Journal of Theoretical Probability (2019) 32:1051–1075
https://doi.org/10.1007/s10959-018-0840-2

Asymptotic Expansion of Spherical Integral

Jiaoyang Huang[1]

Received: 12 June 2017 / Revised: 23 November 2017 / Published online: 29 June 2018

Abstract
We consider the spherical integral of real symmetric or Hermitian matrices when the rank of one matrix is one. We prove the existence of the full asymptotic expansions of these spherical integrals and derive the first and the second term in the asymptotic expansion.

Keywords Random matrices · Spherical integral · Asymptotic expansion · Free probability

Mathematics Subject Classification (2010) 15B52 · 46L54

1 Introduction

In this paper, we consider the expansion of the spherical integral

$$I_N^{(\beta)}(D_N, B_N) = \int \exp\{NTr(D_N U^* B_N U)\} dm_N^{(\beta)}(U), \tag{1}$$

where $m_N^{(\beta)}$ is the Haar measure on orthogonal group $O(N)$ if $\beta = 1$, on unitary group $U(N)$ if $\beta = 2$, and D_N, B_N are deterministic $N \times N$ real symmetric or Hermitian matrices, that we can assume diagonal without loss of generality. We follow [4] to investigate the asymptotics of the spherical integrals under the case $D_N = \mathrm{diag}(\theta, 0, 0, 0, 0 \ldots 0)$:

$$I_N^{(\beta)}(D_N, B_N) = I_N^{(\beta)}(\theta, B_N) = \int \exp\{\theta N(e_1^* B_N e_1)\} dm_N^{(\beta)}(U), \tag{2}$$

where e_1 is the first column of U.

The main result of this paper can be stated as follows,

✉ Jiaoyang Huang
jiaoyang@math.harvard.edu

1 Department of Mathematics, Harvard, Cambridge, MA 02138, USA

Theorem *If* $\sup_N \|B_N\|_\infty < M$*, then for any* $\theta \in \mathbb{R}$ *such that* $|\theta| < \frac{1}{4M^2+10M+1}$*, the spherical integral has the following asymptotic expansion (up to* $O(N^{-n-1})$ *for any given* n*):*

$$e^{-N(\theta v - \frac{1}{2N}\sum_{i=1}^{N}\log(1-2\theta\lambda_i(B_N)+2\theta v))} I_N(\theta, B_N) \\ = m_0 + \frac{m_1}{N} + \frac{m_2}{N^2} + \cdots + \frac{m_n}{N^n} + O(N^{-n-1}) \tag{3}$$

where v *and the coefficients* $\{m_i\}_{i=0}^n$ *depend on* θ *and the derivatives of the Hilbert transform of the empirical spectral distribution of* B_N*.*

The spherical integral provides a finite-dimensional analogue of the R-transform in free probability [8,10], which states that if X and Y are two freely independent self-adjoint noncommutative random variables, then their R-transforms satisfy the following additive formula:

$$R_{X+Y} = R_X + R_Y.$$

In the scenario of random matrices, let $\{B_N\}$ and $\{\tilde{B}_N\}$ be sequences of uniformly bounded real symmetric (or Hermitian) matrices whose empirical spectral distributions converge in law toward τ_B and $\tau_{\tilde{B}}$ respectively. Let V_N be a sequence of independent orthogonal (or unitary) matrices following the Haar measure. Then the noncommutative variables $\{B_N\}$ and $\{V_N^* \tilde{B}_N V_N\}$ are asymptotically free. And the law of their sum $\{B_N + V_N^* \tilde{B}_N V_N\}$ converges toward $\tau_{B+V^*\tilde{B}V}$, which is characterized by the following additive formula,

$$R_{\tau_{B+V^*\tilde{B}V}} = R_{\tau_B} + R_{\tau_{\tilde{B}}}. \tag{4}$$

We refer to [1, Section 5] for a proof of this.

For the spherical integral, using the same notation as above, we have the following additive formula

$$\frac{1}{N}\log \mathbb{E}_{V_N}\left[I_N^{(\beta)}(\theta, B_N + V_N^*\tilde{B}_N V_N)\right] = \frac{1}{N}\log I_N^{(\beta)}(\theta, B_N) + \frac{1}{N}\log I_N^{(\beta)}(\theta, \tilde{B}_N).$$

So if we define the finite-dimensional R-transform for a random or deterministic matrix B_N as follows

$$R_N(B_N) = \frac{1}{N}\log \mathbb{E}_{B_N}\left[I_N^{(\beta)}(\theta, B_N)\right],$$

then the additive law can be formulated in a more concise way,

$$R_N(B_N + V_N^*\tilde{B}_N V_N) = R_N(B_N) + R_N(\tilde{B}_N). \tag{5}$$

Compare with (4), which takes advantage of the fact that $\{B_N\}$ and $\{V_N^*\tilde{B}_N V_N\}$ are asymptotically free, and which holds after we take the limit as N goes to infinity.

However, additive formula (5) holds for any N, which provides a finite-dimensional analogue of the additivity of the R-transform. Indeed, the R-transform is some sort of limit of our R_N:

$$\lim_{N\to\infty} R_N(B_N) = \lim_{N\to\infty} \frac{1}{N} \log I_N^{(\beta)}(\theta, B_N) = \frac{\beta}{2} \int_0^{\frac{2\theta}{\beta}} R_{\tau_B}(s)\mathrm{d}s.$$

Combining this and some concentration of measure estimates, i.e., almost surely, as N goes to infinity,

$$\left| \frac{1}{N} \log I_N^{(\beta)}(\theta, B_N + V_N^* \tilde{B}_N V_N) - R_N(B_N + V_N^* \tilde{B}_N V_N) \right| \to 0, \tag{6}$$

Guionnet and Maïda in [4] showed that the additivity of R-transform (4) is a direct consequence of (5). Moreover, with a more quantitative version of (6), the finite-dimensional analogue of R-transform may be used to study the rate of convergence of sums of asymptotically free random matrices.

The asymptotic expansion and some related properties of spherical integrals were thoroughly studied by Guionnet et al. [4–6]. However, in their paper, they only studied the first term in the asymptotic expansion. We here derive the higher-order asymptotic expansion terms. The proofs are different from those in [4]; Guionnet and Maïda relied on large deviation techniques and used central limit theorem to derive the first-order term. In this paper, we express the spherical integral as an integral over only two Gaussian variables, hence allowing for easier asymptotic analysis.

The spherical integral can also be used to study the Schur polynomials. Spherical integral (1) can be expressed in terms of Schur polynomials. The Harish-Chandra-Itzykson-Zuber integral formula [2,7,9] gives an explicit form for integral (1) in the case $\beta = 2$ and all the eigenvalues of D and B are simple:

$$I_N^{(2)}(D, B) = \prod_{i=1}^{N} i! \frac{\det(e^{N\lambda_i(D)\lambda_j(B)})_{1\le i,j\le N}}{N^{\frac{N^2-N}{2}} \Delta(\lambda(D))\Delta(\lambda(B))}, \tag{7}$$

where Δ denotes the Vandermonde determinant,

$$\Delta(\lambda(D)) = \prod_{1\le i<j\le N} (\lambda_i(D) - \lambda_j(D)),$$

and $\lambda_i(\cdot)$ is the i-th eigenvalue. If we define the N-tuple $\mu = (\lambda_i(B_N) - N + i)_{i=1}^N$, then above expression (7) is the normalized Schur polynomial times an explicit factor,

$$\begin{aligned} I_N^{(2)}(D, B) &= \frac{\prod_{i=1}^N i!}{N^{\frac{N^2-N}{2}}} \frac{S_\mu(e^{N\lambda_1(D)}, e^{N\lambda_2(D)}, \ldots, e^{N\lambda_N(D)})}{\prod_{i<j}((\mu_i - i) - (\mu_j - j))} \frac{\prod_{i<j}(e^{N\lambda_i(D)} - e^{N\lambda_j(D)})}{\prod_{i<j}(\lambda_i(D) - \lambda_j(D))} \\ &= \frac{S_\mu(e^{N\lambda_1(D)}, e^{N\lambda_2(D)}, \ldots, e^{N\lambda_N(D)})}{S_\mu(1, 1, \ldots, 1)} \frac{\prod_{i<j}(e^{N\lambda_i(D)} - e^{N\lambda_j(D)})}{\prod_{i<j}(N\lambda_i(D) - N\lambda_j(D))}. \end{aligned}$$

In [3], Gorin and Panova studied the asymptotic expansion of the normalized Schur polynomial $\frac{S_\mu(x_1,x_2,\dots x_k,1,1,1\dots 1)}{S_\mu(1,1,\dots,1)}$, for fixed k. Its asymptotic expansion can be obtained from a limit formula [3, Proposition 3.9], combining with the asymptotic results for $S_\mu(x_i, 1, 1, \dots 1)$, which corresponds to the spherical integral where D is of rank one. Therefore, our methods give a new proof of [3, Proposition 4.1].

2 Some Notations

Throughout this paper, N is a parameter going to infinity. We use $O(N^{-l})$ to denote any quantity that is bounded in magnitude by CN^{-l} for some constant $C > 0$. We use $O(N^{-\infty})$ to denote a sequence of quantities that are bounded in magnitude by $C_l N^{-l}$ for any $l > 0$ and some constant $C_l > 0$ depending on l.

Given a real symmetric or Hermitian matrix B, with eigenvalues $\{\lambda_i\}_{i=1}^N$. Let $\lambda_{\min}(B)$ ($\lambda_{\max}(B)$) be the minimal (maximal) eigenvalue. We denote the Hilbert transform of its empirical spectral distribution by H_B:

$$\begin{aligned} H_B(z) : \mathbb{R} \setminus [\lambda_{\min}(B), \lambda_{\max}(B)] &\mapsto \mathbb{R} \\ z &\mapsto \frac{1}{N}\sum_{i=1}^{N} \frac{1}{z-\lambda_i}. \end{aligned}$$

On intervals $(-\infty, \lambda_{\min}(B))$ and $(\lambda_{\max}(B), +\infty)$, H_B is monotonous. The R-transform of empirical spectral distribution of B is

$$R_B(z) := H_B^{-1}(z) - \frac{1}{z},$$

on $\mathbb{R}/\{0\}$, where H_B^{-1} is the functional inverse of H_B from $\mathbb{R}/\{0\}$ to $(-\infty, \lambda_{\min}(B)) \cup (\lambda_{\max}(B), +\infty)$. In the following paper, for given nonzero real number θ, we denote $v(\theta) = R_B(2\theta)$, for simplicity we will omit the variable θ in the expression of $v(\theta)$.

Remark 1 Notice v, defined above, satisfies the following explicit formulas:

$$\frac{1}{N}\sum_{i=1}^{N} \frac{1}{1-2\theta\lambda_i+2\theta v} = 1, \qquad \frac{1}{N}\sum_{i=1}^{N} \frac{\lambda_i}{1-2\theta\lambda_i+2\theta v} = v. \tag{8}$$

We denote the normalized k-th derivative of Hilbert transform

$$A_k := \frac{(-1)^{k-1}}{(k-1)!(2\theta)^k} \frac{d^{k-1} H_B}{dz^{k-1}}\left(v + \frac{1}{2\theta}\right) = \frac{1}{N}\sum_{i=1}^{N} \frac{1}{(1-2\theta\lambda_i+2\theta v)^k}. \tag{9}$$

Notice from the definition of v, $A_1 = 1$. The coefficients in asymptotic expansion (3) can be represented in terms of these A_k's.

Also we denote

$$F := \frac{1}{N}\sum_{i=1}^{N}\frac{\lambda_i}{(1-2\theta\lambda_i+2\theta v)^2} = -\frac{1}{2\theta}+\frac{1+2\theta v}{2\theta}A_2,$$

$$G := \frac{1}{N}\sum_{i=1}^{N}\frac{\lambda_i^2}{(1-2\theta\lambda_i+2\theta v)^2} = -\frac{1+4\theta v}{4\theta^2}+\frac{(1+2\theta v)^2}{4\theta^2}A_2.$$

3 Orthogonal Case

3.1 First-Order Expansion

In this section, we consider the real case, $\beta = 1$. For simplicity, in this section we will omit the superscript (β) in all the notations. We can assume B_N is diagonal $B_N = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_N)$. Notice e_1 is the first column of U, which follows Haar measure on orthogonal group $O(N)$. e_1 is uniformly distributed on S^{N-1}, the surface of the unit ball in $\mathbb{R}^N$, and can be represented as the normalized Gaussian vector,

$$e_1 = \frac{g}{\|g\|},$$

where $g = (g_1, g_2, \ldots, g_N)^T$ is the standard Gaussian vector in $\mathbb{R}^N$, and $\|\cdot\|$ is the Euclidean norm in $\mathbb{R}^N$. Plug them into (2)

$$I_N(\theta, B_N) = \int \exp\left\{\theta N \frac{\lambda_1 g_1^2 + \lambda_2 g_2^2 \cdots + \lambda_N g_N^2}{g_1^2 + g_2^2 + \cdots g_N^2}\right\}\prod_{i=1}^{N}\mathrm{d}P(g_i). \tag{10}$$

where $P(\cdot)$ is the standard Gaussian probability measure on $\mathbb{R}$. Following the paper [4], define

$$\gamma_N := \frac{1}{N}\sum_{i=1}^{N} g_i^2 - 1, \qquad \hat{\gamma}_N := \frac{1}{N}\sum_{i=1}^{N}\lambda_i g_i^2 - v.$$

Then (10) can be rewritten in the following form

$$I_N(\theta, B_N) = \int \exp\left\{\theta N \frac{\hat{\gamma}_N + v}{\gamma_N + 1}\right\}\prod_{i=1}^{N}\mathrm{d}P(g_i).$$

Next we will do the following change of measure

$$P_N(\mathrm{d}g_1, \mathrm{d}g_2, \ldots, \mathrm{d}g_N) = \frac{1}{\sqrt{2\pi}^N}\prod_{i=1}^{N}\left(\sqrt{1+2\theta v-2\theta\lambda_i}\, e^{-\frac{1}{2}(1+2\theta v-2\theta\lambda_i)g_i^2}\mathrm{d}g_i\right).$$

With this measure, we have $\mathbb{E}_{P_N}(\gamma_N) = 0$ and $\mathbb{E}_{P_N}(\hat{\gamma}_N) = 0$. Therefore, intuitively from the law of large number, as N goes to infinity, γ_N and $\hat{\gamma}_N$ concentrate at origin. This is exactly what we will do in the following.

Let us also define

$$I_N^{\kappa_1,\kappa_2}(\theta, B_N) = \int_{\substack{|\gamma_N| \le N^{-\kappa_1}, \\ |\hat{\gamma}_N| \le N^{-\kappa_2}}} \exp\left\{\theta N \frac{\hat{\gamma}_N + v}{\gamma_N + 1}\right\} \prod_{i=1}^N \mathrm{d}P(g_i),$$

where constants κ_1 and κ_2 satisfy $\frac{1}{2} > \kappa_1 > 2\kappa_2$ and $2\kappa_1 + \kappa_2 > 1$. We prove the following proposition, then the difference between $I_N(\theta, B_N)$ and $I_N^{\kappa_1,\kappa_2}(\theta, B_N)$ is of order $O(N^{-\infty})$. A weaker form of this proposition also appears in [4]. With this proposition, for the asymptotic expansion, we only need to consider $I_N^{\kappa_1,\kappa_2}(\theta, B_N)$.

Proposition 1 *Given constants κ_1 and κ_2 satisfying $\frac{1}{2} > \kappa_1 > 2\kappa_2$ and $2\kappa_1 + \kappa_2 > 1$. There exist constants c, c', depending on κ_1,κ_2 and* $\sup \|B_N\|_\infty$, *such that for N large enough*

$$|I_N(\theta, B_N) - I_N^{\kappa_1,\kappa_2}(\theta, B_N)| \le ce^{-c'N^{1-2\kappa_1}} I_N(\theta, B_N). \tag{11}$$

Proof We can split (11) into two parts

$$I_N(\theta, B_N) - I_N^{\kappa_1,\kappa_2}(\theta, B_N) = \underbrace{\left(\int - \int_{|\gamma_N| \le N^{-\kappa_1}}\right)}_{E_1} + \underbrace{\left(\int_{|\gamma_N| \le N^{-\kappa_1}} - \int_{\substack{|\gamma_N| \le N^{-\kappa_1}, \\ |\hat{\gamma}_N| \le N^{-\kappa_2}}}\right)}_{E_2}.$$

Following the same argument as in [4, Lemma 14],

$$|E_1| \le ce^{-c'N^{1-2\kappa_1}} I_N(\theta, B_N).$$

For E_2, notice that

$$\begin{aligned}
&\exp\left\{\theta N \frac{\hat{\gamma}_N + v}{\gamma_N + 1}\right\} \prod_{i=1}^N \mathrm{d}P(g_i) \\
&= \exp\left\{\theta N v + \theta N(\hat{\gamma}_N - v\gamma_N) - \theta N \gamma_N \frac{\hat{\gamma}_N - v\gamma_N}{\gamma_N + 1}\right\} \prod_{i=1}^N \mathrm{d}P(g_i) \\
&= e^{\theta N v} \exp\left\{-\theta N \gamma_N \frac{\hat{\gamma}_N - v\gamma_N}{\gamma_N + 1}\right\} \prod_{i=1}^N \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}g_i^2(1 - 2\theta\lambda_i + 2\theta v)} \mathrm{d}g_i.
\end{aligned}$$

With the change of variable $g_i = \frac{\tilde{g}_i}{\sqrt{1-2\theta\lambda_i+2\theta v}}$,

$$\exp\left\{\theta N \frac{\hat{\gamma}_N + v}{\gamma_N + 1}\right\} \prod_{i=1}^{N} \mathrm{d}P(g_i)$$
$$= e^{\theta N v - \frac{1}{2}\sum_{i=1}^{N}\log(1-2\theta\lambda_i+2\theta v)} \exp\left\{-\theta N \gamma_N \frac{\hat{\gamma}_N - v\gamma_N}{\gamma_N + 1}\right\} \prod_{i=1}^{N} \mathrm{d}P(\tilde{g}_i),$$

where in terms of the new variables $\tilde{g}_i$'s, we have

$$\gamma_N = \frac{1}{N}\sum_{i=1}^{N} \frac{\tilde{g}_i^2}{1-2\theta\lambda_i+2\theta v} - 1, \quad \hat{\gamma}_N = \frac{1}{N}\sum_{i=1}^{N} \frac{\lambda_i \tilde{g}_i^2}{1-2\theta\lambda_i+2\theta v} - v. \tag{12}$$

We can bound $|E_2|$ as,

$$\left|\int_{\substack{|\gamma_N|\le N^{-\kappa_1},\\ |\hat{\gamma}_N|>N^{-\kappa_2}}} \exp\left\{\theta N \frac{\hat{\gamma}_N + v}{\gamma_N + 1}\right\} \prod_{i=1}^{N} \mathrm{d}P(g_i)\right|$$
$$= e^{\theta N v - \frac{1}{2}\sum_{i=1}^{N}\log(1-2\theta\lambda_i+2\theta v)} \left|\int_{\substack{|\gamma_N|\le N^{-\kappa_1},\\ |\hat{\gamma}_N|>N^{-\kappa_2}}} \exp\left\{\theta N \gamma_N \frac{\hat{\gamma}_N - v\gamma_N}{\gamma_N + 1}\right\} \prod_{i=1}^{N} \mathrm{d}P(\tilde{g}_i)\right|$$
$$\le e^{\theta N v - \frac{1}{2}\sum_{i=1}^{N}\log(1-2\theta\lambda_i+2\theta v)} e^{|\theta|(\|B_N\|_\infty+|v|)N^{1-\kappa_1}} P(|\hat{\gamma}_N| > N^{-\kappa_2}),$$

where

$$P(|\hat{\gamma}_N| > N^{-\kappa_2}) = P\left(\left|\frac{1}{N}\sum_{i=1}^{N} \frac{\lambda_i \tilde{g}_i^2}{1-2\theta\lambda_i+2\theta v} - v\right| > N^{-\kappa_2}\right)$$
$$= P\left(\left|\frac{1}{\sqrt{N}}\sum_{i=1}^{N} \frac{\lambda_i (\tilde{g}_i^2 - 1)}{1-2\theta\lambda_i+2\theta v}\right| > N^{\frac{1}{2}-\kappa_2}\right)$$
$$\le 2e^{-c' N^{1-2\kappa_2}}.$$

The last inequality comes from the concentration measure inequality in Lemma A1 and the uniformly boundedness of $\|B_N\|_\infty$. Moreover from [4, Lemma 14], we have the lower bound for $I_N(\theta, B_N)$,

$$I_N(\theta, B_N) \ge c e^{\theta N v - \frac{1}{2}\sum_{i=1}^{N}\log(1-2\theta\lambda_i+2\theta|v|)} e^{-|\theta|(\|B_N\|_\infty+v)N^{1-\kappa_1}}.$$

From our assumption $\kappa_1 > 2\kappa_2$, so $e^{|\theta|(\|B_N\|_\infty+v)N^{1-\kappa_1}} \ll e^{c' N^{1-2\kappa_2}}$. Therefore for N large enough

$$|E_2| \leq c e^{-c' N^{1-2\kappa_2}} I_N(\theta, B_N).$$

This finishes the proof of (11). □

Since the difference between $I_N(\theta, B_N)$ and $I_N^{\kappa_1,\kappa_2}(\theta, B_N)$ is of order $O(N^{-\infty})$, for asymptotic expansion, we only need to consider $I_N^{\kappa_1,\kappa_2}(\theta, B_N)$.

$$I_N^{\kappa_1,\kappa_2}(\theta, B_N) = e^{N\left(\theta v - \frac{1}{2N}\sum_{i=1}^N \log(1-2\theta\lambda_i(B_N)+2\theta v)\right)} \int_{\substack{|\gamma_N|\leq N^{-\kappa_1},\\ |\hat{\gamma}_N|\leq N^{-\kappa_2}}} \exp\left\{-\theta N \gamma_N \frac{\hat{\gamma}_N - v\gamma_N}{\gamma_N + 1}\right\} \prod_{i=1}^N \mathrm{d}P(\tilde{g}_i),$$

The next thing is to expand the following integral,

$$\int_{\substack{|\gamma_N|\leq N^{-\kappa_1},\\ |\hat{\gamma}_N|\leq N^{-\kappa_2}}} \exp\left\{-\theta N \gamma_N \frac{\hat{\gamma}_N - v\gamma_N}{\gamma_N + 1}\right\} \prod_{i=1}^N \mathrm{d}P(\tilde{g}_i). \tag{13}$$

We recall the definition of γ_N and $\hat{\gamma}_N$ from (12). Since the denominator $\gamma_N + 1$ in the exponent has been restricted in a narrow interval centered at 1, we can somehow "ignore" it by Taylor expansion, which results in the following error R_1,

$$\begin{aligned} R_1 := &\left| \int_{\substack{|\gamma_N|\leq N^{-\kappa_1},\\ |\hat{\gamma}_N|\leq N^{-\kappa_2}}} \exp\left\{-\theta N \gamma_N \frac{\hat{\gamma}_N - v\gamma_N}{\gamma_N + 1}\right\} \prod_{i=1}^N \mathrm{d}P(\tilde{g}_i) \right. \\ &\left. - \int_{\substack{|\gamma_N|\leq N^{-\kappa_1},\\ |\hat{\gamma}_N|\leq N^{-\kappa_2}}} e^{-\theta N \gamma_N(\hat{\gamma}_N - v\gamma_N)} \prod_{i=1}^N \mathrm{d}P(\tilde{g}_i) \right| \\ = &\left| \int_{\substack{|\gamma_N|\leq N^{-\kappa_1},\\ |\hat{\gamma}_N|\leq N^{-\kappa_2}}} e^{-\theta N \gamma_N(\hat{\gamma}_N - v\gamma_N)} \right. \\ &\left. \left(\exp\left\{\theta N \gamma_N^2 \frac{\hat{\gamma}_N - v\gamma_N}{\gamma_N + 1}\right\} - 1\right) \prod_{i=1}^N \mathrm{d}P(\tilde{g}_i) \right|. \end{aligned} \tag{14}$$

Under our assumption $2\kappa_1 + \kappa_2 > 1$, $\theta N \gamma_N^2 \frac{\hat{\gamma}_N - v\gamma_N}{\gamma_N + 1} = O(N^{1-2\kappa_1-\kappa_2}) = o(1)$, which means the above error R_1 is of magnitude $o(1)$. Therefore, the error R_1 won't contribute to the first term in asymptotic expansion of integral (13). And the first term of asymtotics of the spherical integral comes from the following integral

$$\int_{\substack{|\gamma_N|\leq N^{-\kappa_1},\\ |\hat{\gamma}_N|\leq N^{-\kappa_2}}} e^{-\theta N \gamma_N(\hat{\gamma}_N - v\gamma_N)} \prod_{i=1}^N \mathrm{d}P(\tilde{g}_i). \tag{15}$$

We will revisit this for higher-order expansion in next section. In the remaining part of this section, we prove the following theorem about the first term of asymptotic expansion of spherical integral. This strengthens [4, Theorem 3], where they require additional conditions on the convergence speed of spectral measure).

Theorem 1 *If $\sup_N \|B_N\|_\infty < M$, then for any $\theta \in \mathbb{R}$ such that $|\theta| < \frac{1}{4M^2+10M+1}$, the spherical integral has the following asymptotic expansion,*

$$e^{-N\left(\theta v-\frac{1}{2N}\sum_{i=1}^{N}\log(1-2\theta\lambda_i(B_N)+2\theta v)\right)}I_N(\theta,B_N)=\frac{1}{\sqrt{A_2}}+o(1), \tag{16}$$

where $A_2=\frac{1}{N}\sum_{i=1}^{N}\frac{1}{(1+2\theta v-2\theta\lambda_i)^2}$.

Notice (16) is an integral of N Gaussian variables and the exponent $-\theta N\gamma_N(\hat{\gamma}_N - v\gamma_N)$ is a quartic polynomial in terms of g_i's; So, one cannot compute the above integral directly. Using the following lemma, by introducing two more Gaussian variables x_1 and x_2, we can reduce the degree of the exponent to two, and it turns to be ordinary Gaussian integral. Then we can directly compute it, and obtain the higher order asymptotic expansion.

Write the exponent in integral (15) as sum of two squares, then we can implement Lemma A2, to reduce its degree to two. Let $b^2=\theta/2$, $I_1=[-N^{\frac{1}{2}-\kappa_1+\epsilon}, N^{\frac{1}{2}-\kappa_1+\epsilon}]$ and $I_2=[-N^{\frac{1}{2}-\kappa_2+\epsilon}, N^{\frac{1}{2}-\kappa_2+\epsilon}]$, for some $\epsilon>0$. Then on the region $\{|\gamma_N|\le N^{-\kappa_1}, |\hat{\gamma}_N|\le N^{-\kappa_2}\}$

$$\begin{aligned}
&e^{-\theta N\gamma_N(\hat{\gamma}_N-v\gamma_N)}\\
&\quad=\exp\left\{-\frac{\theta}{2}\frac{(\sqrt{N}((1-v)\gamma_N+\hat{\gamma}_N))^2-(\sqrt{N}((1+v)\gamma_N-\hat{\gamma}_N))^2}{2}\right\}\\
&\quad=\frac{1}{(\sqrt{2\pi})^2}\int_{I_2}\int_{I_1}e^{-\mathrm{i}bx_1\sqrt{N}((1-v)\gamma_N+\hat{\gamma}_N)}e^{-bx_2\sqrt{N}((1+v)\gamma_N-\hat{\gamma}_N)}e^{-\frac{x_1^2}{2}}\\
&\qquad e^{-\frac{x_2^2}{2}}\mathrm{d}x_1\mathrm{d}x_2+O(N^{-\infty}).
\end{aligned} \tag{17}$$

Plug it back into (15), ignoring the $O(N^{-\infty})$ error,

$$\begin{aligned}
&\int_{\substack{|\gamma_N|\le N^{-\kappa_1},\\ |\hat{\gamma}_N|\le N^{-\kappa_2}}}e^{-\theta N\gamma_N(\hat{\gamma}_N-v\gamma_N)}\prod_{i=1}^{N}\mathrm{d}P(\tilde{g}_i)\\
&=\int_{\substack{|\gamma_N|\le N^{-\kappa_1},\\ |\hat{\gamma}_N|\le N^{-\kappa_2}}}\frac{1}{(\sqrt{2\pi})^2}\left\{\int_{I_2}\int_{I_1}e^{-\mathrm{i}bx_1\sqrt{N}((1-v)\gamma_N+\hat{\gamma}_N)}e^{-bx_2\sqrt{N}((1+v)\gamma_N-\hat{\gamma}_N)}\right.\\
&\quad\left.e^{-\frac{x_1^2}{2}}e^{-\frac{x_2^2}{2}}\mathrm{d}x_1\mathrm{d}x_2\right\}\prod_{i=1}^{N}\mathrm{d}P(\tilde{g}_i)
\end{aligned}$$

$$= \frac{1}{(\sqrt{2\pi})^2} \int_{I_1 \times I_2} e^{\mathrm{i}bx_1\sqrt{N}} e^{bx_2\sqrt{N}} \prod_{i=1}^{2} e^{-\frac{x_i^2}{2}} \mathrm{d}x_i$$
$$\int_{\substack{|\gamma_N| \le N^{-\kappa_1} \\ |\hat{\gamma}_N| \le N^{-\kappa_2}}} \prod_{i=1}^{N} \frac{1}{\sqrt{2\pi}} e^{-\frac{\tilde{g}_i^2}{2}\left(1+\frac{2b(\mathrm{i}(1-v+\lambda_i)x_1+(1+v-\lambda_i)x_2)}{\sqrt{N}(1-2\theta\lambda_i+2\theta v)}\right)} \prod_{i=1}^{N} \mathrm{d}\tilde{g}_i. \tag{18}$$

Notice here in the inner integral, the integral domain is the region $\mathcal{D} = \{|\gamma_N| \le N^{-\kappa_1}, |\hat{\gamma}_N| \le N^{-\kappa_2}\}$ and the Gaussian variables $\tilde{g}_i$ are located in this region with overwhelming probability. It is highly likely that if we instead integrate over the whole space $\mathbb{R}^N$, the error is exponentially small. We will first compute the integral under the belief that the integral outside this region $\mathcal{D}$ is negligible, then come back to this point later. Replace the integral region $\mathcal{D}$ by $\mathbb{R}^N$,

$$e^{\mathrm{i}bx_1\sqrt{N}} e^{bx_2\sqrt{N}} \left\{ \int_{\mathbb{R}^N} \prod_{i=1}^{N} \frac{1}{\sqrt{2\pi}} e^{-\frac{\tilde{g}_i^2}{2}\left(1+\frac{2b(\mathrm{i}(1-v+\lambda_i)x_1+(1+v-\lambda_i)x_2)}{\sqrt{N}(1-2\theta\lambda_i+2\theta v)}\right)} \prod_{i=1}^{N} \mathrm{d}\tilde{g}_i \right\}$$
$$= \prod_{i=1}^{N} \frac{\exp\left\{\frac{b(\mathrm{i}(1-v+\lambda_i)x_1+(1+v-\lambda_i)x_2)}{\sqrt{N}(1-2\theta\lambda_i+2\theta v)}\right\}}{\sqrt{1+\frac{2b(\mathrm{i}(1-v+\lambda_i)x_1+(1+v-\lambda_i)x_2)}{\sqrt{N}(1-2\theta\lambda_i+2\theta v)}}}, \tag{19}$$

where for the numerator we used the definition (8) of v, such that

$$\frac{1}{N}\sum_{i=1}^{N} \frac{1-v+\lambda_i}{1-2\theta\lambda_i+2\theta v} = 1, \quad \frac{1}{N}\sum_{i=1}^{N} \frac{1+v-\lambda_i}{1-2\theta\lambda_i+2\theta v} = 1. \tag{20}$$

Let $\mu_i = \frac{b(\mathrm{i}(1-v+\lambda_i)x_1+(1+v-\lambda_i)x_2)}{\sqrt{N}(1-2\theta\lambda_i+2\theta v)}$. Then $\mu_i = O(N^{\epsilon-\kappa_2})$, where $0 < \epsilon < \kappa_2$. So we have the Taylor expansion

$$\prod_{i=1}^{N} \frac{e^{\mu_i}}{\sqrt{1+2\mu_i}} = \prod_{i=1}^{N} e^{\mu_i^2+o(\mu_i^2)} = \prod_{i=1}^{N} e^{\mu_i^2}\left(1+o\left(\sum_{i=1}^{N}\mu_i^2\right)\right),$$

later we will see $\int e^{\sum_{i=1}^N \mu^2} \sum_{i=1}^N \mu_i^2 \mathrm{d}P(x_1)\mathrm{d}P(x_2) = \mathrm{O}(1)$. Thus, the first term in the asymptotics of integral (18) comes from the integral of $e^{\sum_{i=1}^N \mu_i^2}$, which is

$$\frac{1}{(\sqrt{2\pi})^2} \int_{I_2}\int_{I_1} e^{\frac{\theta}{2N}\sum_{i=1}^{N} \frac{(\mathrm{i}(1-v+\lambda_i)x_1+(1+v-\lambda_i)x_2)^2}{(1-2\theta\lambda_i+2\theta v)^2} - \frac{x_1^2+x_2^2}{2}} \mathrm{d}x_1\mathrm{d}x_2. \tag{21}$$

The exponent is a quadratic form and can be written as $-\frac{1}{2}\langle x, Kx\rangle$ where K is the following 2×2 symmetric matrices,

$$K := \begin{bmatrix} 1+\theta\left((1-v)^2A_2+2(1-v)F+G\right) & -\theta\mathrm{i}\left((1-v^2)A_2+2vF-G\right) \\ -\theta\mathrm{i}\left((1-v^2)A_2+2vF-G\right) & 1-\theta\left((1+v)^2A_2-2(1+v)F+G\right) \end{bmatrix},$$

where A_2, F, G are respectively $\frac{1}{N}\sum_{i=1}^N \frac{1}{(1-2\theta\lambda_i+2\theta v)^2}$, $\frac{1}{N}\sum_{i=1}^N \frac{\lambda_i}{(1-2\theta\lambda_i+2\theta v)^2}$ and $\frac{1}{N}\sum_{i=1}^N \frac{\lambda_i^2}{(1-2\theta\lambda_i+2\theta v)^2}$. With this notation, above integral (21) can be rewritten as

$$\frac{1}{(\sqrt{2\pi})^2}\int_{I_2}\int_{I_1} e^{-\frac{1}{2}\langle x, Kx\rangle}\mathrm{d}x_1\mathrm{d}x_2. \tag{22}$$

To deal with this complex Gaussian integral, we will use Lemma A3. Now we need to verify that our matrix K satisfies the condition of Lemma A3. Write K as the following sum

$$\begin{aligned} K = &\begin{bmatrix} 1+\theta\left((1-v)^2A_2+2(1-v)F+G\right) & 0 \\ 0 & 1-\theta\left((1+v)^2A_2-2(1+v)F+G\right)\end{bmatrix} \\ &+ \mathrm{i}\begin{bmatrix} 0 & -\theta\left((1-v^2)A_2+2vF-G\right) \\ -\theta\left((1-v^2)A_2+2vF-G\right) & 0\end{bmatrix}. \end{aligned}$$

Since $\min\{\lambda_i\} \le v(\theta) \le \max\{\lambda_i\}$, $|v(\theta)| < \max|\lambda_i| < M$. If $\theta < \frac{1}{4M^2+10M+1}$, it is easy to check the real part of matrix K is positive definite. To use Lemma A3, the only thing is to compute the determinant of matrix K. Notice the algebraic relations between parameters A_2, F, G, $\det(K) = A_2$. Therefore, we obtain the first-order asymptotic of the spherical integral,

$$\begin{aligned} &e^{-N(\theta v-\frac{1}{2N}\sum_{i=1}^N \log(1-2\theta\lambda_i(B_N)+2\theta v))} I_N(\theta, B_N) \\ &\quad = \frac{1}{(\sqrt{2\pi})^2}\int_{I_2}\int_{I_1} e^{-\frac{1}{2}\langle x, Kx\rangle}(1+o(1))\mathrm{d}x = \frac{1}{\sqrt{A_2}} + o(1). \end{aligned}$$

Go back to integral (18), we need to prove that the integral outside the region $\mathcal{D}$ is of order $O(N^{-\infty})$, then replacing the integral domain $\mathcal{D} = \{|\hat{\gamma}_N| \le N^{-\kappa_1}, |\gamma_N| \le N^{-\kappa_2}\}$ by the whole space $\mathbb{R}^N$ won't affect the asymptotic expansion.

Lemma 1 *Consider integral (18) on the complement of $\mathcal{D}$, i.e., on $\{|\hat{\gamma}_N| \ge N^{-\kappa_1}$ or $|\gamma_N| \ge N^{-\kappa_2}\}$,*

$$\begin{aligned} R := \Bigg| &\int_{I_2}\int_{I_1} e^{\mathrm{i}bx_1\sqrt{N}+bx_2\sqrt{N}} e^{-\frac{x_1^2}{2}} e^{-\frac{x_2^2}{2}}\mathrm{d}x_1\mathrm{d}x_2 \\ &\int_{\substack{|\hat{\gamma}_N|\ge N^{-\kappa_1} \\ \text{or}|\gamma_N|\ge N^{-\kappa_2}}} \prod_{i=1}^N \frac{1}{\sqrt{2\pi}} e^{-\frac{\tilde{g}_i^2}{2}\left(1+\frac{2b(\mathrm{i}(1-v+\lambda_i)x_1+(1+v-\lambda_i)x_2)}{\sqrt{N}(1-2\theta\lambda_i+2\theta v)}\right)} \prod_{i=1}^N \mathrm{d}\tilde{g}_i \Bigg|. \end{aligned}$$

Then for N large enough, the error $R \le c' e^{-cN^{\zeta}}$, for some constant $c, c', \zeta > 0$ depending on κ_1, κ_2 and $\|B_N\|_\infty$.

Proof

$$R \le \int_{I_2}\int_{I_1} e^{\Re\{\mathrm{i}bx_1\sqrt{N}+bx_2\sqrt{N}\}} e^{-\frac{x_1^2}{2}} e^{-\frac{x_2^2}{2}}\mathrm{d}x_1\mathrm{d}x_2$$

$$\int_{\substack{|\hat{\gamma}_N|\geq N^{-\kappa_1}\\ \text{or } |\gamma_N|\geq N^{-\kappa_2}}} \prod_{i=1}^{N} \frac{1}{\sqrt{2\pi}} e^{-\Re\left\{\frac{\tilde{g}_i^2}{2}\left(1+\frac{2b(\mathrm{i}(1-v+\lambda_i)x_1+(1+v-\lambda_i)x_2)}{\sqrt{N}(1-2\theta\lambda_i+2\theta v)}\right)\right\}} \prod_{i=1}^{N} \mathrm{d}\tilde{g}_i.$$

Since $b=\sqrt{\theta/2}$, b is either real or imaginary. For simplicity, here we only discuss the case when b is real. The case when b is imaginary can be proved in the same way. The above integral can be simplified as

$$R \leq \int_{I_2} e^{bx_2\sqrt{N}} \left\{ \int_{\substack{|\hat{\gamma}_N|\geq N^{-\kappa_1}\\ \text{or } |\gamma_N|\geq N^{-\kappa_2}}} \prod_{i=1}^{N} \frac{1}{\sqrt{2\pi}} e^{-\frac{\tilde{g}_i^2}{2}\left(1+\frac{2b(1+v-\lambda_i)x_2}{\sqrt{N}(1-2\theta\lambda_i+2\theta v)}\right)} \prod_{i=1}^{N} \mathrm{d}\tilde{g}_i \right\} e^{-\frac{x_2^2}{2}} \mathrm{d}x_2.$$

To simplify it, we perform a change of measure. Let $h_i = \tilde{g}_i\sqrt{1+\frac{2b(1+v-\lambda_i)x_2}{\sqrt{N}(1-2\theta\lambda_i+2\theta v)}}$,

$$R \leq \int_{I_2} \frac{e^{bx_2\sqrt{N}}}{\prod_{i=1}^{N}\sqrt{1+\frac{2b(1+v-\lambda_i)x_2}{\sqrt{N}(1-2\theta\lambda_i+2\theta v)}}} P_h(\mathcal{D}^c) e^{-\frac{x_2^2}{2}} \mathrm{d}x_2. \tag{23}$$

Here $P_h(\cdot)$ is the Gaussian measure of $(h_i)_{i=1}^N$. Take $0<\epsilon<\frac{1}{2}-\kappa_1$, we can separate the above integral into two parts,

$$\begin{aligned} R \leq\ & \underbrace{\int_{[-N^\epsilon, N^\epsilon]} \frac{e^{bx_2\sqrt{N}}}{\prod_{i=1}^{N}\sqrt{1+\frac{2b(1+v-\lambda_i)x_2}{\sqrt{N}(1-2\theta\lambda_i+2\theta v)}}} P_h(\mathcal{D}^c) e^{-\frac{x_2^2}{2}} \mathrm{d}x_2}_{E_1} \\ & + \underbrace{\int_{I_2\cap[-N^\epsilon, N^\epsilon]^c} \frac{e^{bx_2\sqrt{N}}}{\prod_{i=1}^{N}\sqrt{1+\frac{2b(1+v-\lambda_i)x_2}{\sqrt{N}(1-2\theta\lambda_i+2\theta v)}}} e^{-\frac{x_2^2}{2}} \mathrm{d}x_2}_{E_2}. \end{aligned} \tag{24}$$

For E_2, it is of the same form as (19). Use the same argument, the main contribution from E_2 is a Gaussian integral on $[-N^\epsilon, N^\epsilon]^c$. So it stretched exponential decays, $E_2 \leq c' e^{-cN^\zeta}$. For E_1,

$$E_1 \leq \int \frac{e^{bx_2\sqrt{N}}}{\prod_{i=1}^{N}\sqrt{1+\frac{2b(1+v-\lambda_i)x_2}{\sqrt{N}(1-2\theta\lambda_i+2\theta v)}}} e^{-\frac{x_2^2}{2}} \mathrm{d}x_2 \sup_{x_2\in[-N^\epsilon, N^\epsilon]} \{P_h(\mathcal{D}^c)\}.$$

If we can show that on the interval $[-N^\epsilon, N^\epsilon]$, $P_h(\mathcal{D}^c)$ is uniformly exponentially small, independent of x_2, then E_1 is exponentially small. For the upper bound of $P_h(\mathcal{D}^c)$, first by the union bound,

$$P_h(\mathcal{D}^c) \leq P_h(|\gamma_N| > N^{-\kappa_1}) + P_h(|\hat{\gamma}_N| > N^{-\kappa_2}).$$

For simplicity here I will only bound the first term, the second term can be bounded in exactly the same way.

$$\begin{aligned}
&P_h(|\gamma_N| > N^{-\kappa_1})\\
&= P_h\left(\left|\frac{1}{N}\sum_{i=1}^{N}\frac{h_i^2}{1-2\lambda_i\theta+2\theta v+2bx_2(1+v-\lambda_i)/\sqrt{N}}-1\right|\geq N^{-\kappa_1}\right)\\
&= P_h\left(\left|\frac{1}{N}\sum_{i=1}^{N}\frac{h_i^2-1}{1-2\lambda_i\theta+2\theta v+2bx_2(1+v-\lambda_i)/\sqrt{N}}+O(N^{\epsilon-\frac{1}{2}})\right|\geq N^{-\kappa_1}\right)\\
&\leq c'e^{-cN^{\zeta}},
\end{aligned} \tag{25}$$

For the second to last line, we used the defining relation (8) of v, such that $\frac{1}{N}\sum_{i=1}^{N}\frac{1}{1-2\lambda_i\theta+2v\theta} = 1$. And for the last inequality, since $\epsilon-\frac{1}{2} < -\kappa_2$, the term $O(N^{\epsilon-\frac{1}{2}})$ is negligible compared with $N^{-\kappa_1}$. Then the concentration measure inequality in Lemma A1 implies (25). □

3.2 Higher-Order Expansion

We recall the definition of γ_N and $\hat{\gamma}_N$ in terms of the new variables $\tilde{g}_i$'s,

$$\gamma_N = \frac{1}{N}\sum_{i=1}^{N}\frac{\tilde{g}_i^2}{1-2\theta\lambda_i+2\theta v}-1,\quad \hat{\gamma}_N = \frac{1}{N}\sum_{i=1}^{N}\frac{\lambda_i\tilde{g}_i^2}{1-2\theta\lambda_i+2\theta v}-v. \tag{26}$$

To compute the higher-order expansion of the spherical integral $I_N(\theta, B_N)$, we need to obtain a full asymptotic expansion of (13).

$$\int_{\substack{|\gamma_N|\leq N^{-\kappa_1},\\ |\hat{\gamma}_N|\leq N^{-\kappa_2}}} \exp\left\{-\theta N\gamma_N\frac{\hat{\gamma}_N-v\gamma_N}{\gamma_N+1}\right\}\prod_{i=1}^{i=N}\mathrm{d}P(\tilde{g}_i) \tag{27}$$

$$\begin{aligned}
&= \int_{\substack{|\gamma_N|\leq N^{-\kappa_1},\\ |\hat{\gamma}_N|\leq N^{-\kappa_2}}} e^{-\theta N\gamma_N(\hat{\gamma}_N-v\gamma_N)}\exp\left\{\theta N\gamma_N^2\frac{\hat{\gamma}_N-v\gamma_N}{\gamma_N+1}\right\}\prod_{i=1}^{N}\mathrm{d}P(\tilde{g}_i)\\
&= \int_{\substack{|\gamma_N|\leq N^{-\kappa_1},\\ |\hat{\gamma}_N|\leq N^{-\kappa_2}}} e^{-\theta N\gamma_N(\hat{\gamma}_N-v\gamma_N)}\left\{\sum_{k=0}^{\infty}\frac{1}{k!}\left(\theta N\gamma_N^2\frac{\hat{\gamma}_N-v\gamma_N}{\gamma_N+1}\right)^k\right\}\prod_{i=1}^{N}\mathrm{d}P(\tilde{g}_i)\\
&= \int_{\substack{|\gamma_N|\leq N^{-\kappa_1},\\ |\hat{\gamma}_N|\leq N^{-\kappa_2}}} e^{-\theta N\gamma_N(\hat{\gamma}_N-v\gamma_N)}\prod_{i=1}^{N}\mathrm{d}P(\tilde{g}_i)+\sum_{l=1}^{+\infty}\int_{\substack{|\gamma_N|\leq N^{-\kappa_1},\\ |\hat{\gamma}_N|\leq N^{-\kappa_2}}} e^{-\theta N\gamma_N(\hat{\gamma}_N-v\gamma_N)}\\
&\quad\times\left\{\sum_{k=1}^{l}\binom{l-1}{k-1}\frac{(-1)^k\theta^k}{k!}(-\gamma_N)^l(\sqrt{N}\gamma_N)^k\left(\sqrt{N}(\hat{\gamma}_N-v\gamma_N)\right)^k\right\}\prod_{i=1}^{N}\mathrm{d}P(\tilde{g}_i).
\end{aligned} \tag{28}$$

We consider the l-th summand in expression (28). As proved in [4, Theorem 3], the distribution of $(\sqrt{N}\gamma_N, \sqrt{N}\hat{\gamma}_N)$ tends to Γ, which is a centered two-dimensional Gaussian measure on $\mathbb{R}^2$ with covariance matrix

$$R = 2\begin{bmatrix} A_2 & F \\ F & G \end{bmatrix}.$$

Moreover, the matrix R is non-degenerate. Therefore, by the central limit theorem, we can obtain the asymptotic expression of the l-th term in (28)

$$\begin{aligned} &\int_{\substack{|\gamma_N|\le N^{-\kappa_1},\\ |\hat{\gamma}_N|\le N^{-\kappa_2}}} e^{-\theta N\gamma_N(\hat{\gamma}_N - v\gamma_N)} \sum_{k=1}^{l} \binom{l-1}{k-1} \frac{(-1)^k\theta^k}{k!} (-\gamma_N)^l (\sqrt{N}\gamma_N)^k \\ &\left(\sqrt{N}(\hat{\gamma}_N - v\gamma_N)\right)^k \prod_{i=1}^{N} \mathrm{d}P(\tilde{g}_i) \\ &= \left(\frac{-1}{\sqrt{N}}\right)^l \left\{\int e^{-\theta N x(y-vx)} \left\{\sum_{k=1}^{l} \binom{l-1}{k-1} \frac{(-1)^k\theta^k}{k!} x^{l+k}(y-vx)^k\right\} d\Gamma(x,y) + o(1)\right\} \\ &= O(N^{-l/2}). \end{aligned}$$

Therefore, if we cut off the infinite sum (28) at the l-th term, then the error terms are of magnitude $O(N^{-(l+1)/2})$. To obtain the full expansion, we need to understand each term in expansion (28).

$$\begin{aligned} &\int_{\substack{|\gamma_N|\le N^{-\kappa_1},\\ |\hat{\gamma}_N|\le N^{-\kappa_2}}} e^{-\theta N\gamma_N(\hat{\gamma}_N - v\gamma_N)} \binom{l-1}{k-1} \frac{(-1)^k\theta^k}{k!} (-\gamma_N)^l N^k (\gamma_N)^k \\ &(\hat{\gamma}_N - v\gamma_N)^k \prod_{i=1}^{N} \mathrm{d}P(\tilde{g}_i). \end{aligned} \tag{29}$$

Define

$$f_l(t) = \int_{\substack{|\gamma_N|\le N^{-\kappa_1},\\ |\hat{\gamma}_N|\le N^{-\kappa_2}}} e^{-tN\gamma_N(\hat{\gamma}_N - v\gamma_N)} \gamma_N^l \prod_{i=1}^{N} \mathrm{d}P(\tilde{g}_i). \tag{30}$$

By the dominated convergence theorem, we can interchange derivative and integral. Integral (29) can be written as

$$(-1)^l \binom{l-1}{k-1} \frac{\theta^k}{k!} \frac{\mathrm{d}^k f_l(t)}{\mathrm{d}t^k}\Big|_{t=\theta}.$$

Therefore to understand asymptotic expansion of (29), we only need to compute the asymptotic expansion of $f_l(t)$, for $l = 0, 1, 2, \ldots$. We have the following proposition

Proposition 2 *If* $\sup_N \|B_N\|_\infty < M$, *then for any* $\theta \in \mathbb{R}$ *such that* $|\theta| < \frac{1}{4M^2+10M+1}$, f_l *has the following asymptotic expansion (up to* $O(N^{-n-1})$ *for any given* n*)*

$$f_l(t) = m_0 + \frac{m_1}{N} + \frac{m_2}{N^2} + \cdots + \frac{m_n}{N^n} + O(N^{-n-1}), \tag{31}$$

where $\{m_i\}_{i=0}^n$ *depends explicitly on* t, θ, v *and the derivative of the Hilbert transform of the empirical spectral distribution of* B_N, *namely,* $A_2, A_3, A_4 \dots A_{2n+2}$, *as defined in* (9).

Proof First we show that f_l has asymptotic expansion in form (31), then we show those m_i's depend only explicitly on t, θ and $\{A_k\}_{k=2}^{+\infty}$. We introduce two Gaussian random variables x_1 and x_2, the same as in (17) and (18), we obtain the following expression of $f_l(t)$,

$$f_l(t) = \frac{1}{(\sqrt{2\pi})^2} \int_{I_1 \times I_2} e^{\mathrm{i}bx_1\sqrt{N} + bx_2\sqrt{N}} \prod_{i=1}^{2} e^{-\frac{x_i^2}{2}} \mathrm{d}x_i$$
$$\int_{\substack{|\hat{\gamma}_N| \le N^{-\kappa_1} \\ |\gamma_N| \le N^{-\kappa_2}}} \prod_{i=1}^{N} \frac{1}{\sqrt{2\pi}} e^{-\frac{\tilde{g}_i^2}{2}\left(1+\frac{2b(\mathrm{i}(1-v+\lambda_i)x_1+(1+v-\lambda_i)x_2)}{\sqrt{N}(1-2\theta\lambda_i+2\theta v)}\right)} \gamma_N^l \prod_{i=1}^{N} \mathrm{d}\tilde{g}_i + O(N^{-\infty}),$$

where $b^2 = t/2$, $I_1 = [-N^{\frac{1}{2}-\kappa_1+\epsilon}, N^{\frac{1}{2}-\kappa_1+\epsilon}]$ and $I_2 = [-N^{\frac{1}{2}-\kappa_2+\epsilon}, N^{\frac{1}{2}-\kappa_2+\epsilon}]$, for some $\epsilon > 0$. The same argument, as in Lemma 1, about replacing integral domain in the inner integral can be implemented here without too much change. So we can replace the integral domain $\{|\hat{\gamma}_N| \le N^{-\kappa_1}, |\gamma_N| \le N^{-\kappa_2}\}$ by $\mathbb{R}^N$.

$$e^{\mathrm{i}bx_1\sqrt{N}+bx_2\sqrt{N}} \int_{\mathbb{R}^N} \prod_{i=1}^{N} \frac{1}{\sqrt{2\pi}} e^{-\frac{\tilde{g}_i^2}{2}\left(1+\frac{2b(\mathrm{i}(1-v+\lambda_i)x_1+(1+v-\lambda_i)x_2)}{\sqrt{N}(1-2\theta\lambda_i+2\theta v)}\right)} \tag{32}$$
$$\left\{\frac{1}{N}\sum_{i=1}^{N} \frac{(\tilde{g}_i^2-1)}{1-2\theta\lambda_i+2\theta v}\right\}^l \prod_{i=1}^{N} \mathrm{d}\tilde{g}_i$$

$$= \prod_{i=1}^{N} \frac{\exp\left\{\frac{b(\mathrm{i}(1-v+\lambda_i)x_1+(1+v-\lambda_i)x_2)}{\sqrt{N}(1-2\theta\lambda_i+2\theta v)}\right\}}{\sqrt{1+\frac{2b(\mathrm{i}(1-v+\lambda_i)x_1+(1+v-\lambda_i)x_2)}{\sqrt{N}(1-2\theta\lambda_i+2\theta v)}}} \int_{\mathbb{R}^N} \prod_{i=1}^{N} \frac{1}{\sqrt{2\pi}} e^{-\frac{\tilde{g}_i^2}{2}} \tag{33}$$
$$\left\{\frac{1}{N}\sum_{i=1}^{N} \frac{\frac{\tilde{g}_i^2}{1+\frac{2b(\mathrm{i}(1-v+\lambda_i)x_1+(1+v-\lambda_i)x_2)}{\sqrt{N}(1-2\theta\lambda_i+2\theta v)}} - 1}{1-2\theta\lambda_i+2\theta v}\right\}^l \prod_{i=1}^{N} \mathrm{d}\tilde{g}_i$$

$$= \underbrace{\prod_{i=1}^{N} \frac{e^{\mu_i}}{\sqrt{1+2\mu_i}}}_{E_1} \underbrace{\int \prod_{i=1}^{N} \frac{1}{\sqrt{2\pi}} e^{-\frac{\tilde{g}_i^2}{2}} \left\{\frac{1}{N}\sum_{i=1}^{N} \frac{\nu_i}{1+2\mu_i}(\tilde{g}_i^2 - (1+2\mu_i))\right\}^l \prod_{i=1}^{N} \mathrm{d}\tilde{g}_i}_{E_2}, \tag{34}$$

where we used (20) for the first equality, and

$$\nu_i := \frac{1}{1-2\theta\lambda_i+2\theta v}, \quad \mu_i := \frac{b(\mathrm{i}(1-v+\lambda_i)x_1+(1+v-\lambda_i)x_2)}{\sqrt{N}(1-2\theta\lambda_i+2\theta v)}.$$

Notice here μ_i can be written as a linear function of ν_i

$$\mu_i = \left(i\left(1+\frac{1}{2\theta}\right)\frac{x_1}{\sqrt{N}}+\left(1-\frac{1}{2\theta}\right)\frac{x_2}{\sqrt{N}}\right)b\nu_i+\left(\frac{x_2}{\sqrt{N}}-i\frac{x_1}{\sqrt{N}}\right)\frac{b}{2\theta}. \quad (35)$$

Formula (34) consists of two parts: a product factor E_1 and a Gaussian integral E_2. For E_1 we can obtain the following explicit asymptotic expansion

$$\prod_{i=1}^{N}\frac{e^{\mu_i}}{\sqrt{1+2\mu_i}} = \prod_{i=1}^{N} e^{\mu_i-\frac{1}{2}\log(1+2\mu_i)} = \prod_{i=1}^{N} e^{\mu_i^2+\sum_{k=3}^{\infty}\frac{(-1)^k 2^{k-1}}{k}\mu_i^k}$$
$$= e^{\sum_{i=1}^{N}\mu_i^2} e^{\sum_{k=3}^{\infty}\frac{1}{N^{\frac{k}{2}-1}}\left\{\frac{(-1)^k 2^{k-1}}{k}\sum_{i=1}^{N}\frac{b^k(\mathrm{i}(1-v+\lambda_i)x_1+(1+v-\lambda_i)x_2)^k}{N(1-2\theta\lambda_i+2\theta v)^k}\right\}}. \quad (36)$$

Notice

$$\sum_{i=1}^{N}\frac{(\mathrm{i}(1-v+\lambda_i)x_1+(1+v-\lambda_i)x_2)^k}{N(1-2\theta\lambda_i+2\theta v)^k}$$
$$= \sum_{m=0}^{k}\left\{\binom{k}{m}(ix_1-x_2)^m(\mathrm{i}(1-v)x_1+(1+v)x_2)^{k-m}\sum_{i=1}^{N}\frac{\lambda_i^m}{N(1+2\theta\lambda_i+2\theta v)^k}\right\}.$$

If we regard v and θ as constants (since they are of magnitude $O(1)$), then the sum $\sum_{i=1}^{N}\frac{\lambda_i^m}{N(1+2\theta\lambda_i+2\theta v)^k}$ can be written as a linear combination of $A_2, A_3 \ldots, A_k$ for any $0 \le m \le k$. Thus we can expand (36), to obtain

$$\prod_{i=1}^{N}\frac{e^{\mu_i}}{\sqrt{1+2\mu_i}} = e^{\sum_{i=1}^{N}\mu_i^2}\left\{1+\sum_{k=1}^{\infty}\frac{1}{N^{\frac{k}{2}}}g_k(x_1,x_2)\right\},$$

where $g_k(x_1,x_2)$'s are polynomials of x_1 and x_2. Consider t, θ and v as constants, the coefficients of $g_k(x_1,x_2)$ are polynomials in terms of $A_2, A_3, \ldots A_{k+2}$. Moreover the degree of each monomial of $g_k(x_1,x_2)$ is congruent to k modulo 2.

Next we compute the Gaussian integral E_2 in (34). Expand the l-th power, we obtain

$$\left\{\frac{1}{N}\sum_{i=1}^{N}\frac{\nu_i}{1+2\mu_i}(\tilde{g}_i^2-(1+2\mu_i))\right\}^l$$
$$= \frac{1}{N^l}\sum_{\substack{k_1\ge k_2\cdots\ge k_m\\ k_1+k_2\ldots k_m=l}}\left\{\frac{l!}{k_1!k_2!\cdots k_m!}\sum_{\substack{1\le i_1,i_2,\ldots i_m\le N\\ \text{distinct}}}\prod_{j=1}^{m}(\frac{\nu_{i_j}}{1+2\mu_{i_j}})^{k_j}(\tilde{g}_{i_j}^2-(1+2\mu_{i_j}))^{k_j}\right\}.$$

Denote

$$p_{k_j}(\mu_{i_j}) = \int \frac{1}{\sqrt{2\pi}} e^{-\frac{\tilde{g}_{i_j}^2}{2}} (\tilde{g}_{i_j}^2 - (1 + 2\mu_{i_j}))^{k_j} \mathrm{d}\tilde{g}_{i_j},$$

then p_{k_j} is a k_j-th degree polynomial, which only depends on k_j. With this notation, the Gaussian integral E_2 can be written as,

$$\frac{1}{N^l} \sum_{\substack{k_1 \geq k_2 \cdots \geq k_m \\ k_1 + k_2 \cdots k_m = l}} \left\{ \frac{l!}{k_1! k_2! \cdots k_m!} \sum_{\substack{1 \leq i_1, i_2, \ldots i_m \leq N \\ \text{distinct}}} \prod_{j=1}^{m} \left(\frac{\nu_{i_j}}{1 + 2\mu_{i_j}} \right)^{k_j} p_{k_j}(\mu_{i_j}) \right\}.$$

By the following lemma, the above expression can be expressed in a more symmetric way, as a sum in terms of

$$\frac{1}{N^{l-m}} \prod_{j=1}^{m} \left\{ \frac{1}{N} \sum_{i=1}^{N} \left(\frac{\nu_i}{1 + \mu_i} \right)^{k_j} q_{k_j}(\mu_i) \right\}, \tag{37}$$

where $k_1 \geq k_2 \cdots \geq k_m$, $k_1 + k_2 + \cdots k_m = l$ and q_{k_j}'s are some polynomials depending only on k_j. □

Lemma 2 *Given integers $s_1, s_2 \ldots, s_m$ and polynomials $q_1, q_2, \ldots q_m$, consider the following polynomial in terms of $2N$ variables $x_1, x_2 \ldots x_N, y_1, y_2 \ldots y_N$*

$$h = \sum_{\substack{1 \leq i_1, i_2 \ldots i_m \leq N \\ \textit{distinct}}} \prod_{j=1}^{m} x_{i_j}^{s_j} q_j(y_{i_j}).$$

Then h can be expressed as sum of terms in the following form

$$\prod_{j=1}^{l} \left\{ \sum_{i=1}^{N} x_i^{t_j} \tilde{q}_j(y_i) \right\}, \tag{38}$$

where l, $\{t_i\}_{i=1}^N$ and polynomials $\{\tilde{q}_i\}_{i=1}^N$ are to be chosen.

Proof We prove this by induction on m. If $m = 1$ then h itself is of form (38). We assume the statement holds for $1, 2, 3 \ldots, m-1$, then we prove it for m.

$$h - \prod_{j=1}^{m} \left(\sum_{i=1}^{N} x_i^{s_j} q_j(y_i) \right)$$
$$= -\sum_{d=1}^{m-1} \sum_{\substack{\pi_1, \pi_2 \cdots \pi_d \\ \text{a partition of } \{1,2,\ldots,m\}}} \sum_{\substack{1 \leq i_1, i_2 \ldots, i_d \leq N \\ \text{distinct}}} \prod_{j=1}^{d} x_{i_j}^{\sum_{l \in \pi_j} s_l} \prod_{l \in \pi_j} q_l(y_{i_j}). \tag{39}$$

Notice the summands of (39) are $\sum_{\substack{1\le i_1,i_2\dots,i_d\le N\\ \text{distinct}}}\prod_{j=1}^{d} x_{i_j}^{\sum_{l\in\pi_j} s_l}\prod_{l\in\pi_j} q_l(y_{i_j})$, which are of the same form as h but with less m. Thus, by induction, each term in (39) can be expressed as a sum of terms in form (38), so does h. □

In view of (37), since we have $\mu_i = O(N^{\epsilon-\kappa_2})$, we can Taylor expand $1/(1+\mu_i)$ in (37). Also notice (35), the relation between μ_i and ν_i, (37) has the following full expansion

$$\frac{1}{N^{l-m}}\sum_{k=0}^{\infty}\frac{1}{N^{\frac{k}{2}}}h_k(x_1,x_2), \tag{40}$$

where $h_k(x_1,x_2)$ are k-th degree polynomials of x_1 and x_2. Consider t, θ and v as constant (since they are of magnitude $O(1)$), the coefficients of $h_k(x_1,x_2)$ are polynomials of $A_2, A_3, \dots A_{k+l-m+1}$. Since the Gaussian integral E_2 is the sum of terms which has the asymptotic expansion (40), itself has the full asymptotic expansion,

$$E_2=\sum_{k=0}^{\infty}\frac{1}{N^{\frac{k}{2}}}s_k(x_1,x_2), \tag{41}$$

where the coefficients of $s_k(x_1,x_2)$ are polynomials of $A_2, A_3, \dots A_{k+1}$. And the degree of each monomial of s_k is congruent to k modulo 2. Combine the asymptotic expansions of E_1 and E_2, we obtain the following expansion of $f_l(t)$ (up to an error of order $O(N^{-\infty})$),

$$\begin{aligned}&\frac{1}{2\pi}\int_{I_1\times I_2} e^{-\frac12\langle x,\tilde{K}x\rangle}\left\{\sum_{k=0}^{\infty}\frac{1}{N^{\frac{k}{2}}}g_k(x_1,x_2)\right\}\left\{\sum_{k=0}^{\infty}\frac{1}{N^{\frac{k}{2}}}s_k(x_1,x_2)\right\}\mathrm{d}x_1\mathrm{d}x_2,\\ &=\frac{1}{2\pi}\int_{I_1\times I_2} e^{-\frac12\langle x,\tilde{K}x\rangle}\left\{\sum_{k=0}^{\infty}\frac{\sum_{l=0}^{k} g_l(x_1,x_2)s_{k-l}(x_1,x_2)}{N^{\frac{k}{2}}}\right\}\mathrm{d}x_1\mathrm{d}x_2.\end{aligned} \tag{42}$$

where $\tilde{K}$ is the following 2×2 matrix

$$\tilde{K}:=\begin{bmatrix} 1+t\left((1-v)^2A_2+2(1-v)F+G\right) & -t\mathrm{i}\left((1-v^2)A_2+2vF-G\right)\\ -t\mathrm{i}\left((1-v^2)A_2+2vF-G\right) & 1-t\left((1+v)^2A_2-2(1+v)F+G\right)\end{bmatrix}.$$

Formula (42) is a Gaussian integral in terms of x_1 and x_2. If we cut off at $k=m$, this will result in an error term $O(N^{-\frac{m+1}{2}})$. Now the integrand is a finite sum. The integral is $O(N^{-\infty})$ outside the region $I_1\times I_2$; thus, we obtain the following asymptotic expansion,

$$f_l(t)=\sum_{k=0}^{m}\frac{1}{N^{\frac{k}{2}}}\frac{1}{2\pi}\int e^{-\frac12\langle x,\tilde{K}x\rangle}\left\{\sum_{l=0}^{k} g_l(x_1,x_2)s_{k-l}(x_1,x_2)\right\}\mathrm{d}x_1\mathrm{d}x_2+O\left(N^{-\frac{m+1}{2}}\right). \tag{43}$$

Notice the degree of each monomial of $g_l(x_1,x_2)$ is congruent to l modulo 2, and the degree of each monomial of $s_{k-l}(x_1,x_2)$ is congruent to $k-l$ modulo 2. For any odd

k, $\sum_{l=0}^{k} g_l(x_1, x_2) s_{k-l}(x_1, x_2)$ is sum of monomials of odd degree. Since here x_1, x_2 are centered Gaussian variables, the integral of $\sum_{l=0}^{k} g_l(x_1, x_2) s_{k-l}(x_1, x_2)$ vanishes. Using Lemma A5, (43) can be rewritten as

$$\sum_{k=0}^{[\frac{m}{2}]} \frac{1}{N^k} \frac{1}{2\pi} \int e^{-\frac{1}{2}\langle x, \tilde{K} x\rangle} \left\{ \sum_{l=0}^{2k} g_l(x_1, x_2) s_{2k-l}(x_1, x_2) \right\} \mathrm{d}x_1 \mathrm{d}x_2 + O(N^{-[\frac{m}{2}]-1})$$
$$= \sum_{k=0}^{[\frac{m}{2}]} \frac{1}{N^k} \frac{1}{\sqrt{\det(\tilde{K})}} \left\{ \sum_{l=0}^{2k} g_l(\partial_{\xi_1}, \partial_{\xi_2}) s_{2k-l}(\partial_{\xi_1}, \partial_{\xi_2}) e^{\frac{1}{2}\langle \xi, \tilde{K}^{-1}\xi\rangle} \right\} \Big|_{\xi=0} + O(N^{-[\frac{m}{2}]-1}).$$

Since the entries of matrix $\tilde{K}$ and the coefficients of $\sum_{l=0}^{2k} g_l(x_1, x_2) s_{2k-l}(x_1, x_2)$ depend only on t, θ, v and $A_2, A_3, \ldots, A_{2k+2}$. This implies $f_l(t)$ has the expansion (31). □

From the argument above, the asymptotic expansion of (27) up to $O(N^{-l/2})$ is a finite sum in terms of derivatives of $f_k(t)$'s at $t = \theta$. Differentiate $f_l(t)$ term by term, we arrive at our main theorem of this paper,

Theorem 2 *If* $\sup_N \|B_N\|_\infty < M$, *then for any* $\theta \in \mathbb{R}$ *such that* $|\theta| < \frac{1}{4M^2+10M+1}$, *the spherical integral has the following asymptotic expansion (up to* $O(N^{-n-1})$ *for any given* n*)*

$$e^{-N(\theta v - \frac{1}{2N}\sum_{i=1}^{N} \log(1-2\theta\lambda_i(B_N)+2\theta v))} I_N(\theta, B_N)$$
$$= m_0 + \frac{m_1}{N} + \frac{m_2}{N^2} + \cdots + \frac{m_n}{N^n} + O(N^{-n-1}),$$

where $v = R_{B_N}(\theta)$ *and* $\{m_i\}_{i=0}^{n}$ *depends on* θ, v *and the derivatives of Hilbert transform of the empirical spectral distribution of* B_N *at* $v + 1/2\theta$, *namely* $A_2, A_3, \ldots A_{2n+2}$ *as defined in* (9). *Especially we have*

$$m_0 = \frac{1}{\sqrt{A_2}}, \quad m_1 = \frac{1}{\sqrt{A_2}} \left(\frac{3}{2}\frac{A_4}{A_2^2} - \frac{5}{3}\frac{A_3^2}{A_2^3} + \frac{1}{6} \right).$$

Proof In the last section, Theorem 1, we have computed the first term in the expansion $m_0 = \frac{1}{\sqrt{A_2}}$. We only need to figure out the second term m_1. For this, we cut off (31) at $l = 2$,

$$\int_{\substack{|\gamma_N| \le N^{-\kappa_1}, \\ |\hat{\gamma}_N| \le N^{-\kappa_2}}} \exp\left\{ -\theta N \gamma_N \frac{\hat{\gamma}_N - v\gamma_N}{\gamma_N + 1} \right\} \prod_{i=1}^{i=N} \mathrm{d}P(\tilde{g}_i)$$
$$= f_0\Big|_{t=\theta} - \theta \frac{\mathrm{d}}{\mathrm{d}t} f_1\Big|_{t=\theta} + \left(\frac{\theta^2}{2}\frac{d^2}{dt^2} + \theta \frac{\mathrm{d}}{\mathrm{d}t} \right) f_2\Big|_{t=\theta} + O(N^{-2}). \qquad (44)$$

Take $l = 0, 1, 2$ in (30), we obtain the asymptotic expansion of f_0, f_1 and f_2 (we put the detailed computation in the appendix),

$$\begin{aligned} f_0\Big|_{t=\theta} &= \frac{1}{\sqrt{A_2}} + \frac{1}{N}\frac{1}{\sqrt{A_2}}\left(\frac{7}{6} - \frac{3}{A_2} + \frac{2A_3}{A_2^2} - \frac{5}{3}\frac{A_3^2}{A_2^3} + \frac{3}{2}\frac{A_4}{A_2^2}\right) + O(N^{-2}), \\ -\theta\frac{\mathrm{d}}{\mathrm{d}t} f_1\Big|_{t=\theta} &= \frac{1}{N}\frac{1}{\sqrt{A_2}}\left(-4 + \frac{6}{A_2} - \frac{2A_3}{A_2^2}\right) + O(N^{-2}), \\ \left(\frac{\theta^2}{2}\frac{d^2}{dt^2} + \theta\frac{\mathrm{d}}{\mathrm{d}t}\right) f_2\Big|_{t=\theta} &= \frac{1}{N}\frac{1}{\sqrt{A_2}}\left(3 - \frac{3}{A_2}\right) + O(N^{-2}). \end{aligned}$$

Plug them back to (44), we get

$$m_0 = \frac{1}{\sqrt{A_2}}, \quad m_1 = \frac{1}{\sqrt{A_2}}\left(\frac{3}{2}\frac{A_4}{A_2^2} - \frac{5}{3}\frac{A_3^2}{A_2^3} + \frac{1}{6}\right).$$

□

4 Unitary Case

In this section, we consider the unitary case, $\beta = 2$. As we will see soon that the unitary case is a special case of orthogonal case. With the same notation as before, let $B_N = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_N)$. and U follows the Haar measure on unitary group $U(N)$. The first column e_1 of U can be parametrized as the normalized complex Gaussian vector,

$$e_1 = \frac{g^{(1)} + ig^{(2)}}{\|g^{(1)} + ig^{(2)}\|},$$

where $g^{(1)} = (g_1, g_3 \ldots, g_{2N-3}, g_{2N-1})^T$ and $g^{(2)} = (g_2, g_4, \ldots, g_{2N-2}, g_{2N})^T$ are independent Gaussian vectors in $\mathbb{R}^N$. Then the spherical integral has the following form

$$I_N^{(2)}(\theta, B_N) = \int \exp\left\{N\theta\frac{\lambda_1(g_1^2 + g_2^2) + \cdots + \lambda_N(g_{2N-1}^2 + g_{2N}^2)}{g_1^2 + g_2^2 + \cdots + g_{2N-1}^2 + g_{2N}^2}\right\}\prod_{i=1}^{2N}\mathrm{d}P(g_i).$$

Consider the $2N \times 2N$ diagonal matrix $D_{2N} = \mathrm{diag}\{\lambda_1, \lambda_1, \lambda_2, \lambda_2, \ldots, \lambda_N, \lambda_N\}$ with each λ_i appearing twice. Then we have the following relation,

$$I_N^{(2)}(\theta, B_N) = I_{2N}^{(1)}(\theta/2, D_{2N}).$$

Define $\tilde{v}$ and $\{\tilde{A}_i\}_{i=1}^{\infty}$ as in the notation section but replace θ by $\theta/2$ and replace B_N by D_{2N}, namely $\tilde{v} = R_{B_N}(\theta)$ and

$$\tilde{A}_k = \frac{(-1)^{k-1}}{(k-1)!\theta^k} \frac{d^{k-1} H_{B_N}}{dz^{k-1}} \left(\tilde{v} + \frac{1}{\theta} \right) = \frac{1}{N} \sum_{i=1}^{N} \frac{1}{(1 - \theta\lambda_i + \theta\tilde{v})^k}. \tag{45}$$

Then from Theorem 2, we have the following theorem for unitary case.

Theorem 3 *If* $\sup_N \|B_N\|_\infty < M$, *then for any* $\theta \in \mathbb{R}$ *such that* $|\theta| < \frac{1}{4M^2+10M+1}$, *the spherical integral* $I_N^{(2)}(\theta, B_N)$ *has the following asymptotic expansion (up to* $O(N^{-n-1})$ *for any given* n*)*

$$\begin{aligned} I_N^{(2)}(\theta, B_N) &= I_{2N}^{(1)}(\theta/2, D_{2N}) \\ &= e^{N(\theta v - \frac{1}{N}\sum_{i=1}^N \log(1-\theta\lambda_i(B_N)+\theta v))} \left\{ m_0 + \frac{m_1}{N} + \frac{m_2}{N^2} + \cdots + \frac{m_n}{N^n} + O(N^{-n-1}) \right\}, \end{aligned}$$

where $\tilde{v} = R_{B_N}(\theta)$ *and* $\{m_i\}_{i=0}^n$ *depends on* θ, $\tilde{v}$, $\{\tilde{A}_i\}_{i=2}^{2n+2}$, *and the derivatives of Hilbert transform of the empirical spectral distribution of* B_N *at* $\tilde{v} + 1/\theta$. *Especially we have*

$$m_0 = \frac{1}{\sqrt{\tilde{A}_2}}, \quad m_1 = \frac{1}{2\sqrt{\tilde{A}_2}} \left(\frac{3}{2} \frac{\tilde{A}_4}{\tilde{A}_2^2} - \frac{5}{3} \frac{\tilde{A}_3^2}{\tilde{A}_2^3} + \frac{1}{6} \right).$$

Acknowledgements This research was conducted at the Undergraduate Research Opportunities Program of the MIT Mathematics Department, under the direction of Prof. Alice Guionnet. I would like to express to her my warmest thanks both for introducing me to this problem and for her dedicated guidance throughout the research process. I want to also thank the anonymous reviewer for careful reading of the manuscript and helpful suggestions.

Appendix A. Properties of Gaussian Random Variables

The following is a useful lemma on the concentration of measure for the sum of squares of Gaussian random variables.

Lemma A1 *Given independent Gaussian random variables* $\{g_i\}_{i=1}^N$, *consider the weighted sum* $\sum_{i=1}^N a_i g_i^2$, *where the coefficients* $\{a_i\}_{i=1}^N$ *depend on* N. *If there exists some constant* $c > 0$, *such that* $\max\{|a_i|\} \le \frac{c}{\sqrt{N}}$, *then for* N *large enough, the weighted sum satisfies the following concentration inequality,*

$$\mathbb{P}\left(|\sum_{i=1}^{N} a_i (g_i^2 - 1)| \ge N^\kappa \right) \le 2e^{-c' N^{2\kappa}}, \qquad 0 < \kappa < \frac{1}{2}.$$

Proof This can be proved by applying Markov's inequality to $\exp\{t \sum_{i=1}^N a_i (g_i^2 - 1)\}$ for some well-chosen value of t. □

The following is a useful trick, which states that we can express $e^{\alpha^2/2}$ as a Gaussian integral. Then the exponents are all linear in α.

Lemma A2 *For any $\alpha \in \mathbb{C}$ with $|\alpha| \le CN^{\kappa}$,*

$$e^{\frac{\alpha^2}{2}} = \frac{1}{\sqrt{2\pi}} \int_I e^{-\frac{x^2}{2}} e^{-x\alpha} \mathrm{d}x + O(N^{-\infty}),$$

where interval $I = [-N^{\kappa+\epsilon}, N^{\kappa+\epsilon}]$ for any $\epsilon > 0$.

Proof Recall the formula, for any $\alpha \in \mathbb{C}$

$$\frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{-\frac{(x+\alpha)^2}{2}} = 1.$$

The main contribution of the above integral comes from where $|x+\alpha|$ is small. More precisely,

$$\begin{aligned} &\left| e^{\frac{\alpha^2}{2}} \int_{\mathbb{R}} e^{-\frac{(x+\alpha)^2}{2}} \mathrm{d}x - e^{\frac{\alpha^2}{2}} \int_I e^{-\frac{(x+\alpha)^2}{2}} \right| \le e^{\frac{|\alpha|^2}{2}} \left| \int_{I^c} e^{-\frac{(x+\alpha)^2}{2}} \mathrm{d}x \right| \\ &\le 2e^{\frac{|\alpha|^2}{2}} \int_{N^{\kappa+\epsilon}}^{\infty} e^{-\frac{(x-|\alpha|)^2}{2}} \mathrm{d}x \le 2e^{\frac{C^2N^{2\kappa}}{2}} \int_{N^{\kappa+\epsilon}-CN^{\kappa}}^{\infty} e^{-\frac{x^2}{2}} \mathrm{d}x \le c'e^{-cN^{2\kappa+2\epsilon}} = O(N^{-\infty}), \end{aligned}$$

where c and c' are constants independent of N. Therefore

$$e^{\frac{\alpha^2}{2}} = e^{\frac{\alpha^2}{2}} \frac{1}{\sqrt{2\pi}} \int e^{-\frac{(x+\alpha)^2}{2}} = \frac{1}{\sqrt{2\pi}} \int_I e^{-\frac{x^2}{2}} e^{-x\alpha} \mathrm{d}x + O(N^{-\infty}).$$

□

The following lemmas are useful identities about Gaussian integrals.

Lemma A3 *If an n by n symmetric matrix K can be written as $K = A + iB$, where A is a real positive definite matrix, B is a real symmetric matrix. Then we have the Gaussian integral formula*

$$\frac{1}{(\sqrt{2\pi})^n} \int_{\mathbb{R}^n} e^{-\frac{1}{2}\langle x, Kx\rangle} \mathrm{d}x = \frac{1}{\sqrt{det(K)}}.$$

Proof Since A is positive definite, there exists a positive definite matrix C such that $A = C^T C$. Since CBC^T is symmetric, it can be diagonalized by some special orthogonal matrix U, let $P = UC$. Then we have $A = P^T P$ and $B = P^T DP$, where $D = diag\{d_1, d_2, \ldots, d_n\}$. Thus $K = A + iB = P^T(I + iD)P$. Plug this back to the integral

$$\begin{aligned} &\frac{1}{(\sqrt{2\pi})^n} \int_{\mathbb{R}^N} e^{-\frac{1}{2}\langle x, Kx\rangle} \mathrm{d}x = \frac{1}{(\sqrt{2\pi})^n} \int_{\mathbb{R}^N} e^{-\frac{1}{2}\langle x, P^T(I+iD)Px\rangle} \mathrm{d}x \\ &= \frac{1}{(\sqrt{2\pi})^n \det(P)} \int_{\mathbb{R}^N} e^{-\frac{1}{2}\langle y, (I+iD)y\rangle} \mathrm{d}y = \frac{1}{(\sqrt{2\pi})^n \det(P)} \prod_{k=1}^n \int_{\mathbb{R}^N} e^{-\frac{(1+id_k)}{2} y_k^2} \mathrm{d}y_k \\ &= \frac{1}{(\sqrt{2\pi})^n \det(P)} \prod_{k=1}^n \sqrt{\frac{2\pi}{1+id_k}} = \frac{1}{\sqrt{\det(K)}}, \end{aligned} \tag{46}$$

where the square root in (46) is the branch with positive real part, in our case K is a 2×2 matrix. □

Lemma A4 *Given n by n symmetric matrix K, whose real part is positive definite, then we have the following change of variable formula for Gaussian type integral,*

$$\int_{\mathbb{R}^n} F(x)e^{-\frac{1}{2}\langle x,Kx\rangle}\mathrm{d}x = \det(A)\int_{\mathbb{R}^n} F(Ay+b)e^{-\frac{1}{2}\langle Ay+b,K(Ay+b)\rangle}\mathrm{d}y,$$

where F is a polynomial in terms of $\{x_i\}_{i=1}^n$, $A = diag\{a_1, a_2 \ldots, a_n\}$ and $\Re(a_i) > 0$ for $i = 1, 2 \ldots n$, and $b \in \mathbb{C}^n$.

Proof This can be proved by reducing to one dimensional case. □

Lemma A5 *Given n by n symmetric matrix K, whose real part is positive definite, then we have the following integral formula for any polynomial F(or even infinite power series) in the variables $\{x_i\}_{i=1}^n$,*

$$\int_{\mathbb{R}^n} F(x)e^{-\frac{1}{2}\langle x,Kx\rangle}\mathrm{d}x = \frac{\sqrt{2\pi}^n}{\sqrt{\det(K)}}\left\{F(\partial_\xi)e^{\frac{1}{2}\langle \xi,K^{-1}\xi\rangle}\right\}\Big|_{\xi=0}.$$

Proof Consider the Laplacian transform,

$$\int_{\mathbb{R}^n} e^{\langle x,\xi\rangle}F(x)e^{-\frac{1}{2}\langle x,Kx\rangle}\mathrm{d}x = \int F(\partial_\xi)e^{\langle x,\xi\rangle}e^{-\frac{1}{2}\langle x,Kx\rangle}\mathrm{d}x.$$

Since $e^{-\frac{1}{2}\langle x,Kx\rangle}$ is a Schwartz function, with decaying speed faster than $e^{\langle x,\xi\rangle}$. By dominated convergence theorem, we can interchange the integral and differential.

$$\begin{aligned}\int_{\mathbb{R}^n} e^{\langle x,\xi\rangle}F(x)e^{-\frac{1}{2}\langle x,Kx\rangle}\mathrm{d}x &= F(\partial_\xi)\left\{\int_{\mathbb{R}^n} e^{\langle x,\xi\rangle}e^{-\frac{1}{2}\langle x,Kx\rangle}\mathrm{d}x\right\}\\ &= F(\partial_\xi)\left\{e^{\frac{1}{2}\langle \xi,K^{-1}\xi\rangle}\int_{\mathbb{R}^n} e^{-\frac{1}{2}\langle (x-K^{-1}\xi),K(x-K^{-1}\xi)\rangle}\mathrm{d}x\right\}.\end{aligned}$$

From Lemma A3 and Lemma A4, we get

$$\begin{aligned}\int_{\mathbb{R}^n} F(x)e^{-\frac{1}{2}\langle x,Kx\rangle}\mathrm{d}x &= \int_{\mathbb{R}^n} e^{\langle x,\xi\rangle}F(x)e^{-\frac{1}{2}\langle x,Kx\rangle}\mathrm{d}x\Big|_{\xi=0}\\ &= \frac{\sqrt{2\pi}^n}{\sqrt{\det(K)}}\left\{F(\partial_\xi)e^{\frac{1}{2}\langle \xi,K^{-1}\xi\rangle}\right\}\Big|_{\xi=0}.\end{aligned}$$

□

Appendix B. Detailed Computation for m_1

In this section, we give the detailed computation for coefficient m_1. To compute f_0, f_1, f_2, take $l = 0, 1, 2$ in (30) and follow the process in Page 14, 15. It is not hard to derive,

$$\begin{aligned} f_0(\theta) &= \frac{1}{2\pi}\int \left(1+\frac{1}{N}p_0(x_1,x_2)\right)e^{-\frac{1}{2}\langle x,Kx\rangle}\mathrm{d}x + O(N^{-2}) \\ &= \frac{1}{\sqrt{\det(K)}}\left(1+\frac{1}{N}p_0(\partial_{\xi_1},\partial_{\xi_2})e^{\frac{1}{2}\langle \xi,K^{-1}\xi\rangle}\right)\Big|_{\xi=0} + O(N^{-2}), \end{aligned} \tag{47}$$

where

$$\begin{aligned} p_0(x_1,x_2) &= \frac{\theta^2}{2}\sum_{i=1}^{N}\frac{(\mathrm{i}(1-v+\lambda_i)x_1+(1+v-\lambda_i)x_2)^4}{N(1-2\theta\lambda_i+2\theta v)^4} \\ &\quad + \frac{\theta^3}{9}\left\{\sum_{i=1}^{N}\frac{(\mathrm{i}(1-v+\lambda_i)x_1+(1+v-\lambda_i)x_2)^3}{N(1-2\theta\lambda_i+2\theta v)^3}\right\}^2. \end{aligned}$$

And for f_1 and f_2,

$$\begin{aligned} f_i(t) &= \frac{1}{N}\frac{1}{2\pi}\int p_i(x_1,x_2)e^{-\frac{1}{2}\langle x,\tilde{K}x\rangle}\mathrm{d}x + O(N^{-2}) \\ &= \frac{1}{N}\frac{1}{\sqrt{\det(\tilde{K})}}p_i(\partial_{\xi_1},\partial_{\xi_2})e^{\frac{1}{2}\langle \xi,\tilde{K}^{-1}\xi\rangle}\Big|_{\xi=0} + O(N^{-2}), \end{aligned} \tag{48}$$

where

$$\begin{aligned} p_1(x_1,x_2) &= \frac{2t^2}{3}\left\{\sum_{i=1}^{N}\frac{(\mathrm{i}(1-v+\lambda_i)x_1+(1+v-\lambda_i)x_2)}{N(1-2\theta\lambda_i+2\theta v)^2}\right\} \\ &\quad \left\{\sum_{i=1}^{N}\frac{(\mathrm{i}(1-v+\lambda_i)x_1+(1+v-\lambda_i)x_2)^3}{N(1-2\theta\lambda_i+2\theta v)^3}\right\} \\ &\quad + 2t\sum_{i=1}^{N}\frac{(\mathrm{i}(1-v+\lambda_i)x_1+(1+v-\lambda_i)x_2)^2}{N(1-2\theta\lambda_i+2\theta v)^3}, \\ p_2(x_1,x_2) &= \sum_{i=1}^{N}\frac{2}{N(1-2\theta\lambda_i+2\theta v)^2} + 2t\left\{\sum_{i=1}^{N}\frac{(\mathrm{i}(1-v+\lambda_i)x_1+(1+v-\lambda_i)x_2)}{N(1-2\theta\lambda_i+2\theta v)^2}\right\}^2. \end{aligned}$$

For integral (47) and (48), it is merely symbolic computation, and we can easily do it by some mathematic software, like Mathematica, then get explicit formula for f_0, f_1 and f_2.

$$f_0(\theta) = \frac{1}{\sqrt{A_2}} + \frac{1}{N}\frac{1}{\sqrt{A_2}}\left(\frac{7}{6} - \frac{3}{A_2} + \frac{2A_3}{A_2^2} - \frac{5}{3}\frac{A_3^2}{A_2^3} + \frac{3}{2}\frac{A_4}{A_2^2}\right) + O(N^{-2}),$$

$$f_1(t) = \frac{1}{N}\left\{\frac{2t(t-\theta)(2tA_2^2 - A_3(t-2\theta) - A_2(t+2\theta))}{\theta^3(\det(\tilde{K}))^{\frac{5}{2}}}\right\} + O(N^{-2}),$$

$$f_2(t) = \frac{1}{N}\frac{2A_2}{(\det(\tilde{K}))^{\frac{3}{2}}} + O(N^{-2}).$$

Actually in this way, we can obtain any higher expansion terms of the spherical integral.

References

1. Anderson, G.W., Guionnet, A., Zeitouni, O.: An Introduction to Random Matrices, volume 118 of Cambridge Studies in Advanced Mathematics. Cambridge University Press, Cambridge (2010)
2. Duistermaat, J.J., Heckman, G.J.: On the variation in the cohomology of the symplectic form of the reduced phase space. Invent. Math. **69**(2), 259–268 (1982)
3. Gorin, V., Panova, G.: Asymptotics of symmetric polynomials with applications to statistical mechanics and representation theory. Ann. Probab. **43**(6), 3052–3132 (2015)
4. Guionnet, A., Maï da, M.: A Fourier view on the R-transform and related asymptotics of spherical integrals. J. Funct. Anal. **222**(2), 435–490 (2005)
5. Guionnet, A., Zeitouni, O.: Large deviations asymptotics for spherical integrals. J. Funct. Anal. **188**(2), 461–515 (2002)
6. Guionnet, A., Zeitouni, O.: Addendum to: "Large deviations asymptotics for spherical integrals" [J. Funct. Anal. **188** (2002), no. 2, 461–515; mr1883414]. J. Funct. Anal. **216**(1), 230–241 (2004)
7. Itzykson, C., Zuber, J.B.: The planar approximation. II. J. Math. Phys. **21**(3), 411–421 (1980)
8. Maassen, H.: Addition of freely independent random variables. J. Funct. Anal. **106**(2), 409–438 (1992)
9. Mehta, M.L.: Random Matrices, volume 142 of Pure and Applied Mathematics (Amsterdam), third edn. Elsevier/Academic Press, Amsterdam (2004)
10. Voiculescu, D.: Addition of certain noncommuting random variables. J. Funct. Anal. **66**(3), 323–346 (1986)