Random sparse generators of Markovian evolution and their spectral properties

Goran Nakerst,[1] Sergey Denisov,[2, 3] and Masudul Haque[1, 4]

[1]*Institut für Theoretische Physik, Technische Universität Dresden, D-01062 Dresden, Germany*
[2]*NordSTAR - Nordic Center for Sustainable and Trustworthy AI Research, Pilestredet 52, N-0166, Oslo, Norway*
[3]*Department of Computer Science, Oslo Metropolitan University, N-0130 Oslo, Norway*
[4]*Max-Planck-Institut für Physik komplexer Systeme, D-01187 Dresden, Germany*

The evolution of a complex multi-state system is often interpreted as a continuous-time Markovian process. To model the relaxation dynamics of such systems, we introduce an ensemble of random sparse matrices which can be used as generators of Markovian evolution. The sparsity is controlled by a parameter φ, which is the number of non-zero elements per row and column in the generator matrix. Thus, a member of the ensemble is characterized by the Laplacian of a directed regular graph with D vertices (number of system states) and $2\varphi D$ edges with randomly distributed weights. We study the effects of sparsity on the spectrum of the generator. Sparsity is shown to close the large spectral gap that is characteristic of non-sparse random generators. We show that the first moment of the eigenvalue distribution scales as $\sim \varphi$, while its variance is $\sim \sqrt{\varphi}$. By using extreme value theory, we demonstrate how the shape of the spectral edges is determined by the tails of the corresponding weight distributions, and clarify the behavior of the spectral gap as a function of D. Finally, we analyze complex spacing ratio statistics of ultra-sparse generators, $\varphi = $ const, and find that starting already at $\varphi \geqslant 2$, spectra of the generators exhibit universal properties typical of Ginibre's Orthogonal Ensemble.

I. INTRODUCTION

Continuous-time Markov chains (CTMCs) [1] provide a popular framework to model dynamics of multi-state systems in diverse fields ranging from physics, chemistry, and biology [2–4] to economics [5, 6] and game theory [7, 8]. CTMCs are used to model chemical reactions [9–15], gene regulation processes [16–20], quantum dynamics (approximated by rate equations) [21–25], evolutionary game dynamics [8, 26, 27], and many other processes. CTMCs are also the key element of such celebrated models of statistical physics as contact processes [28–30], zero-range processes [31, 32] and exclusion processes like ASEP [2, 31, 33–39]. In some fields, CTMCs are known under the names 'classical Markovian master equations' or 'rate equations'.

A continuous-time Markovian evolution in finite discrete space consisting of D states can be specified with a transition rate matrix $\mathcal{K}$ [1], which is a generator of Markovian evolution. (It is called 'Kolmogorov operator' in Ref. [40]). The equation governing the evolution of a probability vector $P(t)$, defined on the state space,

$$\frac{d}{dt}P(t) = \mathcal{K}P(t), \tag{1}$$

has the formal solution, $P(t) = \exp(t\mathcal{K})P_0$, where $P_0 = P(0)$ is the initial probability vector. The evolution of $P(t)$ is thus fully determined by the generator $\mathcal{K}$, especially by its spectral properties. The fact that the operator $\mathcal{M}_t = \exp(t\mathcal{K})$ should map a non-negative vector onto another non-negative vector while preserving ℓ_1-norm, means that $\mathcal{K}$ satisfies a set of constraints and these constraints have an effect on its spectral properties [1].

In order to model the evolution of a complex system with CTMCs, we would have to first design a specific Kolmogorov operator. Taking into account the large variety of existing models, it would be beneficial to figure out universal properties of $\mathcal{K}$-generators, i.e., properties that are typical rather than specific to a particular model. The first step in this direction is to define *random* ensembles of generators. A similar situation arises in the case of unitary time-continuous evolution, where the corresponding generators (quantum Hamiltonians) were explored and classified by using the powerful toolbox of random matrix theory (RMT) [41–45]. Implementation of this idea resulted in the creation of Quantum Chaos theory [46–48] which made - and is still making – a strong impact on many-body quantum physics, both theoretical [49] and experimental (see, e.g., Ref. [50]).

Recently, RMT-based approaches were developed to analyze spectral properties of random generators of open quantum (Lindblad operators) [51–54] and classical (Kolmogorov operators) [40, 55, 56] Markovian evolution. The considered generators, both quantum and classical, were on purpose constructed in a completely random way - up to the constraints that make them legitimate generators. In the case of Kolmogorov operators, this means that they are represented by dense matrices [40, 55]. It was shown that the spectral density of such operators represents a free sum of a uniform disc and a Gaussian distribution which results in a distinctive spindle-like shape [40], as shown in Figure 2 (a). This density is universal, in the sense that the particular way the operators are sampled does not affect the shape of the spindle (but may affect its position on the real axis and its overall scaling) [40].

In contrast to the random Kolmogorov operators, for most applications and known models, the corresponding $\mathcal{K}$-generators are represented by *sparse* matrices. This is a consequence of locality and other topological constraints imposed on the allowed transitions in the state spaces of the models. For many-component or many-

particle systems, elements of the generator matrix typically represent changing multiple (or all) components of the system simultaneously, e.g., for an exclusion process, a generic matrix element could represent correlated hopping of many particles. Since such processes are usually absent in physically (biologically, economically,...)-motivated models, most elements of the corresponding $\mathcal{K}$-matrices are zeros.

Sparsity affects the spectra, $\{\lambda_i\}$, $i = 1, 2, ..., D$, of the corresponding generators. Most noticeably, the spectral gap, i.e., the distance between $\lambda_1 = 0$ and the eigenvalue closest to it, $\gamma_* = \min\{|\operatorname{Re}\lambda_i|\}$, does not grow with the increase of number of states D. This is in sharp contrast to the case of dense random generators [51, 53, 54] . In Figure 1 we contrast the case of dense random stochastic matrices (a) with various model systems described by sparse Markov generators (b-f).

The large gap of dense random generators implies that even the slowest decaying mode of a generic initial probability vector converges rapidly to the equilibrium state, the relaxation time (inverse of the spectral gap) decreasing inversely in the state space size D. In contrast, physical generators of CTMCs in general have spectral gaps and relaxation times functionally depending on D very differently than (anti-)linearly; see, e.g. [57–60] for the example of the exclusion process. This difference of behavior suggests that it is more suitable to model physical CTMC generators by sparse rather than dense random matrices.

Our motivation is to refine the RMT approach to random Kolmogorov operators by including sparsity which is characteristic to physically relevant $\mathcal{K}$-generators. We specify an ensemble of random matrices of fixed sparsity φ as an ensemble of negative combinatorial Laplacians of random regular directed graphs. The sparsity is controlled by the vertex degree φ which is equal to the number of non-zero elements per row and column of the generator matrix. In graph terms, this means that each vertex has exactly φ incoming and φ outgoing edges. The non-zero elements (edge weights) are taken to be random, positive, independent, and identically distributed (iid).

A similar setup was studied in Ref. [56], where an ensemble of oriented Erdős-Rényi graphs [62], parameterized with edge probability distribution $p(D)$, was used. The vertex distribution in this case is binomial-distributed [62], and not constant as in our case. However, one might expect similar behavior in the $D \to \infty$ limit with the correspondence $p(D) = \varphi/D$. The authors of Ref. [56] considered the regime $Dp(D) \gg (\log D)^6$, which they found to have the same universal properties as in the non-sparse case. In this work, we consider sparsity beyond this limit, including specifically $\varphi \sim D^0$ (vertex degree not growing with D) and $\varphi \sim \log D$.

In this paper, we investigate the dependence of spectral properties of the sparse Kolmogorov operators on sparsity parameter φ, number of states (dimension of the state space) D, and on the edge weight distribution, i.e., on the distribution of the nonzero elements of matrix $\mathcal{K}$. Explicit results are mostly derived for χ_2^2 and uniform weight distributions; however, these results can be adapted to other weight distributions.

We consider both the bulk of the spectral distribution and its edges.

As for the bulk, we focus on its position μ (the mean of the corresponding eigenvalue distribution) and its variances along the real and imaginary axes (standard deviations of the distribution of the real and imaginary eigenvalue parts, respectively). The first variance estimates the spread of the relaxation rates, while the second one gives the timescales of the oscillations during the relaxation.

As for the edges, we address the spectral gap and the extent of the spectrum along the real axis (the real part of the eigenvalue with largest absolute real part). These determine respectively the slowest and fastest time scales of relaxation to the steady state. The spectral gap is of physical interest for multi-state Markov processes, see, e.g., [57–60, 63–67] for the ASEP and [68, 69] for the contact process. The horizontal extent, in addition to its interpretation as the fastest timescale for CTMCs, is also relevant in the graph theory interpretation, e.g., to quantify the computational complexity of the community detection problem [70, 71] and the max-cut [72, 73] problems.

We demonstrate that the position and variance of the spectral bulk of sparse Kolmogorov operators scale as $\sim \varphi$ and $\sim \sqrt{\varphi}$, respectively. These characteristics do not depend on D but on the first and second moments of the weight distribution. The dependence of the spectral edges on the weight distribution is less straightforward and highly non-universal. In particular, we show that, in the regime of high sparsity, $\varphi \ll D$, the spectral gap (horizontal extent) depends only on the left (right) tail of the weight distribution. We evaluate the dependence of the spectral gap on φ and D for weight distributions with exponential and power-law tails.

We consider the cases of χ_2^2 and uniform weight distributions in detail. For these distributions, we find that the spectral gap and the horizontal extent of the spectrum can be approximated by the largest and smallest diagonal entry of the generator matrix, respectively. Using the conjecture that this correspondence holds in general, we use extreme value theory (EVT) [74, 75] to analytically derive dependencies of the spectral edges on φ and D. In particular, we infer that the distributions of spectral edges only depend on the tails of the weight distributions.

Finally, we analyze correlations of the eigenvalues of sparse Kolmogorov operators. We show that, for $\varphi \geq 2$, the complex spacing ratio distributions [61] of the spectral bulks follow the distribution typical to Ginibre's Orthogonal Ensemble.

The paper is organized as follows. In Section II we introduce an ensemble of sparse random Kolmogorov operators. We analyze the bulk of the spectral distributions of the ensembles in Section III. In Sections IV and V we address the spectral gap and the horizontal extent of the
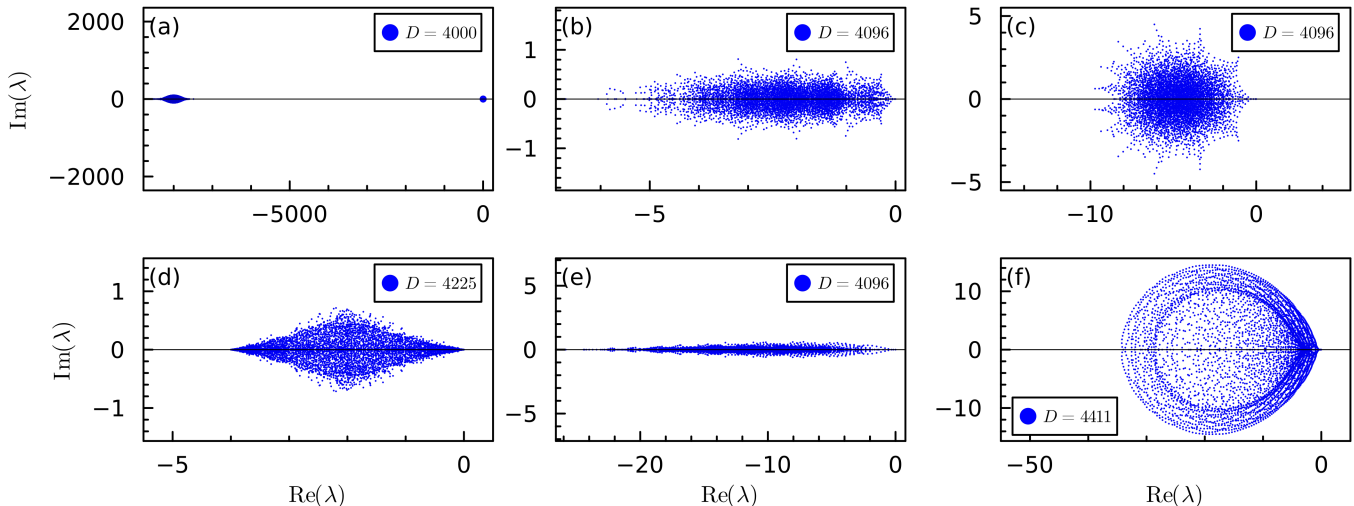
FIG. 1. Spectra of generators of Markovian evolution, Eq. (1). **(a)** Dense (non-sparse) random generator with χ_2^2 edge weight distribution, **(b)** totally asymmetric simple exclusion process (TASEP) on a ring with staggered hopping probabilities [61], **(c)** asymmetric simple exclusion process ASEP on a chain with open boundary conditions and next nearest neighbor terms, **(d)** a process of particle hopping on an open boundary grid with random hopping probabilities, **(e)** a contact process on a chain [28], **(f)** a gene transcription model from Ref. [20]. In each plot the real and imaginary axes have the same scale. The models are described in Appendix A.

spectrum, respectively. A discussion on correlations between eigenvalues in terms of the complex spacing ratio follows in Section VI. We conclude with a summary of our results in Section VII. Appendices contain information on the models whose spectra are presented in Figure 1, details of the sampling of sparse random Kolmogorov operators, and the details of analytical derivations.

II. RANDOM SPARSE KOLMOGOROV OPERATORS

In this section, we first recall some basic properties of Kolmogorov operators and review the case of full random $\mathcal{K}$-matrices. We then define an ensemble of random sparse operators. In what follows, matrices will be referred to by calligraphic letters (e.g., $\mathcal{K}$) while their elements will be denoted by non-calligraphic letters (e.g., K_{ij}).

A. Basic information

In order to be qualified as a Kolmogorov operator, a matrix $\mathcal{K}$ has to fulfill two conditions, (i) all its off-diagonal elements have to be real and non-negative, $K_{ij} \geq 0$, $i \neq j$, and (ii) the sum over every column should be zero. The latter is fulfilled by setting all the diagonal elements as

$$K_{ii} = -\sum_{j \neq i} K_{ji}. \tag{2}$$

The first condition guarantees the preservation of the non-negativity of a vector during the evolution induced by Eq. (1), while the second one guarantees the preservation of the ℓ_1-norm of the vector.

The spectrum of $\mathcal{K}$ is in general complex. Since $\mathcal{K}$ maps real vectors onto real vectors, the spectrum is invariant under complex conjugation, so all complex eigenvalues come in conjugated pairs. The spectrum contains at least one eigenvalue $\lambda_1 = 0$ with right eigenvector corresponding to the steady state. By virtue of the Perron-Frobenius theorem [76–78], the components of the steady state vector can be chosen to be non-negative, which makes it, after normalization, a probability vector. Any Kolmogorov operator can be represented in terms of a real non-negative matrix, $\mathcal{M}$, $M_{ij} \geq 0$,

$$\mathcal{K} = \mathcal{M} - \mathcal{J}, \tag{3}$$

where elements of the diagonal matrix $\mathcal{J}$ are $J_{jj} = \sum_i M_{ij}$.

We now briefly review the case of dense (non-sparse) random Kolmogorov operators [40, 55, 56]. Elements $M_{ij} > 0$ are i.i.d. sampled from a distribution with density $p(x)$ and first two moments $\mu_0 = \int x p(x) dx$ and $\sigma_0^2 = \int (x - \mu_0)^2 p(x) dx$. The particular choice of distribution does not play an essential role (provided that it is not very pathological). For example, we could sample a matrix $\mathcal{Z}$ from Ginibre's Unitary Ensemble (GinUE) and then square its elements, $M_{ij} = |Z_{ij}|^2$ [40]). The matrix $\mathcal{M}$ is then full in the sense that, with probability 1, all its elements are different from zero.

The elements of the matrix $\mathcal{M}$ are i.i.d., thus, in the asymptotic limit, its spectral density is a uniform disk

of radius $\sqrt{D}\sigma_0$, with the center at 0. In the dense limit, the elements of $\mathcal{J}$ are sums of D independent random variables, so its elements can be approximated with Gaussian-distributed random variables with mean $D\mu_0$ and variance $D\sigma_0^2$.

Following the RMT approach [40], the Kolmogorov operator in Eq. (2) can be modelled as

$$\mathcal{K}_{\mathrm{R}M} = -\mu_0 D \cdot \mathbb{1} + \sigma_0\sqrt{D}(\mathcal{G} - \mathcal{D}), \qquad (4)$$

where $\mathcal{G}$ is a member of Ginibre's Orthogonal Ensemble (GinOE) and $\mathcal{D}$ is a diagonal matrix. Elements of $\mathcal{G}$ and $\mathcal{D}$ are sampled from the normal distribution of zero mean and unit variance. Here $\sigma_0\sqrt{D} \cdot \mathcal{G}$ models $\mathcal{M}$ while $\mathcal{J}$ is approximated as $\mu_0 D \cdot \mathbb{1} + \sigma_0\sqrt{D} \cdot \mathcal{D}$.

The spectral density of the non-trivial part, $\mathcal{K}' = \mathcal{G} + \mathcal{D}$, is a free convolution of a disk and a Gaussian distribution along the real axis, which results in a spindle-like shape. Figure 2 (a) presents both the spectrum of a single random dense Kolmogorov operator and histogram obtained with 100 samples.

Alternatively, we can state that the spectral density of the rescaled generator

$$\mathcal{K}' = \frac{1}{\sigma\sqrt{D}}(\mathcal{K} + \mu_0 D \cdot \mathbb{1}) \qquad (5)$$

is expected, in the asymptotic limit, to be the D-independent spindle ("an additive Gaussian deformation of the circular law", according to Ref. [56]).

The spectrum of the random non-sparse generator has a large gap which scales as D, as seen in Figures 1 and 2. We will see that this feature is strongly affected when we introduce sparsity.

B. Ensemble of sparse random Kolmogorov operators as a set of oriented graphs

The operator $\mathcal{K}$ described in the introduction can be considered as the negative Laplacian of a random directed graph with positive, iid edge weights, without self-loops, and with fixed vertex degree equal to φ.

For example, the graph corresponding to the $\mathcal{K}$ generator of a process of a particle hopping on a d-dimensional hypercubic lattice with periodic boundary conditions and random hopping rates is a particular (to the nearest-neighbor connections) realization of the ensemble with $\varphi = 2d$. Figure 1 (d) shows an example spectrum for $d = 2$.

The regularity of the graphs ensures that, with probability $1 - O(D^{-\varphi-1})$, they are strongly connected as long as $\varphi \geq 2$ [79]. Strong connectivity is a good feature because it means that the matrix $\mathcal{K}$ is not of block-diagonal structure and the state space is not partitioned into disconnected subsets. As there is only one strongly connected component, there is only one absorbing component. This implies that the multiplicity of the zero eigenvalue is one and the steady state is unique. Finally,

every state in the state space is reachable from every other state. The steady state, therefore, has all states populated.

Some physical models motivating this study presented in Figure 1 are - except for the contact process - all strongly connected. The contact process is only effectively strongly connected. It has two strongly connected components, where one is a single vertex and the other includes the remaining $D - 1$ vertices. The giant component is the only absorbing component and consequently, the steady state is unique.

We will focus on sparse generators with $\varphi \geq 2$ and will discuss $\varphi = 1$ in Section VI.

The physical models presented in Figure 1 motivate us to focus on two types of dependencies of φ on the matrix size D, namely $\varphi = \mathrm{const}$ and $\varphi \sim \log D$. For generators of single particle hopping models - an example is shown in Figure 1 (d) - the average number of non-zero elements per column and row is constant and independent of D. It increases logarithmically with D in many-body hopping models such as the ASEP or the contact process, Figure 1 (b), (c), and (e). There is no simple dependence of φ on D in the gene transcription model, Figure 1 (f), as the matrix size D is controlled by multiple parameters, see Appendix A.

What can we say about spectral densities of the ultra-sparse $\mathcal{K}$-generators, with $\varphi = \mathrm{const}$? A 'naive' adjustment of the RMT approach, which consists in describing the elements of a sparse matrix $\mathcal{M}$ with probability density function $\widetilde{p}(x) = (1 - \frac{\varphi}{D})\delta(x) + \frac{\varphi}{D}p(x)$, re-scaling the mean and variance accordingly, and then using the RMT model, Eq. (4), would not work here for two reasons. First, the spectral densities of such sparse matrices cannot be approximated with members of 'dense' RMT ensembles. Second, the Central Limit Theorem no longer applies and the entries of matrix $\mathcal{J}$ cannot be approximated with normal random variables (the entries become distribution-specific).

III. POSITION AND WIDTH OF THE BULK OF THE SPECTRUM

In this section, we analyze the dependence of the position and horizontal width of the bulk of the spectrum on the sparsity parameter φ and the matrix dimension D. We first provide (Subsections III A and III B) expressions and bounds for the position and the width, characterized respectively by the mean $\mu(\lambda)$ of all eigenvalues and the standard deviation $\sigma(\mathrm{Re}\,\lambda)$ of the real parts of the eigenvalues. These results are expressed in terms of the mean and standard deviation of the weight distribution (distribution of non-zero elements of the Kolmogorov operator $\mathcal{K}$), denoted by μ_0 and σ_0 respectively.

Since the most prominent effect of sparsity is to reduce the parametrically large gap seen in the full random case, it is instructive to analyze the ratio $\alpha = |\mu(\lambda)|/\sigma(\mathrm{Re}\,\lambda)$. This quantity provides insight into the distance of the
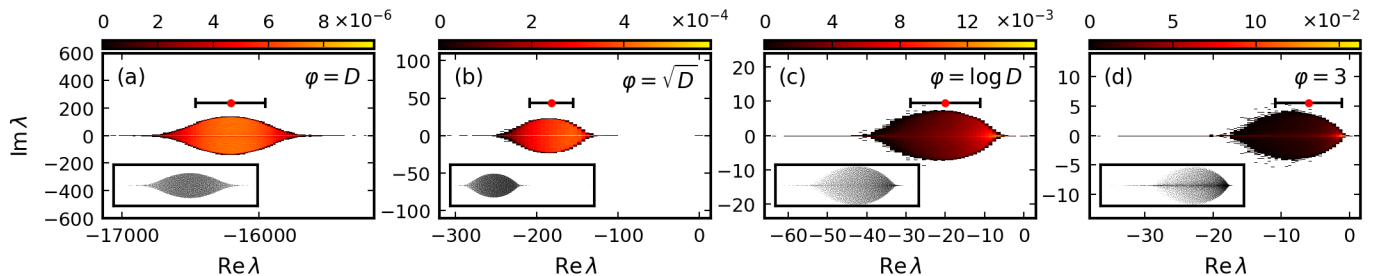
FIG. 2. Spectral densities of random Kolmogorov operators with χ_2^2 weight distribution. The matrix size is $D \approx 8000$ and the densities are estimated with 100 samples. White areas contain no eigenvalues. **(a)** Dense matrix without the zero eigenvalue, **(b)** sparse matrix with $\varphi = \sqrt{D}$ non-zero elements per row and column, **(c)** $\varphi = \log D$ and **(d)** $\varphi = 3$. The insets show spectra of single realizations. In each plot, the real and imaginary axes have the same scale. The red dots mark the location of $\mu(\lambda)$, given by Eq. (6), and the intervals shown in black are $[\mu(\lambda) - \sigma(\lambda), \mu(\lambda) + \sigma(\lambda)]$, where $\sigma(\lambda)$ is given by Eq. (8).

bulk of the spectrum from the origin, relative to the size of the bulk. Subsection III C is devoted to an analysis of the ratio α.

Numerical results presented in this section are obtained by sampling edge weights from the χ_2^2 and the standard uniform distribution.

The spectrum of dense generators ($\varphi = D - 1$) consists of two distinct parts - an eigenvalue $\lambda_1 = 0$ and the rest of the eigenvalues forming the spectral bulk away from the imaginary axis, as shown in Figure 1 (a) and Figure 2 (a). In contrast, the bulk of the spectrum is much closer to the imaginary axis for $\varphi \ll D$, as seen in Figure 2 for (b) $\varphi = \sqrt{D}$, (c) for $\varphi = \log D$ and (d) for $\varphi = 3$. For $\varphi = \sqrt{D}$, the bulk of the spectrum is visibly separated from the zero, as in the dense case. In fact, the spectral boundary is given by the same spindle (properly rescaled). Whether the spectral distribution is separated from zero for $\varphi = \log D$ and $\phi = 3$ is difficult to say with certainty from the available numerical data ($D \approx 8000$).

A. Position

The position of the spectral bulk of $\mathcal{K}$ can be identified with the mean $\mu(\lambda)$ of eigenvalues λ_i,

$$\mu(\lambda) = \left\langle \frac{1}{D} \sum_{i=1}^{D} \lambda_i \right\rangle, \tag{6}$$

where the average $\langle \dots \rangle$ is taken over the ensemble of random Kolmogorov operators described in Section II. Because the eigenvalues are either real or come in complex conjugate pairs, the mean of the spectral bulk is real, $\mu(\lambda) = \mu(\text{Re} \, \lambda)$.

A simple calculation, presented in Appendix C, shows that $\mu(\lambda)$ can be expressed as

$$\mu(\lambda) = \left\langle \frac{1}{D} \text{tr}(\mathcal{K}) \right\rangle = -\varphi \mu_0, \tag{7}$$

The averaging $\langle \dots \rangle$ over the matrix ensemble in Eq. (6) and Eq. (7) is, in principle, not needed since typicality

is expected, i.e., for large enough D, a single sample will display all the spectral features of the ensemble. This is because the quantity $\frac{1}{D} \text{tr}(\mathcal{K})$ is concentrated around its average $\left\langle \frac{1}{D} \text{tr}(\mathcal{K}) \right\rangle$ for increasing D, as shown in Appendix C.

For the four different dependencies of φ on D shown in Figure 2, Eq. (7) implies the following: For $\varphi = \text{const}$, the mean is independent of the matrix size D. For $\varphi = \log D$ ($\varphi = \sqrt{D}$) the mean decreases logarithmically with D (as $\sim \sqrt{D}$) and for $\varphi = D$ the mean decreases linearly with D as is expected for the dense generators [55].

In Figure 2, the location $\mu(\lambda)$ of generator matrices $\mathcal{K}$ is indicated with a red dot in each panel. The real part of the dot resides in the bulk of the spectrum for every dependence of φ on D shown in Figure 2.

B. Horizontal width

In Section III A we investigated where the bulk of the spectrum is located in the complex plane. We now analyze the width of the distribution. We are especially interested in the horizontal width.

We characterize the width of the bulk spectrum, both in the real and imaginary directions, $\text{Re} \, \lambda$ and $\text{Im} \, \lambda$, using the estimated variances

$$\sigma^2(\text{Re} \, \lambda) = \left\langle \frac{1}{D} \sum_{i=1}^{D} \left(\text{Re} \, \lambda_i - \frac{1}{D} \sum_{j=1}^{D} \lambda_j \right)^2 \right\rangle \tag{8}$$

$$\sigma^2(\text{Im} \, \lambda) = \left\langle \frac{1}{D} \sum_{i=1}^{D} (\text{Im} \, \lambda_i)^2 \right\rangle, \tag{9}$$

where we used the fact that $\sum_{j=1}^{D} \lambda_j$ is real.

Because the eigenvalues appear in complex conjugate pairs, $\sigma^2(\text{Re} \, \lambda)$ and $\sigma^2(\text{Im} \, \lambda)$ are related to the estimated

complex pseudo-variance via

$$\sigma^2(\lambda) = \left\langle \frac{1}{D} \sum_{i=1}^{D} \left(\lambda_i - \frac{1}{D} \sum_{j=1}^{D} \lambda_j \right)^2 \right\rangle$$

$$= \sigma^2(\operatorname{Re} \lambda) - \sigma^2(\operatorname{Im} \lambda). \qquad (10)$$

The estimated pseudo variance lower bounds the estimated variance of the real parts of the eigenvalues, $\sigma^2(\lambda) \leq \sigma^2(\operatorname{Re} \lambda)$.

The complex pseudo variance can be analytically calculated for the ensemble of random generator matrices as

$$\sigma^2(\lambda) = \left\langle \frac{1}{D} \operatorname{tr}(\mathcal{K}^2) \right\rangle - \left\langle \frac{1}{D^2} \operatorname{tr}(\mathcal{K})^2 \right\rangle$$

$$= \varphi \left(\sigma_0^2 + \frac{\varphi}{D} \mu_0^2 - \frac{1}{D} \sigma_0^2 \right). \qquad (11)$$

Details of the calculation are provided in Appendix C. The bound of the estimated real variance by the pseudo variance together with Eq. (11) leads to the asymptotic lower bound of $\sigma(\operatorname{Re} \lambda)$ in terms of the sparsity parameter φ. As $1 \leq \varphi \leq D$, the estimated horizontal width of the bulk spectrum cannot grow asymptotically slower than $\sqrt{\varphi}$,

$$\sigma(\operatorname{Re} \lambda) \gtrsim \sqrt{\varphi}. \qquad (12)$$

Numerically, we find that the bound in Eq. (12) is asymptotically sharp for $\varphi \ll D$, as shown in Figure 3 through the ratio α of mean $\mu(\operatorname{Re} \lambda)$ and width $\sigma(\operatorname{Re} \lambda)$. The collapse of the data points in Figure 3 (c) implies that $\sigma(\operatorname{Re} \lambda) \sim \sqrt{\varphi}$.

C. Ratio of mean and horizontal width

In this section, we combine the information of the location of the spectrum given by Eq. (6) and the horizontal width of the bulk given by Eq. (8) into the ratio

$$\alpha = \frac{|\mu(\operatorname{Re} \lambda)|}{\sigma(\operatorname{Re} \lambda)}. \qquad (13)$$

This quantifies how close the bulk spectrum is, relative to its size, to the stationary value $\lambda_1 = 0$. i.e., to the imaginary axis. For $\alpha = O(1)$ the estimated width of the bulk is of the same order as the estimated mean, thus the spectrum is located close to 0. For $\alpha \gg 1$ the estimated mean is much bigger than the horizontal width of the bulk and the bulk of the spectrum is far away from 0.

The analytical result for the estimated mean of the spectrum, Eq. (7), together with the asymptotic bound on the standard deviation of the real parts of the spectrum, Eq. (11), imply the following asymptotic bound on α

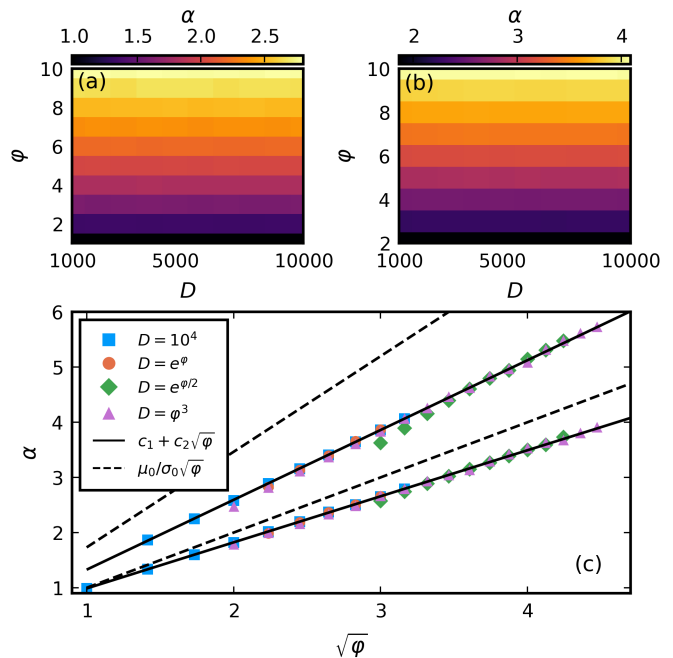$$\alpha \lesssim \sqrt{\varphi}. \qquad (14)$$

FIG. 3. Ratio α of mean $\mu(\operatorname{Re} \lambda)$ and horizontal width $\sigma(\operatorname{Re} \lambda)$ of the bulk of the spectrum of sparse random Kolmogorov operators with (a) χ_2^2 and (b) standard uniform weight distributions. (c) α as a function of $\sqrt{\varphi}$. The bottom markers correspond to χ_2^2 and the top to uniform distribution. Dependencies of φ on D are $\varphi \equiv$ constant, $\varphi = \log D$, $\varphi = 2 \log D$, and $\varphi = D^{1/3}$. The black solid lines correspond to $\alpha = c_1 + c_2 \sqrt{\varphi}$ ($c_{1,2}$ given in the main text) and the dashed lines denote $\alpha = \mu_0/\sigma_0 \sqrt{\varphi}$.

Numerically, we observe that the bound in Eq. (14) is asymptotically tight for $\varphi \ll D$, i.e.

$$\alpha \approx c_1 + c_2 \sqrt{\varphi}, \qquad (15)$$

for constants c_1 and c_2. Since $\mu(\lambda)$ scales linearly with φ, this behavior is consistent with $\sigma(\operatorname{Re} \lambda) \sim \sqrt{\varphi}$, stated previously. The constants are found to be $c_1 \approx 0.15$ (≈ 0.1) and $c_2 \approx 0.84$ (≈ 1.3) for the χ_2^2 (uniform) distribution.

Numerical results for α are summarized in Figure 3. For each combination of φ and D, α is averaged over n samples of random generators such that $nD = 50'000$. The weight distribution is the χ_2^2 distribution in (a) and in the lower part of (c), and is the uniform distribution in $[0, 1]$ in (b) and in the upper part of Figure 3 (c). We have found that these results are qualitatively the same for exponentially distributed edge weights.

In Figure 3 (a,b) we show the value of α as a function of D and φ. On the x-axis D varies in steps of 10^3 between 10^3 and 10^4. We observe that α increases with φ and is independent of D, as predicted by Eq. (15). In Figure 3 (c) we show α as a function of φ for different dependencies of φ on D. In all the cases, values of α collapse onto the black solid line given by Eq. (15).

For $\varphi \sim D$, the ratio α scales as $\sim \sqrt{D}$, thus recovering the parametrically large gap in the non-sparse case. For constant φ, the location of the bulk relative to its size is constant and independent of D, i.e, if measured relative to the size of the bulk, the bulk does not move away from the imaginary axis with increasing D. We have thus quantified how sparsity cures one of the less physical aspects of the non-sparse random model of Markov generators.

IV. SPECTRAL GAP

In this and the following section, we will consider the spectral edges, namely, the locations of the eigenvalues nearest and farthest from the imaginary axis. In this section we will investigate the spectral gap γ_* of $\mathcal{K}$,

$$\gamma_* = \min\{|\operatorname{Re}\lambda_i| : \operatorname{Re}\lambda_i < 0\}. \qquad (16)$$

The spectral gap γ_* is asymptotically, approximately bounded by the right extent of the bulk $|\mu(\lambda)| - \sigma(\lambda)$, which depends on φ as $\sim \varphi - \sqrt{\varphi} \sim \varphi$. So for constant φ, the spectral gap is bounded from above, while for φ increasing with D the spectral gap can increase with D.

Here the edge weights are distributed according to the χ_2^2 and the standard uniform distributions. We first demonstrate that, for $\varphi = $ const, the average spectral gap $\langle\gamma_*\rangle$ decreases as $D^{-1/\varphi}$, while $\langle\gamma_*\rangle$ is constant if φ increases logarithmically with D. We then show that the spectral gap is well approximated by the smallest (in magnitude) diagonal term of $\mathcal{J}(\mathcal{K})$ and use the theory of extreme values to underpin the numerical observations. The results are then generalized to weight distributions with power-law left tails in that for constant φ the average spectral gap decreases as a power-law in D and the crossover from decreasing to increasing $\langle\gamma_*\rangle$ happens when $\varphi \sim \log D$.

A. Numerical results

In Figure 4 we show the average spectral gap $\langle\gamma_*\rangle$ for edge weights distributed as χ_2^2 (a-c) and according to the standard uniform distribution (d-f). For every combination of φ and D, the average of the spectral gap is estimated with 100 samples. In Figure 4 (a) and (d) we show $\langle\gamma_*\rangle$ as a function of D for different dependencies of φ on D. The average spectral gaps for constant $\varphi = 3, 5, 8, 13$ (presented with colored circles) clearly follow a power-law scaling with D.

In Figure 4 (b) and (e) we show the average spectral gap $\langle\gamma_*\rangle$ as a function of φ and D. The black dashed lines are contour lines of constant $\langle\gamma_*\rangle$. They are near straight lines, showing that for a logarithmic increase of φ in D the spectral gap is constant.

We show the average spectral gap $\langle\gamma_*\rangle$ as a function of D for $\varphi = \frac{4}{5}\log D + 8$ in Figure 4 (a) and $\varphi = \frac{7}{10}\log D +$

8 in (d) as black diamonds. These dependencies of φ on D agree well with the top dashed contour lines in (b) and (e), respectively. The average spectral gap of φ depending logarithmically on D is constant in Figure 4 (a) and (d).

B. Gap $\approx$ minimum of $\mathcal{J}$

Let us assume for a moment that the generator matrix $\mathcal{K}$ is hermitian with eigenvalues $\lambda_D \leq \cdots \leq \lambda_2 < \lambda_1 = 0$. Then $\mathbb{1} = (1, \ldots, 1)^t$ is the eigenvector with eigenvalue 0 and all other eigenvectors are orthogonal to it. By the Courant-Fischer theorem [80]

$$\gamma_* = -\lambda_2 = \min_{|v|=1, v \perp \mathbb{1}} v^t(-\mathcal{K})v, \qquad (17)$$

where the minimum runs over all vectors $v \in \mathbb{R}^D$, which have Euclidean norm $|v| = 1$ and are perpendicular to $\mathbb{1}$. Choosing $1 \leq l \leq D$ arbitrary and v as (see Appendix D for more details)

$$v_i = \begin{cases} \sqrt{1 - \frac{1}{D}} & i = l \\ -\frac{1}{\sqrt{D(D-1)}} & i \neq l, \end{cases} \qquad (18)$$

a simple calculation shows that (at least for $\varphi \ll D$)

$$\gamma_* \leq \min_{1 \leq l \leq D} J_{ll} + O(D^{-1}). \qquad (19)$$

Similarly, by using the Courant-Fisher theorem, for the eigenvalue with largest magnitude λ_D we find

$$-\lambda_D = \max_{|v|=1} v^t(-\mathcal{K})v, \qquad (20)$$

and with v as the l-th vector of the standard basis of $\mathbb{R}^D$

$$-\lambda_D \geq \max_{1 \leq l \leq D} J_{ll}. \qquad (21)$$

Under some mild conditions on random weights K_{ij}, a result from Ref. [71] shows that the inequality Eq. (21) becomes an equality in the large D limit with probability approaching 1. Motivated by this observation and the bound from Eq. (19), we expect a similar asymptotic tightness for Eq. (19). However, it is an open question whether the result from Ref. [71] applies to the bound of the spectral gap, Eq. (19). Further, the proof presented in Ref. [71] makes use of the Central Limit Theorem for the diagonal elements J_{ll} of $\mathcal{J}$, and so the corresponding result does not apply to the case of constant or logarithmically increasing (with D) sparsity parameter φ.

Nevertheless, the above arguments allow us to conjecture that in the limit of large D and $\varphi \ll D$ the following

$$\gamma_* \approx \min_{1 \leq l \leq D} J_{ll}, \qquad (22)$$

holds for general, non-hermitian random generator matrices $\mathcal{K}$, with iid and non-exotic weight distributions. We
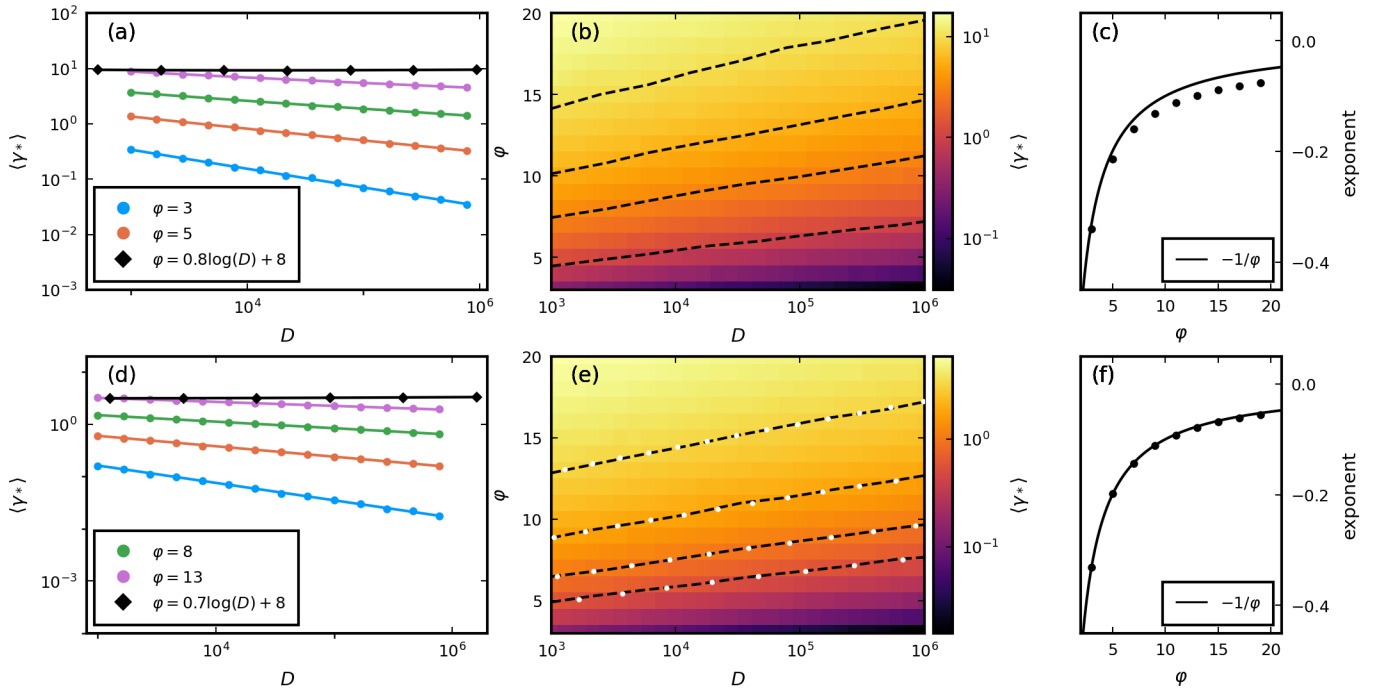
FIG. 4. The average spectral gap $\langle \gamma_* \rangle$ with χ_2^2 (top) and standard uniform (bottom) weight distributions. Solid lines in the log-log plots are analytical predictions from Eq. (25) in **(a)** and Eq. (28) in **(d)**. Black dashed lines in the heatmaps denote contours of constant gap. White circles in the heatmap in **(e)** are given by Eq. (30).
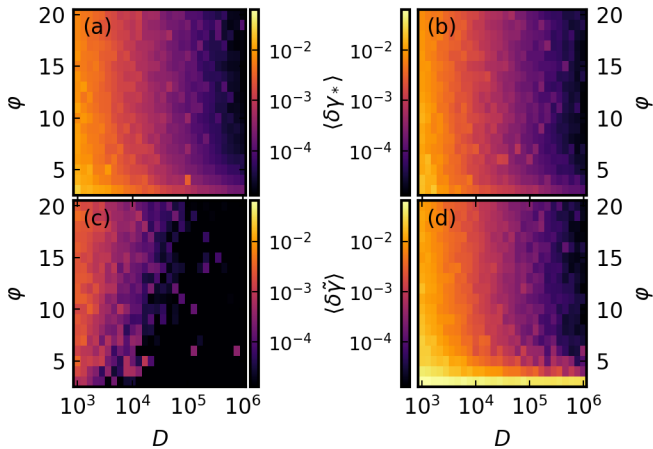

FIG. 5. The average relative error between the spectral gap and the minimal value of $\mathcal{J}$ in the top row (a) and (b) and between the horizontal extent and the maximum of $\mathcal{J}$ in the bottom row (c) and (d). The weight distribution is χ_2^2 on the left and the standard uniform distribution on the right. Averages are over 100 samples. See Eq. (23) and Eq. (38) for the definition of the relative errors $\delta\gamma_*$ and $\delta\tilde{\gamma}$, respectively.

support our conjecture with numerical data presented in Figures 5 (a) and (b). We quantify the approximation in Eq. (22) by the relative error between the spectral gap

γ_* and the minimum $\min_{1 \leq l \leq D} J_{ll}$ of the diagonal of $\mathcal{J}$,

$$\delta\gamma_* = \frac{|\gamma_* - \min_{1 \leq l \leq D} J_{ll}|}{\gamma_*}. \tag{23}$$

Figure 5 shows $\langle \delta\gamma_* \rangle$ as a function of φ and D for the χ_2^2 distribution and the standard uniform distribution. The average relative error is at least two orders of magnitude smaller than the average spectral gap shown in Figure 4 (b) and (e). For increasing D, the approximation in Eq. (22) improves. Thus, the approximation in Eq. (22), works well in the case $\varphi \ll D$.

C. Extreme value theory

The distribution of the right-hand side of Eq. (22) can be tackled with the theory of extreme values. As all non-zero entries of $\mathcal{M}$ (edge weights) are identically and independently distributed random variables, so are the diagonal entries of $\mathcal{J}$. Let the cumulative distribution function (cdf) of the diagonal entries J_{ll} of $\mathcal{J}$ be denoted by F and its probability density function by $f(x) = \frac{d}{dx}F(x)$. If the edge weights are distributed according to a χ^2 distribution (or any gamma distribution) the cdf F of J_{ll} is a gamma distribution function. If the edge weights are uniformly distributed, F is an Irwin-Hall distribution function, see Table I. The expected value of $\min_{1 \leq l \leq D} J_{ll}$

off-diag. $\mathcal{K} = M_{ij}$	χ^2_k	uniform
diag. $\mathcal{K} = J_{ll}$	gamma$\left(\frac{k\varphi}{2}, 2\right)$	Irwin-Hall
C	$\frac{2^\varphi}{\varphi!}$ *	$\frac{1}{\varphi!}$
β	$\frac{k}{2}\varphi$ *	φ

TABLE I. The distributions of the off-diagonal elements M_{ij} of $\mathcal{K}$ (edge weights) and the corresponding distributions of the diagonal elements J_{ll} of $\mathcal{K}$ and the corresponding constants C and β for the convergence of J_{ll} to the Weibull distribution Ψ_β in Eq. (27). (*) constants obtained by a power-law approximation of the left tail of the gamma distribution.

is given in terms of F (and f) by

$$\left\langle \min_{1 \leq l \leq D} J_{ll} \right\rangle = D \int dx\, x f(x)(1 - F(x))^{D-1}. \quad (24)$$

Eq. (22) and Eq. (24) imply that

$$\langle \gamma_* \rangle \approx D \int dx\, x f(x)(1 - F(x))^{D-1}. \quad (25)$$

We demonstrate the validity of Eq. (25) with Figure 4 (a), where the solid lines, given by Eq. (25), perfectly match numerically sampled average spectral gap $\langle \gamma_* \rangle$. In the next section, we will use the theory of extreme values to handle the integral in Eq. (25).

1. Power-law tail distributions

Let us consider first the case $\varphi = \text{const}$ and increasing D. By the Fisher-Tippet-Gnedenko (or 'extreme value') theorem [75], $\min_{1 \leq l \leq D} J_{ll}$ converges in law, under some mild assumptions on the distribution of J_{ll} and properly renormalization, to the Weibull distribution. The Weibull cumulative distribution function is given by $\Psi_\beta(x) = e^{-x^\beta}$, where $\beta > 0$ and the support is on the positive real line.

For distributions of J_{ll} with power-law left tail, the renormalization of $\min_{1 \leq l \leq D} J_{ll}$ for convergence to the Weibull distribution is well known, see e.g. Theorem 3.3.2, page 137 in Ref. [75]. We use a version modified to our case. Let a positive random variable X have cdf F with β-power left tail, i.e.

$$F(x) = Cx^\beta \quad \text{for } 0 \leq x \leq C^{1/\beta}, \quad (26)$$

where $C > 0$ is a constant. Further, let $m_D = \min_{1 \leq l \leq D} X_l$, where the X_l are iid copies of X. Then

$$(DC)^{1/\beta} m_D \to \Psi_\beta \quad \text{in law.} \quad (27)$$

The Irwin-Hall distribution has a left power-law tail given by $F(x) = \frac{x^\varphi}{\varphi!}$ for $0 \leq x \leq 1$. The constants for the Irwin-Hall distribution are listed in table I.

We assume that the convergence in Eq. (27) is not only in distribution but that the renormalized moments of m_D

converge as well. If the convergence of the moments is sufficiently fast, then Eq. (27) together with Eq. (22) imply

$$\langle \gamma_* \rangle \approx \langle m_D \rangle \approx \Gamma\left(1 + \frac{1}{\varphi}\right)(\varphi!)^{1/\varphi} D^{-1/\varphi} \quad (28)$$

when the weight distribution (distribution of non-zero off-diagonal elements of $\mathcal{K}$) is such that the diagonal of $\mathcal{J}$ has a power-law left tail and the coefficients C and β are given by $C = 1/\varphi!$ and $\beta = \varphi$.

Finally, we consider the case that the weight distribution is uniform. We observe that the approximation in Eq. (28) works very well in this case. The solid lines in Figure 4 (d) are given by the right-hand side of Eq. (28) and they match the numerically calculated average spectral gap.

Eq. (28) implies that, for constant $\varphi = \text{const}$ and increasing D, the average spectral gap decreases as

$$\langle \gamma_* \rangle \sim D^{-1/\varphi}. \quad (29)$$

In Figure 4 (f) we show that the numerically retrieved power-law exponents of the average spectral gap, Figure 4 (d), match the scaling in Eq. (29).

We find that the large deviation result is not only valid for constant φ and increasing D but also for φ increasing logarithmically with D; see Figure 4 (d). This allows us to estimate the crossover from decreasing to increasing spectral gap. Let c denote a constant and let $\langle \gamma_* \rangle = c$. Then by Eq. (28)

$$D \approx \left[\frac{\Gamma\left(1 + \frac{1}{\varphi}\right)}{c}\right]^\varphi \varphi!. \quad (30)$$

In Figure 4 (e) the contour lines of constant average spectral gap c perfectly line up with the functional dependence of D on φ through Eq. (30) shown as white dots.

To find φ as a function of D such that the average spectral gap is constant, we assume that φ is reasonably large and approximate $\Gamma\left(1 + \frac{1}{\varphi}\right) \approx 1$ and by Stirling's formula $(\varphi!)^{1/\varphi} \approx \frac{\varphi}{e}$. Denoting $y = \log\frac{\varphi}{ce}$ and rearranging Eq. (30) gives us

$$\frac{\log D}{ce} \approx y e^y, \quad (31)$$

which can be inverted by the Lambert W function. Re-substituting $\varphi = cee^y$ we arrive at

$$\varphi \approx ce \cdot e^{W\left(\frac{\log D}{ce}\right)}, \quad (32)$$

which for $\log D \geq ce^2$ behaves as [81]

$$\varphi \approx \frac{\log D}{(\log\log D - \log c - 1)^{1-\eta(D)}}, \quad (33)$$

where $\eta(D) \to 0$ slowly, as $\eta(D) \sim (\log\log D)^{-1}$. So in the limit $1 \ll \varphi \ll D$ the crossover from decreasing to

increasing spectral gap happens at $\varphi \sim \log D$ with corrections of the order $\log \log D$. This confirms our numerical observations that the average spectral gap $\langle \gamma_* \rangle$ appears to be constant for $\varphi \sim \log D$ in the range of matrix sizes D we considered.

2. Approximate power-law distributions

If the weight distribution is a χ^2 or exponential distribution, the diagonal elements of $\mathcal{J}$ are distributed according to Gamma distribution, see table I. The left tail of the Gamma distribution only follows approximately a power-law. Approximating the left tail by a Taylor expansion, we obtain constants C and β presented in Table I. Especially, for the χ_2^2 distribution, we presented so far in the main text the power-law approximation of the gamma distribution and the large deviation result in the previous subsection suggest that the average spectral gap $\langle \gamma_* \rangle$ decreases for constant φ and increasing D as a power in D with exponent given $-1/\varphi$, see Eq. (29).

In Figure 4 (c) we present the numerically calculated exponents of the power-law decrease of $\langle \gamma_* \rangle$, for χ_2^2 weight distribution, with D and compare it to the prediction $-1/\varphi$. We find excellent agreement for small $\varphi \leq 5$. For larger φ the deviation between the numerical exponent and $-1/\varphi$ is visible but the agreement is still good.

A quantitative comparison between the numerically calculated spectral gap $\langle \gamma_* \rangle$ and the EVT prediction by a power-law approximation of the left tail of the gamma distribution resulted in poor agreement. As the expected minimum value of the diagonal of $\mathcal{K}$ perfectly agrees with $\langle \gamma_* \rangle$, we attribute the disagreement to the power-law approximation of the left tail and slow convergence of Eq. (27) for diagonal elements of $\mathcal{J}$ distributed according to the gamma distribution.

D. Summary

We presented numerical and analytical arguments that, for the weight distributions considered, the average spectral gap decreases as a power-law for constant φ and increasing D with exponent given (approximately) by $-1/\varphi$. The crossover between decreasing and increasing spectral gap happens at $\varphi \sim \log D$, with $\log \log D$ corrections, for uniform weight distribution. For χ_2^2 distributed edge weights the crossover was observed at $\varphi \sim \log D$. If φ increases with D faster than $\log D$ then the average spectral gap increases.

The presented results generalize. Let us assume that the spectral gap is well approximated by the smallest (in magnitude) diagonal of $\mathcal{J}$, at least in the regime of large D and $\varphi \ll D$. Then, after appropriate renormalization, the distribution of the spectral gap is given by the limiting extreme value distribution of the diagonal elements of $\mathcal{J}$. Thus the classification of functional dependencies of the spectral gap on φ and D with respect to
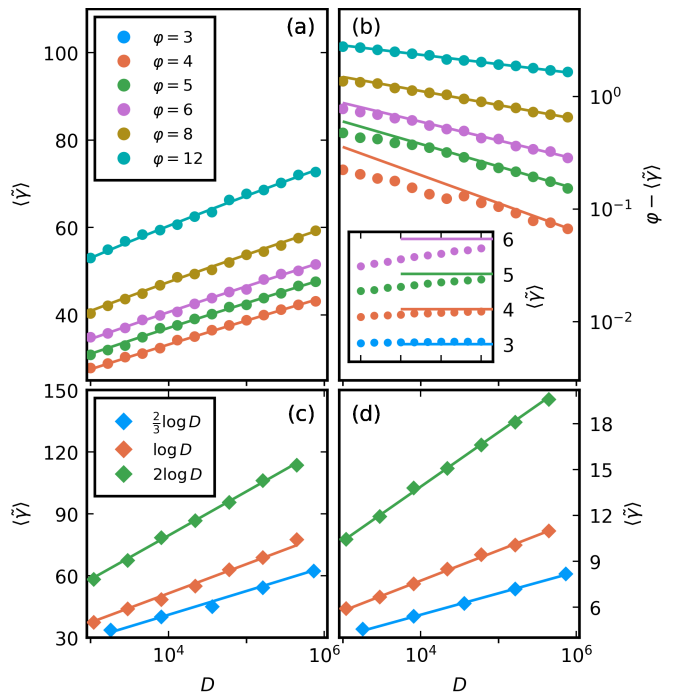

FIG. 6. Average horizontal extent $\langle \tilde{\gamma} \rangle$ with χ_2^2 (left) and standard uniform (right) weight distributions. φ is constant (top) and $\varphi \sim \log D$ (bottom). Solid lines are given by Eq. (39) (left) and Eq. (44) (right).

weight distributions reduces to the classification of limiting extreme value distributions and renormalizations. Extensive research has been conducted on the latter and the renormalizations of a lot of common distributions are well known [74, 75]. Thus the presented approach allows the calculation of the distribution of the spectral gap for broad classes of weight distributions.

V. HORIZONTAL EXTENT (LARGEST ABSOLUTE REAL PART)

In this section we investigate the horizontal extent $\tilde{\gamma}$ of the spectrum given by the eigenvalue with largest absolute real part

$$\tilde{\gamma} = \max_{1 \leq i \leq D} |\operatorname{Re} \lambda_i|. \qquad (34)$$

We focus on the averaged horizontal extent $\langle \tilde{\gamma} \rangle$. We show that for $\varphi \sim \log D$ the average horizontal extent increases logarithmically with D for χ_2^2 or uniformly distributed edge weights. For constant φ and increasing D the dependence of $\langle \tilde{\gamma} \rangle$ is qualitatively very different for the two distributions. For the χ_2^2 distribution $\langle \tilde{\gamma} \rangle$ increases logarithmically, while for the uniform distribution, the average horizontal extent converges to φ as a power-law in D. Ultimately, this is because the support of the uniform distribution is bounded, while the right tail of the χ_2^2 distribution extends to infinity.

The structure of this section follows closely the one from Section IV. We first present numerical results demonstrating the above statements. We then argue that the horizontal extent is given by the largest, in magnitude, diagonal element of $\mathcal{K}$ and invoke again EVT to analytically underpin the functional dependencies of $\langle \tilde{\gamma} \rangle$ on φ and D.

A. Numerical results

In Figure 6 we show the average horizontal extent as a function of D for constant φ and $\varphi \sim \log D$ for edge weights distributed according to a χ_2^2 (a,c) and the standard uniform distribution (b,d). In (a) the dependence of $\langle \tilde{\gamma} \rangle$ on D for constant φ shows a clear logarithmic increase with D for χ_2^2 distributed edge weights. In contrast, for the uniform distribution, the average horizontal extent increases with φ as a power-law, see (b). The power-law behavior sets in for small φ only for larger D. For $\varphi = 4$ and $\varphi = 5$ deviations from the straight lines in Figure 6 (b) are visible for $D < 10^5$ and $D < 10^4$, respectively. The average horizontal extent for constant $\varphi = 2$ and $\varphi = 3$ is not shown. We found that it does not converge to φ in the range of matrix sizes D we investigated.

For $\varphi \sim \log D$ the dependence of $\langle \tilde{\gamma} \rangle$ on D is logarithmic for both the χ_2^2 and the uniform distribution, as shown in Figure 6 (c) and (d).

In the remainder of this section, we will present analytic arguments similar to Section IV. We will explain the difference of the dependence of $\langle \tilde{\gamma} \rangle$ on D for constant φ between χ_2^2 and uniform-like distributions. We show that $\langle \tilde{\gamma} \rangle \sim \log D$ for both distributions and $\varphi \sim \log D$.

B. Extent ≈ maximum of $\mathcal{J}$

By the Perron-Frobenius theorem the spectrum of $\mathcal{K}$ is confined to the ball centered around $\min_i K_{ii} < 0$ with radius $r = |\min_i K_{ii}|$. Thus $2 \max_{1 \leq l \leq D} J_{ll} \geq |\operatorname{Re} \lambda|$ for all eigenvalues λ, so

$$\tilde{\gamma} \leq 2 \max_{1 \leq l \leq D} J_{ll}. \tag{35}$$

For symmetric generator matrices $\mathcal{K}$ we showed in Section IV B that

$$\max_{1 \leq l \leq D} J_{ll} \leq \tilde{\gamma} \tag{36}$$

and stated a result from [71] that for symmetric random generator matrices under mild conditions on the weights K_{ij}, $\max_{1 \leq l \leq D} J_{ll}$ concentrates around the largest eigenvalue in magnitude, $\tilde{\gamma}$. This together with the upper bound by the Perron-Frobenius theorem Eq. (35) leads to our conjecture that the concentration of $\max_{1 \leq l \leq D} J_{ll}$ around $\tilde{\gamma}$ in the symmetric case extends to the non-hermitian case as well

$$\tilde{\gamma} \approx \max_{1 \leq l \leq D} J_{ll}. \tag{37}$$

A concentration result similar to the one in [71] for non-hermitian random generator matrices M has to the best of our knowledge not appeared in the literature.

To quantify the deviation in Eq. (37) we introduce the relative error of $\tilde{\gamma}$ and $\max_{1 \leq l \leq D} J_{ll}$

$$\delta\tilde{\gamma} = \frac{|\tilde{\gamma} - \max_{1 \leq l \leq D} J_{ll}|}{\tilde{\gamma}}. \tag{38}$$

In Figure 5 (c) and (d) we show the average relative error $\langle \delta\tilde{\gamma} \rangle$ as a function of D and φ. If the edge weights are χ_2^2 distributed then for $2 \leq \varphi \leq 20$ and $10^3 \leq D \leq 10^5$ the average relative error is smaller than $\approx 10^{-3}$ and decreases with increasing D. Thus Eq. (37) is a good approximation for large D and $\varphi \ll D$ and the error appears negligible in the limit of large D. For uniformly distributed edge weights the average relative error $\langle \delta\tilde{\gamma} \rangle$ is smaller than 10^{-1} for $2 \leq \varphi \leq 20$ and $10^3 \leq D \leq 10^5$ and for $\varphi \geq 4$ decreases with D. For $2 \leq \varphi \leq 3$, the error does not seem to decrease for increasing D. We conclude that Eq. (37) is an excellent approximation for large D and $4 \leq \varphi \ll D$.

C. Extreme value theory

Recall that the diagonal elements of $\mathcal{J}$ are iid random variables. Similar to the minimum extreme value statistics, the expected value of $\max_{1 \leq l \leq D} J_{ll}$ is

$$\left\langle \max_{1 \leq l \leq D} J_{ll} \right\rangle = D \int dx\, x f(x) F(x)^{D-1}, \tag{39}$$

where we denoted the cdf of the diagonal elements J_{ll} of $\mathcal{J}$ by F and the pdf by $f = \frac{d}{dx} F$. A numerical calculation of the integral in Eq. (39) for χ_2^2 distributed edge weights is shown in Figure 6 (a) and (c) and compared to the average horizontal extent $\langle \tilde{\gamma} \rangle$. The quantities agree excellently.

The remainder of this section is devoted to employing the Fisher-Tippet Gnedenko or extreme value theorem to $\max_{1 \leq l \leq D} J_{ll}$ and thus analytically calculate the integral in Eq. (39).

1. Gamma distribution

Recall that if the edge weights are distributed according to a χ^2 distribution then the diagonal elements of $\mathcal{J}$ are gamma distributed. The maximum of D gamma distributed iid random variables X_l converges in law to a standard Gumbel distribution Gum [71],

$$c \left[\max_{1 \leq l \leq D} X_l - d(D) \right] \to \text{Gum} \quad \text{in law}, \tag{40}$$

with parameters c and $d(D)$ given in table II for the gamma and χ^2 distributions. The cdf of the Gumbel distribution is $x \to e^{-e^{-x}}$ with mean γ, where γ denotes

	gamma(k,θ)
c	$\frac{1}{\theta}$
$d(D)$	$\theta(\log D + (k-1)\log\log D - \log\Gamma(k))$
χ_n^2	gamma$(k=\frac{n}{2}\varphi, \theta=2)$

TABLE II. (top) The normalizing parameters c and $d(D)$ for $\max_{1\leq l\leq D} J_{ll}$ to converge to the Gumbel distribution, where J_{ll} is gamma distributed with shape and rate parameter k and θ, respectively, see Eq. (40). (bottom) The relation between the χ^2 and the gamma distribution.

the Euler-Mascheroni constant, not to be confused with the horizontal extent $\tilde{\gamma}$. The assumption that the first moment converges and the convergence is fast enough together with Eq. (37) yields

$$\langle\tilde{\gamma}\rangle \approx \left\langle \max_{1\leq l\leq D} J_{ll}\right\rangle \approx \frac{\gamma}{c} + d(D). \quad (41)$$

For constant φ and increasing D the dominant contribution of $d(D)$ is $2\log D$ for the χ^2 distribution. Thus the increase is expected to be logarithmic. This is qualitatively consistent with numerical calculations of the average horizontal extent of random generator matrices $\mathcal{K}$ with χ_2^2 distributed edge weights and constant φ shown in Figure 6 (a). There $\tilde{\gamma}$ increases logarithmically with D. Quantitatively, the deviation between the average horizontal extent and the right-hand side of Eq. (41) is not small. The deviation decreases for increasing D (not shown). We attribute the slow convergence to a suboptimal choice of parameters c and $d(D)$, as the right-hand side of Eq. (39) agrees perfectly with the numerically calculated $\langle\tilde{\gamma}\rangle$.

Let us assume that Eq. (41) is valid for φ increasing logarithmically. Note that for the χ^2 distribution, the rate parameter of the corresponding gamma distribution is linear in φ. Thus for large enough φ by Stirling's formula, the dominant term in Eq. (41) is logarithmic in D. Hence the average horizontal extent $\langle\tilde{\gamma}\rangle$ should increase logarithmically for $\varphi \sim \log D$. This is again qualitatively confirmed by numerical results shown in Figure 6 (c), where $\langle\tilde{\gamma}\rangle$ as a function of D for $\varphi \sim \log D$ increases logarithmically with D.

2. Power-law tail distributions with bounded support

If the distribution of the diagonal of $\mathcal{J}$ has bounded support and the right tail decreases as a power-law, then we can reuse the extreme value result from Section IV C. For a random variable X with right support endpoint x_0 and cdf F with power-law right tail, i.e.

$$F(x) = C(x_0 - x)^\beta \quad \text{for } x_0 - C^{1/\beta} \leq x \leq x_0, \quad (42)$$

then $m_D = \max_{1\leq i\leq D} X_l$, where X_l are iid copies of X, converges, properly renormalized, in law to a Weibull

distribution

$$(DC)^{1/\beta}(x_0 - m_D) \to \Psi_\beta \quad \text{in law.} \quad (43)$$

Again, assuming that the first moment converges as well and the convergence is fast enough we get for edge weights distributed according to the standard uniform distribution,

$$\langle\tilde{\gamma}\rangle \approx \left\langle \max_{1\leq l\leq D} J_{ll}\right\rangle \approx \varphi - \Gamma\left(1 + \frac{1}{\varphi}\right)(\varphi!)^{1/\varphi} D^{-1/\varphi}. \quad (44)$$

We find excellent numerical agreement of the right-hand side of Eq. (44) with the average horizontal extent $\langle\tilde{\gamma}\rangle$ for $\varphi \geq 4$. In Figure 6 (b) we show $\langle\tilde{\gamma}\rangle$ as a function of D for fixed φ. The solid lines denote the right-hand side of Eq. (44). They agree perfectly for large enough D and $\varphi \geq 4$. For $4 \leq \varphi \lesssim 6$ and small D the agreement is still reasonable but deviations are clearly visible. Thus for fixed $\varphi \geq 4$ and increasing D, $\langle\tilde{\gamma}\rangle$ converges to φ as a power in D with exponent $-1/\varphi$,

$$\varphi - \langle\tilde{\gamma}\rangle \sim D^{-1/\varphi}. \quad (45)$$

Numerically we find that Eq. (44) is valid for φ increasing with D logarithmically, see Figure 6 (d). There we show the average horizontal extent as a function of D for $\varphi = \log D$. It increases logarithmically with D. The logarithmic increase can be justified analytically by extending Eq. (44) beyond constant φ. In the limit of large enough φ we approximate $\Gamma(1 + 1/\varphi) \approx 1$ and by Stirling's formula $(\varphi!)^{1/\varphi} \approx \varphi/e$ and get

$$\langle\tilde{\gamma}\rangle \approx \varphi(1 - D^{-1/\varphi}) \sim \varphi. \quad (46)$$

Thus in the limit of large φ the average horizontal extent increases as $\sim \varphi \sim \log D$.

D. Summary

We showed numerically and analytically that the horizontal extent increases logarithmically for χ_2^2 and uniformly distributed edge weights if $\varphi \sim \log D$. For constant $\varphi \geq 4$ and uniformly distributed edge weights the horizontal extent increases to φ as $\sim \varphi - D^{-1/\varphi}$, while $\langle\tilde{\gamma}\rangle$ increases logarithmically for constant φ and χ_2^2 distributed edge weights.

The difference of the dependence of the average horizontal extent on φ between the χ_2^2 and uniform distribution goes back to the difference of the right tails. When edge weights are uniformly distributed the diagonal has bounded support and a power-law right tail, while it has unbounded support and an exponentially decaying right tail for χ_2^2 distributed edge weights.

Similar to the spectral gap the limiting distribution of the horizontal extent is given by the limiting extreme value distribution of the diagonal elements of $\mathcal{K}$, under the assumption that the largest (in magnitude) diagonal
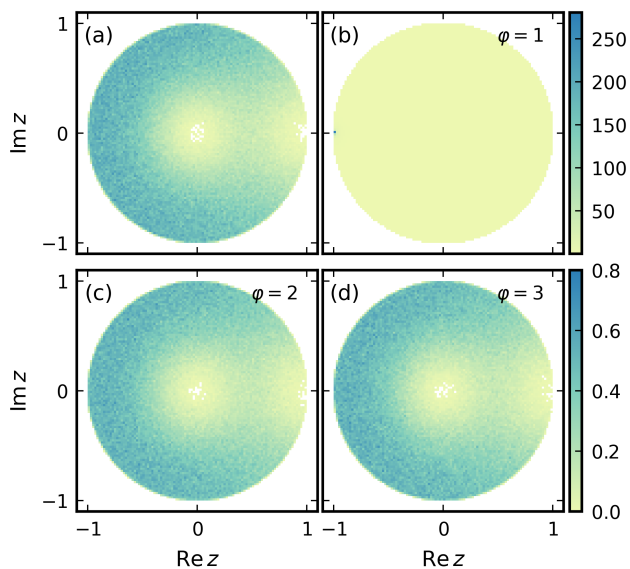
FIG. 7. Density of complex spacing ratios for **(a)** real Ginibre ensemble and **(b)-(d)** sparse Kolmogorov operators with $\varphi = 1, 2, 3$. The number of states $D = 10^4$ and densities are obtained from 10^2 samples. Edge weights are distributed according to the χ_2^2 distribution. The color range is from 0 to 0.8 in (a), (c), and (d) and from 0 to 260 in (b).

of $\mathcal{K}$ is approximating $\tilde{\gamma}$ well enough. Thus the classification of the horizontal extent with respect to weight distributions reduces to the classifications of convergence in extreme value theory.

VI. COMPLEX SPACING RATIOS

So far we considered the marginal distribution of eigenvalues of sparse random generator matrices. But correlations between the eigenvalues are also of interest. Correlations between eigenvalues of real spectra are often quantified with the distribution of consecutive level spacings or their ratios. The latter avoids the need to unfold the corresponding spectrum [82, 83] and has been generalized to complex eigenvalues in the recent work [61]. The complex spacing ratio (CSR) of eigenvalue λ of matrix $\mathcal{K}$ is defined as

$$z = \frac{\lambda^{NN} - \lambda}{\lambda^{NNN} - \lambda}, \quad (47)$$

where λ^{NN} and λ^{NNN} denote the closest, by the Euclidean distance, and second closest eigenvalue (of $\mathcal{K}$) to λ, respectively. By definition, the density of CSRs is supported on the unit disk on the complex plane.

If eigenvalues λ are uncorrelated, the CSR density is uniform. Eigenvalues of generic random matrix ensembles are typically correlated and feature mutual repulsion. This leads to vanishing CSR density at $z = 0$ and $z = 1$. According to [84], complex level spacings categorize random matrix ensembles in three universality

classes. Generic random matrices fall into one of these classes according to their symmetries. The random generators considered in this paper have real entries so they obey the same symmetry as real Ginibre matrices (GinOE).

In Figure 7 we show the CSR densities of (a) GinOE members with Gaussian entries (b-d) and sparse random generators with χ_2^2 distributed edge weights and $\varphi = 1, 2, 3$. The densities are estimated from 100 samples for $D = 10^4$. We also checked that the obtained densities are independent of the weight distribution. As suggested in Ref. [61], we avoid eigenvalues close to the real line (by excluding all eigenvalues from the strip $\mathrm{Im}\,\lambda < 10^{-14}$) when sampling CSR densities.

The CSR density of GinOE matrices shown on Figure 7 (a) exhibit typical depletion at $z = 0$ and $z = 1$. In Ref. [40], it was shown that the CSR density obtained for dense random Kolmogorov operators agrees well with the distribution shown in Figure 7 (a). The CSR density of sparse generators with sparsity $\varphi \geq 2$ (c,d) agrees remarkably well with the GinOE case.

The CSR density for $\varphi = 1$ is anomalous, see Figure 7 (b). It has an extremely high density around $z = -1$ while being nearly flat on the rest of the unit disk. In this ultimate case, the operator can be presented as

$$\mathcal{K} = \mathcal{V} \cdot (\mathcal{P} - \mathbb{1}), \quad (48)$$

where $\mathcal{V}$ is a diagonal matrix (with elements distributed according to, e.g., χ_2^2) and $\mathcal{P}$ is a circulant permutation matrix corresponding to a cyclic unit shift. The spectrum of $\mathcal{P} - \mathbb{1}$ lies on a circle of unit radius centered at $\lambda = -1$ and constitutes a set of equidistant roots of unity. This spectrum is slightly deformed and split into several loops by multiplication of $\mathcal{P} - \mathbb{1}$ with $\mathcal{V}$. Away from $\lambda = -1$, $\mathcal{V}$ dominates which results in the appearance of a real 'tail'; see Fig. 8.

In graph theory terms, such a sparse random graph fragments into a set of disjoint elementary cycle graphs. The independence of the spectra of these cycles leads to the flatness of the density away from $z = -1$, while the elementary cycle structure of the connected components is responsible for the CSR peak at $z = -1$.

To quantify the distance between CSR densities, we use the average length $\langle r \rangle$ and the average cosine of the angle $-\langle \cos\theta \rangle$ of spacing ratios, where $\langle \dots \rangle$ again denotes the average over the random matrix ensemble [61]. We numerically estimate $\langle r \rangle_{\mathrm{GinOE}} \approx 0.7379$ and $-\langle \cos\theta \rangle_{\mathrm{GinOE}} \approx 0.2347$ for 100 $10^4 \times 10^4$-matrices. These agree well with $\langle r \rangle$ and $-\langle \cos\theta \rangle$ for $\varphi = 2$ and $\varphi = 3$, as shown in Table III. We found similar results for $\varphi > 3$ (not shown). In contrast, the corresponding quantities for $\varphi = 1$ deviate substantially from $\langle r \rangle_{\mathrm{GinOE}}$ and $-\langle \cos\theta \rangle_{\mathrm{GinOE}}$, as also shown in Table III. We conclude that, for $\varphi \geq 2$, correlations between eigenvalues of sparse random Kolmogorov operators agree with correlations of eigenvalues of GinOE matrices.

	GinOE	$\varphi = 1$	$\varphi = 2$	$\varphi = 3$
$-\langle\cos\theta\rangle$	0.7379	0.7871	0.7359	0.7372
$\langle r\rangle$	0.2347	0.3516	0.2225	0.2284

TABLE III. Mean and angle of spacing ratio distributions obtained with 10^2 samples of random $10^4 \times 10^4$ matrices rounded to the 4th digit. The matrix ensembles correspond to the ones shown in Figure 7.

VII. DISCUSSION

A. Summary of results

Motivated by the incapability of dense random Kolmogorov operators to capture spectral features of model Markov processes, we introduced and analyzed an ensemble of sparse random Kolmogorov operators. We showed that, if the number of non-zero elements per column (and row) φ increases with the matrix size D, the bulk of the spectrum is shifted away from the stationary eigenvalue 0 in the limit of large matrix size D. This is independent of the weight distribution, i.e. of the distribution of the nonzero matrix elements.

In contrast, the spectral edges depend on the tails of the weight distribution. The tails of the weight distribution determine, together with φ, the tails of diagonal elements of generator matrices. We numerically showed that the spectral edges are well approximated by the extremes of the diagonal elements. From extreme value theory it follows that for diagonal distributions with power-law left tails (this includes among others edge weights being uniform, exponential, χ^2, gamma or beta distributed), the average spectral gap decreases as a power-law in D for fixed φ, is constant for $\varphi \sim \log D$ and increases, whenever φ increases with D substantially faster than $\log D$.
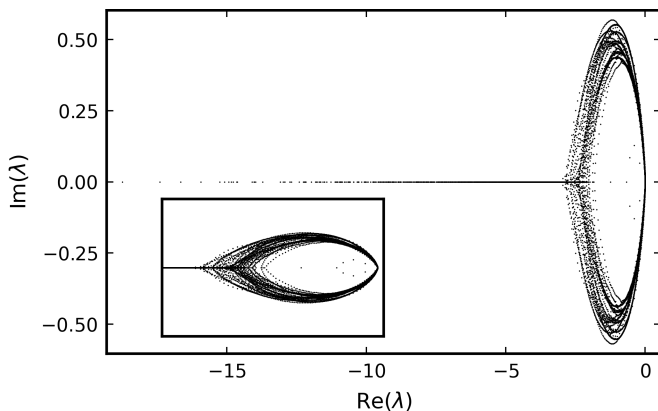

FIG. 8. Spectrum of a random Kolmogorov operator with $\varphi = 1$ and χ_2^2 weight distribution. The matrix size is $D = 10^3$. Inset: same data plotted with both axes having the same scale.

A similar approach was used to calculate the horizontal extent, given by the eigenvalue with the smallest real part. We linked the horizontal extent to the largest diagonal element (in magnitude) of the generator matrix and used extreme value theory to calculate the latter.

Finally, we showed that complex spacing ratio distributions of generator matrices with $\varphi \geq 2$ follow the distribution typical of Ginibre's Orthogonal Ensemble, while there is a strong anomaly for $\varphi = 1$.

B. Open questions

(1) We have introduced sparsity to model $\mathcal{K}$-generators of physical Markov processes, and have used the sparsity to tune spectral features of the generators. There are other ways of providing random matrices with a structure that models physical constraints (e.g., locality). E.g., one could consider banded matrices [85–94] or matrices with decaying off-diagonal terms [93, 95, 96] or temperature based models [97]. These are alternate routes to tuning spectral features. To the best of our knowledge, generators of CTMCs with such structures have not yet been considered.

(2) The application of extreme value theory to find the limiting distribution of the spectral edges relied on the observation that the spectral edges are well approximated by the minimum and the maximum of the diagonal of the generator matrix. By the Courant-Fisher theorem, the extremes of the diagonal are upper and lower bounds, respectively, for symmetric generators. In this case, a concentration of the largest eigenvalue in magnitude around the maximum of the diagonal was shown in [71]. An analytical treatment of general non-symmetric generators and the spectral gap is to the best of our knowledge not known. We hope that our results motivate a rigorous investigation of the connection between the spectral edges and the diagonal of the generator matrix.

(3) Generators of CTMCs have real entries and thus their eigenvalues are real or come in complex conjugate pairs. In the investigation of correlations between eigenvalues, we left out real eigenvalues. The appearance of a large number of real eigenvalues in the spectrum of non-Hermitian matrices is a phenomenon of wide interest [55, 98–107]. For real Ginibre matrices, the average number of real eigenvalues is $\sim D^{-1/2}$ [99–101] while for dense generators of CTMCs, it is substantially larger [55]. We observed that the fraction of real eigenvalues is larger for small φ and smaller for larger φ (not presented). Understanding of the functional dependence of the number of real eigenvalues for sparse CTMC generators is an interesting problem.

(4) We focused on the location and extent of the bulk spectrum as well as the spectral edges. One could inquire about the evolution of other features of the spectral distribution as a function of sparsity, e.g., about the envelope of the spectral distribution. In [40] the spectral density of dense random CTMCs was described

by the convolution of two asymptotically free matrices, leading to the prominent spindle shape. Free probability arguments break down for sparse random CTMCs. Analytical tools which have been employed to calculate the spectral density of sparse, random matrices include replica tricks [108–112], single defect and effective medium approximations [113–115], supersymmetry-based techniques [116, 117] and the cavity approach [112, 118–120]. Spectral properties of symmetric, sparse, random CTMCs have been investigated with the cavity method [121–123] and with supersymmetric approaches [117]. Investigations of the spectral density of non-symmetric sparse, random Kolmogorov operators with the above methods might be an interesting objective.

(5) We have considered sparse generators of CTMCs based on strongly connected, sparse random graphs. It is an open question whether our results can be generalized to other sparse graph ensembles. One potential avenue to explore are directed Erdös-Renyi (dER) graphs.

In dER graphs, the probability of an edge connecting any two vertices is $0 < p \leq 1$. For a dER graph to be strongly connected with a high probability, the value of p must exceed $\sim \log D/D$ [124, 125]. As a result, the average degree of the vertices must increase logarithmically with D to ensure strong connectivity. Consequently, the range of constant average vertex degree and increasing vertex number D is excluded.

Nonetheless, modifying the dER graph by enforcing a minimum (in- and out-) degree ≥ 2 guarantees strong connectivity with high probability [79]. Exploring the spectral properties of CTMC generators based on dER graphs may represent a promising next step towards generalizing our results.

(6) Finally, there is an interesting question: What could 'sparsity' mean in the quantum limit? Namely, what is 'sparsity' for Lindblad operators?

Here we start from the genetic link which allows to obtain a generator of a classical (quantum) Markovian evolution as a properly normalized non-probability (trace) preserving stochastic map (channel). Eq. (3), where $\mathcal{M}$ is a no-probability-preserving map and $\mathcal{J}$ takes care of normalization, illustrates this link in the case of Kolmogorov operators. It seems to be intuitive that sparsity φ of a stochastic matrix can be associated with rank r of a quantum channel [126].

Thus, in the ultimate limit $\varphi = r = 1$, quantum versions of stochastic maps – that are permutations now – are rank-one channels – that are unitaries. It is tempting to extend this analogy beyond the limit $\varphi = r = 1$ and state that mixed-unitary channels (convex combinations of unitaries) are quantum versions of bistochastic matrices (that are, according to Birkhoff, convex combinations of permutations [127]).

However, there is also notion of double stochastic (or "unital") channels [126]. The two classes, mixed unitaries and bistochastic channels, are not identical: there are double stochastic channels that are outside of the convex hull of unitaries [128]. What class to associate with clas-

sical bistochastic maps is then 'a matter of taste' [129]. To resolve the dichotomy, one could rely on the concept of super-decoherence [40] and state that all channels which have classical bistochastic matrices as their fully decohered versions, are quantum analogues of the matrices. In this case the broader class of double stochastic channels is chosen [129].

The superdecoherence-based reasoning can also be applied to generators. In this case the unitary (Hamiltonian) part of a Lindblad operator does not play any role since it vanishes in the limit of complete decoherence [40] and the quantum 'sparsity' is defined by the rank of the dissipative part (the minimal number of jump operators).

Interestingly, Lindblad operators of ultra-low rank $r = 1$ were considered in Refs. [52] and [130]. Features similar to ones we detected for ultra-sparse Kolmogorov operators were reported (e.g., the spectral gap is defined by a real-valued outlier with position independent of D).

Appendix A: Models in Figure 1

In the following, we denote spin creation and annihilation operators as σ^+ and σ^-, respectively. We split the generator matrix $\mathcal{K}$ into an off-diagonal matrix $\mathcal{M}$ and a diagonal matrix $\mathcal{J}$ such that $\mathcal{K} = \mathcal{M} - \mathcal{J}$ and the diagonal entries of $\mathcal{J}$ are the sums of the columns of $\mathcal{M}$.

In Figure 1 (b) we show the TASEP on a ring with $L = 12$ sites and staggered hopping amplitudes. The $\mathcal{M}$ matrix is given by [61]

$$\mathcal{M} = \frac{1}{2}\sigma_1^+ + \frac{1}{2}\sigma_1^- + \sum_{j=1}^{L} p_j \sigma_j^- \sigma_{j+1}^+, \qquad \text{(A1)}$$

where $p_j = 1$ if j is even and $p_j = 0.2$ if j is odd.

In Figure 1 (c) we show the ASEP on a chain of length $L = 12$ with open boundary conditions and next nearest neighbor hopping. The $\mathcal{M}$ matrix is given by

$$\mathcal{M} = \sigma_1^+ + \sigma_L^- + \sum_{j=1}^{L} \sigma_j^- \sigma_{j+1}^+ + \sum_{j=1}^{L/2} \sigma_{2j}^- \sigma_{2j+2}^+. \qquad \text{(A2)}$$

In Figure 1 (d) we show the spectrum of a single particle hopping on a 65×65 grid with periodic boundary

conditions and random hopping amplitudes. The $\mathcal{M}$ matrix is given by

$$\mathcal{M} = \sum_{\langle (i,j),(i',j') \rangle} p_{(i,j)\to(i',j')} \sigma_{i,j}^- \sigma_{i',j'}^+ \qquad \text{(A3)}$$

where $\langle \ldots \rangle$ denotes summation over nearest neighbors and $p_{(i,j)\to(i',j')}$ are randomly uniformly chosen between 0 and 1 under the constraint that $p_{(i,j)\to(i',j')} = 1 - p_{(i',j')\to(i,j)}$. This diffusion model can of course be extended to many particles, but we choose to show the single-particle sector here.

In Figure 1 (e) we show the spectrum of a contact process [28] on a chain with $L = 12$ sites and open boundary conditions. The master equation is generated by $-H$, where H is given by

$$H = \sum_{i=1}^{L} M_i + \sum_{i=1}^{L-1} \left[n_i Q_{i+1} + Q_i n_{i+1} \right], \qquad \text{(A4)}$$

and

$$M = \begin{pmatrix} 0 & -1 \\ 0 & 1 \end{pmatrix}, \quad n = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \quad Q = \begin{pmatrix} 1 & 0 \\ -1 & 0 \end{pmatrix}. \quad \text{(A5)}$$

Finally, in Figure 1 (f) we show the spectrum of the generator matrix $\mathcal{K}$ of a gene transcription model taken from [20]. The following master equations model the accumulation and release of mechanical strain of DNA during transcription. The parameters chosen for the spectral data in Figure 1 (f) are the mRNA transcription rate $r = 2$ and decay rate $\lambda = 0.05$, the maximum number of transcripts until no further strain can be put on the DNA $m_c = 10$, the relaxation rate of the DNA string $g = 0.05$ and a maximum number of transcription events $m_{\max} = 400$ to make the generator matrix M finite. By m we denote the number of current transcripts and by α the number of transcripts made since the last relaxation event. Then for $0 \leq m \leq m_{\max}$ and $1 \leq \alpha \leq m_c - 1$ the master equation reads

$$\frac{d}{dt} P_\alpha = -(r + g + \lambda m) P_\alpha(m,t) + \lambda(m+1) P_\alpha(m+1,t)$$
$$+ r P_{\alpha-1}(m-1,t) \qquad \text{(A6)}$$

while for $\alpha = 0$ we have

$$\frac{d}{dt} P_0 = -(r + g + \lambda m) P_0(m,t) + \lambda(m+1) P_0(m+1,t)$$
$$+ g \sum_{\alpha=0}^{m_c} P_\alpha(m,t) \qquad \text{(A7)}$$

and for $\alpha = m_c$

$$\frac{d}{dt} P_{m_c} = -(g + \lambda m) P_{m_c}(m,t) + \lambda(m+1) P_{m_c}(m+1,t)$$
$$+ r P_{m_c-1}(m-1,t). \qquad \text{(A8)}$$

Appendix B: Sampling of sparse random generators

In order to obtain a sparse random generator matrix, our approach involves first sampling a random directed graph with D vertices and both in- and out-vertex degrees of φ. Subsequently, the non-zero elements of the corresponding adjacency matrix are sampled from a common positive distribution. This procedure results in the off-diagonal matrix $\mathcal{M}$. The random Markov generator matrix is then constructed as $\mathcal{K} = \mathcal{M} - \mathcal{J}$, where $\mathcal{J}$ is a diagonal matrix with diagonal elements equal to the sums of the columns of $\mathcal{M}$.

The random directed graph is generated by iteratively connecting each vertex to φ other vertices, while rejecting edges if the corresponding vertex already has φ incoming edges. For the final vertices, it may not be feasible to connect to other vertices without violating the constraint of φ incoming edges for each vertex. In such cases, the entire process is restarted. To mitigate the risk of restarting the procedure, we reduce the probability of connecting to a vertex that already has a high degree. Following this approach, we find that we rarely need to restart the algorithm for the matrix sizes and vertex degrees φ examined in this study.

To compute the eigenvalues of the Markov matrices, we utilize an exact diagonalization method, while the Arnoldi method is employed to calculate the spectral gap. We deem an eigenvalue to have converged once the norm of the residuals of the Schur vectors is less than 10^{-12}.

Appendix C: Analytical results for the bulk spectrum

In this section, we will derive the analytical results of the estimated mean $\mu(\lambda)$ in Eq. (7) and the estimated pseudo-variance in Eq. (11) in the main text and show that $\frac{1}{D} \sum_{j=1} \lambda_j$ concentrates around its average $\langle \ldots \rangle$.

Denote by ι the function $\iota : \{1, \ldots, \varphi\} \times \{1, \ldots, D\} \to \{1, \ldots, D\}^2$ with $\iota(l,j) = (i,j)$ where i is the lth non-zero index in column j in $\mathcal{M}$. Note that $\iota(l,j) = (i,j)$ implies $i \neq j$ and $l \to \iota(l,j)$ is injective for fixed j. Further, let in this appendix the location of the bulk be denoted as

$$\mu(\lambda) = \frac{1}{D} \sum_{j=1}^{D} \lambda_j = \frac{1}{D} \operatorname{tr}(\mathcal{K}).$$

and the pseudo-variance as

$$\sigma^2(\lambda) = \frac{1}{D} \sum_{j=1}^{D} \lambda_j^2 - \left(\frac{1}{D} \sum_{j=1}^{D} \lambda_j \right)^2$$
$$= \frac{\operatorname{tr}(\mathcal{K}^2)}{D} - \frac{\operatorname{tr}(\mathcal{K})^2}{D^2}. \qquad \text{(C1)}$$

Here we explicitly do not include the averaging over the matrix ensemble $\langle \ldots \rangle$ in contrast to the main text.

1. Location

The average value with respect to $\langle \ldots \rangle$ of the location $\mu(\lambda)$ can then be computed as

$$\langle \mu(\lambda) \rangle = \left\langle \frac{1}{D} \operatorname{tr}(\mathcal{K}) \right\rangle = \frac{1}{D} \sum_{j=1}^{D} \langle K_{jj} \rangle$$

$$= \frac{1}{D} \sum_{j=1}^{D} \sum_{l=1}^{\varphi} \langle K_{\iota(l,j)} \rangle = -\varphi \mu_0,$$

where we used that $\langle K_{\iota(l,j)} \rangle = -\mu_0$. This is Eq. (7) in the main text. Similar,

$$\langle \operatorname{tr}(\mathcal{K})^2 \rangle = \sum_{j_1,j_2=1}^{D} \sum_{l_1,l_2=1}^{\varphi} \langle K_{\iota(l_1,j_1)} K_{\iota(l_2,j_2)} \rangle$$

$$= \sum_{j=1}^{D} \left[\sum_{l=1}^{\varphi} \left\langle K_{\iota(l,j)}^2 \right\rangle + \sum_{l_1 \neq l_2} \left\langle K_{\iota(l_1,j)} K_{\iota(l_2,j)} \right\rangle \right]$$

$$+ \sum_{j_1 \neq j_2} \sum_{l_1,l_2=1}^{\varphi} \left\langle K_{\iota(l_1,j_1)} K_{\iota(l_2,j_2)} \right\rangle.$$

Although the off-diagonal elements of $\mathcal{K}$ are weakly dependent because of the constraint that the number of non-zero elements per row and column has to equal φ, the non-zero elements $K_{\iota(l,j)}$ are independent. Hence, $\langle K_{\iota(l_1,j)} K_{\iota(l_2,j)} \rangle = \langle K_{\iota(l_1,j)} \rangle \langle K_{\iota(l_2,j)} \rangle$ and $\langle K_{\iota(l_1,j_1)} \rangle \langle K_{\iota(l_2,j_2)} \rangle$, so

$$\langle \operatorname{tr}(\mathcal{K})^2 \rangle = D\varphi(\sigma_0^2 + \mu_0^2) + D\varphi(\varphi-1)\mu_0^2 + D(D-1)\varphi^2 \mu_0^2$$

$$= D\varphi \sigma_0^2 + (D\varphi \mu_0)^2,$$

where we used that the second moment $\left\langle K_{\iota(l,j)}^2 \right\rangle$ equals $\sigma_0^2 + \mu_0^2$. This implies that

$$\langle \mu(\lambda)^2 \rangle - \langle \mu \rangle^2 = \left\langle \frac{\operatorname{tr}(\mathcal{K})^2}{D^2} \right\rangle - \left\langle \frac{\operatorname{tr}(\mathcal{K})}{D} \right\rangle^2 = \frac{\varphi \sigma_0^2}{D}.$$

The right-hand side vanishes for increasing D and φ growing slower with D than linear. Relatively to $\langle \mu(\lambda) \rangle$ the typical deviation of $\mu(\lambda)$ from its average value always vanishes for either increasing D or φ, as

$$\frac{\sqrt{\langle \mu(\lambda)^2 \rangle - \langle \mu \rangle^2}}{|\langle \mu(\lambda) \rangle|} = \frac{\sigma_0}{\mu_0} (\varphi D)^{-1/2}.$$

2. Complex pseudo-variance

The first term in the averaged pseudo-variance given by Eq. (C1) can be calculated as

$$\langle \operatorname{tr}(\mathcal{K}^2) \rangle = \sum_{i,j=1}^{D} \langle K_{ij} K_{ji} \rangle$$

$$= \sum_{i=1}^{D} \langle K_{ii}^2 \rangle + \sum_{i \neq j} \langle K_{ij} K_{ji} \rangle. \qquad (C2)$$

We proceed with $\sum_{i=1}^{D} \langle K_{ii}^2 \rangle$ in Eq. (C2) and get

$$\sum_{i=1}^{D} \langle K_{ii}^2 \rangle = \sum_{i=1}^{D} \left\langle \left(-\sum_{j \neq i} K_{ji} \right)^2 \right\rangle$$

$$= \sum_{i=1}^{D} \sum_{j,l \neq i} \langle K_{ji} K_{li} \rangle$$

$$= \sum_{i=1}^{D} \sum_{j \neq i} \langle K_{ji}^2 \rangle + \sum_{i=1}^{D} \sum_{j,l \neq i; j \neq l} \langle K_{ji} \rangle \langle K_{li} \rangle. \qquad (C3)$$

The former sum in Eq. (C3) is given by

$$\sum_{i=1}^{D} \sum_{j \neq i} \langle K_{ji}^2 \rangle = \sum_{i=1}^{D} \sum_{l=1}^{\varphi} \langle K_{\iota(l,i)}^2 \rangle = D\varphi(\sigma_0^2 + \mu_0^2), \quad (C4)$$

where again we used that $\langle K_{\iota(l,i)}^2 \rangle = \sigma_0^2 + \mu_0^2$, while the latter sum in Eq. (C3) is

$$\sum_{i=1}^{D} \sum_{j,l \neq i; j \neq l} \langle K_{ji} \rangle \langle K_{li} \rangle$$

$$= \sum_{i=1}^{D} \sum_{k=1}^{\varphi} \sum_{n=1; \iota(n,i) \neq \iota(k,i)}^{\varphi} \langle K_{\iota(k,i)} \rangle \langle K_{\iota(n,i)} \rangle$$

$$= D\varphi(\varphi-1)\mu_0^2. \qquad (C5)$$

Combining Eq. (C4) and Eq. (C5) we get

$$\sum_{i=1}^{D} \langle K_{ii}^2 \rangle = D\varphi(\sigma_0^2 + \mu_0^2) + D\varphi(\varphi-1)\mu_0^2$$

$$= D\varphi \sigma_0^2 + D\varphi^2 \mu_0^2.$$

Now, we are left with calculating $\sum_{i \neq j} \langle K_{ij} K_{ji} \rangle$, the second term in Eq. (C2),

$$\sum_{i \neq j} \langle K_{ij} K_{ji} \rangle = \sum_{i=1}^{D} \sum_{l=1}^{\varphi} \left\langle K_{\overline{\iota(l,i)}} M_{\iota(l,i)} \right\rangle,$$

where the $\bar{\iota}$ denotes swapping the first and second component. Note that $K_{\overline{\iota(l,i)}}$ is not necessarily a non-zero entry of $\mathcal{K}$, hence $K_{\overline{\iota(l,i)}}$ and $K_{\iota(l,i)}$ depend weakly on each other. In the large D limit we can assume that the dependence is sufficiently weak and we treat $K_{\overline{\iota(l,i)}}$ and $K_{\iota(l,i)}$ as independent, thus $\left\langle K_{\overline{\iota(l,i)}} K_{\iota(l,i)} \right\rangle = \mu_0 \left\langle K_{\overline{\iota(l,i)}} \right\rangle$. By the assumed independence the mean of *every* entry in the ith row, except the diagonal, is $\left\langle K_{\overline{\iota(l,i)}} \right\rangle = \frac{\varphi}{D} \mu_0$. Hence,

$$\sum_{i \neq j} \langle K_{ij} K_{ji} \rangle = \sum_{i=1}^{D} \frac{1}{D} \varphi^2 \mu_0^2 = \varphi^2 \mu_0^2.$$

Collecting the above results we arrive at

$$\langle \mathrm{tr}(\mathcal{K}^2) \rangle = D\varphi\sigma_0^2 + D\varphi^2\mu_0^2 + \varphi^2\mu_0^2$$
$$= D\varphi\sigma_0^2 + (D+1)\varphi^2\mu_0^2$$

The second term of the averaged pseudo-variance in Eq. (C1) has been calculated in the previous subsection,

$$\langle \mathrm{tr}(\mathcal{K})^2 \rangle = D\varphi\sigma_0^2 + (D\varphi\mu_0)^2$$

Finally, we can evaluate

$$\langle \sigma^2(\lambda) \rangle = \left\langle \frac{\mathrm{tr}(\mathcal{K}^2)}{D} \right\rangle - \left\langle \frac{\mathrm{tr}(\mathcal{K})^2}{D^2} \right\rangle$$
$$= \varphi\sigma_0^2 + \varphi^2\mu_0^2 + \frac{1}{D}\varphi^2\mu_0^2 - \frac{1}{D}\varphi\sigma_0^2 - \varphi^2\mu_0^2$$
$$= \varphi\left(\sigma_0^2 + \frac{\varphi}{D}\mu_0^2 - \frac{1}{D}\sigma_0^2\right),$$

which is Eq. (11) in the main text.

Appendix D: Bound of spectral gap for symmetric M

In this section, we give the proof of Eq. (19). Let $\mathcal{K} = \mathcal{M} - \mathcal{J}$ be a symmetric generator matrix. By Eq. (17) we have to show that $v^t \mathcal{K} v \leq \min_{1 \leq l \leq D} J_{ll} + O\left(D^{-1}\right)$ for the vector v given

$$v_i = \begin{cases} \sqrt{1 - \frac{1}{D}} & i = l \\ -\frac{1}{\sqrt{D(D-1)}} & i \neq l, \end{cases}$$

where $1 \leq l \leq D$ is arbitrary. It is easy to see that $|v| = 1$ and $v \perp v_1$. So we proceed with

$$\gamma_* \leq v^t(\mathcal{J} - \mathcal{M})v = \sum_{i,j=1}^{D} v_i v_j (\mathcal{J} - \mathcal{M})_{ij}$$
$$= \sum_{i=1}^{D} v_j^2 J_{jj} - \sum_{i,j=1}^{D} v_i v_j M_{ij}$$
$$= \sum_{i,j=1}^{D} v_j^2 M_{ij} - \sum_{i,j=1}^{D} v_i v_j M_{ij}$$
$$= \sum_{i,j=1}^{D} v_j M_{ij}(v_j - v_i). \tag{D1}$$

Note that any summand in Eq. (D1) where either $i = j = l$ or $i \neq l$ and $j \neq l$ is zero. Inserting the definition of v we get

$$\gamma_* \leq \sum_{i \neq l} v_l M_{il}(v_l - v_i) + \sum_{j \neq l} v_j M_{lj}(v_j - v_l)$$
$$= \sum_{i \neq l} \sqrt{1 - \frac{1}{D}} M_{il}\left(\sqrt{1 - \frac{1}{D}} + \frac{1}{\sqrt{D(D-1)}}\right)$$
$$- \sum_{j \neq l} \frac{1}{\sqrt{D(D-1)}} M_{lj}\left(-\frac{1}{\sqrt{D(D-1)}} - \sqrt{1 - \frac{1}{D}}\right)$$
$$= \left(\sqrt{1 - \frac{1}{D}} + \frac{1}{\sqrt{D(D-1)}}\right)$$
$$\times \sum_{i \neq l} \left[\sqrt{1 - \frac{1}{D}} M_{il} + \frac{1}{\sqrt{D(D-1)}} M_{li}\right]. \tag{D2}$$

After collecting all the prefactors in Eq. (D2) the spectral gap is upper-bounded by

$$\gamma_* \leq \sum_{i \neq l} \left[M_{il} + \frac{1}{D-1} M_{li}\right] = J_{ll} + \frac{1}{D-1}\tilde{J}_{ll},$$

where we denote $\tilde{J}_{ll} = \sum_{i \neq l} M_{il}$. As the number of non-zero elements of $\mathcal{M}$ in every row and column is the same, the distribution of J_{ll} and $\tilde{J}_{ll}$ coincide. In the limit of large D, J_{ll} and $\tilde{J}_{ll}$ are independent. Thus we can approximate $\gamma_* \leq J_{ll} + O\left(D^{-1}\right)$ at least for $\varphi \ll D$. As the index l was chosen arbitrarily we get

$$\gamma_* \leq \min_{1 \leq l \leq D} J_{ll} + O\left(D^{-1}\right),$$

which is Eq. (19) in the main text.

[1] W. J. Anderson, *Continuous-Time Markov Chains* (Springer, 1991).

[2] T. M. Liggett, *Interacting Particle Systems*, Vol. 2 (Springer, 1985).

[3] A. T. Bharucha-Reid, *Elements of the Theory of Markov Processes and their Applications* (Courier Corporation, 1997).

[4] N. Van Kampen, *Stochastic Processes in Physics and Chemistry*, 3rd ed., North-Holland Personal Library (Elsevier Science, 2007).

[5] T. Lux, A master equation approach to the modelling of financial markets microstructure, IFAC Proceedings Volumes **28**, 409 (1995).

[6] M. R. Grasselli and P. X. Li, A stock-flow consistent macroeconomic model with heterogeneous agents: The master equation approach, Journal of Network Theory in Finance **4**, 47 (2018).

[7] J. Hofbauer and K. Sigmund, *The Theory of Evolution and Dynamical Systems. Mathematical Aspects of Selection.* (Cambridge University Press, 1988).

[8] G. Szabó and G. Fáth, Evolutionary games on graphs, Physics Reports **446**, 97 (2007).

[9] D. A. McQuarrie, Stochastic approach to chemical kinetics, Journal of Applied Probability **4**, 413 (1967).

[10] J. Schnakenberg, Network theory of microscopic and macroscopic behavior of master equation systems, Reviews of Modern Physics **48**, 571 (1976).

[11] M. J. Pilling and S. H. Robertson, Master equation models for chemical reactions of importance in combustion, Annual Review of Physical Chemistry **54**, 245 (2003).

[12] D. T. Gillespie, Stochastic simulation of chemical kinetics, Annual Review of Physical Chemistry **58**, 35 (2007).

[13] D. F. Anderson and T. G. Kurtz, Continuous time Markov chain models for chemical reaction networks (Springer New York, 2011) pp. 3–42.

[14] P. Lecca, Stochastic chemical kinetics, Biophysical Reviews **5**, 323 (2013).

[15] M. Rajvanshi and K. V. Venkatesh, Chemical master equation, Encyclopedia of Systems Biology , 396 (2013).

[16] H. D. Jong, Modeling and simulation of genetic regulatory systems: A literature review, Journal of Computational Biology **9**, 67 (2002).

[17] M. Hegland, C. Burden, L. Santoso, S. MacNamara, and H. Booth, A solver for the stochastic master equation applied to gene regulatory networks, Journal of Computational and Applied Mathematics **205**, 708 (2007).

[18] H. S. Booth, C. J. Burden, M. Hegland, and L. Santoso, A stochastic model of gene regulation using the chemical master equation, Modeling and Simulation in Science, Engineering and Technology **38**, 71 (2007).

[19] O. Lipan, Differential equations and chemical master equation models for gene regulatory networks, Molecular Life Sciences , 1 (2014).

[20] S. A. Sevier, D. A. Kessler, and H. Levine, Mechanical bounds to transcriptional noise, Proceedings of the National Academy of Sciences **113**, 13983 (2016).

[21] J. E. Carroll, *Rate Equations in Semiconductor Electronics* (Cambridge University Press, 1986).

[22] D. Vorberg, W. Wustmann, R. Ketzmerick, and A. Eckardt, Generalized Bose-Einstein condensation into multiple states in driven-dissipative systems, Physical Review Letters **111**, 240405 (2013).

[23] P. Prelovšek, J. Bonča, and M. Mierzejewski, Transient and persistent particle subdiffusion in a disordered chain coupled to bosons, Phys. Rev. B **98**, 125119 (2018).

[24] M. Mierzejewski, P. Prelovšek, and J. Bonča, Einstein relation for a driven disordered quantum chain in the subdiffusive regime, Phys. Rev. Lett. **122**, 206601 (2019).

[25] M. P. Liul and S. N. Shevchenko, Rate-equation approach for multi-level quantum systems, Low Temperature Physics **49**, 96 (2023).

[26] J. Knebel, M. F. Weber, T. Krüger, and E. Frey, Evolutionary games of condensates in coupled birth–death processes, Nature Communications **6**, 6977 (2015).

[27] C. E. Overton, M. Broom, C. Hadjichrysanthou, and K. J. Sharkey, Methods for approximating stochastic evolutionary dynamics on graphs, Journal of Theoretical Biology **468**, 45 (2019).

[28] T. E. Harris, Contact interactions on a lattice, The Annals of Probability **2**, 969 (1974).

[29] J. Marro and R. Dickman, *Nonequilibrium Phase Transitions in Lattice Models* (Cambridge University Press, 1999).

[30] M. Henkel, H. Hinrichsen, and S. Lübeck, *Non-Equilibrium Phase Transitions: Volume I: Absorbing Phase Transitions* (Springer, 2008).

[31] F. Spitzer, Interaction of Markov processes, Advances in Mathematics **5**, 246 (1970).

[32] M. R. Evans and T. Hanney, Nonequilibrium statistical mechanics of the zero-range process and related models, Journal of Physics A: Mathematical and General **38**, R195 (2005).

[33] H. Spohn, *Large Scale Dynamics of Interacting Particles* (Springer, 1991).

[34] G. Schütz and E. Domany, Phase transitions in an exactly soluble one-dimensional exclusion process, Journal of Statistical Physics **72**, 277 (1993).

[35] B. Derrida, An exactly soluble non-equilibrium system: The asymmetric simple exclusion process, Physics Reports **301**, 65 (1998).

[36] T. M. Liggett, *Stochastic Interacting Systems: Contact, Voter and Exclusion Processes*, Vol. 324 (Springer, 1999).

[37] G. M. Schütz, Exactly solvable models for many-body systems far from equilibrium, Phase Transitions and Critical Phenomena **19**, 1 (2001).

[38] O. Golinelli and K. Mallick, The asymmetric simple exclusion process: an integrable model for non-equilibrium statistical mechanics, Journal of Physics A: Mathematical and General **39**, 12679 (2006).

[39] T. Chou, K. Mallick, and R. K. P. Zia, Non-equilibrium statistical mechanics: from a paradigmatic model to biological transport, Reports on Progress in Physics **74**, 116601 (2011).

[40] W. Tarnowski, I. Yusipov, T. Laptyeva, S. Denisov, D. Chruściński, and K. Życzkowski, Random generators of Markovian evolution: A quantum-classical transition by superdecoherence, Physical Review E **104**, 034118 (2021).

[41] E. P. Wigner, Results and theory of resonance absorption, in *Conference on neutron physics by time-of-flight* (1956) pp. 1–2.

[42] E. P. Wigner, Characteristic vectors of bordered matrices with infinite dimensions II (Annals of Mathematics, 1957) pp. 203–207.

[43] F. J. Dyson, Statistical theory of the energy levels of complex systems. I, Journal of Mathematical Physics **3**, 140 (1962).

[44] F. J. Dyson, Statistical theory of the energy levels of complex systems. II, Journal of Mathematical Physics **3**, 157 (1962).

[45] F. J. Dyson, Statistical theory of the energy levels of complex systems. III, Journal of Mathematical Physics **3**, 166 (1962).

[46] F. Haake, *Quantum Signatures of Chaos*, Vol. 54 (Springer Berlin Heidelberg, 2010).

[47] H.-J. Stöckmann, *Quantum Chaos* (Cambridge University Press, 1999).

[48] D. Braun, *Dissipative Quantum Chaos and Decoherence* (Springer, 2001).

[49] F. Borgonovi, F. Izrailev, L. Santos, and V. Zelevinsky, Quantum chaos and thermalization in isolated systems of interacting particles, Physics Reports **626**, 1 (2016).

[50] P. Roushan, C. Neill, J. Tangpanitanon, V. M. Bastidas, A. Megrant, R. Barends, Y. Chen, Z. Chen, B. Chiaro, A. Dunsworth, A. Fowler, B. Foxen, M. Giustina, E. Jeffrey, J. Kelly, E. Lucero, J. Mutus, M. Neeley, C. Quintana, D. Sank, A. Vainsencher, J. Wenner, T. White, H. Neven, D. G. Angelakis, and J. Martinis, Spectroscopic signatures of localization with interacting photons in superconducting qubits, Science **358**, 1175 (2017).

[51] S. Denisov, T. Laptyeva, W. Tarnowski, D. Chruściński, and K. Życzkowski, Universal spectra of random Lindblad operators, Phys. Rev. Lett. **123**, 140403 (2019).

[52] T. Can, Random Lindblad dynamics, Journal of Physics A: Mathematical and Theoretical **52**, 485302 (2019).

[53] S. Lange and C. Timm, Random-matrix theory for the Lindblad master equation, Chaos: An Interdisciplinary Journal of Nonlinear Science **31**, 023101 (2021).

[54] L. Sá, P. Ribeiro, and T. Prosen, Spectral and steady-state properties of random Liouvillians, Journal of Physics A: Mathematical and Theoretical **53**, 305303 (2020).

[55] C. Timm, Random transition-rate matrices for the master equation, Phys. Rev. E **80**, 021140 (2009).

[56] C. Bordenave, P. Caputo, and D. Chafaï, Spectrum of Markov generators on sparse random graphs, Communications on Pure and Applied Mathematics **67**, 621 (2014).

[57] L.-H. Gwa and H. Spohn, Bethe solution for the dynamical-scaling exponent of the noisy Burgers equation, Physical Review A **46**, 844 (1992).

[58] D. Kim, Bethe ansatz solution for crossover scaling functions of the asymmetric XXZ chain and the Kardar-Parisi-Zhang-type growth model, Physical Review E **52**, 3512 (1995).

[59] O. Golinelli and K. Mallick, Bethe ansatz calculation of the spectral gap of the asymmetric exclusion process, Journal of Physics A: Mathematical and General **37**, 3321 (2004).

[60] O. Golinelli and K. Mallick, Spectral gap of the totally asymmetric exclusion process at arbitrary filling, Journal of Physics A: Mathematical and General **38**, 1419 (2005).

[61] L. Sá, P. Ribeiro, and T. Prosen, Complex spacing ratios: A signature of dissipative quantum chaos, Physical Review X **10**, 021019 (2020).

[62] B. Bollobás, *Random Graphs* (Cambridge University Press, 2001).

[63] J. de Gier and F. H. L. Essler, Exact spectral gaps of the asymmetric exclusion process with open boundaries, Journal of Statistical Mechanics: Theory and Experiment **2006**, P12011 (2006).

[64] J. de Gier and F. H. L. Essler, Slowest relaxation mode of the partially asymmetric exclusion process with open boundaries, Journal of Physics A: Mathematical and Theoretical **41**, 485002 (2008).

[65] J. de Gier, C. Finn, and M. Sorrell, The relaxation rate of the reverse-biased asymmetric exclusion process, Journal of Physics A: Mathematical and Theoretical **44**, 405002 (2011).

[66] S. Prolhac, Spectrum of the totally asymmetric simple exclusion process on a periodic lattice-first excited states, Journal of Physics A: Mathematical and Theoretical **47**, 375001 (2014).

[67] S. Prolhac, Perturbative solution for the spectral gap of the weakly asymmetric exclusion process, Journal of Physics A: Mathematical and Theoretical **50**, 315001 (2017).

[68] M. Henkel and U. Schollwöck, Universal finite-size scaling amplitudes in anisotropic scaling, Journal of Physics A: Mathematical and General **34**, 3333 (2001).

[69] J. Hooyberghs, F. Iglói, and C. Vanderzande, Absorbing state phase transitions with quenched disorder, Phys. Rev. E **69**, 066140 (2004).

[70] E. Abbe, A. S. Bandeira, A. Bracher, and A. Singer, Decoding binary node labels from censored edge measurements: Phase transition and efficient recovery, IEEE Transactions on Network Science and Engineering **1**, 10 (2014).

[71] A. S. Bandeira, Random Laplacian matrices and convex relaxations, Foundations of Computational Mathematics **18**, 345 (2018).

[72] C. Delorme and S. Poljak, Laplacian eigenvalues and the maximum cut problem, Mathematical Programming **62**, 557 (1993).

[73] M. X. Goemans and D. P. Williamson, Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming, Journal of the ACM **42**, 1115 (1995).

[74] L. De Haan and A. Ferreira, *Extreme Value Theory: an Introduction* (Springer, 2006).

[75] P. Embrechts, C. Klüppelberg, and T. Mikosch, *Modelling Extremal Events*, Vol. 33 (Springer, 1997).

[76] O. Perron, Zur Theorie der Matrices, Mathematische Annalen 1907 64:2 **64**, 248 (1907).

[77] G. Frobenius, Über Matrizen aus nicht negativen Elementen, Sitzungsber. Königl. Preuss. Akad. Wiss. (1912).

[78] J. Keizer, On the solutions and the steady states of a master equation, Journal of Statistical Physics **6**, 67 (1972).

[79] B. Pittel, Counting strongly connected (k1, k2)-directed cores, Random Structures & Algorithms **53**, 3 (2018).

[80] R. A. Horn and C. R. Johnson, *Matrix Analysis* (Cambridge University Press, 1985).

[81] A. Hoorfar and M. Hassani, Inequalities on the Lambert W function and hyperpower function, J. Inequal. Pure and Appl. Math **9**, 5 (2008).

[82] V. Oganesyan and D. A. Huse, Localization of interacting fermions at high temperature, Phys. Rev. B **75**, 155111 (2007).

[83] Y. Y. Atas, E. Bogomolny, O. Giraud, and G. Roux, Distribution of the ratio of consecutive level spacings in random matrix ensembles, Phys. Rev. Lett. **110**, 084101 (2013).

[84] R. Hamazaki, K. Kawabata, N. Kura, and M. Ueda, Universality classes of non-Hermitian random matrices, Phys. Rev. Research **2**, 023286 (2020).

[85] G. Casati, L. Molinari, and F. Izrailev, Scaling properties of band random matrices, Physical Review Letters **64**, 1851 (1990).

[86] G. Casati, F. Izrailev, and L. Molinari, Scaling properties of the eigenvalue spacing distribution for band random matrices, Journal of Physics A: Mathematical and General **24**, 4755 (1991).

[87] Y. V. Fyodorov and A. D. Mirlin, Scaling properties of localization in random band matrices: A σ-model approach, Physical Review Letters **67**, 2405 (1991).

[88] Y. V. Fyodorov and A. D. Mirlin, Statistical properties of random banded matrices with strongly fluctuating diagonal elements, Physical Review B **52**, R11580 (1995).

[89] S. Sodin, The spectral edge of some random band matrices, Annals of Mathematics **172**, 2223 (2010).

[90] L. Erdős and A. Knowles, Quantum diffusion and eigenfunction delocalization in a random band matrix model, Communications in Mathematical Physics **303**, 509 (2011).

[91] L. Erdős, A. Knowles, H.-T. Yau, and J. Yin, Delocalization and diffusion profile for random band matrices, Communications in Mathematical Physics **323**, 367 (2013).

[92] L. Erdős, A. Knowles, and H.-T. Yau, Averaging fluctuations in resolvents of random band matrices, Annales Henri Poincaré **14**, 1837 (2013).

[93] T. Spencer, Random banded and sparse matrices (Oxford University Press, 2015) pp. 470–488.

[94] P. Bourgade, Random band matrices, in *Proceedings of the International Congress of Mathematicians: Rio de Janeiro 2018* (World Scientific, 2018) pp. 2759–2783.

[95] A. D. Mirlin, Y. V. Fyodorov, F.-M. Dittes, J. Quezada, and T. H. Seligman, Transition from localized to extended eigenstates in the ensemble of power-law random banded matrices, Physical Review E **54**, 3221 (1996).

[96] Y. V. Fyodorov, A. Ossipov, and A. Rodriguez, The anderson localization transition and eigenfunction multifractality in an ensemble of ultrametric random matrices, Journal of Statistical Mechanics: Theory and Experiment **2009**, L12001 (2009).

[97] F. Mosam, D. Vidaurre, and E. D. Giuli, Breakdown of random matrix universality in Markov models, Physical Review E **104**, 024305 (2021).

[98] N. Lehmann and H.-J. Sommers, Eigenvalue statistics of random real matrices, Physical Review Letters **67**, 941 (1991).

[99] A. Edelman, E. Kostlan, and M. Shub, How many eigenvalues of a random matrix are real?, Journal of the American Mathematical Society **7**, 247 (1994).

[100] A. Edelman, The probability that a random real Gaussian matrix has k real eigenvalues, related distributions, and the circular law, Journal of Multivariate Analysis **60**, 203 (1997).

[101] E. Kanzieper and G. Akemann, Statistics of real eigenvalues in Ginibre's ensemble of random real matrices, Physical Review Letters **95**, 230201 (2005).

[102] P. J. Forrester and T. Nagao, Eigenvalue statistics of the real Ginibre ensemble, Physical Review Letters **99**, 050603 (2007).

[103] H.-J. Sommers and W. Wieczorek, General eigenvalue correlations for the real Ginibre ensemble, Journal of Physics A: Mathematical and Theoretical **41**, 405003 (2008).

[104] B. A. Khoruzhenko, H.-J. Sommers, and K. Życzkowski, Truncations of random orthogonal matrices, Physical Review E **82**, 040106 (2010).

[105] T. Tao and V. Vu, Random matrices: Universality of local spectral statistics of non-Hermitian matrices, The Annals of Probability **43**, 782 (2015).

[106] W. Tarnowski, Real spectra of large real asymmetric random matrices, Physical Review E **105**, L012104 (2022).

[107] S.-S. Byun, N.-G. Kang, J. O. Lee, and J. Lee, Real eigenvalues of elliptic random matrices, International Mathematics Research Notices **2023**, 2243 (2023).

[108] A. J. Bray and G. J. Rodgers, Diffusion in a sparsely connected space: A model for glassy relaxation, Physical Review B **38**, 11461 (1988).

[109] G. J. Rodgers and A. J. Bray, Density of states of a sparse random matrix, Physical Review B **37**, 3557 (1988).

[110] A. Cavagna, I. Giardina, and G. Parisi, Analytic computation of the instantaneous normal modes spectrum in low-density liquids, Physical Review Letters **83**, 108 (1999).

[111] R. Kühn, Spectra of sparse random matrices, Journal of Physics A: Mathematical and Theoretical **41**, 295002 (2008).

[112] V. A. R. Susca, P. Vivo, and R. Kühn, Cavity and replica methods for the spectral density of sparse symmetric random matrices, SciPost Physics Lecture Notes **33**, 33 (2021).

[113] G. Biroli and R. Monasson, A single defect approximation for localized states on random lattices, Journal of Physics A: Mathematical and General **32**, L255 (1999).

[114] G. Semerjian and L. F. Cugliandolo, Sparse random matrices: the eigenvalue spectrum revisited, Journal of Physics A: Mathematical and General **35**, 303 (2002).

[115] S. N. Dorogovtsev, A. V. Goltsev, J. F. F. Mendes, and A. N. Samukhin, Spectra of complex networks, Physical Review E **68**, 046109 (2003).

[116] A. D. Mirlin and Y. V. Fyodorov, Universality of level correlation function of sparse random matrices, Journal of Physics A: Mathematical and General **24**, 2273 (1991).

[117] P. Akara-Pipattana and O. Evnin, Random matrices with row constraints and eigenvalue distributions of graph Laplacians 10.48550/arxiv.2212.06499 (2022).

[118] T. Rogers, I. P. Castillo, R. Kühn, and K. Takeda, Cavity approach to the spectral density of sparse symmetric random matrices, Physical Review E **78**, 031116 (2008).

[119] T. Rogers and I. P. Castillo, Cavity approach to the spectral density of non-Hermitian sparse matrices, Physical Review E **79**, 012101 (2009).

[120] V. A. R. Susca, P. Vivo, and R. Kühn, Top eigenpair statistics for weighted sparse graphs, Journal of Physics A: Mathematical and Theoretical **52**, 485002 (2019).

[121] R. G. Margiotta, R. Kühn, and P. Sollich, Glassy dynamics on networks: local spectra and return proba-

bilities, Journal of Statistical Mechanics: Theory and Experiment **2019**, 093304 (2019).

[122] D. Tapias, E. Paprotzki, and P. Sollich, From entropic to energetic barriers in glassy dynamics: the Barrat–Mézard trap model on sparse networks, Journal of Statistical Mechanics: Theory and Experiment **2020**, 093302 (2020).

[123] D. Tapias and P. Sollich, Localization properties of the sparse Barrat-Mézard trap model, Physical Review E **105**, 054109 (2022).

[124] I. Palásti, On the strong connectedness of directed random graphs, Studia Sci. Math. Hungar **1**, 205 (1966).

[125] A. J. Graham and D. A. Pike, A note on thresholds and connectivity in random directed graphs, Atl. Electron. J. Math **3**, 1 (2008).

[126] J. Watrous, The Theory of Quantum Information (Cambridge University Press, 2018).

[127] G. Birkhoff, Tres observaciones sobre el algebra lineal, Univ. Nac. Tucuman, Ser. A **5**, 147 (1946).

[128] L. Landau and R. Streater, On Birkhoff's theorem for doubly stochastic completely positive maps of matrix algebras, Linear Algebra and its Applications **193**, 107 (1993).

[129] K. Życzkowski, personal communication.

[130] T. Can, V. Oganesyan, D. Orgad, and S. Gopalakrishnan, Spectral gaps and midgap states in random quantum master equations, Physical Review Letters **123**, 234103 (2019).