Spectral density of sparse sample covariance matrices

2007 J. Phys. A: Math. Theor. 40 4973

(http://iopscience.iop.org/1751-8121/40/19/003)

IOP PUBLISHING JOURNAL OF PHYSICS A: MATHEMATICAL AND THEORETICAL

J. Phys. A: Math. Theor. **40** (2007) 4973–4987 doi:10.1088/1751-8113/40/19/003

Spectral density of sparse sample covariance matrices

Taro Nagao[1] **and Toshiyuki Tanaka**[2]

[1] Graduate School of Mathematics, Nagoya University, Chikusa-ku, Nagoya 464-8602, Japan
[2] Graduate School of Informatics, Kyoto University, Sakyo-ku, Kyoto 606-8501, Japan

Received 22 December 2006, in final form 7 March 2007
Published 24 April 2007
Online at stacks.iop.org/JPhysA/40/4973

Abstract
Applying the replica method of statistical mechanics, we evaluate the eigenvalue density of the large random matrix (sample covariance matrix) of the form $J = A^{\mathrm{T}}A$, where A is an $M \times N$ real sparse random matrix. The difference from a dense random matrix is the most significant in the tail region of the spectrum. We compare the results of several approximation schemes, focusing on the behaviour in the tail region.

PACS numbers: 02.50.−r, 02.10.Yn

1. Introduction

The investigation of sample covariance matrices has a long history. It was originally introduced in multivariate statistical analysis [1]. Let us consider N-dimensional sample vectors $X^{(m)}$, $m = 1, 2, \ldots, L$, where L is the number of samples. The sample covariance matrix S is then defined as

$$S_{jl} = \frac{1}{L-1}\sum_{m=1}^{L}\big(X_j^{(m)} - \bar{X}_j\big)\big(X_l^{(m)} - \bar{X}_l\big), \tag{1.1}$$

where the sample mean vector $\bar{X}$ is

$$\bar{X} = \frac{1}{L}\sum_{m=1}^{L} X^{(m)}. \tag{1.2}$$

Let us suppose that $X^{(m)}$ are random vectors: the components $X_j^{(m)}$, $j = 1, 2, \ldots, N$, are independent and identically distributed Gaussian random variables. Then the distribution of $(L-1)S$ is the same as the distribution of the $N \times N$ matrix

$$J = A^{\mathrm{T}}A \tag{1.3}$$

(A^{T} is the transpose of A), where the elements of the $(L-1) \times N$ matrix A are independent and identically distributed Gaussian random variables. The ensemble of sample covariance matrices is sometimes called ‘chiral Gaussian’ or ‘Laguerre’ ensemble and finds applications in neural-network learning [2], quantum chromodynamics [3], mesoscopic physics [4], finance [5, 6] and wireless communication [7].

In multivariate statistics one is usually interested in the limit $L \to \infty$ with fixed N. On the other hand, motivated by Wigner's celebrated work [8] on real symmetric random matrices, Marčenko and Pastur studied another limit $N, L \to \infty$ with $L/N \to \alpha$ $(0 < \alpha \leqslant 1)$ and derived the asymptotic density for the scaled eigenvalues of J as [9, 10]

$$\rho(x) = (1-\alpha)\delta(x) + \frac{1}{2\pi x}\sqrt{(x-(\sqrt{\alpha}-1)^2)((\sqrt{\alpha}+1)^2-x)}. \tag{1.4}$$

This asymptotic result is valid for more general distributions of the matrix A, and is sometimes called Marčenko–Pastur law.

One of the simplest ways to modify the random matrix J so that Marčenko–Pastur law breaks down is to make the matrix A sparse. An example demonstrating the significance of considering random matrix ensembles defined on the basis of sparse A can be found in communication theory: the information-theoretic channel capacity of a randomly spread code-division multiple-access (CDMA) channel is evaluated in terms of eigenvalue distribution of the matrix $J = A^{\mathrm{T}}A$ with A defining the random spreading. It has been argued [11] that some wideband CDMA schemes can be modelled as a sparsely spread system, where deviations from Marčenko–Pastur law may affect the performance of such systems. Sparse random matrices in general are also of interest in many branches of applications. In particular, the eigenvector localization expected to occur in sparse random matrices is one of the most outstanding phenomena in disordered systems. The appearance of isolated eigenvalue spectra in the tail region is another interesting feature. Such features are also observed in heavy-tailed random matrices and were recently studied in detail [12–14].

In this paper, as a continuation of a brief report by one of the authors [15], we investigate the spectral (eigenvalue) density of sparse sample covariance matrices. In the case of real symmetric sparse random matrices R whose elements R_{jl} $(j \leqslant l)$ are independently distributed, the asymptotic spectral density was already studied by several authors [16–19]. Field theoretic (replica and supersymmetry) methods were usually employed. For example, Semerjian and Cugliandolo utilized a sophisticated version of the replica method and discussed the spectral density in connection with the percolation problem [19]. In the next section, we present a similar replica method developed for the analysis of sparse sample covariance matrices. In sections 3 and 4, the effective medium approximation (EMA) and its symmetrized version are introduced. In section 5, we consider the case of binary distribution and evaluate the asymptotic spectral density. In section 6, the tail behaviour of the spectral density is analysed by means of the single defect approximation (SDA), which improves the symmetrized version of EMA.

2. Replica method

We are interested in the asymptotic spectral density of the $N \times N$ matrix

$$J = A^{\mathrm{T}}A, \tag{2.1}$$

where A is an $M \times N$ real sparse random matrix $(M \leqslant N)$. The elements of A are independently distributed according to the probability distribution function

$$P(x) = \left(1 - \frac{p}{N}\right)\delta(x) + \frac{p}{N}\Pi(x) \tag{2.2}$$

with a fixed positive number $p < N$. Here $\delta(x)$ is Dirac's delta function. The asymptotic limit $N \to \infty$ with p and $\alpha = M/N$ fixed will be considered. We assume that $\Pi(x)$ does not depend on N and has zero measure on the origin:

$$\lim_{\epsilon \searrow 0} \int_{-\epsilon}^{\epsilon} \Pi(x)\,\mathrm{d}x = 0. \tag{2.3}$$

The Fourier transform of $P(x)$ is given by

$$\hat{P}(k) = \int_{-\infty}^{\infty} P(x)\,\mathrm{e}^{-\mathrm{i}kx}\,\mathrm{d}x = 1 + \frac{p}{N} f(k), \tag{2.4}$$

where

$$f(k) = \int_{-\infty}^{\infty} \Pi(x)\,\mathrm{e}^{-\mathrm{i}kx}\,\mathrm{d}x - 1. \tag{2.5}$$

Let us denote the average over MN copies of $P(x)$ by brackets $\langle\cdot\rangle$. Then the average spectral density of J is defined as

$$\rho(\mu) = \left\langle \frac{1}{N} \sum_{j=1}^{N} \delta(\mu - \mu_j) \right\rangle, \tag{2.6}$$

where $\mu_1, \mu_2, \ldots, \mu_N$ are the eigenvalues of J. We can rewrite $\rho(\mu)$ as

$$\begin{aligned} \rho(\mu) &= \frac{1}{N\pi} \operatorname{Im} \operatorname{Tr}\langle (J - (\mu + \mathrm{i}\epsilon)I)^{-1} \rangle \\ &= \frac{2}{N\pi} \operatorname{Im} \frac{\partial}{\partial\mu} \langle \ln Z(\mu + \mathrm{i}\epsilon) \rangle, \end{aligned} \tag{2.7}$$

where I is an $N \times N$ identity matrix, ϵ is a positive infinitesimal number and

$$Z(\mu') = \int \prod_{j=1}^{N} \mathrm{d}\phi_j \exp\left\{ \frac{\mathrm{i}\mu'}{2} \sum_{j=1}^{N} \phi_j^2 - \frac{\mathrm{i}}{2} \sum_{j=1}^{N} \sum_{l=1}^{N} J_{jl} \phi_j \phi_l \right\}, \qquad \mu' = \mu + \mathrm{i}\epsilon \tag{2.8}$$

is called the partition function. In order to evaluate the average over $P(x)$, using the relation

$$\lim_{n \to 0} \frac{\partial}{\partial n} \ln\langle Z^n \rangle = \langle \ln Z \rangle, \tag{2.9}$$

we deduce

$$\begin{aligned} \rho(\mu) &= \frac{2}{N\pi} \operatorname{Im} \frac{\partial}{\partial\mu} \left\{ \lim_{n \to 0} \frac{\partial}{\partial n} \ln\langle (Z(\mu + \mathrm{i}\epsilon))^n \rangle \right\} \\ &= \lim_{n \to 0} \frac{2}{\pi n} \operatorname{Im} \frac{\partial}{\partial\mu} \frac{1}{N} \ln\langle (Z(\mu + \mathrm{i}\epsilon))^n \rangle. \end{aligned} \tag{2.10}$$

Thus we find that the spectral density can be written in terms of the average $\langle Z^n \rangle$.

Noting

$$\int \prod_{k=1}^{M} \mathrm{d}\psi_k \exp\left\{ \frac{\mathrm{i}}{2} \sum_{k=1}^{M} \left(\psi_k - \sum_{j=1}^{N} \phi_j A_{kj} \right) \left(\psi_k - \sum_{l=1}^{N} \phi_l A_{kl} \right) \right\} = (2\pi\mathrm{i})^{M/2}, \tag{2.11}$$

we find

$$\begin{aligned} &\frac{1}{(2\pi\mathrm{i})^{M/2}} \int \prod_{k=1}^{M} \mathrm{d}\psi_k \exp\left\{ \frac{\mathrm{i}}{2} \sum_{k=1}^{M} \psi_k^2 - \mathrm{i} \sum_{k=1}^{M} \sum_{j=1}^{N} \psi_k A_{kj} \phi_j \right\} \\ &\quad = \exp\left\{ -\frac{\mathrm{i}}{2} \sum_{k=1}^{M} \sum_{j=1}^{N} \sum_{l=1}^{N} \phi_j A_{kj} \phi_l A_{kl} \right\} = \exp\left\{ -\frac{\mathrm{i}}{2} \sum_{j=1}^{N} \sum_{l=1}^{N} J_{jl} \phi_j \phi_l \right\}, \end{aligned} \tag{2.12}$$

from which it follows that

$$Z(\mu') = \frac{1}{(2\pi \mathrm{i})^{M/2}} \int \prod_{j=1}^{N} \mathrm{d}\phi_j \prod_{k=1}^{M} \mathrm{d}\psi_k \exp\left\{ \frac{\mathrm{i}\mu'}{2} \sum_{j=1}^{N} \phi_j^2 + \frac{\mathrm{i}}{2} \sum_{k=1}^{M} \psi_k^2 \right\}$$
$$\times \exp\left\{ -\mathrm{i} \sum_{k=1}^{M} \sum_{j=1}^{N} \psi_k A_{kj} \phi_j \right\}. \tag{2.13}$$

Now we introduce the vectors (replica variables)

$$\vec{\phi}_j = \left(\phi_j^{(1)}, \phi_j^{(2)}, \ldots, \phi_j^{(n)}\right), \qquad \vec{\psi}_k = \left(\psi_k^{(1)}, \psi_k^{(2)}, \ldots, \psi_k^{(n)}\right) \tag{2.14}$$

and the corresponding measures

$$\mathrm{d}\vec{\phi}_j = \mathrm{d}\phi_j^{(1)}\,\mathrm{d}\phi_j^{(2)} \cdots \mathrm{d}\phi_j^{(n)}, \qquad \mathrm{d}\vec{\psi}_k = \mathrm{d}\psi_k^{(1)}\,\mathrm{d}\psi_k^{(2)} \cdots \mathrm{d}\psi_k^{(n)} \tag{2.15}$$

so that the average $\langle Z^n \rangle$ can be rewritten as

$$\langle Z^n \rangle = \frac{1}{(2\pi \mathrm{i})^{Mn/2}} \int \prod_{j=1}^{N} \mathrm{d}\vec{\phi}_j \prod_{k=1}^{M} \mathrm{d}\vec{\psi}_k \exp\left\{ \frac{\mathrm{i}\mu'}{2} \sum_{j=1}^{N} \vec{\phi}_j^{\,2} + \frac{\mathrm{i}}{2} \sum_{k=1}^{M} \vec{\psi}_k^{\,2} \right\}$$
$$\times \left\langle \exp\left\{ -\mathrm{i} \sum_{k=1}^{M} \sum_{j=1}^{N} A_{kj} \vec{\psi}_k \cdot \vec{\phi}_j \right\} \right\rangle, \tag{2.16}$$

where $\vec{\phi}_j^{\,2} = \vec{\phi}_j \cdot \vec{\phi}_j$ and $\vec{\psi}_k^{\,2} = \vec{\psi}_k \cdot \vec{\psi}_k$.

Using the Fourier transform (2.4) of the probability distribution function $P(x)$, we obtain

$$\left\langle \exp\left\{ -\mathrm{i} \sum_{k=1}^{M} \sum_{j=1}^{N} A_{kj} \vec{\psi}_k \cdot \vec{\phi}_j \right\} \right\rangle = \prod_{k=1}^{M} \prod_{j=1}^{N} \left(\int \mathrm{d}x\, P(x) \exp\{-\mathrm{i}\vec{\psi}_k \cdot \vec{\phi}_j x\} \right)$$
$$\sim \prod_{k=1}^{M} \prod_{j=1}^{N} \exp\left\{ \frac{p}{N} f(\vec{\psi}_k \cdot \vec{\phi}_j) \right\} \tag{2.17}$$

in the limit $N \to \infty$. It follows that

$$\langle Z^n \rangle \sim \frac{1}{(2\pi \mathrm{i})^{Mn/2}} \int \prod_{j=1}^{N} \mathrm{d}\vec{\phi}_j \prod_{k=1}^{M} \mathrm{d}\vec{\psi}_k \exp\left\{ \frac{\mathrm{i}\mu'}{2} \sum_{j=1}^{N} \vec{\phi}_j^{\,2} + \frac{\mathrm{i}}{2} \sum_{k=1}^{M} \vec{\psi}_k^{\,2} \right\}$$
$$\times \exp\left\{ \frac{p}{N} \sum_{k=1}^{M} \sum_{j=1}^{N} f(\vec{\psi}_k \cdot \vec{\phi}_j) \right\}. \tag{2.18}$$

3. Effective medium approximation

We are in a position to explain how to evaluate the spectral density by means of a scheme called the effective medium approximation (EMA) [19]. We first rewrite (2.18) as

$$\langle Z^n \rangle \sim \int \prod_{j=1}^{N} \mathrm{d}\vec{\phi}_j \exp\left\{ \frac{\mathrm{i}\mu'}{2} \sum_{j=1}^{N} \vec{\phi}_j^{\,2} \right\} \left[\frac{1}{(2\pi \mathrm{i})^{n/2}} \int \mathrm{d}\vec{\psi} \exp\left\{ \frac{\mathrm{i}}{2} \vec{\psi}^{\,2} + \frac{p}{N} \sum_{j=1}^{N} f(\vec{\psi} \cdot \vec{\phi}_j) \right\} \right]^M . \tag{3.1}$$

Introducing

$$\tilde{\xi}(\vec{\phi}) = \frac{1}{N}\sum_{j=1}^{N}\delta(\vec{\phi}-\vec{\phi}_j), \tag{3.2}$$

we find

$$\langle Z^n\rangle \sim \int \prod_{j=1}^{N} \mathrm{d}\vec{\phi}_j \exp\left\{N\frac{\mathrm{i}\mu'}{2}\int \mathrm{d}\vec{\phi}\,\tilde{\xi}(\vec{\phi})\vec{\phi}^2\right\} \times\left[\frac{1}{(2\pi\mathrm{i})^{n/2}}\int \mathrm{d}\vec{\psi}\exp\left\{\frac{\mathrm{i}}{2}\vec{\psi}^2 + p\int \mathrm{d}\vec{\phi}\,\tilde{\xi}(\vec{\phi})f(\vec{\psi}\cdot\vec{\phi})\right\}\right]^M. \tag{3.3}$$

Let us introduce an order parameter function $\xi(\vec{\phi})$, which is normalized as

$$\int \xi(\vec{\phi})\,\mathrm{d}\vec{\phi} = 1. \tag{3.4}$$

Then, using the functional integral of the delta function

$$\int \mathcal{D}\xi(\vec{\phi})\prod_{\vec{\phi}}\delta(\xi(\vec{\phi})-\tilde{\xi}(\vec{\phi})) = 1, \tag{3.5}$$

we obtain

$$\langle Z^n\rangle \sim \int \mathcal{D}\xi(\vec{\phi})\int\prod_{j=1}^{N}\mathrm{d}\vec{\phi}_j\prod_{\vec{\phi}}\delta(\xi(\vec{\phi})-\tilde{\xi}(\vec{\phi}))Q[\xi(\vec{\phi})], \tag{3.6}$$

where

$$Q[\xi(\vec{\phi})] = \exp\left\{N\frac{\mathrm{i}\mu'}{2}\int \mathrm{d}\vec{\phi}\,\xi(\vec{\phi})\vec{\phi}^2\right\} \times\left[\frac{1}{(2\pi\mathrm{i})^{n/2}}\int \mathrm{d}\vec{\psi}\exp\left\{\frac{\mathrm{i}}{2}\vec{\psi}^2 + p\int \mathrm{d}\vec{\phi}\,\xi(\vec{\phi})f(\vec{\psi}\cdot\vec{\phi})\right\}\right]^M. \tag{3.7}$$

Moreover, we can utilize the Fourier transform of the delta function

$$\prod_{\vec{\phi}}\delta(\xi(\vec{\phi})-\tilde{\xi}(\vec{\phi})) = \int \mathcal{D}c(\vec{\phi})\exp\left\{2\pi\mathrm{i}\int \mathrm{d}\vec{\phi}\,c(\vec{\phi})(\xi(\vec{\phi})-\tilde{\xi}(\vec{\phi}))\right\} \tag{3.8}$$

in order to rewrite (3.6) as

$$\langle Z^n\rangle \sim \int \mathcal{D}\xi(\vec{\phi})\int \mathcal{D}c(\vec{\phi})\exp\left\{2\pi\mathrm{i}\int \mathrm{d}\vec{\phi}\,c(\vec{\phi})\xi(\vec{\phi}) - NF[c(\vec{\phi})]\right\}Q[\xi(\vec{\phi})], \tag{3.9}$$

where

$$\begin{aligned} F[c(\vec{\phi})] &= -\frac{1}{N}\ln\int\prod_{j=1}^{N}\mathrm{d}\vec{\phi}_j\exp\left\{-2\pi\mathrm{i}\int \mathrm{d}\vec{\phi}\,c(\vec{\phi})\tilde{\xi}(\vec{\phi})\right\} \\ &= -\frac{1}{N}\ln\int\prod_{j=1}^{N}\mathrm{d}\vec{\phi}_j\exp\left\{-\frac{2\pi\mathrm{i}}{N}\int \mathrm{d}\vec{\phi}\,c(\vec{\phi})\sum_{j=1}^{N}\delta(\vec{\phi}-\vec{\phi}_j)\right\} \\ &= -\ln\int \mathrm{d}\vec{\varphi}\exp\left\{-\frac{2\pi\mathrm{i}}{N}\int \mathrm{d}\vec{\phi}\,c(\vec{\phi})\delta(\vec{\phi}-\vec{\varphi})\right\}. \end{aligned} \tag{3.10}$$

The major contribution to the functional integral over $c(\vec{\phi})$ in the limit $N \to \infty$ comes from the stationary point $c^*(\vec{\phi})$ satisfying

$$\frac{\delta}{\delta c}\left(2\pi \mathrm{i} \int \mathrm{d}\vec{\phi}\, c(\vec{\phi})\xi(\vec{\phi}) - NF[c(\vec{\phi})]\right)\Bigg|_{c=c^*} = 2\pi \mathrm{i}(\xi(\vec{\phi}) - \mathrm{e}^{F-(2\pi \mathrm{i}/N)c^*(\vec{\phi})}) = 0. \tag{3.11}$$

Hence we arrive at

$$\begin{aligned}\langle Z^n\rangle &\sim \int \mathcal{D}\xi(\vec{\phi}) \exp\left\{2\pi \mathrm{i} \int \mathrm{d}\vec{\phi}\, c^*(\vec{\phi})\xi(\vec{\phi}) - NF[c^*(\vec{\phi})]\right\} Q[\xi(\vec{\phi})] \\ &= \int \mathcal{D}\xi(\vec{\phi}) \exp\left\{-N \int \mathrm{d}\vec{\phi}\, \xi(\vec{\phi}) \ln \xi(\vec{\phi})\right\} Q[\xi(\vec{\phi})].\end{aligned} \tag{3.12}$$

Let us rewrite (3.12) as

$$\langle Z^n\rangle \sim \frac{1}{(2\pi \mathrm{i})^{Mn/2}} \int \mathcal{D}\xi(\vec{\phi})\, \mathrm{e}^{NS[\xi]}, \tag{3.13}$$

where

$$\begin{aligned}S[\xi] = &-\int \mathrm{d}\vec{\phi}\, \xi(\vec{\phi}) \ln \xi(\vec{\phi}) + \frac{\mathrm{i}\mu'}{2} \int \mathrm{d}\vec{\phi}\, \xi(\vec{\phi})\vec{\phi}^2 \\ &+ \alpha \ln\left[\int \mathrm{d}\vec{\psi} \exp\left\{\frac{\mathrm{i}}{2}\vec{\psi}^2 + p \int \mathrm{d}\vec{\phi}\, \xi(\vec{\phi}) f(\vec{\psi}\cdot\vec{\phi})\right\}\right]\end{aligned} \tag{3.14}$$

with $\alpha = M/N$.

In the limit $N \to \infty$, the stationary point determined by the equation

$$\frac{\delta S[\xi]}{\delta \xi(\vec{\phi})} = 0 \tag{3.15}$$

gives the major contribution. However, it is hard to solve it directly. Instead, we introduce a Gaussian ansatz

$$\xi^{\mathrm{EMA}}(\vec{\phi}) = \frac{1}{(2\pi \mathrm{i}\sigma)^{n/2}} \exp\left\{-\frac{\vec{\phi}^2}{2\mathrm{i}\sigma}\right\}, \qquad \mathrm{Im}\,\sigma(\mu') \leqslant 0, \tag{3.16}$$

with a parameter σ, and look for a solution of the modified problem

$$\frac{\partial S[\xi^{\mathrm{EMA}}]}{\partial \sigma} = 0. \tag{3.17}$$

This scheme is called the effective medium approximation (EMA) [19]. The condition $\mathrm{Im}\,\sigma(\mu') \leqslant 0$ is necessary to ensure the convergence of the integral over $\vec{\phi}$.

Putting the Gaussian ansatz into (3.17) and taking the limit $n \to 0$, we find

$$1 - \sigma\mu - \alpha + \alpha \sum_{k=0}^{\infty} \frac{\mathrm{e}^{-p} p^k}{k!} \int \prod_{j=1}^{k} (\Pi(x_j)\, \mathrm{d}x_j) \frac{1}{1 - \sigma \sum_{j=1}^{k} x_j^2} = 0. \tag{3.18}$$

Let us suppose that σ satisfies this equation. Then the spectral density is evaluated as

$$\begin{aligned}\rho(\mu) &= \lim_{n\to 0} \frac{2}{\pi n} \mathrm{Im} \frac{\partial}{\partial \mu} \frac{1}{N} \ln\langle (Z(\mu))^n\rangle \\ &\sim \lim_{n\to 0} \frac{2}{\pi n} \mathrm{Im} \frac{\partial}{\partial \mu} S[\xi^{\mathrm{EMA}}] \\ &= \lim_{n\to 0} \frac{1}{\pi n} \mathrm{Re} \int \mathrm{d}\vec{\phi}\, \xi(\vec{\phi})\vec{\phi}^2 \\ &= -\frac{1}{\pi} \mathrm{Im}\,\sigma.\end{aligned} \tag{3.19}$$

The spectral density can be calculated by numerically solving the stationary point equation (3.18) and by putting the solution into (3.19).

We remark that there is another approximation scheme similar to EMA, which we call the dual effective medium approximation (DEMA). Rewriting (2.18) as

$$\langle Z^n\rangle \sim \int \prod_{k=1}^{M} \mathrm{d}\vec{\psi}_k \exp\left\{\frac{\mathrm{i}}{2}\sum_{k=1}^{M}\vec{\psi}_k^2\right\}\left[\frac{1}{(2\pi\mathrm{i})^{\alpha n/2}}\int \mathrm{d}\vec{\phi}\exp\left\{\frac{\mathrm{i}\mu'}{2}\vec{\phi}^2+\frac{p}{N}\sum_{k=1}^{M} f(\vec{\phi}\cdot\vec{\psi}_k)\right\}\right]^N \tag{3.20}$$

and introducing

$$\tilde{\eta}(\vec{\psi}) = \frac{1}{M}\sum_{k=1}^{M}\delta(\vec{\psi}-\vec{\psi}_k), \tag{3.21}$$

we obtain

$$\begin{aligned}\langle Z^n\rangle \sim \int \prod_{k=1}^{M} \mathrm{d}\vec{\psi}_k \exp\left\{M\frac{\mathrm{i}}{2}\int \mathrm{d}\vec{\psi}\,\tilde{\eta}(\vec{\psi})\vec{\psi}^2\right\}\\ \times\left[\frac{1}{(2\pi\mathrm{i})^{\alpha n/2}}\int \mathrm{d}\vec{\phi}\exp\left\{\frac{\mathrm{i}\mu'}{2}\vec{\phi}^2+\alpha p\int \mathrm{d}\vec{\psi}\,\tilde{\eta}(\vec{\psi})f(\vec{\psi}\cdot\vec{\phi})\right\}\right]^N.\end{aligned} \tag{3.22}$$

As before, introducing an order parameter function $\eta(\vec{\psi})$ satisfying the normalization

$$\int \eta(\vec{\psi})\,\mathrm{d}\vec{\psi} = 1, \tag{3.23}$$

we can derive

$$\langle Z^n\rangle \sim \frac{1}{(2\pi\mathrm{i})^{Mn/2}}\int \mathcal{D}\eta(\vec{\psi})\,\mathrm{e}^{NS[\eta]}. \tag{3.24}$$

Here

$$\begin{aligned}S[\eta] = -\alpha\int \mathrm{d}\vec{\psi}\,\eta(\vec{\psi})\ln\eta(\vec{\psi}) + \frac{\mathrm{i}\alpha}{2}\int \mathrm{d}\vec{\psi}\,\eta(\vec{\psi})\vec{\psi}^2\\ +\ln\left[\int \mathrm{d}\vec{\phi}\exp\left\{\frac{\mathrm{i}\mu'}{2}\vec{\phi}^2+\alpha p\int \mathrm{d}\vec{\psi}\,\eta(\vec{\psi})f(\vec{\psi}\cdot\vec{\phi})\right\}\right].\end{aligned} \tag{3.25}$$

In order to approximately evaluate the asymptotic behaviour in the limit $N\to\infty$, we again introduce a Gaussian ansatz

$$\eta^{\mathrm{DEMA}}(\vec{\psi}) = \frac{1}{(2\pi\mathrm{i}\tau)^{n/2}}\exp\left\{-\frac{\vec{\psi}^2}{2\mathrm{i}\tau}\right\},\qquad \mathrm{Im}\,\tau(\mu')\leqslant 0, \tag{3.26}$$

where τ is a parameter, and solve the stationary point equation

$$\frac{\partial S[\eta^{\mathrm{DEMA}}]}{\partial\tau} = 0. \tag{3.27}$$

We shall call this scheme the dual effective medium approximation (DEMA).

Putting the Gaussian ansatz into (3.27) and taking the limit $n\to 0$, we find

$$\alpha-\alpha\tau-1+\mu\sum_{k=0}^{\infty}\frac{\mathrm{e}^{-\alpha p}(\alpha p)^k}{k!}\int\prod_{j=1}^{k}(\Pi(x_j)\,\mathrm{d}x_j)\frac{1}{\mu-\tau\sum_{j=1}^{k}x_j^2} = 0. \tag{3.28}$$

If τ satisfies this equation, we can write the spectral density as

$$\begin{aligned}\rho(\mu) &= \lim_{n\to 0}\frac{2}{\pi n}\,\mathrm{Im}\frac{\partial}{\partial\mu}\frac{1}{N}\ln\langle(Z(\mu))^n\rangle \sim \lim_{n\to 0}\frac{2}{\pi n}\,\mathrm{Im}\frac{\partial}{\partial\mu}S[\eta^{\mathrm{DEMA}}] \\ &= -\frac{1}{\pi}\,\mathrm{Im}\left[\sum_{k=0}^{\infty}\frac{\mathrm{e}^{-\alpha p}(\alpha p)^k}{k!}\int\prod_{j=1}^{k}(\Pi(x_j)\,\mathrm{d}x_j)\frac{1}{\mu-\tau\sum_{j=1}^{k}x_j^2}\right] \\ &= -\frac{1}{\pi}\,\mathrm{Im}\frac{\alpha\tau-\alpha+1}{\mu} = -\frac{\alpha}{\mu\pi}\,\mathrm{Im}\,\tau. \end{aligned} \tag{3.29}$$

Let us suppose that we calculate the spectral density of $\tilde{J} = AA^{\mathrm{T}}$ as well. Although the exact density of the nonzero eigenvalues of $\tilde{J}$ must be identical to that of J, an approximate result can be different. In fact, there is a duality relation between EMA and DEMA: the EMA result for J is the same as the DEMA result for $\tilde{J}$, and the DEMA result for J is the same as the EMA result for $\tilde{J}$.

4. Symmetric effective medium approximation

The approximation schemes so far discussed need numerical treatment of the stationary point equations. In this section, we introduce a symmetrized scheme of approximation, which we call the Symmetric Effective Medium Approximation (SEMA). In SEMA analytic solutions can be deduced, if the distribution of the nonzero elements of A is binary.

In terms of the functions

$$\tilde{\xi}(\vec{\phi}) = \frac{1}{N}\sum_{j=1}^{N}\delta(\vec{\phi}-\vec{\phi}_j), \qquad \tilde{\eta}(\vec{\psi}) = \frac{1}{M}\sum_{k=1}^{M}\delta(\vec{\psi}-\vec{\psi}_k), \tag{4.1}$$

(2.18) can be expressed in the form

$$\begin{aligned}\langle Z^n\rangle \sim \frac{1}{(2\pi\mathrm{i})^{Mn/2}}\int\prod_{j=1}^{N}\mathrm{d}\vec{\phi}_j\int\prod_{k=1}^{M}\mathrm{d}\vec{\psi}_k\exp\left\{N\frac{\mathrm{i}\mu'}{2}\int\mathrm{d}\vec{\phi}\,\tilde{\xi}(\vec{\phi})\vec{\phi}^2 + M\frac{\mathrm{i}}{2}\int\mathrm{d}\vec{\psi}\,\tilde{\eta}(\vec{\psi})\vec{\psi}^2\right\} \\ \times\exp\left\{Mp\int\mathrm{d}\vec{\psi}\int\mathrm{d}\vec{\phi}\,\tilde{\eta}(\vec{\psi})\tilde{\xi}(\vec{\phi})f(\vec{\psi}\cdot\vec{\phi})\right\}.\end{aligned} \tag{4.2}$$

As before, it can be further rewritten as

$$\langle Z^n\rangle \sim \frac{1}{(2\pi\mathrm{i})^{Mn/2}}\int\mathcal{D}\xi(\vec{\phi})\mathcal{D}\eta(\vec{\psi})\,\mathrm{e}^{NS[\xi,\eta]}, \tag{4.3}$$

where

$$\begin{aligned}S[\xi,\eta] = &-\int\mathrm{d}\vec{\phi}\,\xi(\vec{\phi})\ln\xi(\vec{\phi}) - \alpha\int\mathrm{d}\vec{\psi}\,\eta(\vec{\psi})\ln\eta(\vec{\psi}) \\ &+\frac{\mathrm{i}\mu'}{2}\int\mathrm{d}\vec{\phi}\,\xi(\vec{\phi})\vec{\phi}^2 + \alpha\frac{\mathrm{i}}{2}\int\mathrm{d}\vec{\psi}\,\eta(\vec{\psi})\vec{\psi}^2 \\ &+\alpha p\int\mathrm{d}\vec{\psi}\int\mathrm{d}\vec{\phi}\,\eta(\vec{\psi})\xi(\vec{\phi})f(\vec{\psi}\cdot\vec{\phi}).\end{aligned} \tag{4.4}$$

It should be noted that one obtains $S[\xi]$ from $S[\xi,\eta]$ if one eliminates $\eta(\vec{\psi})$ from $S[\xi,\eta]$ using the stationary condition with respect to $\eta(\vec{\psi})$. Alternatively, eliminating $\xi(\vec{\phi})$ from $S[\xi,\eta]$ yields $S[\eta]$.

In SEMA we suppose that both of the approximate solutions for ξ and η have Gaussian forms

$$\xi^{\text{SEMA}}(\vec{\phi}) = \frac{1}{(2\pi \mathrm{i}\sigma)^{n/2}} \exp\left\{-\frac{\vec{\phi}^2}{2\mathrm{i}\sigma}\right\}, \qquad \eta^{\text{SEMA}}(\vec{\psi}) = \frac{1}{(2\pi \mathrm{i}\tau)^{n/2}} \exp\left\{-\frac{\vec{\psi}^2}{2\mathrm{i}\tau}\right\} \tag{4.5}$$

$(\mathrm{Im}\,\sigma(\mu') \leqslant 0, \mathrm{Im}\,\tau(\mu') \leqslant 0)$. Solving the modified stationary point equations

$$\frac{\partial S[\xi^{\text{SEMA}}, \eta^{\text{SEMA}}]}{\partial \sigma} = \frac{\partial S[\xi^{\text{SEMA}}, \eta^{\text{SEMA}}]}{\partial \tau} = 0, \tag{4.6}$$

we find the relations

$$\begin{aligned} &1 - \sigma\mu + \alpha p \sigma\tau \int \mathrm{d}x\, \Pi(x) \frac{x^2}{1 - \sigma\tau x^2} = 0, \\ &\alpha - \alpha\tau + \alpha p \sigma\tau \int \mathrm{d}x\, \Pi(x) \frac{x^2}{1 - \sigma\tau x^2} = 0, \end{aligned} \tag{4.7}$$

from which it follows that

$$\tau = \frac{\mu\sigma + \alpha - 1}{\alpha}. \tag{4.8}$$

Using the solution σ of (4.7), as before we obtain the spectral density as

$$\rho(\mu) \sim \lim_{n\to 0} \frac{1}{\pi n} \mathrm{Re} \int \mathrm{d}\vec{\phi}\, \xi(\vec{\phi})\vec{\phi}^2 = -\frac{1}{\pi} \mathrm{Im}\,\sigma. \tag{4.9}$$

5. Binary distribution

In this section, we calculate approximate spectral densities of the matrix $J = A^{\mathrm{T}}A$, employing EMA and SEMA as the approximation schemes. For the nonzero elements of the $M \times N$ real matrix A, we assume in this section the binary distribution

$$\Pi(x) = \tfrac{1}{2}(\delta(x - 1) + \delta(x + 1)). \tag{5.1}$$

In this typical case, the Fourier transform of $P(x)$ is given by

$$\hat{P}(k) = 1 + \frac{p}{N}(\cos k - 1). \tag{5.2}$$

Namely, we have

$$f(k) = \cos k - 1. \tag{5.3}$$

For the binary distribution (5.1), the stationary point equation (3.18) for EMA takes the form

$$1 - \sigma\mu - \alpha + \alpha \sum_{k=0}^{\infty} \frac{\mathrm{e}^{-p} p^k}{k!} \frac{1}{1 - \sigma k} = 0. \tag{5.4}$$

As we mentioned, we are able to numerically solve this equation.

A comparison of the EMA solution and the average spectral density of numerically generated random matrices (NUMERICAL) is shown in figure 1. One can see that the agreement is reasonably good except in the tail region.

On the other hand, for the binary distribution (5.1), the SEMA stationary point equations (4.7) are written as

$$1 - \sigma\mu + \alpha p \frac{\sigma\tau}{1 - \sigma\tau} = 0, \qquad \alpha - \alpha\tau + \alpha p \frac{\sigma\tau}{1 - \sigma\tau} = 0. \tag{5.5}$$

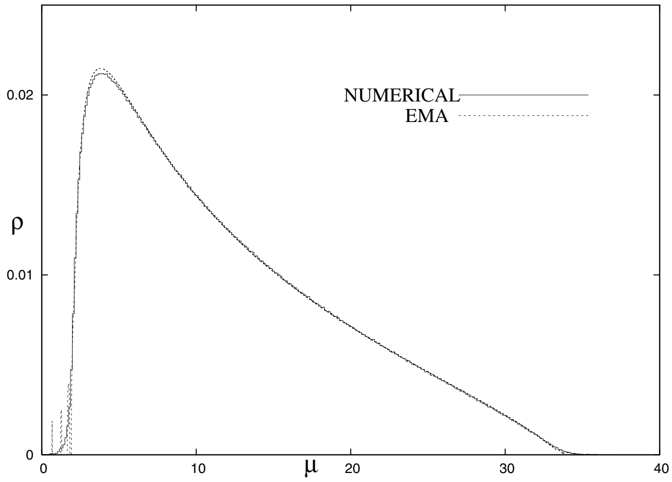

Figure 1. Comparison of the EMA solution ($p = 12, \alpha = 0.3$) for the spectral density and a numerically generated result (average over 1000 samples, $N = 8000$).

Then it is straightforward to derive a cubic equation for σ

$$\sigma^3 + \frac{\alpha(p+1)-2}{\mu}\sigma^2 - \frac{\mu\alpha + (1-\alpha)(\alpha p - 1)}{\mu^2}\sigma + \frac{\alpha}{\mu^2} = 0. \tag{5.6}$$

This cubic equation can be analytically solved. Using the notations

$$a_2 = \frac{\alpha(p+1)-2}{\mu}, \qquad a_1 = -\frac{\mu\alpha + (1-\alpha)(\alpha p - 1)}{\mu^2}, \qquad a_0 = \frac{\alpha}{\mu^2} \tag{5.7}$$

for the coefficients, we define

$$q = \frac{1}{3}a_1 - \frac{1}{9}a_2^2, \qquad r = \frac{1}{6}(a_1 a_2 - 3a_0) - \frac{1}{27}a_2^3. \tag{5.8}$$

Depending on the value of $q^3 + r^2$, the solutions are classified as follows:

(1) If $q^3 + r^2 > 0$, there are one real and two complex solutions;
(2) If $q^3 + r^2 = 0$, there are three real solutions (at least two of them are the same);
(3) If $q^3 + r^2 < 0$, there are three real solutions.

In order to have a nonzero eigenvalue density, we need to have a complex solution with a negative imaginary part. Therefore we focus on the case $q^3 + r^2 > 0$. For a complex number $z = |z|\,\mathrm{e}^{\mathrm{i}\theta}$ $(-\pi < \theta \leqslant \pi)$ and a real number χ, let us define the exponential function as

$$z^\chi = |z|^\chi\,\mathrm{e}^{\mathrm{i}\chi\theta}. \tag{5.9}$$

Then, using real numbers

$$s_1 = -(-r - (q^3 + r^2)^{1/2})^{1/3}, \qquad s_2 = -(-r + (q^3 + r^2)^{1/2})^{1/3}, \tag{5.10}$$

one obtains a real solution

$$z_0 = s_1 + s_2 - \frac{a_2}{3} \tag{5.11}$$

and complex solutions

$$\begin{aligned} z_1 &= -\frac{1}{2}(s_1 + s_2) - \frac{a_2}{3} + \mathrm{i}\frac{\sqrt{3}}{2}(s_1 - s_2), \\ z_2 &= -\frac{1}{2}(s_1 + s_2) - \frac{a_2}{3} - \mathrm{i}\frac{\sqrt{3}}{2}(s_1 - s_2). \end{aligned} \tag{5.12}$$

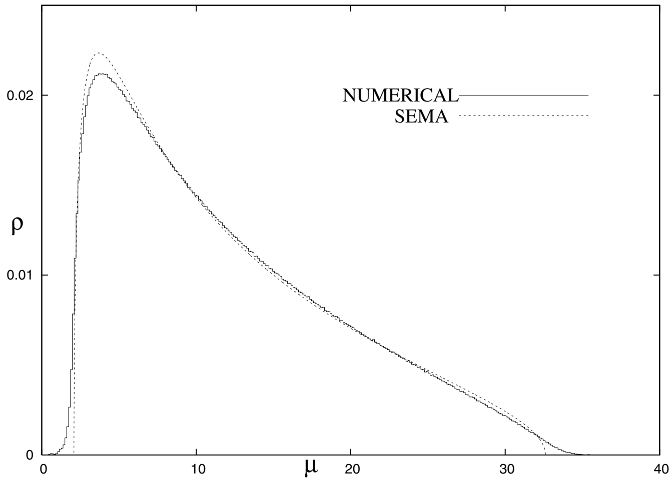

Figure 2. Comparison of the SEMA solution ($p = 12, \alpha = 0.3$) for the spectral density and a numerically generated result (average over 1000 samples, $N = 8000$).

The nonzero eigenvalue density is thus given by

$$\rho(\mu) = -\frac{1}{\pi}\,\mathrm{Im}\, z_2 = \frac{\sqrt{3}}{2\pi}(s_1 - s_2). \tag{5.13}$$

This equation gives a continuous band between the edges μ_- and μ_+ ($\mu_- < \mu_+$), where $q^3 + r^2 = 0$ holds.

We compare the SEMA solution (5.13) and a numerically generated result (NUMERICAL) in figure 2. In spite of the simplification of the scheme, the agreement is still reasonably good. However, as the SEMA spectral density vanishes out of the edges μ_- and μ_+, the discrepancy in the tail region stands out.

In figure 3, we compare the EMA and SEMA solutions and a numerically generated result (NUMERICAL) in the neighbourhood of the outer edge μ_+. One can see that the EMA solution has minibands out of the outer edge, which approximate the tail spectrum of the numerical result. On the other hand, the neighbourhood of the inner edge μ_- is depicted in figure 4. The EMA solution also has minibands within the gap between the origin and the inner edge.

6. Single defect approximation

In SEMA, the spectral density vanishes out of the edges μ_- and μ_+. Therefore, in order to analyse the tail region, we need to improve the approximation. Let us begin with the definition (4.4) of $S[\xi, \eta]$. The exact stationary point equation can be derived from the variational equation

$$\delta\left\{S[\xi,\eta] + a\left(\int \mathrm{d}\vec{\phi}\,\xi(\vec{\phi}) - 1\right) + b\left(\int \mathrm{d}\vec{\psi}\,\eta(\vec{\psi}) - 1\right)\right\} = 0 \tag{6.1}$$

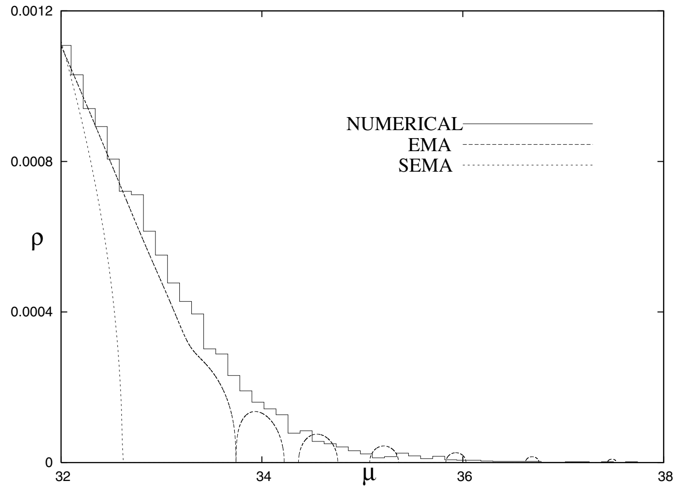

Figure 3. Comparison of the EMA and SEMA solutions ($p = 12, \alpha = 0.3$) and a numerically generated result (average over 1000 samples, $N = 8000$) in the neighbourhood of the outer edge $\mu_+ = 32.611$.

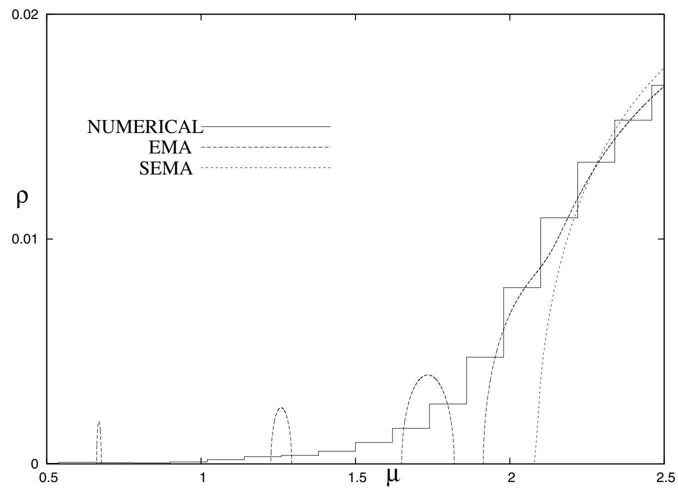

Figure 4. Comparison of the EMA and SEMA solutions ($p = 12, \alpha = 0.3$) and a numerically generated result (average over 1000 samples, $N = 8000$) in the neighbourhood of the inner edge $\mu_- = 2.087$.

as

$$\begin{aligned} \frac{\delta S[\xi,\eta]}{\delta\xi(\vec{\phi})} + a &= -\ln\xi(\vec{\phi}) - 1 + \frac{\mathrm{i}\mu'}{2}\vec{\phi}^2 + \alpha p\int \mathrm{d}\vec{\psi}\,\eta(\vec{\psi})f(\vec{\psi}\cdot\vec{\phi}) + a = 0, \\ \frac{\delta S[\xi,\eta]}{\delta\eta(\vec{\psi})} + b &= -\alpha\ln\eta(\vec{\psi}) - \alpha + \frac{\mathrm{i}\alpha}{2}\vec{\psi}^2 + \alpha p\int \mathrm{d}\vec{\phi}\,\xi(\vec{\phi})f(\vec{\psi}\cdot\vec{\phi}) + b = 0. \end{aligned} \tag{6.2}$$

Here a and b are the Lagrange multipliers which ensure the normalizations of $\xi(\vec{\phi})$ and $\eta(\vec{\psi})$. Then we find

$$\xi(\vec{\phi}) = \mathcal{A}\exp\left[\frac{\mathrm{i}\mu'}{2}\vec{\phi}^2 + \alpha p\int \mathrm{d}\vec{\psi}\,\eta(\vec{\psi})f(\vec{\psi}\cdot\vec{\phi})\right], \tag{6.3}$$

$$\eta(\vec{\psi}) = \mathcal{B}\exp\left[\frac{\mathrm{i}}{2}\vec{\psi}^2 + p\int \mathrm{d}\vec{\phi}\,\xi(\vec{\phi})f(\vec{\psi}\cdot\vec{\phi})\right], \tag{6.4}$$

where $\mathcal{A}$ and $\mathcal{B}$ should be determined so that $\xi(\vec{\phi})$ and $\eta(\vec{\psi})$ are correctly normalized.

We have so far made no approximation. Now, as the first improvement of SEMA, we put the Gaussian ansatz

$$\eta(\vec{\psi}) = \frac{1}{(2\pi\mathrm{i}\tau)^{n/2}}\exp\left\{-\frac{\vec{\psi}^2}{2\mathrm{i}\tau}\right\}, \qquad \mathrm{Im}\,\tau(\mu') \leqslant 0 \tag{6.5}$$

into the RHS of (6.3) and calculate the LHS. Then, assuming that τ is the solution of SEMA, we obtain the improved approximation for $\xi(\vec{\phi})$ as the LHS. This approximation scheme is called the single defect approximation (SDA) [19, 20].

For the binary distribution (5.1), it follows from $f(k) = \cos k - 1$ that

$$\begin{aligned}\xi(\vec{\phi}) &= \mathcal{A}\,\mathrm{e}^{\mathrm{i}\mu'\vec{\phi}^2/2}\sum_{k=0}^{\infty}\frac{\mathrm{e}^{-\alpha p}(\alpha p)^k}{k!}\left[\int \mathrm{d}\vec{\psi}\,\eta(\vec{\psi})\cos(\vec{\psi}\cdot\vec{\phi})\right]^k \\ &= \mathcal{A}\sum_{k=0}^{\infty}\frac{\mathrm{e}^{-\alpha p}(\alpha p)^k}{k!}\exp\left[\frac{\mathrm{i}}{2}(\mu' - k\tau)\vec{\phi}^2\right].\end{aligned} \tag{6.6}$$

Now we are able to determine the normalization constant $\mathcal{A}$. Integrating the both sides of the above equation over $\vec{\phi}$, we find

$$\int \xi(\vec{\phi})\,\mathrm{d}\vec{\phi} = \mathcal{A}\sum_{k=0}^{\infty}\frac{\mathrm{e}^{-\alpha p}(\alpha p)^k}{k!}\left[\frac{\mathrm{i}}{2\pi}(k\tau - \mu')\right]^{-n/2}. \tag{6.7}$$

Hence $\mathcal{A}$ should be set to 1 in the limit $n \to 0$. Then the spectral density can be evaluated as

$$\begin{aligned}\rho(\mu) &= \lim_{n\to 0}\frac{1}{\pi n}\,\mathrm{Re}\int \mathrm{d}\vec{\phi}\,\xi(\vec{\phi})\vec{\phi}^2 \\ &= \lim_{n\to 0}\frac{1}{\pi n}\,\mathrm{Re}\sum_{k=0}^{\infty}\frac{\mathrm{e}^{-\alpha p}(\alpha p)^k}{k!}\int \mathrm{d}\vec{\phi}\,\vec{\phi}^2\exp\left[\frac{\mathrm{i}}{2}(\mu' - k\tau)\vec{\phi}^2\right] \\ &= -\frac{1}{\pi}\,\mathrm{Im}\sum_{k=0}^{\infty}\frac{\mathrm{e}^{-\alpha p}(\alpha p)^k}{k!}\frac{1}{\mu' - k\tau}.\end{aligned} \tag{6.8}$$

Out of the edges μ_- and μ_+, the SEMA solution $\tau(\mu)$ is real. Therefore the spectral density (6.8) has delta peaks

$$\frac{\mathrm{e}^{-\alpha p}(\alpha p)^k}{k!}\frac{1}{1 - k\tau'(\mu_k)}\delta(\mu - \mu_k), \tag{6.9}$$

where μ_k satisfies the equation

$$\mu_k - k\tau(\mu_k) = 0. \tag{6.10}$$

Let us note that the imaginary part of $\sigma(\mu + \mathrm{i}\epsilon)$ and $\tau(\mu + \mathrm{i}\epsilon)$ cannot be positive (see (4.5)). This means that

$$\mathrm{Im}\,\sigma(\mu) < 0 \tag{6.11}$$

or, if $\mathrm{Im}\,\sigma(\mu) = 0$,

$$\mathrm{Re}\frac{\mathrm{d}\sigma}{\mathrm{d}\mu} \leqslant 0, \qquad \mathrm{Re}\frac{\mathrm{d}\tau}{\mathrm{d}\mu} = \frac{1}{\alpha}\mathrm{Re}\frac{\mathrm{d}(\mu\sigma)}{\mathrm{d}\mu} \leqslant 0. \tag{6.12}$$

When μ is larger than the outer edge μ_+ of the band, the solution z_1 satisfies these conditions. The locations of the delta peaks are thus determined by solving the equation

$$\mu - k\frac{\mu z_1(\mu) + \alpha - 1}{\alpha} = 0, \qquad \mu > \mu_+. \tag{6.13}$$

For example, if $p = 12$ and $\alpha = 0.3$, the locations of the delta peaks are

$$\mu = 32.667, 32.926, 33.336, 33.856, 34.456, 35.118, 35.828, \ldots, \tag{6.14}$$

corresponding to $k = 14, 15, 16, 17, 18, 19, 20, \ldots$ ($\mu_+ = 32.611$). We would like to note that the DEMA minibands (not shown) seem to appear around these delta peaks.

From the analytic solution (5.12), we find the asymptotics

$$z_1(\mu) \sim \frac{1}{\mu}, \qquad \mu \to \infty, \tag{6.15}$$

from which it follows that

$$\tau(\mu) = \frac{\mu z_1(\mu) + \alpha - 1}{\alpha} \sim 1, \qquad \mu \to \infty. \tag{6.16}$$

Thus the delta peaks are located at $\mu_k \sim k$, so that a continuous approximation gives

$$\rho(\mu_k)(\mu_k - \mu_{k-1}) \sim \frac{\mathrm{e}^{-\alpha p}(\alpha p)^k}{k!}. \tag{6.17}$$

Therefore, in the limit of large μ, we obtain

$$\rho(\mu) \sim \frac{\mathrm{e}^{-\alpha p}}{\sqrt{2\pi\mu}} \exp\left[-\mu \ln\left(\frac{\mu}{\alpha p\, \mathrm{e}}\right)\right]. \tag{6.18}$$

A similar asymptotic formula is known for real symmetric sparse random matrices R whose elements R_{jl} ($j \leqslant l$) are independently distributed [16, 19]. As argued in [19], SDA is expected to be correct for large μ. However, the weights of the delta peaks (6.9) are too small to account for the numerically calculated cumulative density. In order to improve the agreement, we presumably need to develop higher order extensions of SDA by iterating the approximation scheme.

When μ is smaller than the inner edge μ_-, the solution z_2 satisfies the conditions (6.12). However, as $\mu z_2 + \alpha - 1 \leqslant 0$, there is no SDA delta peak in this region besides at the origin. Therefore the second-order extension of SDA should be applied. Assuming that σ is the solution z_2 of SEMA, we put the Gaussian ansatz

$$\xi(\vec{\phi}) = \frac{1}{(2\pi \mathrm{i}\sigma)^{n/2}} \exp\left\{-\frac{\vec{\phi}^2}{2\mathrm{i}\sigma}\right\} \tag{6.19}$$

into the RHS of (6.4). Then η is calculated as the LHS of (6.4) and put into the RHS of (6.3). Thus the LHS of (6.3) gives the approximate solution for $\xi(\vec{\phi})$. Within this approximation, an argument as before yields an equation,

$$\mu - \sum_{k=0}^{\infty} \frac{n_k}{1 - k z_2(\mu)} = 0, \qquad \mu < \mu_-, \tag{6.20}$$

which determines the locations of the delta peaks. Here n_k are non-negative integers. For example, if $p = 12$ and $\alpha = 0.3$, we find the delta peaks at $\mu = 1, 1.6, 2, 2.035$ besides at the origin ($\mu_- = 2.087$).

7. Summary

In this paper, we evaluated the asymptotic eigenvalue density of the matrix of the form $J = A^{\mathrm{T}}A$, where A is an $M \times N$ real sparse random matrix. We utilized the replica method and developed approximation schemes called the effective medium approximation (EMA) and the dual effective medium approximation (DEMA). Moreover a symmetrized version of EMA (SEMA) was presented. In the case of binary distribution of the nonzero elements of A, analytic solutions were derived for SEMA. We compared the results of EMA and SEMA, focusing on the behaviour in the tail region. In order to analytically improve SEMA, we further developed the single defect approximation (SDA) and evaluated the tail behaviour.

Acknowledgments

One of the authors (TN) is grateful to Dr Laurent Laloux, Dr Marc Potters and Professor Jean-Philippe Bouchaud for suggesting the subject of this article. He also thanks Professor Geoff J Rodgers for valuable discussions. The authors acknowledge support from the Grant-in-Aid for Scientific Research, MEXT, Japan (nos 16740224, 14084209 and 18079010).

References

[1] Wishart J 1928 *Biometrika* **20** 32
[2] Le Cun Y, Kanter I and Solla S A 1991 *Phys. Rev. Lett.* **66** 2396
[3] Verbaarschot J J M 1994 *Acta Phys. Pol.* B **25** 133
[4] Beenakker C W J 1997 *Rev. Mod. Phys.* **69** 731
[5] Laloux L, Cizeau P, Bouchaud J-P and Potters M 1999 *Phys. Rev. Lett.* **83** 1467
[6] Plerou V, Gopikrishnan P, Rosenow B, Nunes Amaral L A and Stanley H E 1999 *Phys. Rev. Lett.* **83** 1471
[7] Tulino A M and Verdú S 2004 *Found. Trends Commun. Inform. Theory* **1** issue 1
[8] Wigner E P 1958 *Ann. Math.* **67** 325
[9] Marčenko V and Pastur L 1967 *Math. USSR-Sb.* **1** 457
[10] Pastur L 2000 *Mathematical Physics 2000* (London: Imperial College Press) p 216
[11] Yoshida M and Tanaka T 2006 Analysis of sparsely spread CDMA via statistical mechanics *Proc. 2006 IEEE Int. Symp. Info. Theory* pp 2378–82
[12] Soshnikov A 2004 *Electron. Commun. Probab.* **9** 82
[13] Soshnikov A and Fyodorov Y V 2005 *J. Math. Phys.* **46** 033302
[14] Biroli G, Bouchaud J-P and Potters M 2006 *Preprint* cond-mat/0609070
[15] Tanaka T 2005 On the eigenvalue spectrum of random matrices (extended abstract) *Randomness and Computation Joint Workshop 'New Horizons in Computing' and 'Statistical Mechanical Approach to Probabilistic Information Processing' (Sendai, Japan, 18–21 July)*
[16] Rodgers G J and Bray A J 1988 *Phys. Rev.* B **37** 3557
[17] Rodgers G J and De Dominicis C 1990 *J. Phys. A: Math. Gen.* **23** 1567
[18] Mirlin A D and Fyodorov Y V 1991 *J. Phys. A: Math. Gen.* **24** 2273
[19] Semerjian G and Cugliandolo L F 2002 *J. Phys. A: Math. Gen.* **35** 4837
[20] Biroli G and Monasson R 1999 *J. Phys. A: Math. Gen.* **32** L255