Journal of Physics A:
Mathematical and
Theoretical

Top eigenpair statistics of diluted Wishart matrices

OPEN ACCESS
IOP Publishing
Journal of Physics A: Mathematical and Theoretical
J. Phys. A: Math. Theor. **58** (2025) 325001 (30pp)
https://doi.org/10.1088/1751-8121/add821

Top eigenpair statistics of diluted Wishart matrices

Barak Budnick[1], Preben Forer[2], Pierpaolo Vivo[2,*], Sabrina Aufiero[3], Silvia Bartolucci[3] and Fabio Caccioli[3,4,5]

[1] Racah Institute of Physics, The Hebrew University, Jerusalem 9190401, Israel
[2] Department of Mathematics, King's College London, Strand, London WC2R 2LS, United Kingdom
[3] Department of Computer Science, University College London, 66-72 Gower Street, London WC1E 6EA, United Kingdom
[4] Systemic Risk Centre, London School of Economics and Political Sciences, London WC2A 2AE, United Kingdom
[5] London Mathematical Laboratory, 8 Margravine Gardens, London WC 8RH, United Kingdom

E-mail: pierpaolo.vivo@kcl.ac.uk

Received 5 February 2025; revised 25 April 2025
Accepted for publication 13 May 2025
Published 6 August 2025

Abstract

Using the replica method, we compute the statistics of the top eigenpair of diluted covariance matrices of the form $\boldsymbol{J} = \boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}$, where $\boldsymbol{X}$ is a $N \times M$ sparse data matrix, in the limit of large N, M with fixed ratio and a bounded number of non-zero entries. We allow for random non-zero weights, provided they lead to an isolated largest eigenvalue. By formulating the problem as the optimisation of a quadratic Hamiltonian constrained to the N-sphere at low temperature, we derive a set of recursive distributional equations for auxiliary probability density functions, which can be efficiently solved using a population dynamics algorithm. The average largest eigenvalue is identified with a Lagrange parameter that governs the convergence of the algorithm, and the resulting stable populations are then used to evaluate the density of the top eigenvector's

* Author to whom any correspondence should be addressed.

components. We find excellent agreement between our analytical results and numerical results obtained from direct diagonalisation.

Keywords: eigenvalues, statistics, Wishart, random matrix theory, diluted covariance matrices

1. Introduction

In recent decades, we have witnessed an unprecedented surge in the amount of information available for processing and forecasting, marking the emergence of the Big Data era. Contemporary data analysis challenges frequently involve processing datasets with numerous variables and observations. This high-dimensional nature of data is particularly evident in fields such as climate studies, genetics, biomedical imaging, and economics [1].

Consider a scenario where one conducts N measurements of M variables that characterise a system. These variables might represent, for instance, assets in a stock market or a collection of climate observables, with measurements taken simultaneously at N different time points. The collected data can be organised into an $N \times M$ matrix $\boldsymbol{X}$, where element X_{ij} represents the ith measurement of the jth variable. From this, we construct the $M \times M$ sample covariance matrix $\boldsymbol{J} = \boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}$, which encodes all possible correlations among the variables. This covariance matrix plays a fundamental role in multivariate statistical analysis, finding applications in dimensional reduction and classification procedures, such as principal component analysis [2] and linear discriminant analysis [3].

A reasonable assumption for many natural phenomena is that each variable exhibits significant correlation with only a limited subset of other variables, resulting in sparse covariance matrices characterised by numerous entries that are either very small or zero. This sparsity is particularly relevant in inferring causal influences among system components from empirical covariance matrices. Notable examples include the experimental reconstruction of interactions in biological systems, such as cellular signalling networks [4], gene regulatory networks [5, 6], and ecological association networks [7, 8]. Similar sparse structures also emerge in other fields: in natural language processing, where word co-occurrence matrices reveal correlations between contextually related words [9]; in finance, where asset correlations tend to cluster within sectors [10]; and in social networks, where relationships between users are captured by sparse covariance matrices [11]. Additionally, working with large, dense covariance matrices is computationally demanding, often requiring regularisation techniques that induce sparsity and improve efficiency [12].

One of the most important observables in the case of random covariance matrix is the top eigenvalue and its associated eigenvector. For instance, in Principal Component Analysis the top eigenvalue and eigenvector capture the most significant variability in data, enabling dimensionality reduction and assisting in signal detection [13–20].

In this paper, we build on the works [21–24] to formulate a replica approach that is well suited to the average largest eigenvalue and the density of its associated eigenvector's components of diluted Wishart matrices. We allow for a large class of weights on non-zero entries that lead to an *isolated* top eigenvalue (see below for more details).

The outline of the paper is as follows. In section 2 we review the relevant literature; in section 3, we formulate the problem, introduce notations, and specify our assumptions; in section 4, we use the replica formalism to compute the largest eigenvalue of sparse random matrices of the form $\boldsymbol{J} = \boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}$; in section 5, we build on the results of the previous section to compute the density of the corresponding top eigenvector components; in section 6, we discuss the population dynamics algorithm used to solve the system of self-consistent equations; in

section 7 we show how taking the dense limit recovers the known noncentral Wishart-ensemble results; finally, in section 8, we summarise our results and conclusions; appendix A provides an upper bound for the largest eigenvalue, which may serve as a suitable starting point for the population dynamics algorithm, while appendix B is devoted to the calculation of a technical average.

2. Literature review

Since Wishart's pioneering work [25], random matrix theory has played a fundamental role in multivariate statistics [26]. Results derived from random matrix models serve as crucial benchmarks for comparison with empirical data. A central focus of this field is the study of eigenvalue and eigenvector statistics, which provide insight into correlations and principal components in complex data.

A common null model for the covariance matrix $\boldsymbol{J} = \boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}$ assumes independent Gaussian random variables, adjusted to have zero mean, as entries of $\boldsymbol{X}$. This model yields an analytically known joint distribution of eigenvalues, completely decoupled from the distribution of eigenvectors, enabling the application of the Coulomb gas technique in the large N, M limit with their ratio fixed [27–29]. This approach has led to extensive results on the eigenvalue statistics of dense covariance matrices [13, 30–33], including a detailed characterisation of both typical and atypical eigenvalue fluctuations [13]. The eigenvectors of this rotationally invariant model are Haar-distributed over the sphere [34], and their associated components follow the Porter-Thomas distribution [35].

While the eigenvalue and eigenvector statistics of dense covariance matrices are well understood, the situation is markedly different for sparse ('diluted') covariance matrices, where many entries are zero. Analytical results in this case are primarily limited to the average spectral density [21, 36] and the number of eigenvalues in a given interval [37]. A key challenge is the absence of an analytical expression for the joint eigenvalue distribution, as the loss of rotational invariance precludes the use of the Coulomb gas approach and other techniques based on orthogonal polynomials [35] or Fredholm determinants and Painlevé transcendents [38]. While novel methods have expanded our understanding of sparse random matrices [36, 39–47], the analytical framework remains less developed compared to the 'classical' dense case.

A particularly important aspect of covariance matrix spectra is the behaviour of the largest eigenvalue and its associated eigenvector, which serves as a key indicator of system-wide correlations. In the *dense* regime, significant progress has been made in characterising the largest eigenvalue distribution. Several works [48–52] have established that, under fairly general conditions, the largest eigenvalue follows the Tracy–Widom law, demonstrating a form of strong universality. The statistics of eigenvectors of non-centred and doubly correlated Gaussian random matrices has been tackled in [53] using a supersymmetric technique.

In contrast, in the *sparse* regime, the largest eigenvalue may exhibit a fundamentally different behaviour, and elementary results remain relatively scarce. While some progress has been made-such as a local Tracy–Widom law for sparse covariance matrices with zero-mean entries [54] and studies on heavy-tailed distributions revealing deviations from classical universality [55] — many open questions remain. In particular, for matrices of the form $\boldsymbol{X}^{\mathrm{T}}\boldsymbol{\Sigma}\boldsymbol{X}$, where $\boldsymbol{\Sigma}$ introduces non-uniform sample couplings, the largest eigenvalue can separate from the bulk spectrum. In the dense regime, this phenomenon corresponds to the well-known *BBP transition* [56], in which the structure of $\boldsymbol{\Sigma}$ drives the detachment. Interestingly, even in the null case, where $\boldsymbol{\Sigma} = \mathbb{1}$, a spectral gap can emerge if the entries of $\boldsymbol{X}$ have a non-zero mean [57, 58]. In

this scenario, the nonzero mean entries can be interpreted as a deterministic signal, to which Gaussian noise is added in the form of a zero-mean random covariance matrix. Under this interpretation, the detachment of the largest eigenvalue marks the point at which the noise level becomes low enough to allow inference of the principal component of the original data. Correspondingly, the detachment is accompanied by a similar transition in its associated eigenvector, which becomes localised on a cone whose axis aligns with the principal component of the deterministic signal [58]. In the sparse regime, a qualitatively similar behaviour is observed, but a precise quantitative characterisation of this transition remains an open problem.

To study spectral properties of large random matrices, various analytical techniques have been developed. Originally introduced in the context of spin glasses [59, 60], the replica method was first applied to random matrices by Edwards and Jones [61] to compute the spectral density of dense matrices. This approach, which relies on the joint distribution of matrix entries rather than eigenvalues, was later extended by Bray and Rodgers [62] to derive an expression for the spectral density of sparse Erdős–Rényi adjacency matrices. However, solving the resulting integral equations remains challenging, with numerical progress made only recently [63].

Alternative functional methods, such as the single defect approximation and effective medium approximation [64, 65], have been developed to tackle these problems. In the context of sparse covariance matrices, these approaches were used in [21] to compute the spectral density. Another promising line of research builds on the replica-symmetric framework of Bray and Rodgers, representing order parameters as continuous superpositions of Gaussians with fluctuating variances [39, 66]. This method was recently applied in [22, 23] to study the typical largest eigenvalue of sparse weighted graphs, leading to non-linear integral equations that can be efficiently solved using a population dynamics algorithm.

These techniques, originally developed in [39, 67, 68], have since found widespread use in random matrix theory [69–72], providing a powerful framework for analysing spectral properties beyond the classical setting.

3. Formulation of the problem

Consider the sparse $N \times M$ matrix $\boldsymbol{X}$, whose entries, X_{ij}, are random variables, defined as

$$X_{ij} = c_{ij} K_{ij} \, . \tag{1}$$

Here, $c_{ij} \in \{0, 1\}$ regulate the density of non-zero elements of $\boldsymbol{X}$ and K_{ij} represent the non-zero elements' weights, randomly drawn from the pdf $p(K)$. The central object of this study is the $M \times M$ symmetric matrix

$$\boldsymbol{J} = \boldsymbol{X}^{\mathrm{T}} \boldsymbol{X} \, . \tag{2}$$

We work in the regime $N \to \infty$ and $M \to \infty$, but with the ratio

$$\alpha = \sqrt{\frac{N}{M}} \tag{3}$$

kept finite. According to the spectral theorem, the symmetric matrix $\boldsymbol{J}$ can be diagonalised via an orthonormal basis of eigenvectors, $\{\boldsymbol{v}_m\}_{m=1}^{M} \in \mathbb{R}^M$, whose corresponding real eigenvalues are denoted by $\{\lambda_m\}_{m=1}^{M}$. Assuming that the real eigenvalues are not degenerate, we can sort them as $\lambda_1 > \lambda_2 > \ldots > \lambda_M$. The main goal of this work is to evaluate

- The typical value of the largest of them, denoted by $\langle\lambda_1\rangle$ — assumed to be $\sim\mathcal{O}(1)$ in the limit.
- The density of its corresponding eigenvector's components, $T(u)=\left\langle\frac{1}{M}\sum_{i=1}^{M}\delta\left(u-v_1^{(i)}\right)\right\rangle$,

where $\langle\cdot\rangle$ stands for averaging over different realisations of $\boldsymbol{X}$.

By analogy to the standard prototype of sparse random systems (the Erdős–Rényi graph), the model we study is defined by the following probability to draw (independently) the matrix entries X_{ij}

$$P(X_{ij})=\left[\frac{q}{\sqrt{NM}}\delta_{c_{ij},1}+\left(1-\frac{q}{\sqrt{NM}}\right)\delta_{c_{ij},0}\right]p(K_{ij})\ . \tag{4}$$

Indeed, in graph-theoretical terms [73], the random matrix $\boldsymbol{X}$ can be interpreted as the weighted adjacency matrix of a Poissonian bipartite random graph with two distinct node types [36]: i-nodes, corresponding to the rows of $\boldsymbol{X}$, and j-nodes, corresponding to its columns. The matrix $\boldsymbol{X}$ is sparse in the sense that the average number q/α of its non-zero elements per row does not scale with either N or M.

The model defined in (4) suffers from two potential drawbacks, though: (i) without further restrictions on the *maximal* number of nonzero elements allowed in each row and column, the largest eigenvalue may (slowly) grow with N,M, in contrast with our assumption that $\langle\lambda_1\rangle\sim\mathcal{O}(1)$; and (ii) a general and unrestricted weight distribution $p(K)$ may lead to a largest eigenvalue that is not detached from the continuous bulk of the spectrum.

The concern (i) follows from the observation that — in the similar case of (square) adjacency matrices of sparse random graphs — the largest eigenvalue (without further restrictions) indeed grows (slowly) with N, as proven in [74]. Although we are not aware of a similar theorem in the context of diluted correlation matrices, it is a plausible assumption that a similar mechanism may be at work here. To ensure that the largest eigenvalue remains $\mathcal{O}(1)$, we therefore impose the constraints that there be at most C non-zero elements per column and R per row. In appendix A we show that such a constraint indeed results in an $\mathcal{O}(1)$ upper bound for $\langle\lambda_1\rangle$.

The concern (ii) can be allayed more easily by assuming that the weight pdf $p(K)$ is such that the largest eigenvalue is isolated, i.e. there is a macroscopic gap between it and the sea of smaller eigenvalues. This detachment also occurs in the dense regime, where precise relationships between $p(K)$ and the resulting spectral gap can be established [57] — notably, a necessary (though not sufficient) condition is that $p(K)$ has a non-zero mean. The sparse regime exhibits a similar qualitative behaviour, although a complete analytical characterisation of the transition remains an open problem. Throughout the rest of the paper, when referring to $p(K)$ as a 'nonzero mean distribution', we specifically mean it in the sense of it generating a spectral gap.

While the restriction on the maximal number of nonzero elements in $\boldsymbol{X}$ imposes non-trivial couplings between the c_{ij}'s, which would in principle require a re-working of the form (4) of the pdf of entries and introduce an additional analytical burden, we benefit here from a key observation made in [22, 39, 75]: a convenient shortcut for the calculation consists in (i) initially replacing the 'microcanonical' version of the model with the simpler 'canonical' one, in which the c_{ij}'s are *independent* Bernoulli random variables with success probability $q/\sqrt{NM}$ (as given in (4)), and (ii) manually adjusting the Poissonian distribution of 'degrees' of the connectivity matrix — which naturally emerges in the replica calculation — to allow for a finite maximum number of nonzero entries per row and column (see, e.g. equations (55)–(57)

and discussion after equation (68)). For all the technical details that motivate this shortcut, we refer to appendix B in [22].

In the next section, we provide a detailed analysis of the replica calculation for the typical largest eigenvalue of diluted Wishart matrices.

4. Replica analysis of the typical largest eigenvalue

We begin our analysis by noting that the problem of evaluating $\boldsymbol{J}$'s largest eigenvalue can be formulated in terms of the Courant–Fisher maximisation

$$\lambda_1 = \frac{1}{M} \max_{\boldsymbol{v}\in\mathbb{R}^M,\ |\boldsymbol{v}|^2=M} \langle \boldsymbol{v}, \boldsymbol{J}\boldsymbol{v}\rangle \ , \tag{5}$$

where $\langle\cdot,\cdot\rangle$ stands for the standard dot product among vectors in $\mathbb{R}^M$. We now introduce an auxiliary canonical partition function at inverse temperature β

$$Z = \int \mathrm{d}\boldsymbol{v}\ \exp\left(\frac{\beta}{2}\langle \boldsymbol{v}, \boldsymbol{J}\boldsymbol{v}\rangle\right) \delta\left(|\boldsymbol{v}|^2 - M\right) \ . \tag{6}$$

In the zero-temperature limit $\beta \to \infty$, applying the Laplace method to the integral in (6) we obtain using (5) that

$$Z \approx \exp\left(\frac{\beta}{2} \max_{\boldsymbol{v}\in\mathbb{R}^M,\ |\boldsymbol{v}|^2=M} \langle \boldsymbol{v}, \boldsymbol{J}\boldsymbol{v}\rangle\right) = \exp\left(\frac{\beta}{2} M\lambda_1\right) \ . \tag{7}$$

Therefore

$$\langle\lambda_1\rangle = \lim_{\beta\to\infty} \frac{2}{\beta M} \langle \ln Z\rangle \ . \tag{8}$$

To tackle the average on the r.h.s of equation (8) we invoke the replica trick [60]

$$\langle\lambda_1\rangle = \lim_{\beta\to\infty} \frac{2}{\beta M} \lim_{n\to 0} \frac{1}{n} \ln \langle Z^n\rangle \ , \tag{9}$$

where n is initially treated as an integer, and then analytically continued to real values around $n = 0$. The next section is devoted to computing the average of the replicated partition function $\langle Z^n\rangle$.

The evaluation of (9) proceeds through five principal steps: (1) we average the replicated partition function over the disorder (randomness encoded in the matrix entries X_{ij}); (2) we reformulate the replicated partition function in terms of a functional integral, which lends itself to a suitable form for a saddle point analysis; (3) to study the saddle point structure of the replicated partition function, we then invoke a replica symmetric ansatz that recasts the order parameters as a superposition of an uncountably infinite set of Gaussians; (4) we use the replica symmetric ansatz to derive the saddle point equations, and finally, (5) we evaluate the replicated partition function at the saddle point in the limit $\beta \to \infty$.

We will guide the reader through the various steps below.

4.1. Averaging the replicated partition function over the disorder

The first step in our analysis is to write the replicated partition function,

$$\langle Z^n \rangle = \int \left(\prod_{a=1}^{n} \mathrm{d}\boldsymbol{v}_a \right) \left\langle \exp\left[\frac{\beta}{2} \sum_{a=1}^{n} \sum_{i,k=1}^{M} v_{ia} J_{ik} v_{ka} \right] \right\rangle \prod_{a=1}^{n} \delta\left(|\boldsymbol{v}_a|^2 - M \right) , \tag{10}$$

as an integral of an exponential function. Expressing $J_{ik} = \sum_{j=1}^{N} X_{ji} X_{jk}$, we note that

$$\begin{aligned}
\left\langle \exp\left[\frac{\beta}{2} \sum_{a=1}^{n} \sum_{i,k=1}^{M} v_{ia} J_{ik} v_{ka} \right] \right\rangle &= \left\langle \exp\left[\frac{\beta}{2} \sum_{a=1}^{n} \sum_{i,k=1}^{M} \sum_{j=1}^{N} X_{ji} v_{ia} X_{jk} v_{ka} \right] \right\rangle \\
&= \left\langle \prod_{a=1}^{n} \prod_{j=1}^{N} \exp\left[\frac{\beta}{2} \left(\sum_{i=1}^{M} v_{ia} X_{ji} \right)^2 \right] \right\rangle \\
&= \left\langle \prod_{a=1}^{n} \prod_{j=1}^{N} \sqrt{\frac{\beta}{2\pi}} \int \mathrm{d}u \exp\left(-\frac{\beta}{2} u^2 + \beta u \sum_{i=1}^{M} v_{ia} X_{ji} \right) \right\rangle \\
&= \left(\frac{\beta}{2\pi} \right)^{\frac{Nn}{2}} \int \left(\prod_{a=1}^{n} \prod_{j=1}^{N} \mathrm{d}u_{ja} \mathrm{e}^{-\frac{\beta}{2} u_{ja}^2} \right) \left\langle \prod_{i=1}^{M} \prod_{j=1}^{N} \exp\left(\beta X_{ji} \sum_{a=1}^{n} v_{ia} u_{ja} \right) \right\rangle ,
\end{aligned} \tag{11}$$

where we used the Hubbard–Stratonovich transformation,

$$\int_{-\infty}^{\infty} \mathrm{d}x \, \mathrm{e}^{-ax^2 + bx} = \sqrt{\frac{\pi}{a}} \mathrm{e}^{\frac{b^2}{4a}} . \tag{12}$$

In appendix B we show that using the sparsity condition, the average in equation (11) can be performed for large N, M as

$$\begin{aligned}
&\left\langle \prod_{i=1}^{M} \prod_{j=1}^{N} \exp\left(\beta X_{ji} \sum_{a=1}^{n} v_{ia} u_{ja} \right) \right\rangle \\
&\quad \simeq \exp\left\{ \frac{q}{\sqrt{NM}} \sum_{i=1}^{M} \sum_{j=1}^{N} \left[\left\langle \exp\left(\beta K \sum_{a=1}^{n} v_{ia} u_{ja} \right) \right\rangle_K - 1 \right] \right\} ,
\end{aligned} \tag{13}$$

where the average $\langle \cdot \rangle_K$ is over a single realisation of the random variable K drawn from $p(K)$, the weight distribution. Furthermore, we use the Fourier representation of the delta function,

$$\prod_{a=1}^{n} \delta\left(|\boldsymbol{v}_a|^2 - M \right) = \int_{-\infty}^{\infty} \left(\prod_{a=1}^{n} \frac{\beta}{2} \frac{\mathrm{d}\lambda_a}{2\pi} \right) \prod_{a=1}^{n} \exp\left[-\mathrm{i} \frac{\beta}{2} \lambda_a \left(\sum_{i=1}^{M} v_{ia}^2 - M \right) \right] , \tag{14}$$

such that equation (10) takes the form (ignoring pre-factors whose logarithm vanishes in the limit)

$$\langle Z^n\rangle \propto \int\left(\prod_{a=1}^{n}\mathrm{d}\boldsymbol{v}_a\mathrm{d}\boldsymbol{u}_a\mathrm{d}\lambda_a\right)\exp\left(-\frac{\beta}{2}\sum_{a=1}^{n}\sum_{j=1}^{N}u_{ja}^2\right)\exp\left(\mathrm{i}M\frac{\beta}{2}\sum_{a=1}^{n}\lambda_a\right)$$
$$\times\exp\left(-\mathrm{i}\frac{\beta}{2}\sum_{a=1}^{n}\sum_{i=1}^{M}\lambda_a v_{ia}^2\right)\exp\left\{\frac{q}{\sqrt{NM}}\sum_{i=1}^{M}\sum_{j=1}^{N}\left[\left\langle\exp\left(\beta K\sum_{a=1}^{n}v_{ia}u_{ja}\right)\right\rangle_K-1\right]\right\}. \tag{15}$$

Note that in (15), $\{\boldsymbol{u}_a\}_{a=1}^n\in\mathbb{R}^N$ and $\{\boldsymbol{v}_a\}_{a=1}^n\in\mathbb{R}^M$.

4.2. Functional integral representation

Next, we aim at expressing the replicated partition function through a functional integral over the following order parameters

$$\phi(\vec{v})=\frac{1}{M}\sum_{i=1}^{M}\prod_{a=1}^{n}\delta(v_a-v_{ia}) \tag{16}$$

$$\psi(\vec{u})=\frac{1}{N}\sum_{j=1}^{N}\prod_{a=1}^{n}\delta(u_a-u_{ja})\ , \tag{17}$$

where $\vec{v},\vec{u}\in\mathbb{R}^n$ are n-dimensional vectors in replica space. The order parameters were chosen as such since this approach will eventually lead to a symmetric representation of the replicated partition function under the duality transformation $\alpha\to 1/\alpha$. This symmetry reflects the simple fact that the matrix $\boldsymbol{J}=\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}$ shares its largest eigenvalue with its 'dual' $N\times N$ counterpart $\tilde{\boldsymbol{J}}=\boldsymbol{X}\boldsymbol{X}^{\mathrm{T}}$. This approach serves as a starting point for a functional scheme introduced in [21] for the analysis of the spectral density of $\boldsymbol{J}$.

To enforce the definitions given in equations (16) and (17) upon the replicated partition function, we multiply equation (15) by the functional-integral representations of the identity

$$1=\int M\mathcal{D}\phi\,\mathcal{D}\hat{\phi}\exp\left\{-\mathrm{i}\int\mathrm{d}\vec{v}\hat{\phi}(\vec{v})\left[M\phi(\vec{v})-\sum_{i=1}^{M}\prod_{a=1}^{n}\delta(v_a-v_{ia})\right]\right\} \tag{18}$$

$$1=\int N\mathcal{D}\psi\,\mathcal{D}\hat{\psi}\exp\left\{-\mathrm{i}\int\mathrm{d}\vec{u}\hat{\psi}(\vec{u})\left[N\psi(\vec{u})-\sum_{j=1}^{N}\prod_{a=1}^{n}\delta(u_a-u_{ja})\right]\right\}, \tag{19}$$

where $\mathrm{d}\vec{v}=\prod_{a=1}^{n}\mathrm{d}v_a$, and similarly $\mathrm{d}\vec{u}=\prod_{a=1}^{n}\mathrm{d}u_a$. This allows us to rewrite equation (15) as

$$\langle Z^n\rangle\propto\int\mathcal{D}\phi\,\mathcal{D}\hat{\phi}\mathcal{D}\psi\,\mathcal{D}\hat{\psi}\mathrm{d}\vec{\lambda}\exp\left[-\mathrm{i}M\int\mathrm{d}\vec{v}\hat{\phi}(\vec{v})\phi(\vec{v})-\mathrm{i}N\int\mathrm{d}\vec{u}\hat{\psi}(\vec{u})\psi(\vec{u})\right]$$
$$\times\exp\left[q\sqrt{NM}\int\mathrm{d}\vec{v}\mathrm{d}\vec{u}\phi(\vec{v})\psi(\vec{u})\left(\left\langle\mathrm{e}^{\beta K\vec{v}\cdot\vec{u}}\right\rangle_K-1\right)+\mathrm{i}M\frac{\beta}{2}\sum_{a=1}^{n}\lambda_a\right]$$
$$\times\int\left(\prod_{a=1}^{n}\mathrm{d}\boldsymbol{v}_a\right)\exp\left[-\frac{\beta}{2}\sum_{a=1}^{n}\sum_{i=1}^{M}\lambda_a v_{ia}^2+\mathrm{i}\sum_{i=1}^{M}\int\mathrm{d}\vec{v}\hat{\phi}(\vec{v})\prod_{a=1}^{n}\delta(v_a-v_{ia})\right]$$
$$\times\int\left(\prod_{a=1}^{n}\mathrm{d}\boldsymbol{u}_a\right)\exp\left[-\frac{\beta}{2}\sum_{a=1}^{n}\sum_{j=1}^{N}u_{ja}^2+\mathrm{i}\sum_{j=1}^{N}\int\mathrm{d}\vec{u}\hat{\psi}(\vec{u})\prod_{a=1}^{n}\delta(u_a-u_{ja})\right]. \tag{20}$$

Note that the two multiple integrals appearing in the last two lines of equation (20) can be factorised into M and N identical n-fold integrals respectively,

$$I_M = \int \left(\prod_{a=1}^{n} \mathrm{d}\boldsymbol{v}_a\right) \exp\left[-\frac{\beta}{2}\sum_{a=1}^{n}\sum_{i=1}^{M}\lambda_a v_{ia}^2 + \mathrm{i}\sum_{i=1}^{M}\int \mathrm{d}\vec{v}\hat{\phi}(\vec{v})\prod_{a=1}^{n}\delta(v_a - v_{ia})\right]$$
$$= \left\{\int \mathrm{d}\vec{v}\exp\left[-\mathrm{i}\frac{\beta}{2}\sum_{a=1}^{n}\lambda_a v_a^2 + \mathrm{i}\hat{\phi}(\vec{v})\right]\right\}^M , \tag{21}$$

$$I_N = \int \left(\prod_{a=1}^{n} \mathrm{d}\boldsymbol{u}_a\right) \exp\left[-\frac{\beta}{2}\sum_{a=1}^{n}\sum_{j=1}^{N}u_{ja}^2 + \mathrm{i}\sum_{j=1}^{N}\int \mathrm{d}\vec{u}\hat{\psi}(\vec{u})\prod_{a=1}^{n}\delta(u_a - u_{ja})\right]$$
$$= \left\{\int \mathrm{d}\vec{u}\exp\left[-\mathrm{i}\frac{\beta}{2}\sum_{a=1}^{n}u_a^2 + \mathrm{i}\hat{\psi}(\vec{u})\right]\right\}^N , \tag{22}$$

such that (20) can be written as

$$\langle Z^n\rangle \propto \int \mathcal{D}\phi\mathcal{D}\hat{\phi}\mathcal{D}\psi\mathcal{D}\hat{\psi}\mathrm{d}\vec{\lambda}\,\mathrm{e}^{\sqrt{NM}S\left[\phi,\hat{\phi},\psi,\hat{\psi};\vec{\lambda}\right]} . \tag{23}$$

The action $S\left[\phi,\hat{\phi},\psi,\hat{\psi};\vec{\lambda}\right]$ is defined as

$$S\left[\phi,\hat{\phi},\psi,\hat{\psi};\vec{\lambda}\right] = S_1\left[\phi,\hat{\phi}\right] + S_2\left[\hat{\phi};\vec{\lambda}\right] + \tilde{S}_1\left[\psi,\hat{\psi}\right] + \tilde{S}_2\left[\hat{\psi}\right] + S_3\left[\vec{\lambda}\right] + S_{\mathrm{int}}\left[\phi,\psi\right] , \tag{24}$$

where

$$S_1\left[\phi,\hat{\phi}\right] = -\frac{\mathrm{i}}{\alpha}\int \mathrm{d}\vec{v}\hat{\phi}(\vec{v})\,\phi(\vec{v}) \tag{25}$$

$$S_2\left[\hat{\phi};\vec{\lambda}\right] = \frac{1}{\alpha}\mathrm{Log}\int \mathrm{d}\vec{v}\exp\left[-\mathrm{i}\frac{\beta}{2}\sum_{a=1}^{n}\lambda_a v_a^2 + \mathrm{i}\hat{\phi}(\vec{v})\right] \tag{26}$$

$$\tilde{S}_1\left[\psi,\hat{\psi}\right] = -\mathrm{i}\alpha\int \mathrm{d}\vec{u}\hat{\psi}(\vec{u})\,\psi(\vec{u}) \tag{27}$$

$$\tilde{S}_2\left[\hat{\psi}\right] = \alpha\,\mathrm{Log}\int \mathrm{d}\vec{u}\exp\left[-\mathrm{i}\frac{\beta}{2}\sum_{a=1}^{n}u_a^2 + \mathrm{i}\hat{\psi}(\vec{u})\right] \tag{28}$$

$$S_3\left[\vec{\lambda}\right] = \mathrm{i}\frac{\beta}{2\alpha}\sum_{a=1}^{n}\lambda_a \tag{29}$$

$$S_{\mathrm{int}}\left[\phi,\psi\right] = q\int \mathrm{d}\vec{v}\mathrm{d}\vec{u}\phi(\vec{v})\,\psi(\vec{u})\left(\left\langle \mathrm{e}^{\beta K\vec{v}\cdot\vec{u}}\right\rangle_K - 1\right) , \tag{30}$$

and Log is the branch of the complex logarithm such that $\mathrm{Log}\,\mathrm{e}^z = z$.

The form (23) is amenable to a saddle-point evaluation for large N, M. In order to facilitate the $n \to 0$ limit, we will first adopt a replica symmetric ansatz as detailed in the sub-section below.

4.3. Replica-symmetric ansatz

We now employ a replica symmetric ansatz, which assumes that the dependence on the vector arguments $\vec{v}$ and $\vec{u}$ is only through a permutation-symmetric function of the vector components. An even stronger 'rotationally invariant' assumption — namely that such dependence would only be through the modulus $|\vec{v}|$ and $|\vec{u}|$ of the vectors involved — was shown to lead to the correct solution for the *spectra* of sparse random matrices [39, 60–62]. However, for questions related to the largest eigenvalue/eigenvector, the latter assumption was shown to be too restrictive on the space of functions within which to seek for an extremiser of the action [21–24].

The permutation-symmetric ansatz consists in writing the replicated order parameters as a superposition of uncountably infinite Gaussians with non-zero mean. We will follow this prescription, as originally suggested in [39, 67, 68], while noting that it is not the most general possible as it does not include cross-terms.

To this end, we introduce the following normalised densities, $\pi(\omega,h)$, $\hat{\pi}(\hat{\omega},\hat{h})$, $\rho(\sigma,\mu)$, $\hat{\rho}(\hat{\sigma},\hat{\mu})$, and their respective measures, $\mathrm{d}\pi = \mathrm{d}\omega\,\mathrm{d}h\,\pi(\omega,h)$, $\mathrm{d}\hat{\pi} = \mathrm{d}\hat{\omega}\,\mathrm{d}\hat{h}\,\hat{\pi}(\hat{\omega},\hat{h})$, $\mathrm{d}\rho = \mathrm{d}\sigma\mathrm{d}\mu\rho(\sigma,\mu)$ and $\mathrm{d}\hat{\rho} = \mathrm{d}\hat{\sigma}\mathrm{d}\hat{\mu}\hat{\rho}(\hat{\sigma},\hat{\mu})$. We then use these densities to represent the replicated order parameters as

$$\phi(\vec{v}) = \int \mathrm{d}\pi \prod_{a=1}^{n} \frac{1}{\mathcal{Z}_\beta(\omega,h)} \mathrm{e}^{-\frac{\beta}{2}\omega v_a^2 + \beta h v_a} \tag{31}$$

$$\mathrm{i}\hat{\phi}(\vec{v}) = \hat{c} \int \mathrm{d}\hat{\pi} \prod_{a=1}^{n} \mathrm{e}^{\frac{\beta}{2}\hat{\omega} v_a^2 + \beta \hat{h} v_a} \tag{32}$$

$$\psi(\vec{u}) = \int \mathrm{d}\rho \prod_{a=1}^{n} \frac{1}{\mathcal{Z}_\beta(\sigma,\mu)} \mathrm{e}^{-\frac{\beta}{2}\sigma u_a^2 + \beta \mu u_a} \tag{33}$$

$$\mathrm{i}\hat{\psi}(\vec{u}) = \hat{t} \int \mathrm{d}\hat{\rho} \prod_{a=1}^{n} \mathrm{e}^{\frac{\beta}{2}\hat{\sigma} u_a^2 + \beta \hat{\mu} u_a} \tag{34}$$

$$\mathrm{i}\lambda_a = \lambda \qquad \forall\, 1 \leqslant a \leqslant n\,, \tag{35}$$

with

$$\mathcal{Z}_\beta(x,y) = \sqrt{\frac{2\pi}{\beta x}} \mathrm{e}^{\frac{\beta y^2}{2x}}. \tag{36}$$

Note that since $\pi, \hat{\pi}, \rho$ and $\hat{\rho}$ are normalised densities, this representation preserves the normalisation of $\phi(\vec{v})$ and $\psi(\vec{u})$. The constants $\hat{c}$ and $\hat{t}$ are introduced to account for the fact that the conjugate functions $\mathrm{i}\hat{\phi}$ and $\mathrm{i}\hat{\psi}$ do not have the interpretation of a density, therefore they need not be normalised.

This representation allows us to integrate out the $\vec{v}$'s and $\vec{u}$'s and extract the leading $n \to 0$ behaviour, which is currently only implicit in (24) (for full details of how to apply the transformation, see appendix E in [24]). Inserting equations (31)–(35) into (25)–(30) and collecting terms up to $\mathcal{O}(n)$, while introducing Lagrange multipliers that enforce normalisation upon the

densities, the action takes the form of

$$
\begin{aligned}
S[\pi,\hat{\pi},\rho,\hat{\rho};\lambda] \simeq & -\frac{n\hat{c}}{\alpha}\int \mathrm{d}\pi\ \mathrm{d}\hat{\pi}\mathrm{Log}\left[\frac{\mathcal{Z}_\beta\left(\omega-\hat{\omega},h+\hat{h}\right)}{\mathcal{Z}_\beta(\omega,h)}\right] \\
& +\frac{n}{\alpha}\sum_{s=0}^{\infty}p_{\hat{c}}(s)\int\{\mathrm{d}\hat{\pi}\}_s\mathrm{Log}\mathcal{Z}_\beta\left(\lambda-\{\hat{\omega}\}_s,\left\{\hat{h}\right\}_s\right) \\
& -n\hat{t}\alpha\int\mathrm{d}\rho\mathrm{d}\hat{\rho}\mathrm{Log}\left[\frac{\mathcal{Z}_\beta(\sigma-\hat{\sigma},\mu+\hat{\mu})}{\mathcal{Z}_\beta(\sigma,\mu)}\right] \\
& +n\alpha\sum_{s=0}^{\infty}p_{\hat{t}}(s)\int\{\mathrm{d}\hat{\rho}\}_s\mathrm{Log}\mathcal{Z}_\beta\left(1-\{\hat{\sigma}\}_s,\{\hat{\mu}\}_s\right) \\
& +n\frac{\beta}{2\alpha}\lambda+nq\int\mathrm{d}\pi\,\mathrm{d}\rho\left\langle\mathrm{Log}\left[\frac{\mathcal{Z}_\beta\left(\omega-\frac{K^2}{\sigma},h+\frac{K\mu}{\sigma}\right)}{\mathcal{Z}_\beta(\omega,h)}\right]\right\rangle_K \\
& +\gamma\left(\int\mathrm{d}\pi-1\right)+\hat{\gamma}\left(\int\mathrm{d}\hat{\pi}-1\right)+\xi\left(\int\mathrm{d}\rho-1\right)+\hat{\xi}\left(\int\mathrm{d}\hat{\rho}-1\right),
\end{aligned}
\tag{37}
$$

where we introduced the shorthands $\{\mathrm{d}\hat{\pi}\}_s=\prod_{\ell=1}^{s}\mathrm{d}\hat{\pi}_\ell$, $\{\hat{\omega}\}_s=\sum_{\ell=1}^{s}\hat{\omega}_\ell$, $\{\hat{h}\}_s=\sum_{\ell=1}^{s}\hat{h}_\ell$, and similarly with $\hat{\rho}$, $\hat{\sigma}$ and $\hat{\mu}$. Moreover, we denoted by $p_m(s)=\mathrm{e}^{-m}m^s/s!$ the Poisson distribution with mean m. Note that for the $\vec{u}$ and $\vec{v}$ integrals to converge, one has to formally require the following inequalities, $\omega>\hat{\omega}$, $\omega>0$, $\lambda>\{\hat{\omega}\}_s$ and similarly, $\sigma>\hat{\sigma}$, $\sigma>0$, $1>\{\hat{\sigma}\}_s$. Furthermore, if we denote the lower (upper) bound of the support of $p(K)$ by ζ^- (ζ^+), another requirement is $\omega\sigma>\left[\max\left(|\zeta^-|,|\zeta^+|\right)\right]^2$. In practice, to satisfy these constraints, one has to dynamically enforce them while running the population dynamics algorithm (see section 6).

4.4. Saddle point analysis

We now proceed with our fourth step, which involves studying the saddle point structure of the normalised densities introduced in section 4.3 under the replica-symmetric framework.

In the limit of $N,M\to\infty$, equation (23) is evaluated using a saddle-point method to give

$$
\langle\mathcal{Z}^n\rangle\approx\mathrm{e}^{\sqrt{NM}S[\pi^\star,\hat{\pi}^\star,\rho^\star,\hat{\rho}^\star;\lambda^\star]}\,, \tag{38}
$$

where $\pi^\star,\hat{\pi}^\star,\rho^\star,\hat{\rho}^\star$ are the saddle point forms of the densities, obtained from the stationary conditions $\delta S/\delta\pi|_{\pi^\star,\hat{\pi}^\star,\rho^\star,\hat{\rho}^\star;\lambda^\star}=0$ and similar, and '$\approx$' denotes equivalence on a logarithmic scale. To facilitate the notation, from now on we discard the $\star$'s when addressing the saddle point forms of the densities. Consequently, the first stationary condition, $\delta S/\delta\pi=0$, entails

$$
\frac{\hat{c}}{\alpha q}\int\mathrm{d}\hat{\pi}\mathrm{Log}\left[\frac{\mathcal{Z}_\beta\left(\omega-\hat{\omega},h+\hat{h}\right)}{\mathcal{Z}_\beta(\omega,h)}\right]=\int\mathrm{d}\rho\left\langle\mathrm{Log}\left[\frac{\mathcal{Z}_\beta\left(\omega-\frac{K^2}{\sigma},h+\frac{K\mu}{\sigma}\right)}{\mathcal{Z}_\beta(\omega,h)}\right]\right\rangle_K+\frac{\gamma}{q}\,, \tag{39}
$$

where γ is the Lagrange multiplier enforcing the normalisation of π. To match the two sides of equation (39) for all values of the non-integrated variables, ω and h [24], while preserving

normalisation of $\hat{\pi}$, we set

$$\hat{\pi}\left(\hat{\omega},\hat{h}\right) = \int \mathrm{d}\rho \left\langle \delta\left(\hat{\omega} - \frac{K^2}{\sigma}\right)\delta\left(\hat{h} - \frac{K\mu}{\sigma}\right)\right\rangle_K \tag{40}$$

$$\hat{c} = \alpha q \tag{41}$$

$$\gamma = 0\,. \tag{42}$$

To obtain the next stationary condition, $\delta S/\delta\rho = 0$, we first note that the interaction term in (37) was evaluated by integrating out first the u's and then the v's. However, one could have equally well swapped the order of integrations, which results in an equivalent form of S_{int} given by

$$S_{\mathrm{int}}[\pi,\rho] = nq\int \mathrm{d}\pi\,\mathrm{d}\rho \left\langle \mathrm{Log}\left[\frac{\mathcal{Z}_\beta\left(\sigma - \frac{K^2}{\omega}, \mu + \frac{Kh}{\omega}\right)}{\mathcal{Z}_\beta(\sigma,\mu)}\right]\right\rangle_K . \tag{43}$$

Keeping that in mind, the stationary condition $\delta S/\delta\rho = 0$ can be written as

$$\frac{\alpha\hat{t}}{q}\int \mathrm{d}\hat{\rho}\,\mathrm{Log}\left[\frac{\mathcal{Z}_\beta(\sigma-\hat{\sigma},\mu+\hat{\mu})}{\mathcal{Z}_\beta(\sigma,\mu)}\right] = \int \mathrm{d}\pi \left\langle \mathrm{Log}\left[\frac{\mathcal{Z}_\beta\left(\sigma - \frac{K^2}{\omega}, \mu + \frac{Kh}{\omega}\right)}{\mathcal{Z}_\beta(\sigma,\mu)}\right]\right\rangle_K + \frac{\xi}{q}\,, \tag{44}$$

where ξ is the Lagrange multiplier enforcing normalisation of ρ. Using the same argument that led us to equation (40), we find that

$$\hat{\rho}(\hat{\sigma},\hat{\mu}) = \int \mathrm{d}\pi \left\langle \delta\left(\hat{\sigma} - \frac{K^2}{\omega}\right)\delta\left(\hat{\mu} - \frac{Kh}{\omega}\right)\right\rangle_K \tag{45}$$

$$\hat{t} = \alpha^{-1} q \tag{46}$$

$$\xi = 0\,. \tag{47}$$

The next stationary condition, $\delta S/\delta\hat{\pi} = 0$, is given by

$$\begin{aligned}\int \mathrm{d}\pi\,\mathrm{Log}&\left[\frac{\mathcal{Z}_\beta\left(\omega-\hat{\omega}, h+\hat{h}\right)}{\mathcal{Z}_\beta(\omega,h)}\right] \\ &= \sum_{s=0}^{\infty}\frac{sp_{\hat{c}}(s)}{\hat{c}}\int \{\mathrm{d}\hat{\pi}\}_{s-1}\,\mathrm{Log}\mathcal{Z}_\beta\left(\lambda - \{\hat{\omega}\}_{s-1} - \hat{\omega}, \left\{\hat{h}\right\}_{s-1} + \hat{h}\right) + \frac{\hat{\gamma}}{\hat{c}}\,,\end{aligned} \tag{48}$$

where $\hat{\gamma}$ is the Lagrange multiplier enforcing normalisation of $\hat{\pi}$. Using $\hat{c} = \alpha q$ (equation (41)) we thus find that

$$\pi(\omega,h) = \sum_{s=1}^{\infty}\frac{sp_{\alpha q}(s)}{\alpha q}\int \{\mathrm{d}\hat{\pi}\}_{s-1}\,\delta\left(\omega - \left(\lambda - \{\hat{\omega}\}_{s-1}\right)\right)\delta\left(h - \left\{\hat{h}\right\}_{s-1}\right) \tag{49}$$

$$\hat{\gamma} = -\alpha q\int \mathrm{d}\pi\,\mathrm{Log}\mathcal{Z}_\beta(\omega,h)\,. \tag{50}$$

The next stationary condition, $\delta S/\delta\hat{\rho}=0$, reads

$$\int \mathrm{d}\rho \mathrm{Log}\left[\frac{\mathcal{Z}_\beta(\sigma-\hat{\sigma},\mu+\hat{\mu})}{\mathcal{Z}_\beta(\sigma,\mu)}\right] = \sum_{s=0}^{\infty}\frac{sp_{\hat{t}}(s)}{\hat{t}}\int\{\mathrm{d}\hat{\rho}\}_{s-1}\mathrm{Log}\mathcal{Z}_\beta\left(1-\{\hat{\sigma}\}_{s-1}-\hat{\sigma},\{\hat{\mu}\}_{s-1}+\hat{\mu}\right)+\frac{\hat{\xi}}{\hat{t}}\,, \tag{51}$$

where $\hat{\xi}$ is the Lagrange multiplier enforcing normalisation of $\hat{\rho}$. Using $\hat{t}=\alpha^{-1}q$ (equation (46)), the saddle point form of ρ can be expressed as

$$\rho(\sigma,\mu)=\sum_{s=1}^{\infty}\frac{sp_{\alpha^{-1}q}}{\alpha^{-1}q}\int\{\mathrm{d}\hat{\rho}\}_{s-1}\,\delta\left(\sigma-\left(1-\{\hat{\sigma}\}_{s-1}\right)\right)\delta\left(\mu-\{\hat{\mu}\}_{s-1}\right) \tag{52}$$

$$\hat{\xi}=-\alpha^{-1}q\int\mathrm{d}\rho\mathrm{Log}\mathcal{Z}_\beta(\sigma,\mu)\,. \tag{53}$$

Finally, in the $\beta\to\infty$ limit, the condition $\partial S/\partial\lambda=0$ yields

$$\sum_{s=0}^{\infty}p_{\alpha q}(s)\int\{\mathrm{d}\hat{\pi}\}_s\left(\frac{\left\{\hat{h}\right\}_s}{\lambda-\{\hat{\omega}\}_s}\right)^2=1\,. \tag{54}$$

A further simplification can be made by reducing the number of equations. This is done by inserting (45) into (52) to obtain

$$\rho(\sigma,\mu)=\sum_{s=1}^{\infty}\frac{sp_{\alpha^{-1}q}(s)}{\alpha^{-1}q} \times\int\{\mathrm{d}\pi\}_{s-1}\left\langle\delta\left(\sigma-\left(1-\sum_{\ell=1}^{s-1}\frac{K_\ell^2}{\omega_\ell}\right)\right)\delta\left(\mu-\sum_{\ell=1}^{s-1}\frac{K_\ell h_\ell}{\omega_\ell}\right)\right\rangle_{\{K\}_{s-1}}, \tag{55}$$

where $\langle\cdot\rangle_{\{K\}_{s-1}}$ means averaging over $s-1$ random variables drawn from $p(K)$. Then, by substituting (40) into (49) we get

$$\pi(\omega,h)=\sum_{s=1}^{\infty}\frac{sp_{\alpha q}(s)}{\alpha q} \times\int\{\mathrm{d}\rho\}_{s-1}\left\langle\delta\left(\omega-\left(\lambda-\sum_{\ell=1}^{s-1}\frac{K_\ell^2}{\sigma_\ell}\right)\right)\delta\left(h-\sum_{\ell=1}^{s-1}\frac{K_\ell\mu_\ell}{\sigma_\ell}\right)\right\rangle_{\{K\}_{s-1}}. \tag{56}$$

Furthermore, to express (54) in terms of ρ, we substitute (40) into (54) and obtain

$$\sum_{s=0}^{\infty}p_{\alpha q}(s)\int\{\mathrm{d}\rho\}_s\left\langle\left(\frac{\sum_{\ell=1}^{s}\frac{K_\ell\mu_\ell}{\sigma_\ell}}{\lambda-\sum_{\ell=1}^{s}\frac{K_\ell^2}{\sigma_\ell}}\right)^2\right\rangle_{\{K\}_s}=1\,. \tag{57}$$

One can, in principle, substitute (55) into (56) and (57), and obtain self-contained equations for π, but this results in somewhat cumbersome expressions.

4.5. The replicated partition function at the saddle point

The final step in the analysis is to evaluate the saddle point form of the replicated partition function in the $\beta \to \infty$ limit. To this end, we use the saddle point forms of $\hat{\pi}$ and $\hat{c}$, (i.e. (40) and (41) respectively) to obtain (arguments removed for ease of notation)

$$S_1 \sim -\frac{nq\beta}{2}\int \mathrm{d}\pi\,\mathrm{d}\rho\left\langle \frac{\left(h+\frac{K\mu}{\sigma}\right)^2}{\omega-\frac{K^2}{\sigma}}-\frac{h^2}{\omega}\right\rangle_K , \tag{58}$$

where we used the definition of $\mathcal{Z}_\beta$ (equation (36)) and evaluated the $\beta \to \infty$ asymptotic behaviour ($\sim$). Similarly,

$$\tilde{S}_1 \sim -\frac{nq\beta}{2}\int \mathrm{d}\pi\,\mathrm{d}\rho\left\langle \frac{\left(\mu+\frac{Kh}{\omega}\right)^2}{\sigma-\frac{K^2}{\omega}}-\frac{\mu^2}{\sigma}\right\rangle_K . \tag{59}$$

Next, we have

$$S_2 \sim \frac{nq\beta}{2}\int \mathrm{d}\hat{\pi}\sum_{s=0}^{\infty}\frac{sp_{\alpha q}(s)}{\alpha q}\int\{\mathrm{d}\hat{\pi}\}_{s-1}\frac{\left\{\hat{h}\right\}_{s-1}+\hat{h}}{\lambda-\{\hat{\omega}\}_{s-1}-\hat{\omega}}\hat{h} . \tag{60}$$

Multiplying the last line by $1=\int \mathrm{d}\omega\mathrm{d}h\delta\left(\omega-\left(\lambda-\{\hat{\omega}\}_{s-1}\right)\right)\delta\left(h-\{\hat{h}\}_{s-1}\right)$ and using the saddle point form of π (equation (49)), we have

$$S_2 \sim \frac{nq\beta}{2}\int \mathrm{d}\hat{\pi}\mathrm{d}\pi\left(\frac{h+\hat{h}}{\omega-\hat{\omega}}\hat{h}\right) . \tag{61}$$

Then, by using the saddle point form of $\hat{\pi}$ (equation (40)), we can further rewrite (61) as

$$S_2 \sim \frac{nq\beta}{2}\int \mathrm{d}\pi\,\mathrm{d}\rho\left\langle \frac{K\mu}{\sigma}\frac{h+\frac{K\mu}{\sigma}}{\omega-\frac{K^2}{\sigma}}\right\rangle_K . \tag{62}$$

Following similar lines, we also conclude that

$$\tilde{S}_2 \sim \frac{nq\beta}{2}\int \mathrm{d}\pi\,\mathrm{d}\rho\left\langle \frac{Kh}{\omega}\frac{\mu+\frac{Kh}{\omega}}{\sigma-\frac{K^2}{\omega}}\right\rangle_K . \tag{63}$$

Lastly, considering the two equivalent forms of the interaction term (fifth line in equation (37) and (43)), its $\beta \to \infty$ limit can be written as

$$S_{\mathrm{int}} \sim \frac{nq\beta}{4}\int \mathrm{d}\pi\,\mathrm{d}\rho\left\langle \frac{\left(h+\frac{K\mu}{\sigma}\right)^2}{\omega-\frac{K^2}{\sigma}}-\frac{h^2}{\omega}+\frac{\left(\mu+\frac{Kh}{\omega}\right)^2}{\sigma-\frac{K^2}{\omega}}-\frac{\mu^2}{\sigma}\right\rangle_K . \tag{64}$$

Inserting equations (58), (59), (62)–(64) into (37), while noting that

$$\frac{K\mu}{\sigma}\frac{h+\frac{K\mu}{\sigma}}{\omega-\frac{K^2}{\sigma}}+\frac{Kh}{\omega}\frac{\mu+\frac{Kh}{\omega}}{\sigma-\frac{K^2}{\omega}}-\frac{1}{2}\left[\frac{\left(h+\frac{K\mu}{\sigma}\right)^2}{\omega-\frac{K^2}{\sigma}}-\frac{h^2}{\omega}+\frac{\left(\mu+\frac{Kh}{\omega}\right)^2}{\sigma-\frac{K^2}{\omega}}-\frac{\mu^2}{\sigma}\right]=0 , \tag{65}$$

the saddle point action eventually takes the form

$$S[\pi,\hat{\pi},\rho,\hat{\rho};\lambda] \sim \frac{n\beta}{2\alpha}\lambda\,, \tag{66}$$

in the $n\to 0$ and $\beta\to\infty$ limits. Then, by inserting (66) into (38), the replicated partition function at the saddle point becomes

$$\langle Z^n\rangle \approx \mathrm{e}^{\frac{n\beta M}{2}\lambda}\,. \tag{67}$$

Finally, substituting (67) into (9), we obtain

$$\langle\lambda_1\rangle = \lambda\,. \tag{68}$$

We recall at this point that the replica derivation started under the simplifying assumption that the c_{ij}'s are independent Bernoulli random variables (see equation (4)). This implies that the distribution of total number of nonzero elements in each row (column) — $p_{\alpha q}(s)$ ($p_{\alpha^{-1}q}(s)$) — naturally appearing in equations (55)–(57) is a Poisson distribution with unbounded support. However, due to [22], we know that these equations remain formally valid for any connectivity distribution $p(s)$. In our case, it is then necessary to consider the truncated Poisson distribution and manually amend the upper limit of the sums to account for the existence of a maximal number of nonzero elements in each row (column), R (C). Putting everything together, in this section we have shown that by finding λ, π and ρ that solve the following system of recursive distributional equations supplemented by an integral constraint

$$\begin{aligned}
\pi(\omega,h) &= \sum_{s=1}^{R}\frac{sp_{\alpha q}(s)}{\langle s\rangle_{p_{\alpha q}}}\int\{\mathrm{d}\rho\}_{s-1}\left\langle\delta\left(\omega-\left(\lambda-\sum_{\ell=1}^{s-1}\frac{K_\ell^2}{\sigma_\ell}\right)\right)\delta\left(h-\sum_{\ell=1}^{s-1}\frac{K_\ell\mu_\ell}{\sigma_\ell}\right)\right\rangle_{\{K\}_{s-1}}\\
\rho(\sigma,\mu) &= \sum_{s=1}^{C}\frac{sp_{\alpha^{-1}q}(s)}{\langle s\rangle_{p_{\alpha^{-1}q}}}\int\{\mathrm{d}\pi\}_{s-1}\left\langle\delta\left(\sigma-\left(1-\sum_{\ell=1}^{s-1}\frac{K_\ell^2}{\omega_\ell}\right)\right)\delta\left(\mu-\sum_{\ell=1}^{s-1}\frac{K_\ell h_\ell}{\omega_\ell}\right)\right\rangle_{\{K\}_{s-1}}\\
&\sum_{s=0}^{R}p_{\alpha q}(s)\int\{\mathrm{d}\rho\}_s\left\langle\left(\frac{\sum_{\ell=1}^{s}\frac{K_\ell\mu_\ell}{\sigma_\ell}}{\lambda-\sum_{\ell=1}^{s}\frac{K_\ell^2}{\sigma_\ell}}\right)^2\right\rangle_{\{K\}_s} = 1
\end{aligned} \tag{69}$$

the typical largest eigenvalue of $\boldsymbol{J}$ is given by (68). Note that for ease of notation, in (69), we used $p_{\alpha q}(s)$ ($p_{\alpha^{-1}q}(s)$) to denote the *truncated* Poisson distribution with parameter αq ($\alpha^{-1}q$), an upper cutoff R (C), and $\langle s\rangle_{p_{\alpha q}}$ ($\langle s\rangle_{p_{\alpha^{-1}q}}$) denoting its average.

In section 6, we will show that these integral equations can be efficiently solved using a population dynamics algorithm. In the next section, we instead provide the theoretical framework to compute the probability density of the top eigenvector's components for diluted Wishart matrices.

5. Density of the top eigenvector's components

We now demonstrate how the results from the previous section can be applied to compute the average density of the top eigenvector's components for large M,N,

$$T(u) = \left\langle\frac{1}{M}\sum_{i=1}^{M}\delta\left(u-v_1^{(i)}\right)\right\rangle\,, \tag{70}$$

where once again, $\langle\cdot\rangle$ denotes averaging over different realisations of $\boldsymbol{X}$. We begin by outlining the strategy, highlighting its similarities and differences with the analysis in section 4. We then carry it out to derive an expression for $T(u)$, which builds on the solution of equation (69).

To this end, we introduce the auxiliary partition function

$$Z_{\epsilon}^{(\beta)}(t,\boldsymbol{X};u)=\int \mathrm{d}\boldsymbol{v}\exp\left[\frac{\beta}{2}(\boldsymbol{v},\boldsymbol{J}\boldsymbol{v})+\beta t\sum_{i}\delta_{\epsilon}(u-v_i)\right]\delta\left(|\boldsymbol{v}|^2-M\right)\,, \tag{71}$$

where δ_ϵ is a smooth regulariser of the delta function and $\boldsymbol{J}=\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}$. Due to the concentration of the Gibbs measure (see (6)),

$$P_{\beta,\boldsymbol{X}}(\boldsymbol{v})=\frac{\exp\left(\frac{\beta}{2}(\boldsymbol{v},\boldsymbol{J}\boldsymbol{v})\right)\delta\left(|\boldsymbol{v}|^2-M\right)}{\int \mathrm{d}\boldsymbol{v}'\exp\left(\frac{\beta}{2}(\boldsymbol{v}',\boldsymbol{J}\boldsymbol{v}')\right)\delta\left(|\boldsymbol{v}'|^2-M\right)}\,, \tag{72}$$

which localises around $\boldsymbol{J}$'s top eigenvector in the $\beta\to\infty$ limit, we can formally express $T(u)$ as

$$T(u)=\lim_{\beta\to\infty}\lim_{\epsilon\to 0^+}\frac{1}{\beta M}\frac{\partial}{\partial t}\left\langle \mathrm{Log}\, Z_{\epsilon}^{(\beta)}(t,\boldsymbol{X};u)\right\rangle\Big|_{t=0}\,. \tag{73}$$

To evaluate $Z_{\epsilon}^{(\beta)}(t,\boldsymbol{X};u)$, we apply again the replica trick, leading to

$$T(u)=\lim_{\beta\to\infty}\lim_{\epsilon\to 0^+}\lim_{n\to 0}\frac{1}{\beta M}\frac{\partial}{\partial t}\frac{1}{n}\mathrm{Log}\left\langle\left[Z_{\epsilon}^{(\beta)}(t,\boldsymbol{X};u)\right]^n\right\rangle\Big|_{t=0}\,. \tag{74}$$

Since the structure of equation (71) resembles that of equation (6), with an additional t-dependent term, we expect that, in the large N,M limit, the replicated partition function will once again take the form

$$\left\langle\left[Z_{\epsilon}^{(\beta)}(t,\boldsymbol{X};u)\right]^n\right\rangle\propto\int\mathcal{D}\varphi\mathcal{D}\hat{\varphi}\mathcal{D}\psi\mathcal{D}\hat{\psi}\mathrm{d}\vec{\lambda}\exp\left\{\sqrt{NM}S_n^{(\beta)}\left[\phi,\hat{\phi},\psi,\hat{\psi},\vec{\lambda};t,\epsilon;u\right]\right\}\,. \tag{75}$$

The above structure will then enable us to employ the replica-symmetric ansatz (i.e. represent the fields as a superposition of uncountably infinite Gaussians, see equations (31)–(36)) and perform a saddle-point evaluation for large N,M

$$\left\langle\left[Z_{\epsilon}^{(\beta)}(t,\boldsymbol{X};u)\right]^n\right\rangle\approx\exp\left\{\sqrt{NM}S_n^{(\beta)}\left[\pi^\star,\hat{\pi}^\star,\rho^\star,\hat{\rho}^\star,\vec{\lambda}^\star;t,\epsilon;u\right]\right\}\,, \tag{76}$$

where the starred objects represent the saddle point forms of $\pi,\hat{\pi},\rho,\hat{\rho},\vec{\lambda}$, found through the corresponding stationary conditions. Since the partial derivative $\frac{\partial}{\partial t}$ in (73) only acts on terms containing any explicit dependence on t, and not through any other indirect functional dependence, t can be safely set to zero in the resulting saddle-point equations. Consequently, $\pi^\star,\hat{\pi}^\star,\rho^\star,\hat{\rho}^\star,\vec{\lambda}^\star$ satisfy the same saddle-point equations derived in section 4.4.

Inserting (76) into (74) and assuming that the leading $n\to 0$ behaviour of the action at the saddle point is given by

$$S_n^{(\beta)}\left[\pi^\star,\hat{\pi}^\star,\rho^\star,\hat{\rho}^\star,\vec{\lambda}^\star;t,\epsilon;u\right]\sim n s_\beta(t,\epsilon;u)+o(n)\,, \tag{77}$$

the final expression for the average density of top eigenvector's components is obtained by inserting equations (74), (76) and (77) into (73)

$$T(u) = \lim_{\beta\to\infty} \frac{\alpha}{\beta} s'_\beta(0,0;u) \ , \tag{78}$$

where $(\cdot)'$ stands for differentiation with respect to t. Since the saddle-point equations for $\pi, \hat{\pi}, \rho, \hat{\rho}, \vec{\lambda}$ are identical to those derived in section 4.4, the remaining challenge is to identify $s_\beta(t,\epsilon;u)$ and evaluate (78).

To this end, we apply this strategy to our matrix $\boldsymbol{J} = \boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}$, where $\boldsymbol{X}$'s entries follow the distribution given in equation (1). Exponentiating the replicated partition function by following the same lines as in section 4.1, we obtain

$$\begin{aligned}\left\langle \left[Z_\epsilon^{(\beta)}(t,\boldsymbol{X};u)\right]^n \right\rangle \propto \int \left(\prod_{a=1}^{n} \mathrm{d}\boldsymbol{v}_a \mathrm{d}\boldsymbol{u}_a \mathrm{d}\lambda_a\right) \exp\left(-\frac{\beta}{2}\sum_{a=1}^{n}\sum_{j=1}^{N} u_{ja}^2\right) \exp\left(\mathrm{i}M\frac{\beta}{2}\sum_{a=1}^{n}\lambda_a\right) \\ \times \exp\left(-\mathrm{i}\frac{\beta}{2}\sum_{a=1}^{n}\sum_{i=1}^{M}\lambda_a v_{ia}^2\right) \exp\left\{\frac{q}{\sqrt{NM}}\sum_{i=1}^{M}\sum_{j=1}^{N}\left[\left\langle \exp\left(\beta K\sum_{a=1}^{n} v_{ia}u_{ja}\right)\right\rangle_K - 1\right]\right\} \\ \times \exp\left[\beta t\sum_{a=1}^{n}\sum_{i=1}^{M}\delta_\epsilon(u-v_{ia})\right] \ . \end{aligned} \tag{79}$$

Comparing this expression with equation (15), it is natural to define the same functional order parameters as in equation (17),

$$\phi(\vec{v}) = \frac{1}{M}\sum_{i=1}^{M}\prod_{a=1}^{n}\delta(v_a - v_{ia}) \ , \tag{80}$$

$$\psi(\vec{u}) = \frac{1}{N}\sum_{j=1}^{N}\prod_{a=1}^{n}\delta(u_a - u_{ja}) \ . \tag{81}$$

Then, by following the same lines as in section 4.2, we see that the functional-integral form of (79) is identical to the one in (23), except for the term S_2. Hence, the functional integral representation of the replicated partition function can indeed be expressed as

$$\left\langle \left[Z_\epsilon^{(\beta)}(t,\boldsymbol{X};u)\right]^n \right\rangle \propto \int \mathcal{D}\varphi\mathcal{D}\hat{\varphi}\mathcal{D}\psi\mathcal{D}\hat{\psi}\mathrm{d}\vec{\lambda} \exp\left\{\sqrt{NM}S_n^{(\beta)}\left[\phi,\hat{\phi},\psi,\hat{\psi},\vec{\lambda};t,\epsilon;u\right]\right\} \ , \tag{82}$$

with the action given by

$$\begin{aligned} S_n^{(\beta)}\left[\phi,\hat{\phi},\psi,\hat{\psi},\vec{\lambda};t,\epsilon;u\right] = S_1\left[\phi,\hat{\phi}\right] + S_2\left[\hat{\phi},\vec{\lambda};t;\epsilon;u\right] \\ + \tilde{S}_1\left[\psi,\hat{\psi}\right] + \tilde{S}_2\left[\hat{\psi}\right] + S_3\left[\vec{\lambda}\right] + S_{\mathrm{int}}\left[\phi,\psi\right] \ , \end{aligned} \tag{83}$$

where all contributions other than S_2 are identical to those defined in equations (25)–(30), and the t and ϵ dependence is confined to S_2, which is now given by

$$S_2\left[\hat{\phi},\vec{\lambda};t,\epsilon;u\right] = \frac{1}{\alpha}\mathrm{Log}\int \mathrm{d}\vec{v}\exp\left[-\mathrm{i}\frac{\beta}{2}\sum_{a}^{n}\lambda_a v_a^2 + \beta t\sum_{a}^{n}\delta_\epsilon(u-v_a) + \mathrm{i}\hat{\phi}(\vec{v})\right] \ . \tag{84}$$

We then follow the same strategy as in section 4.3, and enforce the replica symmetric ansatz by representing the functional order parameters as a superposition of uncountably infinite Gaussians (see equations (31)–(36)). Specifically, we recall that

$$\mathrm{i}\hat{\phi}\left(\vec{v}\right) = \hat{c}\int \mathrm{d}\hat{\pi}\prod_{a=1}^{n}\mathrm{e}^{\frac{\beta}{2}\hat{\omega}v_a^2+\beta\hat{h}v_a}\ . \tag{85}$$

Substituting this representation into equation (84) yields the following leading $n\to 0$ behaviour,

$$\begin{aligned} S_2 \simeq \frac{\hat{c}}{\alpha} + \frac{n}{\alpha}\sum_{s=0}^{\infty}p_{\hat{c}}\left(s\right)\int\left\{\mathrm{d}\hat{\pi}\right\}_s\ \mathrm{Log}\int \mathrm{d}v\exp\Bigg[-\mathrm{i}\frac{\beta}{2}\lambda v^2 \\ +\beta t\delta_{\epsilon}\left(u-v\right)+\frac{\beta}{2}\left\{\hat{\omega}\right\}_s v^2+\beta\left\{\hat{h}\right\}_s v\Bigg]\ . \end{aligned} \tag{86}$$

Therefore, we can identify the function $s_{\beta}(t,\epsilon;u)$ in (77) as

$$\begin{aligned} s_{\beta}\left(t,\epsilon;u\right) = \frac{1}{\alpha}\sum_{s=0}^{\infty}p_{\hat{c}}\left(s\right)\int\left\{\mathrm{d}\hat{\pi}\right\}_s\ \mathrm{Log}\int \mathrm{d}v\exp\Bigg[-\frac{\beta}{2}\lambda v^2+\beta t\delta_{\epsilon}\left(u-v\right) \\ +\frac{\beta}{2}\left\{\hat{\omega}\right\}_s v^2+\beta\left\{\hat{h}\right\}_s v\Bigg]\ , \end{aligned} \tag{87}$$

with $\mathrm{i}\lambda\equiv\lambda$ solving (69) as before. Taking the t-derivative and setting t and ϵ to zero, while recalling that $\hat{c}=\alpha q$, we obtain

$$s'_{\beta}\left(0,0;u\right) = \frac{\beta}{\alpha}\sum_{s=0}^{\infty}p_{\alpha q}\left(s\right)\int\left\{\mathrm{d}\hat{\pi}\right\}_s\frac{\exp\left[-\frac{\beta}{2}\left(\lambda-\left\{\hat{\omega}\right\}_s\right)u^2+\beta\left\{\hat{h}\right\}_s u\right]}{\int \mathrm{d}v\exp\left[-\frac{\beta}{2}\left(\lambda-\left\{\hat{\omega}\right\}_s\right)v^2+\beta\left\{\hat{h}\right\}_s v\right]}\ .$$

Taking the $\beta\to\infty$ limit and inserting the result into equation (78), we find that

$$T\left(u\right) = \sum_{s=0}^{\infty}p_{\alpha q}\left(s\right)\int\left\{\mathrm{d}\hat{\pi}\right\}_s\delta\left(u-\frac{\left\{\hat{h}\right\}_s}{\lambda-\left\{\hat{\omega}\right\}_s}\right)\ . \tag{88}$$

Finally, using the saddle point form of $\hat{\pi}$ (equation (41)) to express it via ρ, and truncating the Poisson distribution as before, we obtain

$$T\left(u\right) = \sum_{s=0}^{R}p_{\alpha q}\left(s\right)\int\left\{\mathrm{d}\rho\right\}_s\left\langle\delta\left(u-\frac{\sum_{\ell=1}^{s}\frac{K_\ell\mu_\ell}{\sigma_\ell}}{\lambda-\sum_{\ell=1}^{s}\frac{K_\ell^2}{\sigma_\ell}}\right)\right\rangle_{\{K\}_s}\ . \tag{89}$$

Putting everything together, after solving equation (69) for ρ and λ, these can then be used to sample the integral in (89) and obtain the density of the top eigenvector's components. The algorithmic way to do this is explained in the next section.

6. Population dynamics

In this section we briefly present the population dynamics algorithm [67, 76, 77], which can be used to numerically solve the system given by (69). Different incarnations of this algorithm have been used in a number of problems recently [22–24, 78–81]. For a specified set of inputs $q,\alpha,p(K),R,C$ and a target error tolerance Δ, the algorithm outputs the theoretical value of $\langle\lambda_1\rangle$, with an uncertainty $\pm\Delta/2$:

(i) Initialise the real parameter λ to a 'large' value (using the estimate in appendix A).
(ii) Randomly initialise two sets of coupled populations, each of size N_P, $\{(\omega_i,h_i)\}_{1\leqslant i\leqslant N_P}$ and $\{(\sigma_i,\mu_i)\}_{1\leqslant i\leqslant N_P}$.
(iii) Generate a random $s\sim\frac{sp_{\alpha q}(s)}{\langle s\rangle}$, where $p_{\alpha q}(s)$ is a truncated Poisson distribution with parameter αq and upper cutoff R.
(iv) Draw $s-1$ i.i.d. random variables K_ℓ from $p(K)$.
(v) Select $s-1$ random pairs $\{(\sigma_\ell,\mu_\ell)\}_{\ell=1}^{s-1}$ from the population, compute

$$\omega^{(\text{new})}=\lambda-\sum_{\ell=1}^{s-1}\frac{K_\ell^2}{\sigma_\ell}\,, \tag{90}$$

$$h^{(\text{new})}=\sum_{\ell=1}^{s-1}\frac{K_\ell\mu_\ell}{\sigma_\ell}\,, \tag{91}$$

and replace a randomly selected pair (ω_r,h_r) with $(\omega^{(\text{new})},h^{(\text{new})})$.
(vi) Generate a random $s\sim\frac{sp_{\alpha^{-1}q(s)}}{\langle s\rangle}$, where $p_{\alpha^{-1}q}(s)$ is a truncated Poisson distribution with parameter $\alpha^{-1}q$ and upper cutoff C.
(vii) Draw $s-1$ i.i.d. random variables K_ℓ from $p(K)$.
(viii) Select $s-1$ random pairs $\{(\omega_\ell,h_\ell)\}_{\ell=1}^{s-1}$ from the population, compute

$$\sigma^{(\text{new})}=1-\sum_{\ell=1}^{s-1}\frac{K_\ell^2}{\omega_\ell}\,, \tag{92}$$

$$\mu^{(\text{new})}=\sum_{\ell=1}^{s-1}\frac{K_\ell h_\ell}{\omega_\ell}\,, \tag{93}$$

and replace a randomly selected pair (σ_r,μ_r) with $(\sigma^{(\text{new})},\mu^{(\text{new})})$.
(ix) After every sweep, monitor the populations' first moment.
- If any one of them shrinks to zero. Set $\lambda^{(\text{new})}=\lambda-\Delta$ and return to (ii).
- If any one of them explodes, set $\langle\lambda_1\rangle=\lambda+\Delta/2$ and exit the algorithm.

(x) Return to (iii).

The nature of the algorithm ensures that the only value of the (real) parameter λ under which stability is reached is the one corresponding to $\langle\lambda_1\rangle$ [22]. When $\lambda<\langle\lambda_1\rangle$ the h and μ populations will explode, and for $\lambda>\langle\lambda_1\rangle$ they will shrink to zero. Consequently, one can monitor the populations' stability by examining the time-evolution of their first moment, as shown in figure 1. Another observation is that the rates at which the populations diverge and vanish increase as the value of λ deviates from $\langle\lambda_1\rangle$. Furthermore, the stable regime is highly peaked around $\lambda=\langle\lambda_1\rangle$, which allows us to pinpoint the value of $\langle\lambda_1\rangle$ with very high precision.

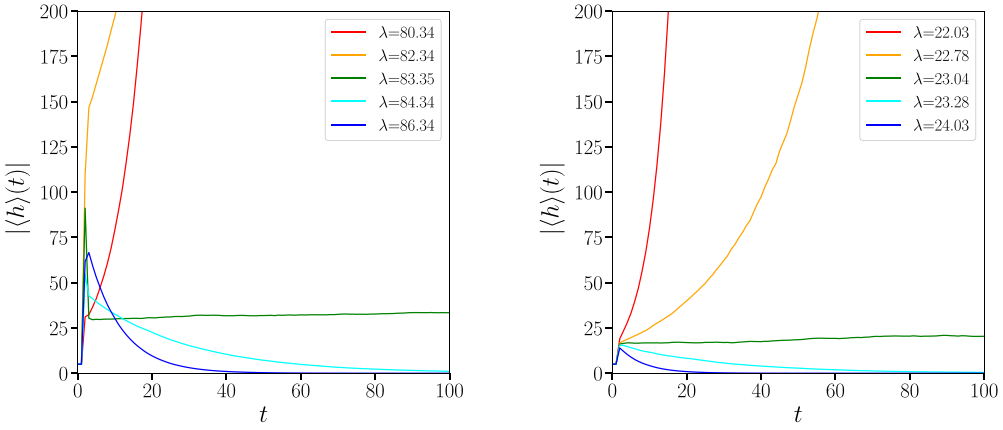

Figure 1. Evolution of the first moment of the h population in absolute value, $|\langle h\rangle|(t)$, according to the population dynamics algorithm as outlined in section 6, with population size of $N_P = 10^5$ and where t is measured in sweeps. The control parameters in this figure are chosen as $q = 8$, $\alpha = \sqrt{5/4}$ and $p(K) = \delta_{K,1}$ for the left figure and $p(K) = \Theta(1 - K)\Theta(K)$ for the right figure. In both figures the maximal number of nonzero elements in each row is set to $R = 70$ and in each column to $C = 60$. The target error tolerance was set to $\Delta = 0.1$. The different curves correspond to ascending values of λ (top to bottom), the parameter that governs the convergence of the algorithm, which was initialised as $\lambda_{\text{initial}} = 250$ in both figures. For $\lambda < \langle\lambda_1\rangle$ (red and orange lines) the population diverges, for $\lambda > \langle\lambda_1\rangle$ (blue and cyan lines) it vanishes, and only when $\lambda = \langle\lambda_1\rangle$ (green line), stability is reached. The rate of divergence/decay depends on the amount by which λ deviates from $\langle\lambda_1\rangle$.

Specifying to the case where $\lambda = \langle\lambda_1\rangle$ and non-trivial stability is achievable, it is possible to identify multiple fixed points for the densities π and ρ that satisfy the first two equations in (69) by adjusting the initial populations. However, incorporating the third equation in (69) uniquely determines the solution. Once the algorithm identifies the value of λ that allows non-trivial stable populations, the third condition in (69) can be fulfilled by rescaling the h and μ populations, yielding a solution that satisfies the full set of equations (69) in its entirety. This rescaling is always allowed due to the linear nature of the recursion governing their updates [22].

Given the behaviour described above, the strategy for pinning down the value of λ under which stability can be reached, is to start with a large value, determined by a proper upper bound for $\langle\lambda_1\rangle$. Then, while running the algorithm, one monitors the time evolution of the populations' first moment, and gradually decreases the value of λ until they stabilise. A plausible upper bound that can be used as a starting point is $\lambda^\star = [\max(|\zeta^-|, |\zeta^+|)]^2 RC$, where ζ^- (ζ^+) is the lower (upper) bound of the support of $p(K)$, while R and C are the maximal numbers of nonzero elements in each row and in each column respectively (see appendix A for a proof). Once the populations stabilise and the typical largest eigenvalue is determined, one can use them to obtain the density of the top eigenvector's components via (89).

In figure 2 we present the scaling of $\langle\lambda_1\rangle$ with the dimensions of the matrix $\boldsymbol{X}$, under the following choice of control parameters: (a) $\alpha = \sqrt{5/4}$, $q = 11.8$ and $p(K) = \delta_{K,1}$; (b) $\alpha = \sqrt{5/4}$, $q = 8$ and $p(K) = \Theta(K)\Theta(1 - K)$, with $\Theta(\cdot)$ being the Heaviside function (i.e. $K \in (0, 1)$ with uniform probability). In both figures the maximal number of non-zero elements in each row is set to $R = 70$ and in each column to $C = 60$. The target error tolerance was set to $\Delta = 0.1$.

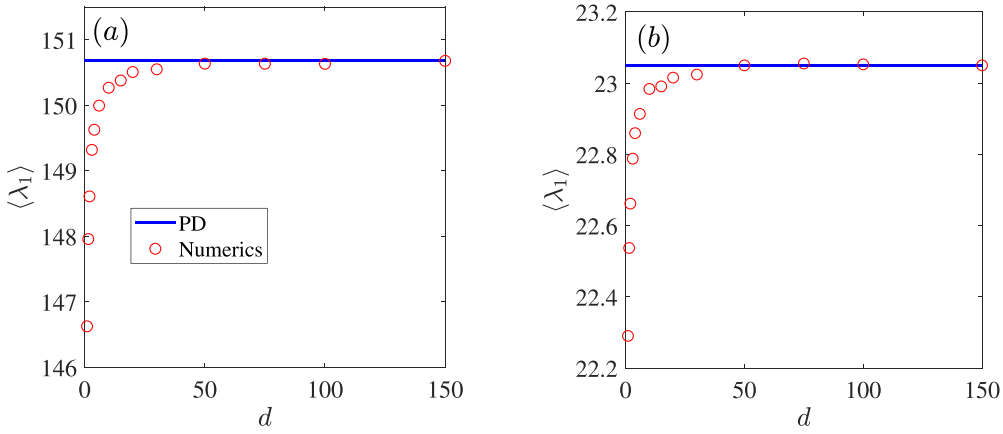

Figure 2. Scaling of $\langle\lambda_1\rangle$ with the dimensions of the matrix $\boldsymbol{X}$. This figure shows $\langle\lambda_1\rangle$, collected from direct numerical diagonalisation of 10^2 realisations of $\boldsymbol{J}$ (circles), as the size of the matrix $\boldsymbol{X}$ is increased, while the ratio $\alpha = \sqrt{N/M}$ is kept fixed. The scaling parameter d is defined such that each data point was obtained using a matrix $\boldsymbol{X}$ of size $(100 \cdot d) \times (80 \cdot d)$. The solid blue line represents the results obtained from the replica analysis, using the population dynamics algorithm, using populations of size $N_P = 10^5$. The set of control parameters used here is (a) $\alpha = \sqrt{5/4}$, $q = 11.8$ and $p(K) = \delta_{K,1}$; (b) $\alpha = \sqrt{5/4}$, $q = 8$ and $p(K) = \Theta(K)\Theta(1-K)$. In both figures the maximal number of nonzero elements in each row is set to $R = 70$ and in each column to $C = 60$. The target error tolerance was set to $\Delta = 0.1$. As can be observed from the figure, even for a relatively small matrix of size 100×80, finite size effects are responsible for a deviation of only up to $\sim 4\%$ from the analytical result. For a matrix ~ 50 times bigger than that, this deviation drops below the measurement's resolution.

As outlined in section 4, we computed the leading behaviour of $\langle\lambda_1\rangle$ as both of $\boldsymbol{J}$'s linear dimensions tend to infinity. Therefore, our analysis does not account for any finite size effects. However, in figure 2 we show that these corrections are negligible compared to the leading behaviour, which is perfectly captured by our analysis. Specifically, even for a relatively small matrix of size 100×80, finite size corrections are responsible for a deviation of merely $\sim 4\%$. When the matrix size is further increased, the numerical results quickly align with our analytical results, to the extent that the two are indistinguishable within our measurement's resolution.

In figure 3 we compare results for $\langle\lambda_1\rangle$ obtained from the replica analysis (solid line) and direct numerical diagonalisation (circles), as a function of q, which regulates the average density of nonzero elements in $\boldsymbol{X}$. In this figure we chose $\alpha = \sqrt{5/4}$ and the weight distributions (a) $p(K) = \delta_{K,1}$; (b) $p(K) = \Theta(1-K)\Theta(K)$. In both figures the maximal number of nonzero elements in each row is set to $R = 70$ and in each column to $C = 60$. The target error tolerance was set to $\Delta = 0.1$. In figure 2, the numerical data was obtained by averaging over 10^2 realisations of $\boldsymbol{X}$ with a fixed size of 5000×4000, such that finite size corrections are negligible. Within this framework, we find excellent agreement between the numerical and analytical results.

Building on the results for $\langle\lambda_1\rangle$, in figure 4, we compare the results for $T(u)$, obtained from equation (89) (red crosses) and direct numerical diagonalisation (green circles). For this

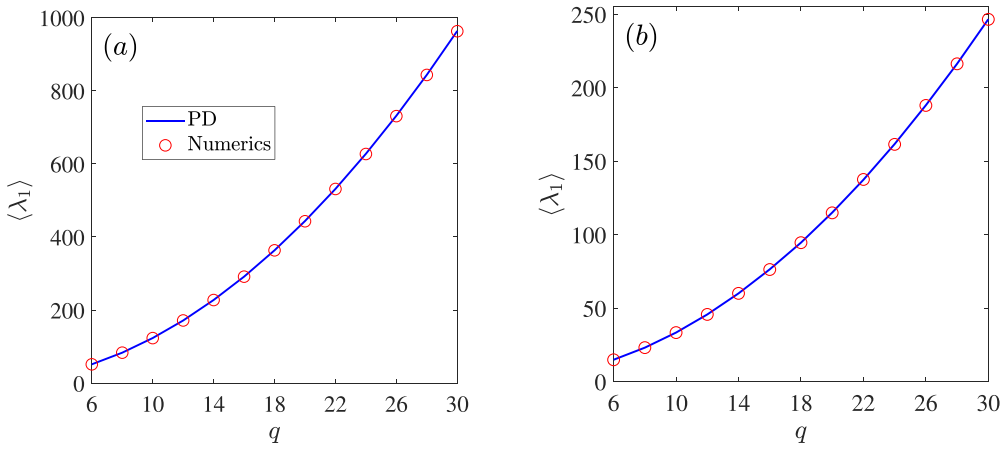

Figure 3. We show $\langle\lambda_1\rangle$ as obtained by both population dynamics (solid line) and direct numerical diagonalisation (circles) as a function of q, which regulates the average density of nonzero elements in $\boldsymbol{X}$. For this analysis, we used $\alpha = \sqrt{5/4}$ and set the weight distribution to (a) $p(K) = \delta_{K,1}$, and (b) $p(K) = \Theta(1-K)\Theta(K)$. In both figures the maximal number of nonzero elements in each row is set to $R = 70$ and in each column to $C = 60$. The target error tolerance was set to $\Delta = 0.1$. The numerical data represents an average over 10^2 realisations of $\boldsymbol{X}$, each of fixed and large dimensions 5000×4000. Under these conditions, the numerical and analytical results are in very strong agreement.

analysis, we used the same settings as in figure 3. The numerical and analytical results are again in very strong agreement.

The density $T(u)$ was numerically evaluated using a procedure based on a population generated by the algorithm outlined in section 6. We initially choose a resolution for our density, denoted by Δu, and split up the interval $[0,3]$ into bins of size Δu. We then generate a stable population following the algorithm of section 6 and randomly sample members of the population in order to evaluate the value

$$\frac{\sum_{\ell=1}^{s} \frac{K_\ell \mu_\ell}{\sigma_\ell}}{\lambda - \sum_{\ell=1}^{s} \frac{K_\ell^2}{\sigma_\ell}} \,. \tag{94}$$

Each time the computed value of (94) fell within a given bin, a count of one was added to that bin. This procedure was performed many times, after which the bin counts were normalised in order to produce the numerical density $T(u)$.

7. The dense limit

Below, we demonstrate how taking the large q limit recovers the familiar results of the non-central Wishart ensemble, by following the same lines as in [22]. To ensure a finite largest eigenvalue in this limit, we rescale the bond weights as

$$K_{ij} = \frac{\tilde{K}_{ij}}{q} \,, \tag{95}$$

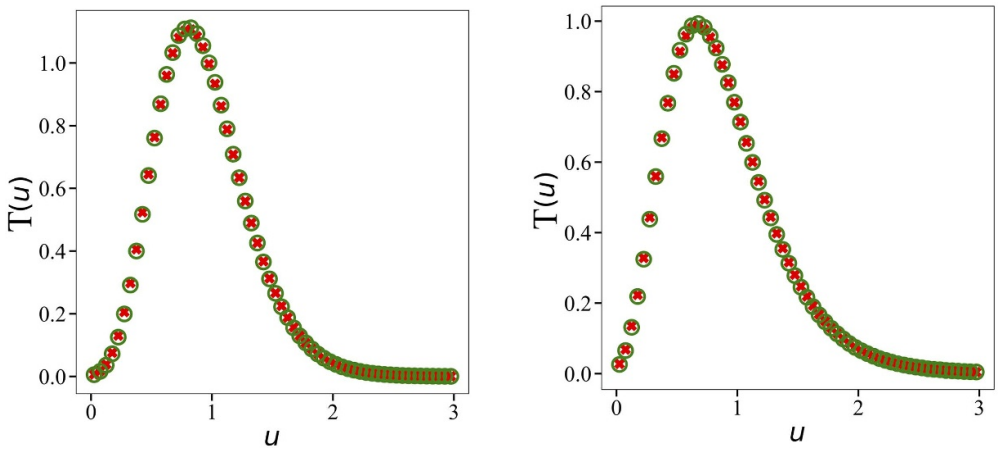

Figure 4. We compare the results for $T(u)$, obtained from (89) (red crosses) and direct numerical diagonalisation (green circles). For this analysis, we used $\alpha = \sqrt{5/4}$ and set the weight distribution to (a) $p(K) = \delta_{K,1}$, and (b) $p(K) = \Theta(1-K)\Theta(K)$ as in figure 3. In both figures the maximal number of nonzero elements in each row is set to $R = 70$ and in each column to $C = 60$. The target error tolerance was set to $\Delta = 0.1$. The numerical data represents an average over 10^2 realisations of $\boldsymbol{X}$, each of fixed and large dimensions 5000×4000. The numerical and analytical results are in very strong agreement. The resolution of the density, Δu, has been set at $\Delta u = 0.05$.

and assume that $\langle \tilde{K} \rangle$ and $\langle \tilde{K}^2 \rangle$ are nonzero and of $\mathcal{O}(1)$. Note that this scaling differs from that used in [22], which is $K = \tilde{K}/\sqrt{q}$, due to a subtle but important difference in the underlying assumptions. In [22], the authors consider a sparse *central* model, in which the nonzero entries satisfy $\langle K \rangle = 0$. In the dense limit, they recover the upper edge of the semicircle law. In contrast, our model is *non-central*, in the sense that $\langle K \rangle \neq 0$, which causes the largest eigenvalue to detach from the bulk of the spectrum. As we will later demonstrate, our scaling ensures that the detached $\langle \lambda_1 \rangle$ remains of $\mathcal{O}(1)$, whereas alternative scalings would yield a vanishing or diverging result in the $q \to \infty$ limit. Inserting (95) into the first two lines of (69), we obtain

$$\pi(\omega, h) = \sum_{s=1}^{\infty} \frac{s p_{\alpha q}(s)}{\alpha q} \int \{\mathrm{d}\rho\}_{s-1} \left\langle \delta\left(\omega - \left(\lambda - \frac{1}{(\alpha q)^2} \sum_{\ell=1}^{s-1} \frac{\alpha^2 \tilde{K}_\ell^2}{\sigma_\ell}\right)\right) \times \delta\left(h - \frac{1}{\alpha q} \sum_{\ell=1}^{s-1} \frac{\alpha \tilde{K}_\ell \mu_\ell}{\sigma_\ell}\right) \right\rangle_{\{\tilde{K}\}_{s-1}} \tag{96}$$

and

$$\rho(\sigma, \mu) = \sum_{s=1}^{\infty} \frac{s p_{\alpha^{-1} q}(s)}{\alpha^{-1} q} \int \{\mathrm{d}\pi\}_{s-1} \left\langle \delta\left(\sigma - \left(1 - \frac{1}{(\alpha^{-1} q)^2} \sum_{\ell=1}^{s-1} \frac{\alpha^{-2} \tilde{K}_\ell^2}{\omega_\ell}\right)\right) \times \delta\left(\mu - \frac{1}{\alpha^{-1} q} \sum_{\ell=1}^{s-1} \frac{\alpha^{-1} \tilde{K}_\ell h_\ell}{\omega_\ell}\right) \right\rangle_{\{\tilde{K}\}_{s-1}} . \tag{97}$$

As $q \to \infty$, the Poissonian weights effectively concentrate around $s = \alpha q \pm \mathcal{O}(\sqrt{\alpha q})$ in (96) and $s = \alpha^{-1} q \pm \mathcal{O}(\sqrt{\alpha^{-1} q})$ in (97). Thus, the quantities that appear in the δ-functions in

equations (96) and (97),

$$\bar{\omega} := \lambda - \frac{1}{(\alpha q)^2} \sum_{\ell=1}^{s-1} \frac{\alpha^2 \tilde{K}_\ell^2}{\sigma_\ell} \,, \tag{98}$$

$$\bar{\sigma} := 1 - \frac{1}{(\alpha^{-1} q)^2} \sum_{\ell=1}^{s-1} \frac{\alpha^{-2} \tilde{K}_\ell^2}{\omega_\ell} \,, \tag{99}$$

$$\bar{h} := \frac{1}{\alpha q} \sum_{\ell=1}^{s-1} \frac{\alpha \tilde{K}_\ell \mu_\ell}{\sigma_\ell} \,, \tag{100}$$

$$\bar{\mu} := \frac{1}{\alpha^{-1} q} \sum_{\ell=1}^{s-1} \frac{\alpha^{-1} \tilde{K}_\ell h_\ell}{\omega_\ell} \,, \tag{101}$$

are non-fluctuating in the limit, due to the law of large numbers. Consequently, the δ-functions force the densities to concentrate around $(\omega, h) = (\bar{\omega}, \bar{h})$ and $(\sigma, \mu) = (\bar{\sigma}, \bar{\mu})$,

$$\pi(\omega, h) = \delta(\omega - \bar{\omega})\, \delta\left(h - \bar{h}\right) \,, \tag{102}$$

$$\rho(\sigma, \mu) = \delta(\sigma - \bar{\sigma})\, \delta(\mu - \bar{\mu}) \,. \tag{103}$$

This fact, in turn, enables us to evaluate $\bar{\omega}, \bar{\sigma}, \bar{h}$ and $\bar{\mu}$ self-consistently, by substituting $\omega_\ell = \bar{\omega}$, $\sigma_\ell = \bar{\sigma}$, $h_\ell = \bar{h}$ and $\mu_\ell = \bar{\mu}$ into equations (98)–(101),

$$\bar{\omega} = \lambda + \mathcal{O}\left(q^{-1}\right) \,, \tag{104}$$

$$\bar{\sigma} = 1 + \mathcal{O}\left(q^{-1}\right) \,, \tag{105}$$

$$\bar{h} = \frac{\alpha \langle \tilde{K} \rangle \bar{\mu}}{\bar{\omega}} \,, \tag{106}$$

$$\bar{\mu} = \frac{\alpha^{-1} \langle \tilde{K} \rangle \bar{h}}{\bar{\sigma}} \,. \tag{107}$$

At this point, it becomes clear why the two cases — central and non-central — require different scalings of the nonzero entries in order to obtain $\langle \lambda_1 \rangle = \mathcal{O}(1)$. In the non-central case, where $\langle K \rangle \neq 0$, the moments of the variables h and μ would diverge in the $q \to \infty$ limit if we were to choose $K = \tilde{K}/\sqrt{q}$. This divergence would, in turn, lead to a diverging $\langle \lambda_1 \rangle$. On the contrary, if $\langle K \rangle = 0$, the scaling $K = \tilde{K}/q$ would lead to a vanishing moments of the variables h and μ in the $q \to \infty$ limit, hence to a vanishing $\langle \lambda_1 \rangle$.

Solving equations (104)–(107) for $\bar{\omega}, \bar{\sigma}$ and λ we obtain

$$\bar{\omega} = \langle \tilde{K} \rangle^2 + \mathcal{O}\left(q^{-1}\right) \,, \tag{108}$$

$$\bar{\sigma} = 1 + \mathcal{O}\left(q^{-1}\right) \,, \tag{109}$$

$$\lambda = \langle \tilde{K} \rangle^2 + \mathcal{O}\left(q^{-1}\right) \,. \tag{110}$$

Recalling that $\langle \lambda_1 \rangle = \lambda$, we finally get

$$\langle \lambda_1 \rangle = \langle \tilde{K} \rangle^2 + \mathcal{O}\left(q^{-1}\right) \,, \tag{111}$$

which coincides with the isolated largest eigenvalue of the non-central Wishart ensemble (see equation (52) in [82], with the appropriate scaling).

To obtain the density of the top eigenvector's components in the dense limit, we start from equation (89). After rescaling the weights as $K_{ij} = \tilde{K}_{ij}/q$, and accounting for the fact that $\rho(\sigma,\mu)$ and $\pi(\omega,h)$ concentrate (see equations (103) and (102)), we obtain

$$T(u) = \sum_{s=0}^{\infty} p_{\alpha q}(s) \left\langle \delta\left(u - \frac{\alpha\bar{\mu}}{\bar{\omega}\bar{\sigma}} \frac{1}{\alpha q} \sum_{\ell=1}^{s} \tilde{K}_{\ell}\right)\right\rangle_{\{K\}_s} . \tag{112}$$

Again, as $q \to \infty$, the Poissonian weights concentrate around $s = \alpha q \pm \mathcal{O}(\sqrt{\alpha q})$. Hence, the quantity

$$\bar{u} = \frac{\alpha\bar{\mu}}{\bar{\omega}\bar{\sigma}} \frac{1}{\alpha q} \sum_{\ell=1}^{s} \tilde{K}_{\ell} \tag{113}$$

is again non-fluctuating, due to the law of large numbers. Consequently, the δ function in equation (112) forces $T(u)$ to concentrate around $u = \bar{u}$, which evaluates to

$$\bar{u} = \frac{\alpha\bar{\mu}}{\bar{\sigma}\bar{\omega}} \langle \tilde{K} \rangle . \tag{114}$$

Combining the concentration of the top eigenvector's components with the normalisation of the eigenvectors, $|\boldsymbol{v}|^2 = M$, we expect that $\bar{u} = 1$. This result indeed follows directly from evaluating $\bar{\mu}$. Since the first two equations of (69) determine the distribution of μ's up to an arbitrary scaling, to fix the value of $\bar{\mu}$ we use the integral normalisation condition (third line in equation (69)). After rescaling the weights and accounting for the fact that $\pi(\sigma,\mu)$ and $\rho(\omega,h)$ concentrate, it takes the form

$$\sum_{s=0}^{\infty} p_{\alpha q}(s) \frac{\alpha^2\bar{\mu}^2}{\bar{\omega}^2\bar{\sigma}^2} \left\langle \frac{1}{(\alpha q)^2} \left(\sum_{\ell=1}^{s} \tilde{K}_{\ell}\right)^2 \right\rangle_{\{K\}_s} = 1 . \tag{115}$$

Evaluating the average over the weights, we obtain

$$\frac{\alpha^2\bar{\mu}^2}{\bar{\omega}^2\bar{\sigma}^2} \sum_{s=0}^{\infty} p_{\alpha q}(s) \frac{s(s-1)\langle \tilde{K} \rangle^2 + s\langle \tilde{K}^2 \rangle}{(\alpha q)^2} = 1 . \tag{116}$$

Using the known moments of the Poisson distribution, we have

$$\frac{\alpha^2\bar{\mu}^2}{\bar{\sigma}^2\bar{\omega}^2} \langle \tilde{K} \rangle^2 + \mathcal{O}\left(q^{-1}\right) = 1 . \tag{117}$$

Substituting equation (117) into (114), we obtain $\bar{u} = 1$ as anticipated, such that

$$T(u) = \delta(u-1) . \tag{118}$$

Hence, in the *dense limit*[6] of our non-central model, the top eigenvector is fully localised around $\mathbf{1} = (1,\ldots,1) \in \mathbb{R}^M$. Incidentally, we note that this localisation phenomenon is very similar to [83], valid for the slightly different setting of dense symmetric random matrices with *independent* entries drawn from an arbitrary distribution.

[6] Note that this corresponds to a sequence of two limits: first, $N, M \to \infty$ (with their ratio fixed), and next $q \to \infty$.

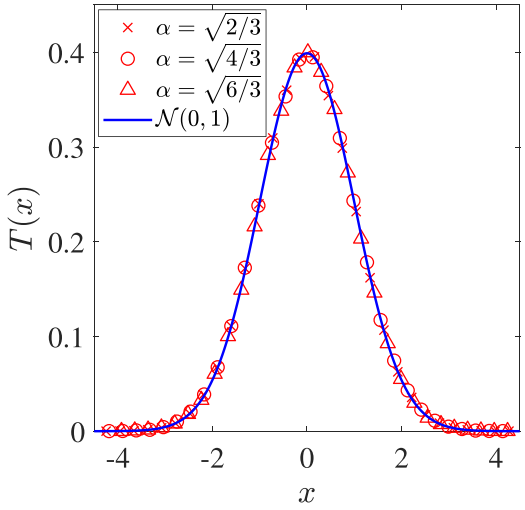

Figure 5. Density of the top eigenvector's components in the dense regime (i.e. $q = \alpha M$), plotted as a function of the rescaled variable $x = \delta\sqrt{M}(u-1)$, where δ is defined in equation (121). The numerical data (symbols) were obtained by diagonalizing 20 independent realizations of the matrix $\boldsymbol{J} = \boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}$, where $\boldsymbol{X}$'s entries are drawn from a uniform distribution $p(K) = \Theta(K)\Theta(1-K)$. We fixed $M = 4\cdot 10^4$ and examined three values of N, corresponding to ($\times$) $\alpha = \sqrt{2/3}$, ($\circ$) $\alpha = \sqrt{4/3}$, and ($\triangle$) $\alpha = \sqrt{6/3}$. The numerical results exhibit excellent agreement with the standard normal distribution $\mathcal{N}(0,1)$ (solid line).

For *finite* N,M, we could have defined the density of top eigenvector's components as $T_{N,M}(u \mid q, p(K))$. Our previous result would then correspond to computing the double limit

$$\lim_{q\to\infty}\lim_{M\to\infty} T_{\alpha^2 M,M}(u \mid q, p(K)) = \delta(u-1)\ . \tag{119}$$

Empirically, numerical diagonalisation on large but finite matrices shows that $T(u)$ indeed becomes narrower as q is gradually increased. However, it would be interesting to study how this complete localisation in the limit is approached on a narrower scale as N,M increase. While the finite N,M density $T_{N,M}(u \mid q, p(K))$ is not attainable via our method, we nevertheless conjecture (on the basis of numerical simulations) that the components of the top eigenvector display Gaussian fluctuations in the double-scaling limit

$$T(x) := \lim_{M\to\infty} \frac{1}{\delta\sqrt{M}} T_{\alpha^2 M,M}\left(u = 1 + \frac{1}{\sqrt{M}}\frac{x}{\delta} \mid q = \alpha M, p(K)\right) = \mathcal{N}(0,1) \tag{120}$$

with

$$\delta = \frac{\alpha}{\sqrt{\frac{\langle \tilde{K}^2\rangle}{\langle \tilde{K}\rangle^2} - 1}}\ . \tag{121}$$

In figure 5, we numerically validate our conjecture. The figure shows the density of the top eigenvector's components in the dense regime (i.e. $q = \alpha M$), plotted as a function of the rescaled variable $x = \delta\sqrt{M}(u-1)$, where δ is defined in equation (121). The numerical data (symbols) were obtained by diagonalizing 20 independent realisations of the matrix $\boldsymbol{J} = \boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}$,

where $\boldsymbol{X}$'s entries are drawn from a uniform distribution $p(K) = \Theta(K)\Theta(1-K)$. We fixed $M = 4 \cdot 10^4$ and examined three values of N, corresponding to ($\times$) $\alpha = \sqrt{2/3}$, ($\circ$) $\alpha = \sqrt{4/3}$, and ($\triangle$) $\alpha = \sqrt{6/3}$. The numerical results exhibit excellent agreement with the standard normal distribution $\mathcal{N}(0,1)$ (solid line). A first-principles proof of this conjecture would be very welcome.

8. Summary and conclusions

In summary, we developed a replica formalism to compute the top eigenpair statistics of sparse correlation matrices of the form $\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}$, where the nonzero entries follow a nonzero mean weight distribution $p(K)$, leading to an isolated largest eigenvalue.

Specifically, we focused on the average largest eigenvalue and the density of its associated eigenvector components. The problem of evaluating the average largest eigenvalue can be reformulated as an optimisation problem involving a quadratic Hamiltonian on the sphere. In the zero-temperature limit $\beta \to \infty$, the Gibbs measure concentrates around the ground state, corresponding to the top eigenvector. Using the replica method, we evaluated the disorder-averaged partition function and derived a system of self-consistent equations governing the order parameter λ (see equation (69)).

We solved these equations via a population dynamics algorithm and identified $\langle \lambda_1 \rangle$, the average largest eigenvalue, as the critical value of λ that determines the convergence of the population dynamics: for $\lambda < \langle \lambda_1 \rangle$, variables diverge, while for $\lambda > \langle \lambda_1 \rangle$, they converge to zero. Numerical simulations confirmed excellent agreement between this critical value and direct numerical diagonalisation, both for the degenerate case $p(K) = \delta_{K,1}$ and for a uniform weight distribution over $K \in [0,1]$.

Building on this, we extended our method to compute the density of top eigenvector components. Again, numerical results showed excellent agreement with diagonalisation for both weight distributions.

Finally, we demonstrated that taking the appropriate dense limit of our model recovers known results from the noncentral Wishart ensemble. Future work should explore the non-gapped regime and the connection between $p(K)$ and the detachment transition.

Data availability statement

No new data were created or analysed in this study.

Acknowledgment

P V acknowledges support from UKRI FLF Scheme (No. MR/X023028/1).

Appendix A. Upper bound for $\langle \lambda_1 \rangle$

In this appendix we show that

$$\langle \lambda_1 \rangle \leqslant \left[\max\left(|\zeta^-|, |\zeta^+|\right)\right]^2 RC\,, \tag{A.1}$$

where ζ^- (ζ^+) is the lower (upper) bound of the support of $p(K)$, while R and C are the maximal numbers of nonzero elements in each row and in each column respectively. Our starting point

is the identification of λ_1 with the square of the spectral norm of the matrix $\boldsymbol{X}$. According to identity 15.511.1 from [84], the spectral norm obeys

$$\lambda_1 \leq \left(\max_{1\leqslant j\leqslant M}\sum_{i=1}^{N}|X_{ij}|\right)\left(\max_{1\leqslant i\leqslant N}\sum_{j=1}^{M}|X_{ij}|\right) . \tag{A.2}$$

Since $X_{ij} = c_{ij}K_{ij}$ (equation (1)), we can use the fact that $p(K)$ has a bounded support and that the number of nonzero elements in each row (column) is restricted by R (C) to write

$$\begin{aligned}\lambda_1 &\leqslant \left[\max\left(|\zeta^-|,|\zeta^+|\right)\right]^2\left(\max_{1\leqslant j\leqslant M}\sum_{i=1}^{N}c_{ij}\right)\left(\max_{1\leqslant i\leqslant N}\sum_{j=1}^{M}c_{ij}\right)\\ &\leqslant \left[\max\left(|\zeta^-|,|\zeta^+|\right)\right]^2 RC\,.\end{aligned} \tag{A.3}$$

Since this inequality holds for every realisation of $\boldsymbol{X}$, the ensemble average of λ_1 clearly satisfies this condition too. Hence, we obtained our desired result, equation (A.1).

Appendix B. Performing the Average in (13)

In this appendix, we show how to compute the average

$$\left\langle\prod_{i=1}^{M}\prod_{j=1}^{N}\exp\left(\beta X_{ji}\sum_{a=1}^{n}v_{ia}u_{ja}\right)\right\rangle , \tag{B.1}$$

in the $q \ll \sqrt{NM}$ limit. This average is performed over different realisations of the $N\times M$ random matrix $\boldsymbol{X}$, whose entries are i.i.d random variables, expressed as $X_{ji} = c_{ji}K_{ji}$, and are drawn from

$$P(X_{ji}) = \left[\frac{q}{\sqrt{NM}}\delta_{c_{ji},1} + \left(1-\frac{q}{\sqrt{NM}}\right)\delta_{c_{ji},0}\right]p(K_{ji}) , \tag{B.2}$$

with $p(K)$ being the weight distribution. First, we use the independence of the entries to factorise the average,

$$\left\langle\prod_{i=1}^{M}\prod_{j=1}^{N}\exp\left(\beta X_{ji}\sum_{a=1}^{n}v_{ia}u_{ja}\right)\right\rangle = \prod_{i=1}^{M}\prod_{j=1}^{N}\left\langle\exp\left(\beta cK\sum_{a=1}^{n}v_{ia}u_{ja}\right)\right\rangle_{c,K} , \tag{B.3}$$

where $\langle\cdot\rangle_{c,K}$ denotes averaging over a single instance of the random variables c and K. Next, we average over the c's and take the $q \ll \sqrt{NM}$ limit to obtain

$$\begin{aligned}\left\langle\prod_{i=1}^{M}\prod_{j=1}^{N}\exp\left(\beta X_{ji}\sum_{a=1}^{n}v_{ia}u_{ja}\right)\right\rangle &= \prod_{i=1}^{M}\prod_{j=1}^{N}\left[1+\frac{q}{\sqrt{NM}}\left(\left\langle e^{\beta K\sum_{a=1}^{n}v_{ia}u_{ja}}\right\rangle_K - 1\right)\right]\\ &\simeq \exp\left[\frac{q}{\sqrt{NM}}\sum_{i=1}^{M}\sum_{j=1}^{N}\left(\left\langle e^{\beta K\sum_{a=1}^{n}v_{ia}u_{ja}}\right\rangle_K - 1\right)\right],\end{aligned} \tag{B.4}$$

which matches the result in (13).

ORCID iDs

Preben Forer 0009-0005-8449-4013
Pierpaolo Vivo 0000-0002-2366-3080
Sabrina Aufiero 0009-0007-5336-4165

References

[1] Fan J, Han F and Liu H 2014 *Natl Sci. Rev.* **1** 293
[2] Jolliffe I 2002 *Principal Component Analysis* (*Springer Series in Statistics*) (Springer)
[3] James G, Witten D, Hastie T and Tibshirani R 2014 *An Introduction to Statistical Learning: With Applications in R* (*Springer Texts in Statistics*) (Springer)
[4] Sachs K, Perez O, Pe'er D, Lauffenburger D A and Nolan G P 2005 *Science* **308** 523
[5] Butte A J, Tamayo P, Slonim D, Golub T R and Kohane I S 2000 *Proc. Natl Acad. Sci.* **97** 12182
[6] Witten D M, Tibshiran R and Hastie T 2009 *Biostatistics* **10** 515
[7] Deng Y, Jiang Y-H, Yang Y, He Z, Luo F and Zhou J 2012 *BMC Bioinform.* **13** 113
[8] Kurtz Z D, Müller C L, Miraldi E R, Littman D R, Blaser M J and Bonneau R A 2015 *PLS Comput. Biol.* **11** 1
[9] Levy O and Goldberg Y 2014 *Neural Inf. Process. Syst.* **27** 2177
[10] Bouchaud J P 2009 arXiv:0903.2428
[11] Newman M E J 2003 *SIAM Rev.* **45** 167
[12] Fan J, Liao Y and Liu H 2016 *Econ. J.* **19** C1
[13] Majumdar S N and Vivo P 2012 *Phys. Rev. Lett.* **108** 200601
[14] Johnstone I M 2001 *Ann. Stat.* **29** 295–327
[15] Mardia K V, Kent J T and Bibby J M 1979 *Multivariate Analysis* (Academic)
[16] Monasson R and Villamaina D 2015 *Europhys. Lett.* **112** 50001
[17] Bai Z and Silverstein J W 2010 *Spectral Analysis of Large Dimensional Random Matrices* (Springer)
[18] Stoica P and Cedervall M 1996 *IFAC Proc.* vol 29 p 4098
[19] Bianchi P, Debbah M, Maida M and Najim J 2011 *IEEE Trans. Inf. Theory* **57** 2400
[20] Nadakuditi R R and Edelman A 2008 *IEEE Trans. Signal Process.* **56** 2625
[21] Nagao T and Tanaka T 2007 *J. Phys. A: Math. Theor.* **40** 4973
[22] Susca V A, Vivo P and Kühn R 2019 *J. Phys. A: Math. Theor.* **52** 485002
[23] Susca V A, Vivo P and Kühn R 2020 *J. Phys. A: Math. Theor.* **54** 015004
[24] Susca V A, Vivo P and Kühn R 2021 *SciPost Phys. Lect. Notes* **33**
[25] Wishart J 1928 *Biometrika* **20A** 32
[26] Gupta A and Nagar D 1999 *Matrix Variate Distributions* (*Monographs and Surveys in Pure and Applied Mathematics*) (Taylor & Francis)
[27] Dyson F J 1962 *J. Math. Phys.* **3** 140
[28] Dyson F J 1962 *J. Math. Phys.* **3** 157
[29] Dyson F J 1962 *J. Math. Phys.* **3** 166
[30] Vivo P, Majumdar S N and Bohigas O 2007 *J. Phys. A: Math. Theor.* **40** 4317
[31] Katzav E and Pérez Castillo I 2010 *Phys. Rev. E* **82** 040104
[32] Majumdar S N and Vergassola M 2009 *Phys. Rev. Lett.* **102** 060601
[33] Zavatone-Veth J A and Pehlevan C 2023 *SciPost Phys. Core* **6** 026
[34] Forrester P J 2010 *Log-Gases and Random Matrices* (Princeton University Press)
[35] Livan G, Novaes M and Vivo P 2018 *Introduction to Random Matrices: Theory and Practice* (Springer)
[36] Rogers T, Pérez Castillo I, Kühn R and Takeda K 2008 *Phys. Rev. E* **78** 031116
[37] Pérez Castillo I and Metz F L 2018 *Phys. Rev. E* **97** 032124
[38] Forrester P J 2015 *Constr. Approx.* **41** 589–613
[39] Kühn R 2008 *J. Phys. A: Math. Theor.* **41** 295002
[40] Rogers T and Pérez Castillo I 2009 *Phys. Rev. E* **79** 012101
[41] Rogers T, Vicente C P, Takeda K and Pérez Castillo I 2010 *J. Phys. A: Math. Theor.* **43** 195002
[42] Metz F L, Neri I and Bollé D 2010 *Phys. Rev. E* **82** 031135
[43] Metz F L, Neri I and Bollé D 2011 *Phys. Rev. E* **84** 055101
[44] Neri I and Metz F L 2012 *Phys. Rev. Lett.* **109** 030602

[45] Bollé D, Metz F L and Neri I 2013 On the spectra of large sparse graphs with cycles *Spectral Analysis, Differential Equations and Mathematical Physics: A Festschrift in Honor of Fritz Gesztesy's 60th Birthday* (American Mathematical Society) pp 35–58
[46] Metz F L, Parisi G and Leuzzi L 2014 *Phys. Rev. E* **90** 052109
[47] Neri I and Metz F L 2016 *Phys. Rev. Lett.* **117** 224101
[48] Péché S 2009 *Probab. Theory Relat. Fields* **143** 481
[49] El Karoui N 2007 *Ann. Probab.* **35** 663–714
[50] Bao Z, Pan G and Zhou W 2015 *Ann. Stat.* **43** 382
[51] Natesh S P and Yin J 2014 *Ann. Appl. Probab.* **24** 935
[52] Ding X and Yang F 2018 *Ann. Appl. Probab.* **28** 1679
[53] Barucca P, Kieburg M and Ossipov A 2020 *Europhys. Lett.* **129** 60003
[54] Hwang J Y, Lee J O and Yang W 2020 *Bernoulli* **26** 2400–35
[55] Auffinger A and Tang S 2015 *Stoch. Process. Appl.* **126** 3310
[56] Baik J, Ben Arous G and Péché S 2005 *Ann. Probab.* **33** 1643
[57] Bassler K E, Forrester P J and Frankel N E 2009 *J. Math. Phys.* **50** 033302
[58] Benaych-Georges F and Nadakuditi R R 2011 *Adv. Math.* **227** 494
[59] Zamponi F 2010 arXiv:1008.4844
[60] Mezard M, Parisi G and Virasoro M 1987 *Spin Glass Theory and Beyond* (*Lecture Notes in Physics Series*) (World Scientific)
[61] Edwards S F and Jones R C 1976 *J. Phys. A: Math. Gen.* **9** 1595
[62] Rodgers G J and Dominicis C D 1990 *J. Phys. A: Math. Gen.* **23** 1567
[63] Akara-pipattana P and Evnin O 2025 *J. Phys. A: Math. Gen.* **58** 035006
[64] Biroli G and Monasson R 1999 *J. Phys. A: Math. Gen.* **32** L255
[65] Semerjian G and Cugliandolo L F 2002 *J. Phys. A: Math. Gen.* **35** 4837
[66] Bianconi G 2008 arXiv:0804.1744
[67] Kühn R, Mourik J v, Weigt M and Zippelius A 2007 *J. Phys. A: Math. Theor.* **40** 9227
[68] Dean D S 2002 *J. Phys. A: Math. Gen.* **35** L153
[69] Metz F L and Stariolo D A 2015 *Phys. Rev. E* **92** 042153
[70] Metz F L and Pérez Castillo I 2016 *Phys. Rev. Lett.* **117** 104101
[71] Metz F L and Pérez Castillo I 2017 *Phys. Rev. B* **96** 064202
[72] López F A and Coolen A C C 2020 *J. Phys. A: Math. Theor.* **53** 065002
[73] Bollobas B 2001 *Random Graphs* 2nd edn (Cambridge University Press)
[74] Krivelevich M and Sudakov B 2003 *Comb. Probab. Comput.* **12** 61–72
[75] Kühn R 2015 *Acta Phys. Polon. B* **46** 1653
[76] Mezard M and Parisi G 2001 *Eur. Phys. J. B* **20** 217
[77] Zdeborová L and Krzakala F 2015 *Adv. Phys.* **65** 453
[78] Kühn R and Rogers T 2017 *Europhys. Lett.* **118** 68003
[79] Kühn R 2024 *Phys. Rev. E* **110** L032301
[80] Bartolucci S, Caccioli F, Caravelli F and Vivo P 2024 *Proc. Natl Acad. Sci. USA* **121** 40
[81] Tapias D, Grüger B, Kühn R and Sollich P 2025 Localization and top eigenvalue detection (arXiv:2507.07310 [cond-mat.dis-nn])
[82] Jagadish V 2014 *Phys. Rev. E* **90** 042144
[83] Lang D W 1964 *Phys. Rev.* **135** B1082
[84] Gradshteyn I S and Ryzhik I M 2007 *Table of Integrals, Series and Products* 7th edn (Academic)