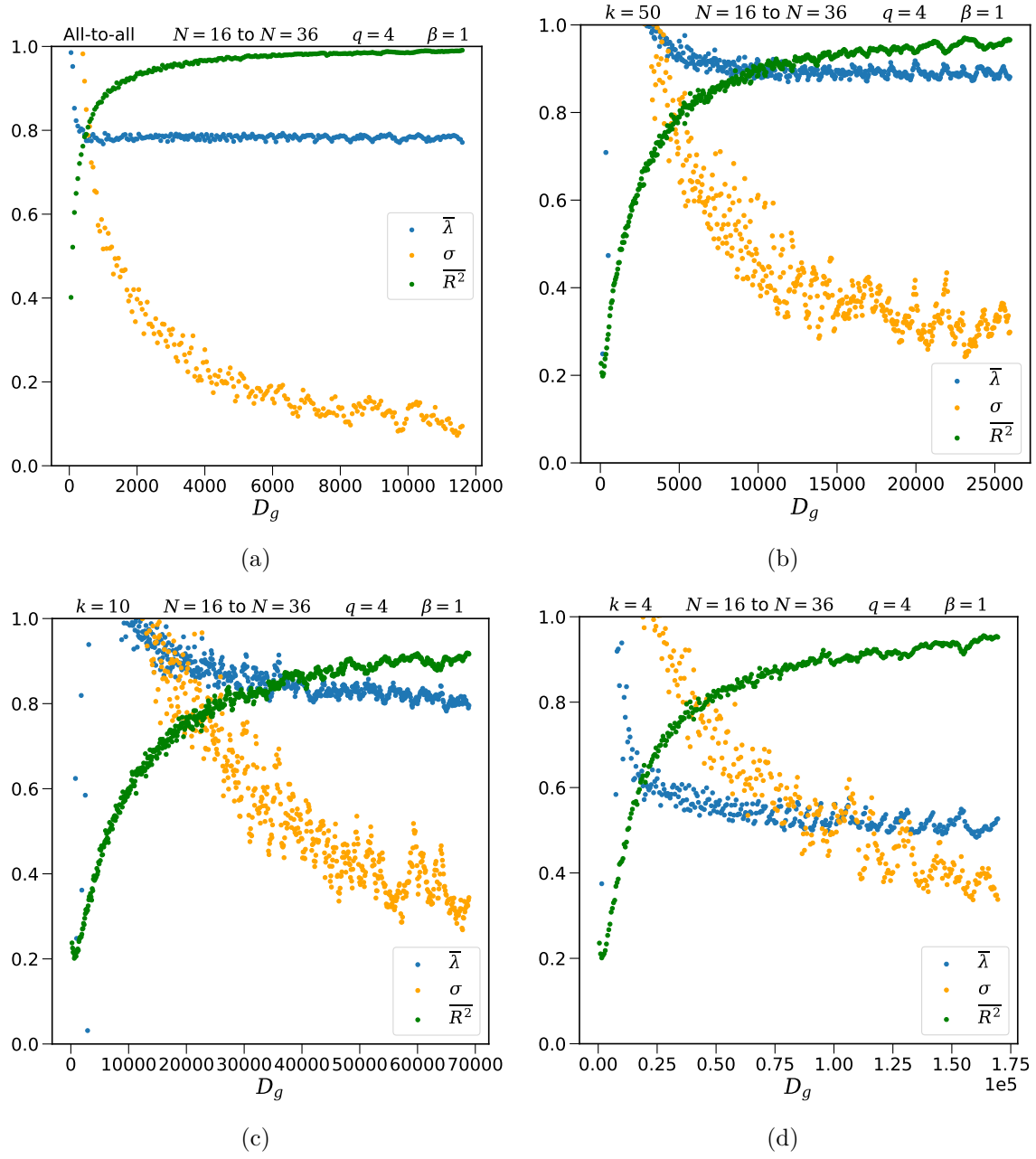
Figure 9. A replication of figure 5 (see corresponding caption for plot details) with a single parameter change: $\mathcal{F}^* = 0.3$. Convergence and behavior is consistent with figure 5 across this range of D_g's.
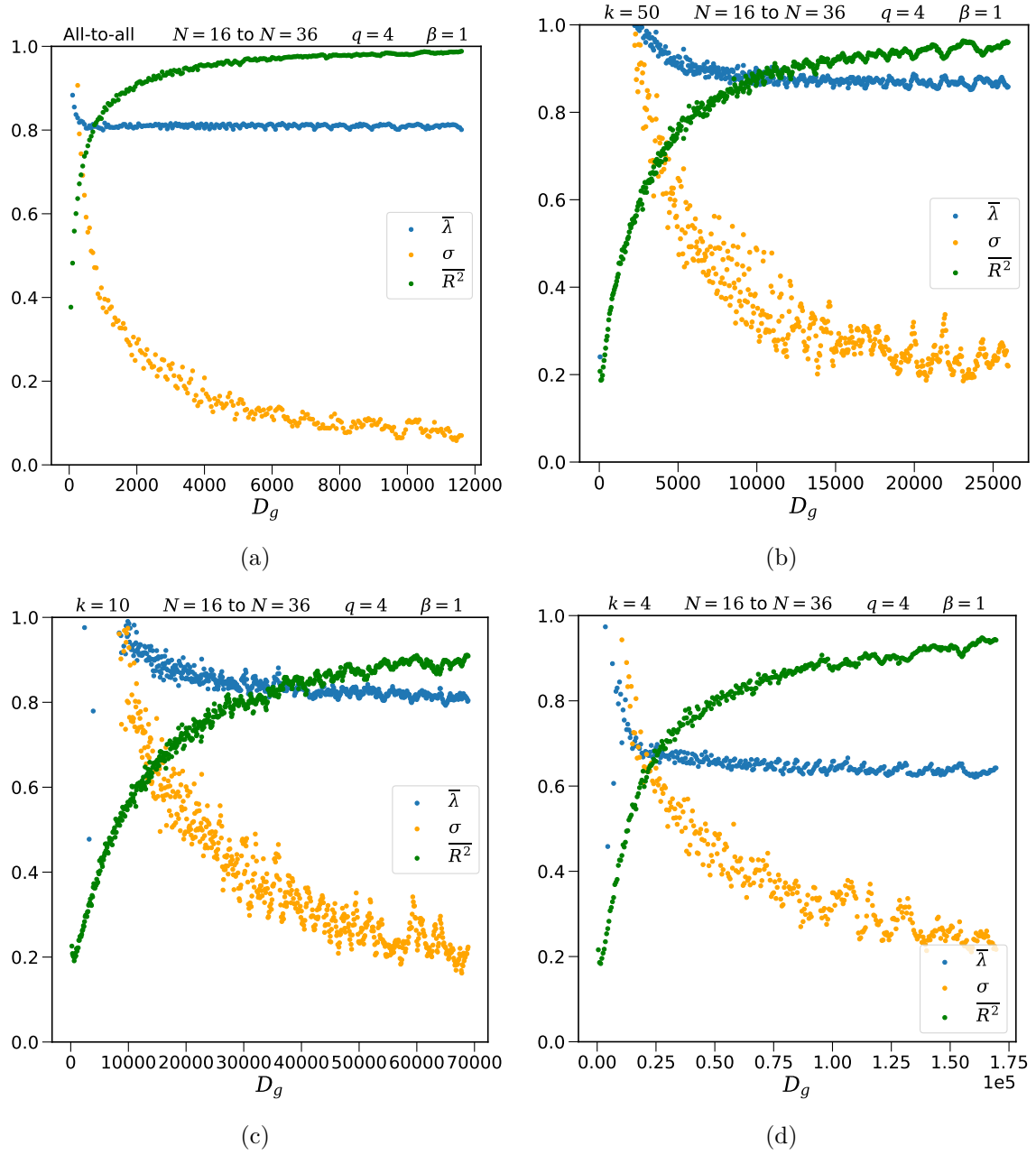
Figure 10. A replication of figure 5 (see corresponding caption for plot details) with a single parameter change: $\mathcal{F}^* = 0.5$. Convergence and behavior is consistent with figure 5 across this range of D_g's.
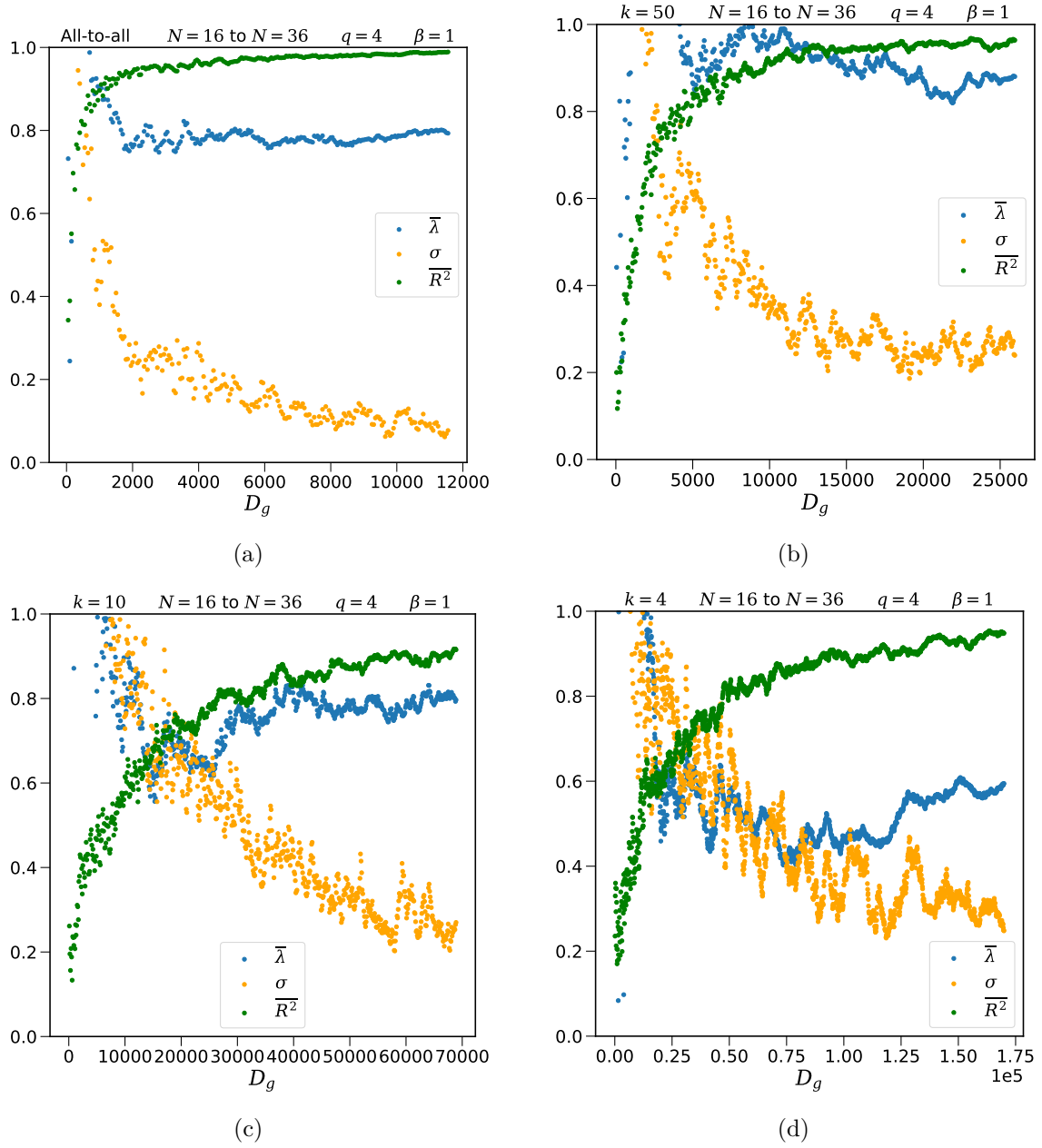
Figure 11. Many realizations are subdivided into 10 samples with D_g number of realizations, from which is extracted a Lyapunov exponent and R^2. For each choice of D_g, Lyapunov exponents are averaged over the 10 samples (*blue*) with standard deviation (*orange*). The R^2 factor for each D_g are also averaged (*green*). Resolution of plots all have $\Delta D_g = 50$. **(a)** all-to-all model. Exhibits convergence for plotted quantities; $R^2 \to 1$ shows agreement with the symmetry. **(b)** $k = 50$, R^2 exhibits some divergence from (a), larger fluctuations. Error appears to asymptote at a higher value compared to (a). **(c,d)** k=10,4 respectively. Random fluctuations have much larger effect, very high values of D_g must used to perform consistent Lyapunov extraction.
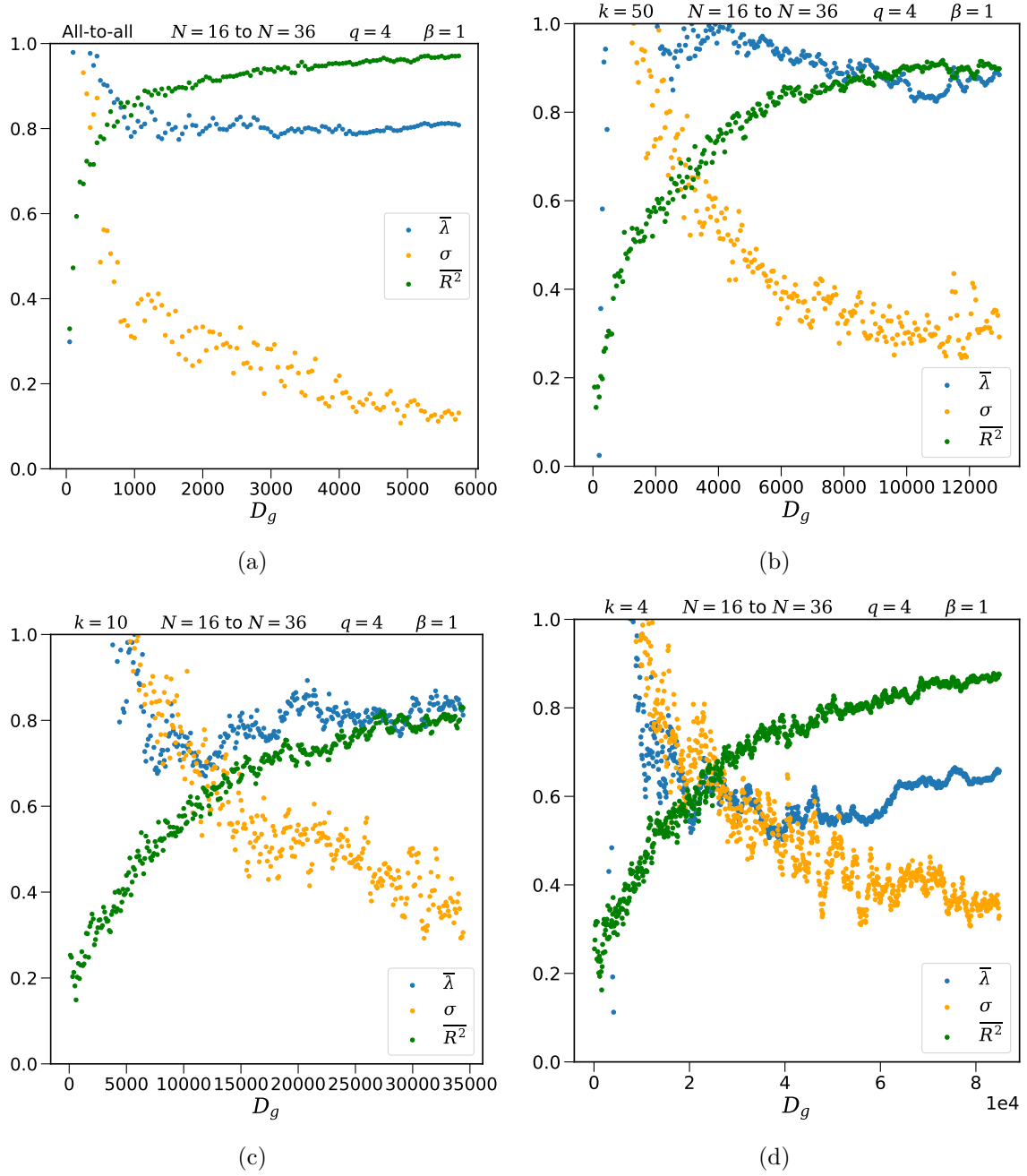
Figure 12. A replication of figure 11 (see corresponding caption for plot details) with a change of grouping sizes for statistics, namely 20 simulations are averaged over for each D_g. This reduces the maximum D_g (and hence the range) by $\frac{1}{2}$. As a result, higher values of R^2 are unable to be attained for the sparse models (i.e. for (b), (c), (d)), requiring probing of higher maximum values of D_g. Within the ranges present, the behavior is consistent with figure 11.
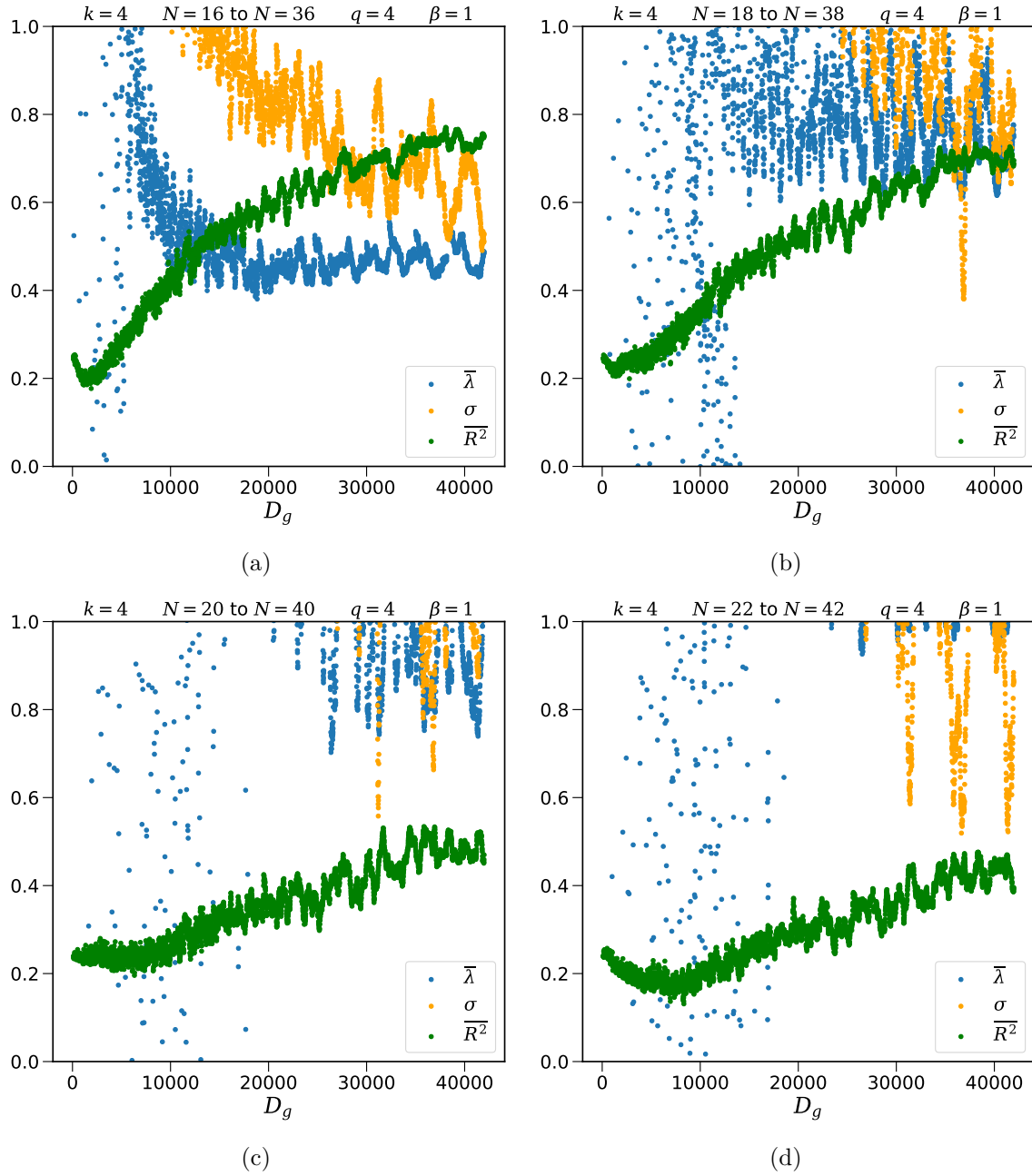
Figure 13. A replication of figure 7 (see corresponding caption for plot details) with changes in the lower bounds of N used to extract the Lyapunov exponent, namely (a) $N = 16$ to $N = 36$ (unchanged) (b) $N = 18$ to $N = 38$ (c) $N = 20$ to $N = 40$ (d) $N = 22$ to $N = 42$. Despite a majority of data overflowing the plot, the choice of ranges for the axes was maintained to demonstrate even further disagreement from figure 7.

References

[1] S.H. Shenker and D. Stanford, *Black holes and the butterfly effect*, *JHEP* **03** (2014) 067 [[arXiv:1306.0622](https://arxiv.org/abs/1306.0622)] [[INSPIRE](https://inspirehep.net/)].

[2] S.H. Shenker and D. Stanford, *Stringy effects in scrambling*, *JHEP* **05** (2015) 132 [[arXiv:1412.6087](https://arxiv.org/abs/1412.6087)] [[INSPIRE](https://inspirehep.net/)].

[3] J. Maldacena, S.H. Shenker and D. Stanford, *A bound on chaos*, *JHEP* **08** (2016) 106 [[arXiv:1503.01409](https://arxiv.org/abs/1503.01409)] [[INSPIRE](https://inspirehep.net/)].

[4] P. Hosur, X.-L. Qi, D.A. Roberts and B. Yoshida, *Chaos in quantum channels*, *JHEP* **02** (2016) 004 [[arXiv:1511.04021](https://arxiv.org/abs/1511.04021)] [[INSPIRE](https://inspirehep.net/)].

[5] D. Stanford, *Many-body chaos at weak coupling*, *JHEP* **10** (2016) 009 [[arXiv:1512.07687](https://arxiv.org/abs/1512.07687)] [[INSPIRE](https://inspirehep.net/)].

[6] K. Jensen, *Chaos in AdS$_2$ Holography*, *Phys. Rev. Lett.* **117** (2016) 111601 [[arXiv:1605.06098](https://arxiv.org/abs/1605.06098)] [[INSPIRE](https://inspirehep.net/)].

[7] J. Maldacena and D. Stanford, *Remarks on the Sachdev-Ye-Kitaev model*, *Phys. Rev. D* **94** (2016) 106002 [[arXiv:1604.07818](https://arxiv.org/abs/1604.07818)] [[INSPIRE](https://inspirehep.net/)].

[8] D.A. Roberts and B. Yoshida, *Chaos and complexity by design*, *JHEP* **04** (2017) 121 [[arXiv:1610.04903](https://arxiv.org/abs/1610.04903)] [[INSPIRE](https://inspirehep.net/)].

[9] B. Swingle, *Unscrambling the physics of out-of-time-order correlators*, *Nature Phys.* **14** (2018) 988 [[INSPIRE](https://inspirehep.net/)].

[10] A.M. García-García, B. Loureiro, A. Romero-Bermúdez and M. Tezuka, *Chaotic-Integrable Transition in the Sachdev-Ye-Kitaev Model*, *Phys. Rev. Lett.* **120** (2018) 241603 [[arXiv:1707.02197](https://arxiv.org/abs/1707.02197)] [[INSPIRE](https://inspirehep.net/)].

[11] Y. Gu and A. Kitaev, *On the relation between the magnitude and exponent of OTOCs*, *JHEP* **02** (2019) 075 [[arXiv:1812.00120](https://arxiv.org/abs/1812.00120)] [[INSPIRE](https://inspirehep.net/)].

[12] W. Fischler, T. Guglielmo and P. Nguyen, *Quantum chaos in a weakly-coupled field theory with nonlocality*, *JHEP* **09** (2022) 097 [[arXiv:2111.10895](https://arxiv.org/abs/2111.10895)] [[INSPIRE](https://inspirehep.net/)].

[13] S. Xu and B. Swingle, *Scrambling Dynamics and Out-of-Time Ordered Correlators in Quantum Many-Body Systems: a Tutorial*, [arXiv:2202.07060](https://arxiv.org/abs/2202.07060) [[INSPIRE](https://inspirehep.net/)].

[14] I. García-Mata, R.A. Jalabert and D.A. Wisniacki, *Out-of-time-order correlators and quantum chaos*, *Scholarpedia* **18** (2023) 55237 [[arXiv:2209.07965](https://arxiv.org/abs/2209.07965)] [[INSPIRE](https://inspirehep.net/)].

[15] A.I. Larkin and Y.N. Ovchinnikov, *Quasiclassical Method in the Theory of Superconductivity*, *JETP* **28** (1969) 1200.

[16] Y. Sekino and L. Susskind, *Fast Scramblers*, *JHEP* **10** (2008) 065 [[arXiv:0808.2096](https://arxiv.org/abs/0808.2096)] [[INSPIRE](https://inspirehep.net/)].

[17] A. Kitaev, *A simple model of quantum holography (part 1)*, talk at *KITP*, April 7, 2015, http://online.kitp.ucsb.edu/online/entangled15/kitaev/.

[18] A. Kitaev, *A simple model of quantum holography (part 2)*, talk at *KITP*, May 27, 2015, http://online.kitp.ucsb.edu/online/entangled15/kitaev2/.

[19] A. Almheiri and J. Polchinski, *Models of AdS$_2$ backreaction and holography*, *JHEP* **11** (2015) 014 [arXiv:1402.6334] [INSPIRE].

[20] S. Xu, L. Susskind, Y. Su and B. Swingle, *A Sparse Model of Quantum Holography*, arXiv:2008.02303 [INSPIRE].

[21] A.M. García-García, Y. Jia, D. Rosa and J.J.M. Verbaarschot, *Sparse Sachdev-Ye-Kitaev model, quantum chaos and gravity duals*, *Phys. Rev. D* **103** (2021) 106002 [arXiv:2007.13837] [INSPIRE].

[22] E. Cáceres, A. Misobuchi and A. Raz, *Spectral form factor in sparse SYK models*, *JHEP* **08** (2022) 236 [arXiv:2204.07194] [INSPIRE].

[23] E. Cáceres, A. Misobuchi and R. Pimentel, *Sparse SYK and traversable wormholes*, *JHEP* **11** (2021) 015 [arXiv:2108.08808] [INSPIRE].

[24] D. Jafferis et al., *Traversable wormhole dynamics on a quantum processor*, *Nature* **612** (2022) 51 [INSPIRE].

[25] B. Kobrin et al., *Many-Body Chaos in the Sachdev-Ye-Kitaev Model*, *Phys. Rev. Lett.* **126** (2021) 030602 [arXiv:2002.05725] [INSPIRE].

[26] M. Tezuka et al., *Binary-coupling sparse Sachdev-Ye-Kitaev model: An improved model of quantum chaos and holography*, *Phys. Rev. B* **107** (2023) L081103 [arXiv:2208.12098] [INSPIRE].

[27] I. Dumitriu and Y. Zhu, *Spectra of Random Regular Hypergraphs*, *Electron. J. Comb.* **28** (2021) P3.36.

[28] N.Y. Yao et al., *Interferometric Approach to Probing Fast Scrambling*, arXiv:1607.01801 [INSPIRE].

[29] É. Lantagne-Hurtubise, S. Plugge, O. Can and M. Franz, *Diagnosing quantum chaos in many-body systems using entanglement as a resource*, *Phys. Rev. Res.* **2** (2020) 013254 [arXiv:1907.01628] [INSPIRE].

[30] H. Shen, P. Zhang, R. Fan and H. Zhai, *Out-of-Time-Order Correlation at a Quantum Phase Transition*, *Phys. Rev. B* **96** (2017) 054503 [arXiv:1608.02438] [INSPIRE].

[31] A. Keleş, E. Zhao and W.V. Liu, *Scrambling dynamics and many-body chaos in a random dipolar spin model*, *Phys. Rev. A* **99** (2019) 053620 [arXiv:1810.03815] [INSPIRE].

[32] G.D. Kahanamoku-Meyer and J. Wei, *GregDMeyer/dynamite: v0.3.1*, (2023), DOI:10.5281/ZENODO.7706785.

[33] T.J. Park and J.C. Light, *Unitary quantum time evolution by iterative Lanczos reduction*, *J. Chem. Phys.* **85** (1986) 5870.

[34] S. Goldstein, J.L. Lebowitz, R. Tumulka and N. Zanghi, *Canonical Typicality*, *Phys. Rev. Lett.* **96** (2006) 050403 [cond-mat/0511091] [INSPIRE].

[35] D.J. Luitz and Y.B. Lev, *The ergodic side of the many-body localization transition*, *Annalen Phys.* **529** (2017) 1600350 [arXiv:1610.08993] [INSPIRE].

[36] A.M. García-García and J.J.M. Verbaarschot, *Spectral and thermodynamic properties of the Sachdev-Ye-Kitaev model*, *Phys. Rev. D* **94** (2016) 126010 [arXiv:1610.03816] [INSPIRE].

[37] A.S. Shankar, M. Fremling, S. Plugge and L. Fritz, *Lyapunov exponents in a Sachdev-Ye-Kitaev-type model with population imbalance in the conformal limit and beyond*, arXiv:2302.08876 [INSPIRE].

[38] R.L. Graham, D.E. Knuth and O. Patashnik, *Concrete Mathematics: A Foundation for Computer Science*, Addison-Wesley, Reading (1989).

[39] D. Stanzione et al., *Frontera: The Evolution of Leadership Computing at the National Science Foundation*, in the proceedings of the *PEARC '20: Practice and Experience in Advanced Research Computing*, New York, NY, U.S.A., Association for Computing Machinery (2020), p. 106–111 [DOI:10.1145/3311790.3396656].