Characteristic polynomials of sparse non-Hermitian random matrices

Ievgenii Afanasiev* Tatyana Shcherbina †

Abstract

We consider the asymptotic local behavior of the second correlation functions of the characteristic polynomials of sparse non-Hermitian random matrices X_n whose entries have the form $x_{jk} = d_{jk}w_{jk}$ with iid complex standard Gaussian w_{jk} and normalised iid Bernoulli(p) d_{jk}. It is shown that, as $p \to \infty$, the local asymptotic behavior of the second correlation function of characteristic polynomials near $z_0 \in \mathbb{C}$ coincides with those for Ginibre ensemble: it converges to a determinant with Ginibre kernel in the bulk $|z_0| < 1$, and it is factorized if $|z_0| > 1$. For the finite $p > 0$, the behavior is different and exhibits the transition between different regimes depending on values p and $|z_0|^2$.

1 Introduction

Introduce $n \times n$ non-Hermitian random matrices

$$X_n = (x_{jk})_{j,k=1}^n, \tag{1.1}$$

whose entries can be written in the form

$$x_{jk} = d_{jk}w_{jk}$$

with i.i.d. complex random variables w_{jk} such that

$$\mathbf{E}\{w_{jk}\} = \mathbf{E}\{w_{jk}^2\} = 0, \quad \mathbf{E}\{|w_{jk}|^2\} = 1, \tag{1.2}$$

and normalized i.i.d. Bernoulli(p), $0 < p \le n$, indicators d_{jk} independent of $\{w_{jk}\}$, i.e.

$$d_{jk} = \frac{1}{\sqrt{p}} \begin{cases} 1, & \text{with probability } \frac{p}{n}, \\ 0, & \text{with probability } 1 - \frac{p}{n}. \end{cases} \tag{1.3}$$

We will refer to this ensemble as to *sparse non-Hermitian random matrices*. Parameter p can be fixed or may depend on n. Clearly, this ensemble interpolates between non-Hermitian random matrices with iid entries for $p = n$ and very sparse matrices (which have, on average, p non-zero entries in a row) for a fixed p.

The limiting empirical spectral distributions for such matrices with p growing together with n was excessively studied in the mathematical literature (see, e.g., [28], [14], [30], [7] and

---

*B. Verkin Institute for Low Temperature Physics and Engineering of the National Academy of Sciences of Ukraine, Kharkiv, Ukraine; e-mail: afanasiev@ilt.kharkov.ua. The author is partially supported by the Grant EFDS-FL2-08 of the found The European Federation of Academies of Sciences and Humanities (ALLEA).

†Department of Mathematics, University of Wisconsin - Madison, USA, e-mail: tshcherbyna@wisc.edu. This material is based upon work supported in part by Alfred P. Sloan Foundation grant FG-2022-18916 and the National Science Foundation under grant DMS-2346379

references therein) with an optimal result obtained by Rudelson and Tikhomirov in [21]: as soon as $p \to \infty$ together with n, the empirical spectral distribution of sparse non-Hermitian random matrices converges weakly in probability to the circular law, i.e. to the uniform distribution on a unit disk. The existence of the limiting empirical spectral distributions for finite $p > 0$ (in this case it is not a circular law anymore) was obtained very recently in [22].

The local eigenvalue statistics of (1.1) is much less studied. For a non-Hermitian matrices with iid random entries (i.e. $p = n$ case) the local eigenvalue statistics in the bulk and at the edge of the spectrum coincide with those of the Ginibre ensemble, i.e. matrices with iid Gaussian entries. This is known as the *universality of non-Hermitian random matrices* (see [29], [11],[17], [20], [12], [4] and references therein). For the sparse case, the universality at the edge of the spectrum for $n^\alpha \ll p \le \frac{1}{2}$ was obtained recently in [15]. The bulk universality with $p \ll n$ is still an open question.

In this paper we are going to study the local behavior of correlation functions of characteristic polynomials. For the non-Hermitian random matrices it can be defined as

$$\mathsf{f}_k(z_1, \dots, z_k) = \mathbf{E}\Big\{ \prod_{s=1}^{k} \det(X_n - z_s) \det(X_n - z_s)^* \Big\}. \tag{1.4}$$

We are interested in the asymptotic behavior of f_2 for matrices (1.1) as $n \to \infty$ and

$$z_j = z_0 + \tfrac{\zeta_j}{\sqrt{n}}, \; j = 1, 2. \tag{1.5}$$

Characteristic polynomials are the objects of independent interest because of their connections to the number theory, quantum chaos, integrable systems, combinatorics, representation theory and others. In additional, although f_k is not a local object in terms of eigenvalue statistics, it is also expected to be universal in some sense. In particular, it was proved in [2] (see also [6] for the Gaussian (Ginibre) case) that for non-Hermitian random matrices H with iid complex entries with mean zero, variance one, and $2k$ finite moments for any $z_j = z_0 + \zeta_j/\sqrt{n}$, $j = 1, .., k$ and $|z_0| < 1$ we get

$$\lim_{n\to\infty} n^{-\frac{k^2-k}{2}} \frac{\mathsf{f}_k(z_1, \dots, z_k)}{\prod_j \mathsf{f}_1(z_j)} = C_k \frac{\det(K(\zeta_i, \zeta_j))_{i,j}^k}{|\Delta(\zeta)|^2}, \tag{1.6}$$

where

$$K(w_1, w_2) = e^{-|w_1|^2/2 - |w_2|^2/2 + w_1 \bar{w}_2}, \tag{1.7}$$

$\Delta(\zeta)$ is a Vandermonde determinant of $\zeta_1, \dots, \zeta_k$, and C_k is constant depending only on the fourth cumulant $\kappa_4 = \mathbf{E}[|H_{11}|^4] - 2$ of the elements distribution, but not on the higher moments. In particular, this means that the local limiting behavior (1.6) of non-Hermitian matrices with iid entries coincides with those for the Ginibre ensemble as soon as the elements distribution has four Gaussian moments, i.e. the local behavior of the correlation functions of characteristic polynomials also exhibits a certain form of universality. Similar results were obtained for many classical Hermitian random matrix ensembles (see, e.g., [9], [10], [27],[24], [25],[1], [26], etc.)

Notice that the local asymptotic behavior of characteristic polynomials of the sparse *Hermitian* random matrices was obtained in [1]. In particular, it was shown that while for $p \to \infty$ the behavior coincides with that for Gaussian Unitary Ensemble (GUE), i.e. Hermitian matrices with iid (up to the symmetry) Gaussian entries. For the finite p the local asymptotic behavior of the second correlation function of characteristic polynomials of sparse Hermitian random matrices demonstrates the transition: when $p < 2$ the second correlation function of

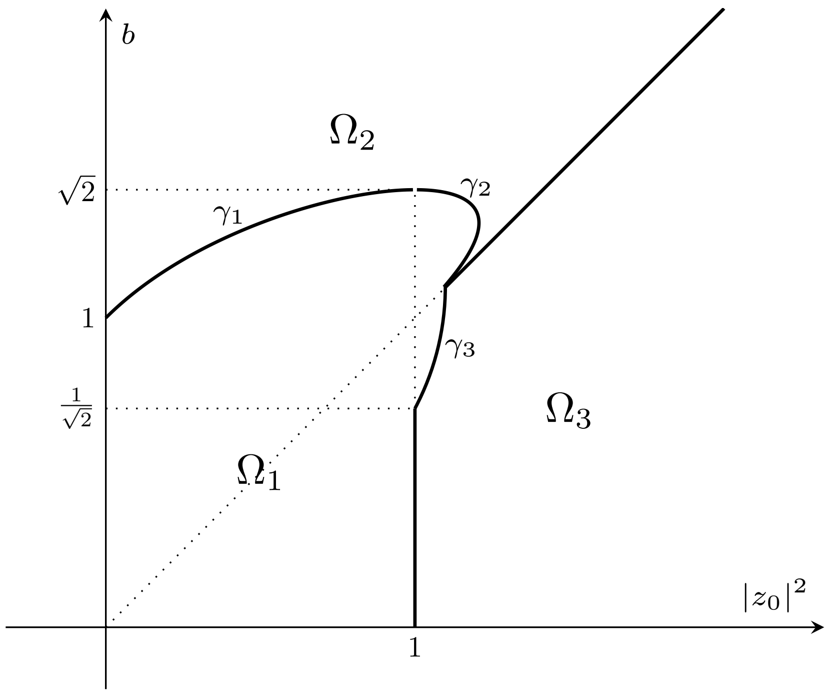

Figure 1: phase diagram of the three different types of behavior of f_2

characteristic polynomials factorizes in the limit $n \to \infty$, while for $p > 2$ there appears an interval $(\lambda_-(p), \lambda_+(p))$ such that inside $(\lambda_-(p), \lambda_+(p))$ the second correlation function behaves like that for GUE, while outside the interval the second correlation function is still factorized.

The goal of the current paper is to establish similar result for the sparse non-Hermitian matrices (1.1). Define

$$b = \sqrt{\tfrac{2(n-p)}{np}} \tag{1.8}$$

with p of (1.3). Notice that if p is finite, then

$$b = \sqrt{\frac{2}{p}} + O(n^{-1}), \quad n \to \infty, \tag{1.9}$$

and

$$b = O(p^{-1/2}), \quad n \to \infty, \tag{1.10}$$

if $p \to \infty$ but $p < (1-\varepsilon)n$ for some $\varepsilon > 0$.

The main result of the paper is the following theorem:

Theorem 1. *Let X_n be the sparse non-Hermitian complex random matrices (1.1) with the standard complex Gaussian w_{jk} and finite fixed $p > 0$. Then for the second correlation function of characteristic polynomials f_2 of (1.4) with z_1, z_2 from (1.5) we have*

(i) if $(b, |z_0|^2) \in \Omega_1$, then

$$\lim_{n\to\infty} \frac{\mathsf{f}_2(z_1, z_2)}{\sqrt{\mathsf{f}_2(z_1, z_1)\mathsf{f}_2(z_2, z_2)}} = e^{-\gamma(\Re(\bar{z}_0(\zeta_1-\zeta_2)))^2} \frac{\det(K(\sqrt{\beta}\zeta_i, \sqrt{\beta}\zeta_j))^2_{i,j=1}}{\beta\,|\triangle(\zeta)|^2}, \tag{1.11}$$

where K is defined in (1.7) and $\beta \in [0,1]$ is a solution (the largest one, if several) to the equation

$$p\beta - p + 2 = p\,|z_0|^2\,(1-\beta)^2(2 - p\,|z_0|^2\,(1-\beta)),$$

and $\gamma > 0$ is a certain constant depending on p and $|z_0|^2$ (see (3.71)).

(ii) if $(b, |z_0|^2) \in \Omega_2$, then

$$\lim_{n\to\infty} \frac{\mathsf{f}_2(z_1, z_2)}{\sqrt{\mathsf{f}_2(z_1, z_1)\mathsf{f}_2(z_2, z_2)}} = e^{-p(\Re(\bar{z}_0(\zeta_1-\zeta_2)))^2}. \tag{1.12}$$

(iii) if $(b, |z_0|^2) \in \Omega_3$, then

$$\lim_{n\to\infty} \frac{\mathsf{f}_2(z_1, z_2)}{\sqrt{\mathsf{f}_2(z_1, z_1)\mathsf{f}_2(z_2, z_2)}} = 1. \tag{1.13}$$

The domains Ω_i, $i = 1, 2, 3$ are shown in Figure 1. Here γ_1 is the curve

$$|z_0|^2 = \frac{b^2 - b\sqrt{2-b^2}}{2}, \quad b \in [1, \sqrt{2}],$$

and γ_2, γ_3 are certain explicit curves that will be defined later (see Lemma 4 and Lemma 7).

Remark 1. The asymptotic behavior of the second correlation function is given by

(i) for $(b, |z_0|^2) \in \Omega_1$

$$\mathsf{f}_2(z_1, z_2) = C(\zeta_1, \zeta_2)n^2 \exp\left\{C_1 n + C_2(\zeta_1, \zeta_2)\sqrt{n}\right\} \frac{\det\left\{e^{\beta\zeta_j\bar{\zeta}_k}\right\}_{j,k=1}^2}{\beta\,|\triangle(\zeta)|^2}(1 + o(1)), \tag{1.14}$$

(ii) for $(b, |z_0|^2) \in \Omega_2$

$$\begin{aligned}\mathsf{f}_2(z_1, z_2) = C^{II}\sqrt{n}\exp&\left\{n\left(-1 + \frac{|z_0|^4}{b^2} + \log b^2\right) + \sqrt{n}\frac{2\,|z_0|^2}{b^2}\Re(\bar{z}_0(\zeta_1 + \zeta_2))\right\}\\ &\times \exp\left\{\frac{p}{2}\left[2(\Re(\bar{z}_0(\zeta_1 + \zeta_2)))^2 - \Re(\bar{z}_0^2(\zeta_1^2 + \zeta_2^2))\right]\right\}(1 + o(1)),\end{aligned} \tag{1.15}$$

(iii) for $(b, |z_0|^2) \in \Omega_3$

$$\begin{aligned}\mathsf{f}_2(z_1, z_2) = &\frac{(1 + o(1))}{\left(1 - \frac{1}{|z_0|^2}\right)^4\left(1 - \frac{b^2}{|z_0|^4}\right)} \exp\left\{n \log|z_0|^4 + \frac{2\sqrt{n}}{|z_0|^2}\Re(\bar{z}_0(\zeta_1 + \zeta_2))\right\}\\ &\times \exp\left\{-\frac{1}{|z_0|^4}\Re(\bar{z}_0^2(\zeta_1^2 + \zeta_2^2))\right\},\end{aligned} \tag{1.16}$$

where C, C_1, C_2, C^{II} will be defined later (see (3.73), (3.74), (3.75), (3.78)).

In the case $p \to \infty$ we can also get

Theorem 2. *Let X_n be the sparse non-Hermitian complex random matrices (1.1) with the standard complex Gaussian w_{jk} and $p \to \infty$ but $p < (1 - \varepsilon)n$ for some $\varepsilon > 0$. Then the limiting behavior of second correlation function of characteristic polynomials f_2 of (1.4) with z_1, z_2 from (1.5) coincides with those for Ginibre ensemble (1.6) in the bulk, i.e. if $|z_0| < 1$. If $|z_0| > 1$, then (1.13) holds.*

In order to prove Theorems 1-2 we are going to apply the supersymmetry techniques (SUSY). SUSY techniques is based on the representation of the determinant as an integral (formal) over the Grassmann variables, which allows to obtain the integral representation for the main spectral characteristics of random matrices (such as density of states, correlation functions, characteristic polynomial, etc.) as an integral containing both complex and Grassmann (anticommuting) variables. Although at a heuristic level SUSY was actively used in theoretical physics literature (see e.g. reviews [13], [19]) for several decades, the rigorous analysis of such integrals poses a very serious challenge. However, the method was successfully applied to the rigorous study of local regime of some random matrix ensembles, including the most successful applications to the Gaussian Hermitian random band matrices (see [23] and reference therein). The asymptotic behavior of characteristic polynomials is known to be especially convenient for the SUSY approach and were successfully studied by the techniques for many Hermitian (see, e.g., [9], [10], [24], [25],[1], [26] etc.) and some non-Hermitian (see, e.g., [5], [2], [3]) ensembles.

The paper organized as follows. In Section 2 we give the brief outline of SUSY techniques and obtain the SUSY integral representation of f_2 of (1.4). Section 3 is devoted to the proof of Theorem 1 by performing the saddle-point analysis of the obtained representation: in Section 3.1 we determine the main saddle-points that can give the leading contribution to the integral; in Section 3.2 we determine the domain of domination of each of the obtained saddle-point; finally, the contribution of each of the saddle-points is computed in Section 3.3.

2 Integral representation for f_2

In this section we obtain a convenient integral representation for the correlation function of characteristic polynomials f_2 defined by (1.4).

Proposition 1. *Let X_n be defined by (1.1) and (1.2). Then the second correlation function of the characteristic polynomials f_2 defined by (1.4) can be represented in the following form*

$$\mathsf{f}_2(z_1, z_2) = \left(\frac{n}{\pi}\right)^5 \int e^{n\hat{f}(Q,v)} dQ\, d\bar{v}\, dv, \tag{2.1}$$

where Q is a complex 2×2 matrix, $v \in \mathbb{C}$,

$$dQ = \prod_{j,k=1}^{2} d\bar{Q}_{jk}\, dQ_{kj}$$

and

$$\hat{f}(Q, v) = -\operatorname{tr} Q^* Q - |v|^2 + \log h(Q, v); \tag{2.2}$$

$$h(Q, v) = \det A + bv \det Q^* + b\bar{v} \det Q + b^2 |v|^2 ; \tag{2.3}$$

$$A = A(Q) = \begin{pmatrix} -Z & Q \\ -Q^* & -Z^* \end{pmatrix}, \quad Z = \begin{pmatrix} z_1 & 0 \\ 0 & z_2 \end{pmatrix} \tag{2.4}$$

with b of (1.8).

2.1 Proof of Proposition 1

To derive the integral representation of f_2 we will use the SUSY. For the reader convenience, we start with a very brief outline of the basic formulas of SUSY techniques we need. More detailed information about the techniques and its applications to random matrix theory can be found, e.g., in [13] or [19].

2.2 SUSY techniques: basic formulas

Let us consider two sets of formal variables $\{\psi_j\}_{j=1}^n, \{\overline{\psi}_j\}_{j=1}^n$, which satisfy the anticommutation relations

$$\psi_j\psi_k + \psi_k\psi_j = \overline{\psi}_j\psi_k + \psi_k\overline{\psi}_j = \overline{\psi}_j\overline{\psi}_k + \overline{\psi}_k\overline{\psi}_j = 0, \quad j,k = 1,\dots,n. \tag{2.5}$$

Note that this definition implies $\psi_j^2 = \overline{\psi}_j^2 = 0$. These two sets of variables $\{\psi_j\}_{j=1}^n$ and $\{\overline{\psi}_j\}_{j=1}^n$ generate the Grassmann algebra $\mathfrak{A}$. Taking into account that $\psi_j^2 = 0$, we have that all elements of $\mathfrak{A}$ are polynomials of $\{\psi_j\}_{j=1}^n$ and $\{\overline{\psi}_j\}_{j=1}^n$ of degree at most one in each variable. We can also define functions of the Grassmann variables. Let χ be an element of $\mathfrak{A}$, i.e.

$$\chi = a + \sum_{j=1}^n (a_j\psi_j + b_j\overline{\psi}_j) + \sum_{j\neq k}(a_{j,k}\psi_j\psi_k + b_{j,k}\psi_j\overline{\psi}_k + c_{j,k}\overline{\psi}_j\overline{\psi}_k) + \dots. \tag{2.6}$$

For any sufficiently smooth function f we define by $f(\chi)$ the element of $\mathfrak{A}$ obtained by substituting $\chi - a$ in the Taylor series of f at the point a:

$$f(\chi) = a + f'(a)(\chi - a) + \frac{f''(a)}{2!}(\chi - a)^2 + \dots$$

Since χ is a polynomial of $\{\psi_j\}_{j=1}^n$, $\{\overline{\psi}_j\}_{j=1}^n$ of the form (2.6), according to (2.5) there exists such l that $(\chi - a)^l = 0$, and hence the series terminates after a finite number of terms and so $f(\chi) \in \mathfrak{A}$.

Following Berezin [8], we define the operation of integration with respect to the anticommuting variables in a formal way:

$$\int d\,\psi_j = \int d\,\overline{\psi}_j = 0, \quad \int \psi_j d\,\psi_j = \int \overline{\psi}_j d\,\overline{\psi}_j = 1,$$

and then extend the definition to the general element of $\mathfrak{A}$ by the linearity. A multiple integral is defined to be a repeated integral. Assume also that the "differentials" $d\,\psi_j$ and $d\,\overline{\psi}_k$ anticommute with each other and with the variables ψ_j and $\overline{\psi}_k$. Thus, according to the definition, if

$$f(\psi_1,\dots,\psi_k) = p_0 + \sum_{j_1=1}^k p_{j_1}\psi_{j_1} + \sum_{j_1<j_2} p_{j_1,j_2}\psi_{j_1}\psi_{j_2} + \dots + p_{1,2,\dots,k}\psi_1\dots\psi_k,$$

then

$$\int f(\psi_1,\dots,\psi_k)d\,\psi_k \dots d\,\psi_1 = p_{1,2,\dots,k}.$$

The key formulas we need in this subsection are the well-known complex Gaussian integration formula for a complex n-dimensional vector $\boldsymbol{t}$

$$\int_{\mathbb{C}^n} \exp\{-\boldsymbol{t}^*B\boldsymbol{t} - \boldsymbol{t}^*\boldsymbol{h}_2 - \boldsymbol{h}_1^*\boldsymbol{t}\}\, d\boldsymbol{t}^*d\boldsymbol{t} = \pi^n \mathrm{det}^{-1}B\exp\{\boldsymbol{h}_1^*B^{-1}\boldsymbol{h}_2\}, \tag{2.7}$$

valid for any positive definite matrix B and its analog for Grassmann n-dimensional vector $\boldsymbol{\tau}$ (see [13], Ch 2.4):

$$\int \exp\{-\boldsymbol{\tau}^+A\boldsymbol{\tau} - \boldsymbol{\tau}^+\boldsymbol{v}_2 - \boldsymbol{v}_1^+\boldsymbol{\tau}\}\, d\boldsymbol{\tau}^+d\boldsymbol{\tau} = \det B\exp\{\boldsymbol{v}_1^+B^{-1}\boldsymbol{v}_2\} \tag{2.8}$$

valid for an arbitrary complex matrix B. Here and below $\tau^+ = (\bar{\tau}_1, \ldots, \bar{\tau}_n)$.

We will need also the following Hubbard-Stratonovich transformation formulas which basically is an employment of (2.7) in the reverse direction:

$$e^{ab} = \pi^{-1} \int_{\mathbb{C}} e^{a\bar{u}+bu-\bar{u}u} d\bar{u}\, du. \tag{2.9}$$

Here a, b can be complex numbers or sums of the products of even numbers of Grassmann variables (i.e. commuting elements of Grassmann algebra).

2.3 Integral representation

Rewrite the expression (1.4) for f_2 using (2.8)

$$\mathsf{f}_2(z_1,z_2) = \mathbf{E}\left\{ \int \exp\left\{ -\sum_{j=1}^{2} \boldsymbol{\phi}_j^+ (X_n - z_j)\, \boldsymbol{\phi}_j - \sum_{j=1}^{2} \boldsymbol{\theta}_j^+ (X_n - z_j)^*\, \boldsymbol{\theta}_j \right\} d\Phi d\Theta \right\},$$

where $\boldsymbol{\phi}_j$, $\boldsymbol{\theta}_j$, $j = 1,2$ are n-dimensional vectors with components ϕ_{kj} and θ_{kj} respectively, Θ and Φ are $n \times 2$ matrices composed of columns $\boldsymbol{\theta}_1, \boldsymbol{\theta}_2$ and $\boldsymbol{\phi}_1, \boldsymbol{\phi}_2$, and

$$d\Phi = \prod_{j=1}^{2} d\boldsymbol{\phi}_j^+ d\boldsymbol{\phi}_j, \quad d\Theta = \prod_{j=1}^{2} d\boldsymbol{\theta}_j^+ d\boldsymbol{\theta}_j.$$

Denoting $\boldsymbol{\varphi}_k = (\phi_{k1}, \phi_{k2})^t$, $\boldsymbol{\vartheta}_k = (\theta_{k1}, \theta_{k2})^t$ we can rewrite the previous formula as

$$\begin{aligned}\mathsf{f}_2(z_1,z_2) = \mathbf{E}\Bigg\{ \int \exp\Bigg\{ &\sum_{k=1}^{n} \boldsymbol{\varphi}_k^+ Z \boldsymbol{\varphi}_k + \sum_{k=1}^{n} \boldsymbol{\vartheta}_k^+ Z^* \boldsymbol{\vartheta}_k \\ &+ \sum_{k,l=1}^{n} (\Phi\Phi^+)_{lk} x_{kl} + \sum_{k,l=1}^{n} (\Theta\Theta^+)_{kl} \bar{x}_{kl} \Bigg\} d\Phi d\Theta \Bigg\},\end{aligned} \tag{2.10}$$

where Z is defined in (2.4) and Θ^+, Φ^+ are $2 \times n$ matrices composed of rows $\boldsymbol{\theta}_1^+, \boldsymbol{\theta}_2^+$ and $\boldsymbol{\phi}_1^+, \boldsymbol{\phi}_2^+$ respectively.

Let us introduce a notation

$$\psi\,(t_1, t_2) := \mathbf{E}\left\{ e^{t_1 x_{11} + t_2 \bar{x}_{11}} \right\}.$$

Then the expectation in (2.10) can be written in the following form

$$\begin{aligned}\mathsf{f}_2(z_1,z_2) &= \int \prod_{k,l=1}^{n} \psi\left((\Phi\Phi^+)_{lk}, (\Theta\Theta^+)_{kl} \right) \\ &\qquad \times \exp\left\{ \sum_{k=1}^{n} \boldsymbol{\varphi}_k^+ Z \boldsymbol{\varphi}_k + \sum_{k=1}^{n} \boldsymbol{\vartheta}_k^+ Z^* \boldsymbol{\vartheta}_k \right\} d\Phi d\Theta \\ &= \int \exp\Bigg\{ \sum_{k=1}^{n} \boldsymbol{\varphi}_k^+ Z \boldsymbol{\varphi}_k + \sum_{k=1}^{n} \boldsymbol{\vartheta}_k^+ Z^* \boldsymbol{\vartheta}_k \\ &\qquad + \sum_{k,l=1}^{n} \log \psi\left((\Phi\Phi^+)_{lk}, (\Theta\Theta^+)_{kl} \right) \Bigg\} d\Phi d\Theta.\end{aligned}$$

Expansion of $\log \Phi$ into series gives us

$$\mathsf{f}_2(z_1, z_2) = \int \exp\left\{ \sum_{k=1}^{n} \boldsymbol{\varphi}_k^+ Z \boldsymbol{\varphi}_k + \sum_{k=1}^{n} \boldsymbol{\vartheta}_k^+ Z^* \boldsymbol{\vartheta}_k + \sum_{k,l=1}^{n} \sum_{p,s=0}^{2} \frac{\kappa_{p,s}}{p!s!} \left((\Phi\Phi^+)_{lk}\right)^p \left((\Theta\Theta^+)_{kl}\right)^s \right\} d\Phi d\Theta, \tag{2.11}$$

with

$$\kappa_{p,s} = \frac{\partial^{p+s}}{\partial^p y_1 \partial^s y_2} \log \psi(y_1, y_2)\Big|_{y_1=y_2=0}.$$

Using (1.1)–(1.3), one can compute

$$\begin{aligned} &\kappa_{0,0} = 0; \\ &\kappa_{1,0} = \overline{\kappa_{0,1}} = \mathbf{E}\{x_{11}\} = 0; \\ &\kappa_{2,0} = \overline{\kappa_{0,2}} = \mathbf{E}\{x_{11}^2\} - \mathbf{E}^2\{x_{11}\} = 0; \\ &\kappa_{1,1} = \mathbf{E}\{|x_{11}|^2\} - |\mathbf{E}\{x_{11}\}|^2 = \frac{1}{n}; \\ &\kappa_{2,2} = \mathbf{E}\{|x_{11}|^4\} - 2\,\mathbf{E}^2\{|x_{11}|^2\} = \frac{2(n-p)}{pn^2}. \end{aligned} \tag{2.12}$$

Let us transform the terms in the exponent again. Denote

$$\mathcal{I}_{2,1} = \{1, 2\}, \quad \mathcal{I}_{2,2} = \{\{1, 2\}\} \tag{2.13}$$

For non-zero terms with $p = s = 1$ or $p = s = 2$ one can write

$$\begin{aligned} &\sum_{k,l=1}^{n} \left((\Phi\Phi^+)_{lk}\right)^p \left((\Theta\Theta^+)_{kl}\right)^s \\ &= \sum_{k,l=1}^{n} \left(\sum_{j=1}^{2} \phi_{lj}\overline{\phi_{kj}}\right)^p \left(\sum_{j=1}^{2} \theta_{kj}\overline{\theta_{lj}}\right)^s = p!s! \sum_{k,l=1}^{n} \sum_{\substack{\alpha\in\mathcal{I}_{2,p} \\ \beta\in\mathcal{I}_{2,s}}} \prod_{q=1}^{p} \phi_{l\alpha_q}\overline{\phi_{k\alpha_q}} \prod_{r=1}^{s} \theta_{k\beta_r}\overline{\theta_{l\beta_r}} \\ &= (-1)^{p^2} p!s! \sum_{k,l=1}^{n} \sum_{\substack{\alpha\in\mathcal{I}_{2,p} \\ \beta\in\mathcal{I}_{2,s}}} \prod_{r=1}^{s} \theta_{k\beta_r} \prod_{q=1}^{p} \overline{\phi_{k\alpha_q}} \prod_{q=1}^{p} \phi_{l\alpha_q} \prod_{r=1}^{s} \overline{\theta_{l\beta_r}} \\ &= p!s! \sum_{\substack{\alpha\in\mathcal{I}_{2,p} \\ \beta\in\mathcal{I}_{2,s}}} \left(\sum_{k=1}^{n} (-1)^p \prod_{r=1}^{s} \theta_{k\beta_r} \prod_{q=1}^{p} \overline{\phi_{k\alpha_q}}\right) \left(\sum_{k=1}^{n} \prod_{q=1}^{p} \phi_{k\alpha_q} \prod_{r=1}^{s} \overline{\theta_{k\beta_r}}\right). \end{aligned} \tag{2.14}$$

At this point the Hubbard–Stratonovich transformation (2.9) is applied. It yields for $p = s = 1$ or $p = s = 2$

$$\begin{aligned} &\exp\left\{\kappa_{p,s}\left(\sum_{k=1}^{n} (-1)^p \prod_{r=1}^{s} \theta_{k\beta_r} \prod_{q=1}^{p} \overline{\phi_{k\alpha_q}}\right)\left(\sum_{k=1}^{n} \prod_{q=1}^{p} \phi_{k\alpha_q} \prod_{r=1}^{s} \overline{\theta_{k\beta_r}}\right)\right\} \\ &\quad = \frac{n}{\pi} \int \exp\left\{ -\sum_{k=1}^{n} \tilde{y}_{\beta\alpha}^{(k,p,s)} q_{\alpha\beta}^{(p,s)} - \sum_{k=1}^{n} \bar{q}_{\alpha\beta}^{(p,s)} y_{\alpha\beta}^{(k,p,s)} - n\left|q_{\alpha\beta}^{(p,s)}\right|^2 \right\} \\ &\quad \times d\bar{q}_{\alpha\beta}^{(p,s)} dq_{\alpha\beta}^{(p,s)}, \end{aligned} \tag{2.15}$$

where

$$\begin{aligned}\tilde{y}^{(k,p,s)}_{\beta\alpha} &= \sqrt{n\kappa_{p,s}}(-1)^p \prod_{r=1}^{s} \theta_{k\beta_r} \prod_{q=1}^{p} \overline{\phi_{k\alpha_q}};\\ y^{(k,p,s)}_{\alpha\beta} &= \sqrt{n\kappa_{p,s}} \prod_{q=1}^{p} \phi_{k\alpha_q} \prod_{r=1}^{s} \overline{\theta_{k\beta_r}}.\end{aligned} \tag{2.16}$$

Then the combination of (2.11), (2.14) and (2.15) gives

$$\mathsf{f}_2(z_1, z_2) = \left(\frac{n}{\pi}\right)^5 \int e^{-\operatorname{tr} Q^*Q - |v|^2} \prod_{k=1}^{n} \mathsf{j}_k \, dQ^* dQ dv^* dv \tag{2.17}$$

where

$$\mathsf{j}_k = \int \exp\left\{a_{k,2} + a_{k,4}\right\} d\boldsymbol{\varphi}_k^+ d\boldsymbol{\varphi}_k d\boldsymbol{\vartheta}_k^+ d\boldsymbol{\vartheta}_k, \tag{2.18}$$

$$a_{k,2} = -\left(\operatorname{tr} \tilde{Y}_{k,1} Q + \operatorname{tr} Q^* Y_{k,1}\right) + \boldsymbol{\varphi}_k^+ Z \boldsymbol{\varphi}_k + \boldsymbol{\vartheta}_k^+ Z^* \boldsymbol{\vartheta}_k,$$

$$a_{k,4} = -\left(v\tilde{Y}_{k,2} + \bar{v} Y_{k,2}\right), \tag{2.19}$$

In the formulas above Q, $\tilde{Y}_{k,1}$ and $Y_{k,1}$ are matrices whose entries are $q^{(1,1)}_{\alpha\beta}$, $\tilde{y}^{(k,1,1)}_{\beta\alpha}$ and $y^{(k,1,1)}_{\alpha\beta}$ with $\alpha, \beta = 1, 2$ respectively, and $v = q^{(2,2)}_{\alpha\alpha}$, $\tilde{Y}_{k,2} = \tilde{y}^{(k,2,2)}_{\alpha\alpha}$, $Y_{k,2} = y^{(k,2,2)}_{\alpha\alpha}$ with $\alpha = \{1, 2\}$. Therefore, $v \in \mathbb{C}$, and Q is a 2×2 complex matrix.

Fortunately, the integral in (2.17) over Φ and Θ factorizes. Therefore, the integration can be performed over $\boldsymbol{\varphi}_k$ and $\boldsymbol{\vartheta}_k$ separately for every k. Lemma 1 provides a corresponding result.

Lemma 1. *Let* j_k *be defined by* (2.18). *Then*

$$\mathsf{j}_k = \det A + bv \det Q^* + b\bar{v} \det Q + b^2 |v|^2, \tag{2.20}$$

where A *is defined in* (2.4) *and* $b = \sqrt{n\kappa_{2,2}} = \sqrt{\frac{2(n-p)}{pn}}$.

Proof. The integral j_k is computed by the expansion of the exponent $e^{a_{k,4}}$ into series. We have

$$\mathsf{j}_k = \int \left(1 + a_{k,4} + \frac{a_{k,4}^2}{2}\right) e^{a_{k,2}} d\boldsymbol{\varphi}_k^+ d\boldsymbol{\varphi}_k d\boldsymbol{\vartheta}_k^+ d\boldsymbol{\vartheta}_k, \tag{2.21}$$

because $a_{k,4}^3 = 0$. Recalling the definition (2.16) of $y^{(k,p,s)}_{\alpha\beta}$ and the values (2.12) of $\kappa_{p,s}$, one can render $a_{k,2}$ and $a_{k,4}$ in the form

$$a_{k,2} = -\boldsymbol{\rho}_k^+ A \boldsymbol{\rho}_k, \tag{2.22}$$

$$a_{k,4} = -b\left(\theta_{k2}\theta_{k1}\overline{\phi_{k2}\phi_{k1}}v + \bar{v}\phi_{k1}\phi_{k2}\overline{\theta_{k1}\theta_{k2}}\right), \tag{2.23}$$

where A is defined in (2.4),

$$\boldsymbol{\rho}_k = \begin{pmatrix} \boldsymbol{\varphi}_k \\ \boldsymbol{\vartheta}_k \end{pmatrix} \tag{2.24}$$

and $b=\sqrt{n\kappa_{2,2}}=\sqrt{\frac{2(n-p)}{pn}}$. Then we can compute the integral (2.21) term-wise using (2.8):

$$\int e^{-\boldsymbol{\rho}_k^+ A\boldsymbol{\rho}_k} d\boldsymbol{\rho}_k^+ d\boldsymbol{\rho}_k = \det A;$$
$$-\int b\left(\theta_{k2}\theta_{k1}\overline{\phi_{k2}\phi_{k1}}v+\bar{v}\phi_{k1}\phi_{k2}\overline{\theta_{k1}\theta_{k2}}\right)e^{-\boldsymbol{\rho}_k^+ A\boldsymbol{\rho}_k} d\boldsymbol{\rho}_k^+ d\boldsymbol{\rho}_k = b(v\det Q^*+\bar{v}\det Q);$$
$$\frac{b^2}{2}\int \left(\theta_{k2}\theta_{k1}\overline{\phi_{k2}\phi_{k1}}v+\bar{v}\phi_{k1}\phi_{k2}\overline{\theta_{k1}\theta_{k2}}\right)^2 e^{-\boldsymbol{\rho}_k^+ A\boldsymbol{\rho}_k} d\boldsymbol{\rho}_k^+ d\boldsymbol{\rho}_k =$$
$$b^2\int |v|^2\,\theta_{k2}\theta_{k1}\overline{\phi_{k2}\phi_{k1}}\phi_{k1}\phi_{k2}\overline{\theta_{k1}\theta_{k2}}d\boldsymbol{\rho}_k^+ d\boldsymbol{\rho}_k = b^2\,|v|^2\,.$$

□

A substitution of (2.20) into (2.17) gives (2.1).

3 Saddle-point analysis

In order to perform the asymptotic analysis of (2.1) let us change the variables

$$Q=UTV^*,\quad v\to v\det UV^*,$$

where $Q=UTV^*$ is the singular value decomposition of the matrix Q, i.e.

$$T=\operatorname{diag}\{t_j\}_{j=1}^2,\quad t_j\ge 0,\quad U,V\in U(2).$$

The Jacobian of such change is $2\pi^4\triangle^2(T^2)\prod\limits_{j=1}^{2}t_j$ (see e.g. [16]) with $\Delta(T^2)=(t_1^2-t_2^2)^2$, and hence we obtain

$$\mathsf{f}_2(z_1,z_2)=\frac{2n^5}{\pi}\int\limits_{\mathcal{D}}\triangle^2(T^2)t_1t_2\exp\{nf(T,U,V,v)\}\times d\mu(U)d\mu(V)dTd\bar{v}dv, \tag{3.1}$$

where

$$\mathcal{D}=\{(T,U,V,v)\mid t_j\ge 0,\ j=1,2,\ U,V\in U(2)\},\quad dT=\prod_{j=1}^{2}dt_j, \tag{3.2}$$

μ is a Haar measure on $U(2)$, and

$$f(T,U,V,v)=f_0(T,v)+\frac{1}{\sqrt{n}}f_r(T,U,V,v); \tag{3.3}$$
$$f_0(T,v)=-\operatorname{tr}T^2-|v|^2+\log h_0(T,v); \tag{3.4}$$
$$h_0(T,v)=\prod_{j=1}^{2}(|z_0|^2+t_j^2)+2bt_1t_2\Re v+b^2|v|^2; \tag{3.5}$$
$$f_r(T,U,V,v)=\sqrt{n}(\hat{f}(UTV^*,v\det UV^*)-f_0(T,v)) \tag{3.6}$$

with $\hat{f}$ of (2.2). Note that h of (2.3) in new variables takes the form

$$h(Q,v)=\det A+bv\det T^*+b\bar{v}\det T+b^2\,|v|^2\,. \tag{3.7}$$

3.1 Saddle-points

Since we use the Laplace method to analyse (3.1), we are interested in the saddle-points where the global maximum of $f_0(T,v) = f_0(T,x,y)$ with $v = x+iy$ is achieved.

We start with the following simple lemma

Lemma 2. *The function $f_0\colon [0,\infty)^2 \times \mathbb{C} \to \mathbb{R}$ defined by (3.4) attains its global maximum value. Moreover, if $(\hat{T},\hat{v})$ is a point of the global maximum then $\hat{t}_1 = \hat{t}_2$.*

Proof. The function $f_0(T,v)$ is continuous and

$$\lim_{\substack{T\to+\infty \\ v\to\infty}} f_0(T,v) = -\infty.$$

Therefore, f_0 attains its global maximum. Next, by AM-GM and QM-GM inequalities

$$(|z_0|^2 + t_1^2)(|z_0|^2+t_2^2) \le \left(|z_0|^2 + \frac{t_1^2+t_2^2}{2}\right)^2, \tag{3.8}$$

$$2bt_1t_2x \le b(t_1^2+t_2^2)\,|x|\,.$$

Note that the inequality (3.8) is strict if $t_1 \ne t_2 \ge 0$. Hence, for $t_1 \ne t_2$ we have

$$f_0(T,v) = -\operatorname{tr} T^2 - |v|^2 + \log\left[(|z_0|^2+t_1^2)(|z_0|^2+t_2^2) + 2bt_1t_2x + b^2\,|v|^2\right]$$

$$< -\operatorname{tr} T^2 - |v|^2 + \log\left[\left(|z_0|^2 + \frac{t_1^2+t_2^2}{2}\right)^2 + b(t_1^2+t_2^2)\,|x| + b^2\,|v|^2\right] = f_0(tI, |x|+iy),$$

where $t = \sqrt{\frac{t_1^2+t_2^2}{2}}$. Thus, a point with $t_1 \ne t_2$ cannot be a global maximum point of f_0. □

Now we are ready to prove

Proposition 2. *Function $f_0(T,v) = f_0(t_1,t_2,x,y)$ with $v = x+iy$ of (3.4) may attain its global maximum only at the saddle-points of these three types:*

I. $*$-saddle point: $t_1 = t_2 = t_* \ne 0$, $x = x_*$, $y = 0$.

Here $x_* = \alpha b$ and $\alpha \in [0,1]$ is a real solution to the equation

$$2\alpha(1-\alpha)^2b^2 + 1 - \alpha = |z_0|^2, \tag{3.9}$$

such that

$$(6\alpha^2 - 8\alpha + 2)b^2 - 1 \le 0, \tag{3.10}$$

and

$$t_*^2 = \frac{\alpha|z_0|^2}{1-\alpha} - \alpha b^2 > 0. \tag{3.11}$$

The value of f_0 at this point is

$$F_I(\alpha,b,|z_0|^2) = f_0(t_1,t_2,x,y)\Big|_I = -\alpha^2b^2 - \frac{2\alpha|z_0|^2}{1-\alpha} + 2\alpha b^2 + \log\frac{|z_0|^2}{1-\alpha}. \tag{3.12}$$

II. v-saddle points: $t_1 = t_2 = 0$, $|v|^2 = x^2+y^2 = 1 - |z_0|^4/b^2$.

The value of f_0 at this point is

$$F_{II}(b,|z_0|^2) = f_0(t_1,t_2,x,y)\Big|_{II} = -(1 - \frac{|z_0|^4}{b^2}) + \log b^2. \tag{3.13}$$

III. Zero saddle-point: $t_1 = t_2 = x = y = 0$.

The value of f_0 *at this point is*

$$F_{III}(b, |z_0|^2) = f_0(t_1, t_2, x, y)\Big|_{III} = \log |z_0|^4. \tag{3.14}$$

Proof. Taking the derivatives of f_0 of (3.4), we get the following system of equations for the stationary points

$$\begin{cases} -x + \dfrac{bt_1t_2 + b^2x}{h_0(T,v)} = 0 \\ -y + \dfrac{b^2y}{h_0(T,v)} = 0 \\ -t_1 + \dfrac{t_1(|z_0|^2 + t_2^2) + bxt_2}{h_0(T,v)} = 0 \\ -t_2 + \dfrac{t_2(|z_0|^2 + t_1^2) + bxt_1}{h_0(T,v)} = 0 \end{cases}$$

According to Lemma 2, we want to consider only points with $t_1 = t_2 = t$, so the system transforms to

$$\begin{cases} -x + \dfrac{bt^2 + b^2x}{h_0(t,v)} = 0 \\ -y + \dfrac{b^2y}{h_0(t,v)} = 0 \\ -t + \dfrac{t(|z_0|^2 + t^2) + bxt}{h_0(t,v)} = 0 \end{cases} \tag{3.15}$$

To find the solutions of (3.15), consider first the case $y = 0$. Then (3.15) can be rewritten as

$$\begin{cases} -x + \dfrac{bt^2 + b^2x}{h_0(T,v)} = 0 \\ -t + \dfrac{t(|z_0|^2 + t^2) + bxt}{h_0(T,v)} = 0 \end{cases}$$

Clearly, if $t = 0$, then we can have $x = 0$, or

$$h_0(0,v) = b^2 \Longrightarrow x^2 = 1 - \frac{|z_0|^4}{b^2}.$$

which gives two solutions:

$$t_1 = t_2 = t = 0, \quad x = y = 0, \tag{3.16}$$

$$t_1 = t_2 = t = 0, \quad x^2 = 1 - \frac{|z_0|^4}{b^2}, \quad y = 0. \tag{3.17}$$

If $t \neq 0$, we get

$$h_0(T,v) = bx + t^2 + |z_0|^2, \tag{3.18}$$

and hence from the first equation of (3.15)

$$bt^2 + b^2x = xh_0(T,v) \Longrightarrow t^2 = \frac{x|z_0|^2}{b - x} - bx \tag{3.19}$$

if $x \neq b$. Substituting this to the r.h.s of (3.18), we get

$$bx + t^2 + |z_0|^2 = \frac{b|z_0|^2}{b-x},$$

and hence

$$h_0(T,v) = \frac{b^2|z_0|^4}{(b-x)^2} - 2bx|z_0|^2 = \frac{b|z_0|^2}{b-x} \Longrightarrow 2x(b-x)^2 + b - x = |z_0|^2 b, \tag{3.20}$$

if $b \neq 0$ and $|z_0| \neq 0$. The last one is a cubic equation with respect to x, which may have three real roots. It is convenient to write

$$x = \alpha b, \tag{3.21}$$

and hence from (3.19), (3.20) we get

$$\begin{aligned} t^2 &= \frac{\alpha|z_0|^2}{1-\alpha} - \alpha b^2, \quad h_0(T,v) = \frac{|z_0|^2}{1-\alpha}, \\ 2\alpha(1-\alpha)^2 b^2 &+ 1 - \alpha = |z_0|^2. \end{aligned} \tag{3.22}$$

Now we need

Lemma 3. *Among all stationary points corresponding to real roots of the equation*

$$2\alpha(1-\alpha)^2 b^2 + 1 - \alpha = |z_0|^2, \tag{3.23}$$

the global maximum of f_0 can be achieved only at the saddle point corresponding to the solution α of (3.23) lying on $[0,1]$ and such that (3.10) holds, i.e.

$$(6\alpha^2 - 8\alpha + 2)b^2 - 1 \leq 0. \tag{3.24}$$

The saddle-point corresponding to this solution exists if and only if one of the following holds

(i) $b \leq \frac{1}{\sqrt{2}}$, $|z_0|^2 \leq 1$;

(ii) $b \in (\frac{1}{\sqrt{2}}, 1]$, $|z_0|^2 \leq z_-(b)$, *where*

$$z_-(b) = \frac{4b^2+9}{27} + \frac{4b^2+6}{27}\sqrt{1+\tfrac{3}{2b^2}}; \tag{3.25}$$

(iii) $b \in (1, \sqrt{\frac{5+\sqrt{5}}{5}})$, $|z_0|^2 \in [\frac{b^2 - b\sqrt{2-b^2}}{2}, z_-(b)]$;

(iv) $b \in [\sqrt{\frac{5+\sqrt{5}}{5}}, \sqrt{2}]$, $|z_0|^2 \in [\frac{b^2 - b\sqrt{2-b^2}}{2}, \frac{b^2 + b\sqrt{2-b^2}}{2}]$;

In addition, at this saddle-point

$$b^2 \leq \frac{1}{1 - 2\alpha(1-2\alpha)}, \tag{3.26}$$

$$h_* = h_0(t_*, x_*, 0) = \frac{|z_0|^2}{1-\alpha} = bx_* + t_*^2 + |z_0|^2. \tag{3.27}$$

The proof of Lemma 3 is given after the proof of Proposition 2.

If $b = 0$, then the first equation of (3.15) implies $x = 0$, and then

$$h_0(T, v) = t^2 + |z_0|^2 \Longrightarrow (t^2 + |z_0|^2)^2 = t^2 + |z_0|^2 \Longrightarrow t = \sqrt{1 - |z_0|^2}$$

for $|z_0| \leq 1$.

If $|z_0| = 0$, then

$$h_0(T, v) = t^2 + bx \Longrightarrow (t^2 + bx)^2 = t^2 + bx \Longrightarrow t^2 + bx = 1,$$

and hence the first equation of (3.15) gives $x = b$, thus

$$t_1 = t_2 = \sqrt{1 - b^2}, \quad x = b, \quad y = 0$$

for $b \leq 1$. Notice also that $x = b \neq 0$ always implies

$$t^2 + b^2 = t^2 + b^2 + |z_0|^2 \Longrightarrow |z_0| = 0.$$

Thus, the case $t_1 = t_2 = t \neq 0$, $y = 0$ for $|z_0| \neq 0$, $b \neq 0$ gives the solution

$$t_1 = t_2 = t_*, \quad x = x_* = \alpha b, \quad y = 0 \tag{3.28}$$

where $\alpha \in [0, 1]$ and

$$2\alpha(1 - \alpha)^2 b^2 + 1 - \alpha = |z_0|^2, \quad t_*^2 = \frac{\alpha |z_0|^2}{1 - \alpha} - \alpha b^2$$

if such solution exists and (3.26) holds (see Lemma 3).

If $z_0 = 0$, but $b \leq 1$, then the solution takes the form

$$t_1 = t_2 = \sqrt{1 - b^2}, \quad x = b, \quad y = 0.$$

Notice that the last solution can be considered as a limiting case of (3.28), so one can consider (3.28) also for $|z_0| = 0$. Same is true for the solution

$$t_1 = t_2 = \sqrt{1 - |z_0|^2}, \quad x = y = 0$$

obtained for $b = 0$, $|z_0| \leq 1$.

Assume now that $y \neq 0$. Then the second equation of (3.15) gives

$$h_0(T, v) = b^2,$$

and hence the first equation gives $t = 0$. Thus the previous equation implies

$$|z_0|^4 + b^2(x^2 + y^2) = b^2 \Longrightarrow x^2 + y^2 = 1 - \frac{|z_0|^4}{b^2}$$

if $|z_0|^2 \leq b$. Hence, another family of solutions takes the form

$$t_1 = t_2 = 0, \quad x^2 + y^2 = 1 - \frac{|z_0|^4}{b^2} \tag{3.29}$$

if $|z_0|^2 \leq b$. Note that this includes (3.17), so one can include here $y = 0$.

□

Proof of Lemma 3.

Notice that as soon as (3.23) is satisfied we get

$$f_0(T,v) = -2t^2 - x^2 + \log h_0(T,v) = \alpha^2 b^2 - \frac{2\alpha|z_0|^2}{1-\alpha} + 2b^2\alpha(1-\alpha) + \log\frac{|z_0|^2}{1-\alpha}$$
$$= \alpha^2 b^2 - \frac{|z_0|^2}{1-\alpha} + \log\frac{|z_0|^2}{1-\alpha} + 2|z_0|^2 - 1.$$

Here we used (3.21) – (3.22) and

$$2b^2\alpha(1-\alpha) = \frac{|z_0|^2}{1-\alpha} - 1.$$

Taking the derivative with respect to α of the function above we get

$$2\alpha b^2 - \frac{|z_0|^2}{(1-\alpha)^2} + \frac{1}{1-\alpha} = \frac{2\alpha(1-\alpha)^2 b^2 + 1 - \alpha - |z_0|^2}{(1-\alpha)^2}. \tag{3.30}$$

Suppose (3.23) has three real roots. Then, according to (3.30), two of them are the local minimums, and the middle one is a local maximum of the function above, and hence the value of f_0 in the saddle-point associated to the middle root of (3.23) is greater then those of other two. Notice that if $p(\alpha) = 2\alpha(1-\alpha)^2 b^2 + 1 - \alpha - |z_0|^2$, then

$$p(0) = 1 - |z_0|^2, \quad p(1) = -|z_0|^2,$$

and hence the biggest root of $p(\alpha)$ is always grater than 1. If $|z_0|^2 \le 1$, then (3.23) must has a non-positive root and the root between 0 and 1, and so the middle root belongs to $[0,1]$. If $|z_0|^2 > 1$, then one root is still bigger than 1, and two other roots, if exist, must be of the same sign (since the product of three roots is $(|z_0|^2 - 1)/2b^2$). It is easy to see that the point with negative x and $t \neq 0$ cannot be a global maximum of f_0 since changing the sign of x evidently increases the value of f_0, thus we interested only in the case when all three roots are positive. As was mentioned, one of them is greater than 1. Two other roots should be at the same side of 1, and since the sum of all three roots is 2 according to Vieta's formula, it implies that two smaller roots must lie on $[0,1]$. In addition, at the middle root p' must be negative which gives (3.10). Notice also that if t^2 of (3.22) corresponding to this middle root is negative, then t^2 corresponding to the smallest root is also negative (since $1/(1-\alpha)$ increases on [0,1)), and the corresponding t^2 is always negative for $\alpha > 1$. This also implies that in the case when $p(\alpha)$ has only one real root $\alpha > 1$ the saddle-point corresponding to the root does not exists since $t^2 < 0$.

Thus, the saddle-point corresponding to the solution of (3.23) exists and can be a global maximum of f_0 only if (3.23) has a solution $\alpha \in [0,1]$ and this solution satisfies (3.10) and

$$\frac{\alpha|z_0|^2}{1-\alpha} \ge \alpha b^2 \Longrightarrow 2\alpha(1-\alpha)b^2 + 1 \ge b^2 \Longleftrightarrow \alpha \in [\tfrac{b-\sqrt{2-b^2}}{2b}, \tfrac{b+\sqrt{2-b^2}}{2b}]. \tag{3.31}$$

The last condition guarantees the existence of t_* (see (3.11)) and implies (3.26). Notice that since $0 \le 2\alpha(1-\alpha) \le 1/2$, we immediately get that $b \le \sqrt{2}$ and (3.31) is always satisfied if $b \le 1$.

From the discussion above, if $|z_0| \le 1$, then the solution α of (3.23) satisfying (3.10) exists. Note that

$$p'(\alpha) = 0 \Longleftrightarrow \alpha = \alpha_{\pm} = \frac{2 \pm \sqrt{1 + \frac{3}{2b^2}}}{3}. \tag{3.32}$$

Therefore, since $\alpha_+ > 1$, if $|z_0| > 1$, then two positive solutions on $[0, 1]$ exist if

$$\alpha_- > 0, \quad p(\alpha_-) \geq 0$$

which is equivalent to $b > \frac{1}{\sqrt{2}}$ and $1 < |z_0|^2 \leq z_-(b) = p(\alpha_-) + |z_0|^2$ with $z_-(b)$ of (3.25). Therefore, if $b \leq 1$, then $*$-saddle point exists if and only if (i) or (ii) holds, and if $b \in (1, \sqrt{2}]$ and $|z_0|^2 \leq z_-(b)$, then the solution α of (3.23) satisfying (3.10) exists.

If $b \in (1, \sqrt{2}]$, then in order to satisfy (3.31) we must have

$$\alpha \in [\tfrac{b-\sqrt{2-b^2}}{2b}, \tfrac{b+\sqrt{2-b^2}}{2b}]. \tag{3.33}$$

Notice that if $b \in [\sqrt{\frac{5+\sqrt{5}}{5}}, \sqrt{2}]$, then it is easy to check that

$$[\tfrac{b-\sqrt{2-b^2}}{2b}, \tfrac{b+\sqrt{2-b^2}}{2b}] \subset [\alpha_-, 1],$$

and hence $p'(\alpha) < 0$ for all α satisfying (3.33). Therefore, if

$$|z_0|^2 \in [p(\tfrac{b+\sqrt{2-b^2}}{2b}) + |z_0|^2, p(\tfrac{b-\sqrt{2-b^2}}{2b}) + |z_0|^2] = [\tfrac{b^2-b\sqrt{2-b^2}}{2}, \tfrac{b^2+b\sqrt{2-b^2}}{2}],$$

then the solution α_* of (3.23) satisfying (3.10) and (3.33) exists, and so does $*$-saddle point, which gives (iv).

If $b \in [1, \sqrt{\frac{5+\sqrt{5}}{5}})$, then it is easy to check that

$$\tfrac{b-\sqrt{2-b^2}}{2b} < \alpha_- < \tfrac{b+\sqrt{2-b^2}}{2b} < 1,$$

and so possible values of $|z_0|^2$ should correspond to $\alpha \in [\alpha_-, \frac{b+\sqrt{2-b^2}}{2b}]$ which gives (iii).

The expression for h_* can be obtained straightforwardly from (3.22). □

3.2 Main saddle-points

Depending on the values of $|z_0|$ and b, the main contribution to (3.1) is given by the different saddle-points.

Notice first that if $|z_0|^2 \leq b$ and so the v-saddle point exists, then

$$F_{II}(b, |z_0|^2) \geq F_{III}(b, |z_0|^2). \tag{3.34}$$

Indeed, according to (3.13) – (3.14)

$$F_{II}(b, |z_0|^2) - F_{III}(b, |z_0|^2) = \frac{|z_0|^4}{b^2} - 1 - \log \frac{|z_0|^4}{b^2} \geq 0.$$

Compare now the values at $*$-saddle point and v-saddle points:

Lemma 4. *Let $|z_0|^2 \leq b$ (i.e. v-saddle point exists). Then*

(i) If one of the following holds

- $b \leq 1$, $|z_0| \leq 1$
- $b \in (1, \sqrt{2}]$ *and*

$$\frac{b^2 - b\sqrt{2-b^2}}{2} \leq |z_0|^2 \leq 1,$$

then $*$-saddle point exists and

$$F_I(\alpha, b, |z_0|^2) \geq F_{II}(b, |z_0|^2),$$

where F_I, F_{II} are defined in (3.12) – (3.13).

(ii) If $b \in (1, \sqrt{2}]$, but $|z_0| > 1$, then there exists a curve $z_1(b)$ such that $*$-saddle point exists for $1 \leq |z_0|^2 \leq \min(z_1(b), b)$ and

$$F_I(\alpha, b, |z_0|^2) \geq F_{II}(b, |z_0|^2), \quad 1 \leq |z_0|^2 \leq \min(z_1(b), b). \tag{3.35}$$

If $\min(z_1(b), b) \leq |z_0|^2 \leq b$, then either $*$-point does not exist, or

$$F_I(\alpha, b, |z_0|^2) < F_{II}(b, |z_0|^2). \tag{3.36}$$

Therefore, the curve γ_2 on Figure 1 is

$$|z_0|^2 = z_1(b), \quad b : z_1(b) < b.$$

It coincides with

$$|z_0|^2 = \frac{b^2 + b\sqrt{2 - b^2}}{2}$$

for all $b \in [\sqrt{\frac{5+\sqrt{5}}{5}}; \sqrt{2}]$.

Proof. We start with (i). It is easy to check that $z_-(b) \geq 1$ (see (3.25)) and, given $b \geq 1$,

$$\frac{b^2 + b\sqrt{2 - b^2}}{2} \geq 1.$$

Thus the existence of $*$-saddle point in the conditions of (i) follows from Lemma 3.

Now, according to (3.9), we have

$$\frac{|z_0|^2}{1 - \alpha} = 2\alpha(1 - \alpha)b^2 + 1.$$

Therefore,

$$\begin{aligned} F_I(\alpha, b, |z_0|^2) - F_{II}(b, |z_0|^2) &= -|z_0|^2(\tfrac{|z_0|^2}{b^2} + \tfrac{2\alpha}{1-\alpha}) + 2\alpha b^2 - \alpha^2 b^2 + 1 \\ + \log(2\alpha(1 - \alpha) + \tfrac{1}{b^2}) &= b^2(2\alpha - 5\alpha^2 + 4\alpha^3 - 4\alpha^2(1 - \alpha)^4) - \frac{(1 - \alpha)^2}{b^2} \\ + (1 - 2\alpha - 4\alpha(1 - \alpha)^3) &+ \log(2\alpha(1 - \alpha) + \tfrac{1}{b^2}), \end{aligned} \tag{3.37}$$

where to obtain the last equality we substitute $|z_0|^2 = 2\alpha(1 - \alpha)^2 b^2 + 1 - \alpha$ and open the parentheses.

Fix α and consider $s = 1/b^2 \geq 1 - 2\alpha(1 - \alpha)$ (recall that if $*$-saddle point exists, then we have (3.26)). Denote

$$\begin{aligned} H(s) = &s^{-1} \cdot (2\alpha - 5\alpha^2 + 4\alpha^3 - 4\alpha^2(1 - \alpha)^4) - (1 - \alpha)^2 s \\ &+ (1 - 2\alpha - 4\alpha(1 - \alpha)^3) + \log(2\alpha(1 - \alpha) + s). \end{aligned} \tag{3.38}$$

One can easily check that

$$H(1 - 2\alpha(1 - \alpha)) = 0. \tag{3.39}$$

Moreover, computing

$$H'(s) = -\frac{2\alpha - 5\alpha^2 + 4\alpha^3 - 4\alpha^2(1-\alpha)^4}{s^2} - (1-\alpha)^2 + \frac{1}{2\alpha(1-\alpha)+s},$$

one can also check

$$H'(1-2\alpha(1-\alpha)) = 0$$

and

$$H'(s) = \frac{s-1+2\alpha(1-\alpha)}{s^2(s+2\alpha(1-\alpha))} \cdot g(s),$$

where

$$g(s) = -(1-\alpha)^2 s^2 + (2\alpha - \alpha^2)s + 4\alpha^2(1-\alpha)^4 + 2\alpha^3(1-\alpha). \tag{3.40}$$

Since

$$\frac{s-1+2\alpha(1-\alpha)}{s^2(s+2\alpha(1-\alpha))} \geq 0$$

for $s \geq 1 - 2\alpha(1-\alpha)$, the sign of $H'(s)$ is determined by the sign of $g(s)$.

Recall that the existence of v-point implies

$$|z_0|^2 = 2\alpha(1-\alpha)^2 b^2 + 1 - \alpha \leq b \Longrightarrow b \in (b_-(\alpha), b_+(\alpha)),$$

where

$$b_\pm(\alpha) = \frac{1 \pm \sqrt{1-8\alpha(1-\alpha)^3}}{4\alpha(1-\alpha)^2}. \tag{3.41}$$

Since it is easy to check that $b_+(\alpha) > \sqrt{2}$ and $b_-(\alpha)^2 \leq (1-2\alpha(1-\alpha))^{-1}$ for any $\alpha \in [0,1]$, we are interested in $b \in [b_-(\alpha), (1-2\alpha(1-\alpha))^{-1/2}]$, i.e.

$$s \in [1-2\alpha(1-\alpha), s_0(\alpha)], \quad s_0(\alpha) = (b_-(\alpha))^{-2}. \tag{3.42}$$

Notice that

$$\begin{aligned} g(1-2\alpha(1-\alpha)) &= -(1-\alpha)^2(1-2\alpha(1-\alpha))^2 + (2\alpha-\alpha^2)(1-2\alpha(1-\alpha)) \\ &+ 4\alpha^2(1-\alpha)^4 + 2\alpha^3(1-\alpha) = -(4\alpha^2 - 6\alpha + 1)(1-2\alpha(1-\alpha)), \end{aligned}$$

and hence we get

$$\begin{aligned} g(1-2\alpha(1-\alpha)) &< 0, \quad \alpha \in [0, \tfrac{3-\sqrt{5}}{4}), \\ g(1-2\alpha(1-\alpha)) &\geq 0, \quad \alpha \in [\tfrac{3-\sqrt{5}}{4}, 1] \end{aligned} \tag{3.43}$$

Now we need the following simple lemma

Lemma 5. *If $|z_0|^2 \leq 1$, $b \leq (1-2\alpha(1-\alpha))^{-1/2}$ and $|z_0|^2 \leq b$, then $\alpha \geq 1 - \frac{1}{\sqrt{2}}$.*

Proof. Indeed, according to (3.9), $|z_0|^2 \leq 1$ implies

$$2\alpha(1-\alpha)^2 b^2 + 1 - \alpha \leq 1 \Longrightarrow \alpha \geq 1 - \frac{1}{b\sqrt{2}}, \tag{3.44}$$

which gives the statement for $b \geq 1$. If $b \leq 1$, then we must have $b_-(\alpha) \leq 1$ (see (3.41)) which also implies $1 - \alpha \leq 1/\sqrt{2}$. □

According to Lemma 5, in the conditions of Lemma 4 (i) we have $\alpha > 1 - 1/\sqrt{2} > \frac{3-\sqrt{5}}{4}$, and thus $g(1-2\alpha(1-\alpha)) \geq 0$. As g is a quadratic polynomial with negative top coefficient, it has exactly one root on $[1-2\alpha(1-\alpha), \infty)$, hence, the same is true for $H'(s)$. Thus $H(s)$ increases for $[1-2\alpha(1-\alpha), a)$ with a certain $a \geq 1-2\alpha(1-\alpha)$, and then decreases.

Therefore, if $H(s_0(\alpha)) \geq 0$, then $H(s) \geq 0$ for all $s \in [1-2\alpha(1-\alpha), s_0(\alpha)]$ and $\alpha \geq 1 - 1/\sqrt{2}$.

It remains to check

$$H(s_0(\alpha)) \geq 0, \quad \alpha \geq 1 - 1/\sqrt{2}. \tag{3.45}$$

If $s = s_0(\alpha)$ (i.e. $b = b_-(\alpha)$), then

$$|z_0|^2 = 2\alpha(1-\alpha)^2 b^2 + 1 - \alpha = b, \tag{3.46}$$

i.e.

$$\begin{aligned} H(s_0(\alpha)) &= -\frac{2\alpha b}{1-\alpha} + 2\alpha b^2 - \alpha^2 b^2 - \log(b(1-\alpha)) \\ &= b - 1 - \alpha^2 b^2 - \alpha + 2\alpha^3 b^2 - \log(b(1-\alpha)) =: r(b,\alpha) \end{aligned}$$

where we have used (3.37), (3.38), (3.46) and

$$\begin{aligned} &2\alpha b^2 - \alpha^2 b^2 = 2\alpha(1-\alpha)b^2 + \alpha^2 b^2 = \frac{b}{1-\alpha} - 1 + \alpha^2 b^2 \\ &\frac{b}{1-\alpha} = 2\alpha(1-\alpha)b^2 + 1. \end{aligned}$$

Taking the derivative with respect to α we get

$$r'_\alpha(b,\alpha) = \frac{\alpha}{1-\alpha}(1 - b^2(2 - 8\alpha + 6\alpha^2)).$$

Taking into account (3.10), $r'_\alpha(b,\alpha) \geq 0$, and hence $r(b,\alpha)$ increases in α. At $\alpha = 1 - 1/\sqrt{2}$ we get

$$r(b, 1 - 1/\sqrt{2}) = b - 2 + \frac{1}{\sqrt{2}} - (\sqrt{2}-1)^3 b^2/2 - \log\frac{b}{\sqrt{2}},$$

which, as one can easily check, is positive for all $b \in (0, \sqrt{2})$. Therefore, $r(b,\alpha) \geq 0$ for all $\alpha \geq 1 - 1/\sqrt{2}$, and hence we obtain (3.45), which finishes the proof of Lemma 4 (i).

Let's now prove (ii). Suppose $|z_0| > 1$. Proceeding similarly to the proof of (i), we want to study when $H(s) \geq 0$ for $1 - 2\alpha(1-\alpha) \leq s \leq s_0(\alpha)$, where H is defined at (3.38) and $s_0(\alpha)$ is defined in (3.42). Since $g(0) \geq 0$, $1 - 2\alpha(1-\alpha) > 0$, and according to (3.43), we have for all $s > 1 - 2\alpha(1-\alpha)$

$$g(s) < 0, \quad \alpha < \frac{3-\sqrt{5}}{4},$$

and hence

$$H'(s) < 0, \quad \alpha < \frac{3-\sqrt{5}}{4} \Longrightarrow H(s) < 0 \text{ if } \alpha < \frac{3-\sqrt{5}}{4}, s > 1 - 2\alpha(1-\alpha).$$

Therefore, if $\alpha < \frac{3-\sqrt{5}}{4}$, and $*$-saddle point exists, then (3.36) holds.

Let $\alpha \geq \frac{3-\sqrt{5}}{4}$. According to (3.43), then

$$\begin{aligned} &H(s) \geq 0, \quad s \in [1-2\alpha(1-\alpha), s_1(\alpha)], \\ &H(s) < 0, \quad s \geq s_1(\alpha), \end{aligned} \tag{3.47}$$

where $s_1(\alpha)$ is a solution of $H(\alpha, s) = 0$ bigger than $1 - 2\alpha(1-\alpha)$ (which exists since $H(s) \to -\infty$, $s \to \infty$) for $\alpha < 1$. Notice that

$$s_1(\tfrac{3-\sqrt{5}}{4}) = 1 - 2\alpha(1-\alpha)\Big|_{\alpha=\frac{3-\sqrt{5}}{4}} = \frac{5-\sqrt{5}}{4} < s_0(\tfrac{3-\sqrt{5}}{4}).$$

Recall that we are interested in s such that $1 - 2\alpha(1-\alpha) \le s \le s_0(\alpha)$. If $H(s_0(\alpha)) \ge 0$ (i.e. $s_1(\alpha) \ge s_0(\alpha)$), then $H(s) \ge 0$ for all such s. Numerically, one can compute that this happens if $\alpha \ge \alpha_0 \approx 0.22$. Note that the exact numerical value of α_0 is not important here and is given for reference only. What we really going to use is that

$$\tfrac{3-\sqrt{5}}{4} < \alpha_0 < 1 - \tfrac{1}{\sqrt{2}} \tag{3.48}$$

which follows from (3.45).

If $\frac{3-\sqrt{5}}{4} \le \alpha \le \alpha_0$, then $H(s) \ge 0$ for $s \in [1-2\alpha(1-\alpha), s_1(\alpha)]$, and $H(s) < 0$ for $s \in [s_1(\alpha), s_0(\alpha)]$.

Notice also that if $\frac{3-\sqrt{5}}{4} \le \alpha \le 1$, then

$$6\alpha^2 - 8\alpha + 2 \le 1 - 2\alpha(1-\alpha) \le s \Longrightarrow (6\alpha^2 - 8\alpha + 2)b^2 - 1 < 0 \tag{3.49}$$

for all $b^2 = 1/s$ with $s \ge 1 - 2\alpha(1-\alpha)$.

Now we will need

Lemma 6. *In the notations above we have*

$$s_1'(\alpha) > 0$$

as soon as $\frac{3-\sqrt{5}}{4} < \alpha < 1 - \frac{1}{\sqrt{2}}$ *(and so, in particular, if* $\frac{3-\sqrt{5}}{4} < \alpha \le \alpha_0$*)*

The proof of Lemma 6 is given after the proof of Lemma 4.

Note that it is easy to check that $s_0'(\alpha) > 0$ for $1 \ge \alpha \ge \frac{3-\sqrt{5}}{4}$. This and the lemma implies that (3.47) can be rewritten as

$$H(\alpha, s) \ge 0$$

if

$$\begin{aligned} &\tfrac{1}{2} \le s \le \tfrac{5-\sqrt{5}}{4}, \quad \alpha \in [\tfrac{1-\sqrt{2s-1}}{2}; \tfrac{1+\sqrt{2s-1}}{2}] \text{ or} \\ &\tfrac{5-\sqrt{5}}{4} \le s \le 1, \quad \alpha \in [\tilde{\alpha}_1(s), \tfrac{1+\sqrt{2s-1}}{2}] \end{aligned} \tag{3.50}$$

and $H(\alpha, s) < 0$ or $*$-saddle point does not exist in the remaining domain. Here $\tilde{\alpha}_1(s)$ is an inverse function to $s_1(\alpha)$ for $\frac{3-\sqrt{5}}{4} \le \alpha < \alpha_0$ and to $s_0(\alpha)$ for $\alpha \in [\alpha_0, 1]$. In terms of $b \in (1, \sqrt{2}]$ the domain (3.50) can be rewritten as

$$\begin{aligned} &\sqrt{\tfrac{5+\sqrt{5}}{5}} \le b \le \sqrt{2}, \quad \alpha \in [\tfrac{b-\sqrt{2-b^2}}{2b}; \tfrac{b+\sqrt{2-b^2}}{2b}] \text{ or} \\ &1 < b \le \sqrt{\tfrac{5+\sqrt{5}}{5}}, \quad \alpha \in [\alpha_1(b), \tfrac{b+\sqrt{2-b^2}}{2b}] \end{aligned} \tag{3.51}$$

where $\alpha_1(b) = \tilde{\alpha}_1(1/b^2)$.

It remains to notice that for $\alpha \ge \frac{3-\sqrt{5}}{4}$, $b^2 \le \frac{1}{1-2\alpha(1-\alpha)}$

$$(2\alpha(1-\alpha)^2 b^2 + 1 - \alpha)'_\alpha = (6\alpha^2 - 8\alpha + 2)b^2 - 1 < 0,$$

and hence in terms of b, $|z_0|^2$ (3.51) takes the form

$$\begin{aligned} &\sqrt{\tfrac{5+\sqrt{5}}{5}} \le b \le \sqrt{2}, \quad |z_0|^2 \in [\tfrac{b^2-b\sqrt{2-b^2}}{2}; \tfrac{b^2+b\sqrt{2-b^2}}{2}] \text{ or} \\ &1 < b \le \sqrt{\tfrac{5+\sqrt{5}}{5}}, \quad |z_0|^2 \in [\tfrac{b^2-b\sqrt{2-b^2}}{2}; z_1(b)] \end{aligned} \tag{3.52}$$

with

$$z_1(b) = 2\alpha(1-\alpha)^2 b^2 + 1 - \alpha\Big|_{\alpha=\alpha_1(b)}.$$

According to the definition of α_0, we get $z_1(b) = b$ for $1 \le b \le 1/\sqrt{s_1(\alpha_0)}$. Notice that because of Lemma 6 and (3.48), $s_0(\alpha_0) \le s_0(1-\frac{1}{\sqrt{2}}) = 1$, and hence $1/\sqrt{s_1(\alpha_0)} > 1$. The exact value of $\sqrt{s_1(\alpha_0)}$ is not important here, but numerically, $1/\sqrt{s_1(\alpha_0)} \approx 1.11$.

The existence of $*$-saddle point in the domain (3.52) follows from (3.49) (which gives $z_1(b) \le z_-(b)$) and Lemma 3. This finishes the proof of Lemma 4. □

Proof of Lemma 6 Taking into account the definition of $s_1(\alpha)$, we get

$$H(\alpha, s_1(\alpha)) = 0 \Longrightarrow H'_\alpha(\alpha, s_1(\alpha)) + s'_1(\alpha)\cdot H'_s(\alpha, s_1(\alpha)) = 0.$$

Since $H'_s(\alpha, s_1(\alpha)) < 0$, it is enough to check that $H'_\alpha(\alpha, s_1(\alpha)) > 0$. Taking the derivative of (3.38) one get

$$\begin{aligned} \frac{1}{2}H'_\alpha &= \frac{1}{s}\cdot(1 - 5\alpha + 6\alpha^2 - 4\alpha(1-\alpha)^4 + 8\alpha^2(1-\alpha)^3) + (1-\alpha)s \\ &\quad - 1 - 2(1-\alpha)^3 + 6\alpha(1-\alpha)^2 + \frac{1-2\alpha}{s + 2\alpha(1-\alpha)} \\ &= \frac{s + 2\alpha(1-\alpha) - 1}{s(s+2\alpha(1-\alpha))}\cdot q(\alpha, s), \end{aligned}$$

where

$$q(\alpha, s) = (1-\alpha)s^2 + s(-\alpha - 2(1-\alpha)^3 + 6\alpha(1-\alpha)^2) - 2\alpha(1-\alpha)(1-3\alpha)(2(1-\alpha)^2 - 1).$$

If $\frac{3-\sqrt{5}}{4} \le \alpha < 1 - \frac{1}{\sqrt{2}}$, then it is easy to check that

$$q(\alpha, 0) < 0.$$

In addition,

$$q(\alpha, 1 - 2\alpha(1-\alpha)) = 8\alpha^3 - 16\alpha^2 + 8\alpha - 1 = (2\alpha - 1)(4\alpha^2 - 6\alpha + 1) > 0$$

for $\frac{3-\sqrt{5}}{4} < \alpha < 1 - \frac{1}{\sqrt{2}}$. Since $q(\alpha, s)$ is a quadratic polynomial in s and $1 - 2\alpha(1-\alpha) > 0$, the consideration above gives

$$q(\alpha, s) > 0$$

for all $s \ge 1 - 2\alpha(1-\alpha)$ including $s = s_1(\alpha)$, which implies the lemma. □

Next we compare the values at v-saddle points and the zero saddle point:

Lemma 7. *(i) If $|z_0| \le 1$, then*

$$F_I(\alpha, b, |z_0|^2) \ge F_{III}(b, |z_0|^2)$$

wherever $*$-saddle point exists, i.e.

- $b \le 1$, $|z_0| \le 1$;
- $b \in (1, \sqrt{2}]$, $\frac{b^2 - b\sqrt{2-b^2}}{2} \le |z_0|^2 \le 1$.

(ii) if $|z_0|^2 > \max(1, b)$, $b \ge 1/\sqrt{2}$, *then there exists a curve* $z_2(b)$ *such that*

$$F_I(\alpha, b, |z_0|^2) \ge F_{III}(b, |z_0|^2), \quad \max(1, b) \le |z_0|^2 \le z_2(b).$$

If $|z_0|^2 > z_2(b)$, *then* $*$*-saddle point does not exist or*

$$F_I(\alpha, b, |z_0|^2) < F_{III}(b, |z_0|^2).$$

(iii) if $|z_0| > 1$, *but* $b < 1/\sqrt{2}$, *then* $*$*-saddle point does not exist.*
Therefore, the curve γ_3 *on Figure 1 is*

$$|z_0|^2 = z_2(b), \quad b : z_2(b) > b.$$

Proof. We start with (i). Notice that if $|z_0|^2 \le b$ and so the v-point exists, then the statement follows from (3.34) and Lemma 4. It is easy to see also that the second derivative with respect to t of $f_0(tI, v)$ at zero saddle-point has the form

$$(f_0'')_t(0,0) = 4\Big(\frac{1}{|z_0|^2} - 1\Big),$$

and hence it is positive for $|z_0|^2 < 1$. Therefore, this stationary point cannot be a point of local maximum. Hence, if $|z_0|^2 > b$ and so the v-point does not exist, the global maximum can be achieved only at the $*$-saddle point (see Proposition 2) which implies (i). (iii) follows from Lemma 3.

It remains to prove (ii). Assume $|z_0|^2 > 1$, $b \ge 1/\sqrt{2}$. Since we are interested in the case $|z_0|^2 > b$ and the $*$-point exists, according to Lemma 3, we need to consider $b \in [\frac{1}{\sqrt{2}}, \sqrt{\frac{5+\sqrt{5}}{5}}]$

$$\max(1, b) \le |z_0|^2 \le z_-(b)$$

with z_- of (3.25). Since

$$z_-(b) > b \Longrightarrow b \le b_0 \approx 1.128, \tag{3.53}$$

we are interested in $b \in [\frac{1}{\sqrt{2}}, b_0]$.

In addition, we get

$$\alpha \ge \alpha_-(b)$$

with α_- of (3.32) and, since $|z_0| > 1$ (see (3.44)),

$$\alpha < 1 - \frac{1}{b\sqrt{2}}.$$

Define

$$\begin{aligned} W(\alpha, b) &= F_I(\alpha, b, |z_0|^2) - F_{III}(\alpha, b, |z_0|^2) \\ &= -\alpha^2 b^2 - \frac{2\alpha|z_0|^2}{1-\alpha} + 2\alpha b^2 - \log(|z_0|^2(1-\alpha)) \\ &= b^2(2\alpha - 5\alpha^2 + 4\alpha^3) - 2\alpha - 2\log(1-\alpha) - \log(1 + 2\alpha(1-\alpha)b^2). \end{aligned} \tag{3.54}$$

Here we used

$$\frac{|z_0|^2}{1-\alpha} = 1 + 2\alpha(1-\alpha)b^2.$$

Consider now

$$
\begin{aligned}
\tfrac{1}{2}\frac{\partial}{\partial\alpha}W(\alpha,b) &= b^2(1-5\alpha+6\alpha^2)-1+\frac{1}{1-\alpha}-\frac{(1-2\alpha)b^2}{1+2\alpha(1-\alpha)b^2}\\
&=\frac{\alpha}{(1-\alpha)(1+2\alpha(1-\alpha)b^2)}(1+b^2(-3+11\alpha-8\alpha^2)+2b^4(1-\alpha)^2(1-5\alpha+6\alpha^2))\\
&=\frac{\alpha}{(1-\alpha)(1+2\alpha(1-\alpha)b^2)}\big(2(1-\alpha)(3\alpha-1)b^2+1\big)\big((1-\alpha)(2\alpha-1)b^2+1\big).
\end{aligned}
$$

According to (3.24),

$$2(1-\alpha)(3\alpha-1)b^2+1 = 1-b^2(6\alpha^2-8\alpha+2)\geq 0.$$

Since $\alpha<1-\frac{1}{b\sqrt{2}}\leq 1/2$, we have for $\alpha\in[\alpha_-(b),1-\frac{1}{b\sqrt{2}}]$

$$(1-\alpha)(2\alpha-1)b^2+1\geq 1-b^2(1-\alpha_-(b))(1-2\alpha_-(b))=\frac{6-b^2-\sqrt{b^4+3b^2/2}}{9}>0$$

for $b<\sqrt{2}$. Therefore, we obtain

$$W'(\alpha,b)\geq 0,\quad \alpha\in[\alpha_-(b),1-\tfrac{1}{b\sqrt{2}}]$$

and hence $W(\alpha,b)$ increases on $\alpha\in[\alpha_-(b),1-\frac{1}{b\sqrt{2}}]$.

In addition, if $\alpha=1-\frac{1}{b\sqrt{2}}$, then

$$|z_0|^2=1,\quad b^2=\tfrac{1}{2(1-\alpha)^2},$$

and so

$$
\begin{aligned}
W(1-\tfrac{1}{b\sqrt{2}},b) &= b^2(2\alpha-5\alpha^2+4\alpha^3)-2\alpha-\log(1-\alpha)\\
&=\frac{2\alpha-5\alpha^2+4\alpha^3}{2(1-\alpha)^2}-2\alpha-\log(1-\alpha)\\
&\geq\frac{2\alpha-5\alpha^2+4\alpha^3}{2(1-\alpha)^2}-2\alpha+\alpha+\frac{\alpha^2}{2}=\frac{\alpha^4}{2(1-\alpha)^2}\geq 0.
\end{aligned}
$$

Therefore,

$$W(\alpha,b)\geq 0,\quad \alpha\in[\alpha_1(b),1-\tfrac{1}{b\sqrt{2}}] \tag{3.55}$$

where $\alpha_1(b)=\alpha_-(b)$ if $W(\alpha_-(b),b)\geq 0$ and $\alpha_1(b)\in[\alpha_-(b),1-\frac{1}{b\sqrt{2}}]$ such that

$$W(\alpha_1(b),b)=0$$

if $W(\alpha_-(b),b)<0$.

According to (3.24), it gives that (3.55) holds for any $b\in[1/\sqrt{2},\sqrt{2}]$, $\max(1,b)\leq|z_0|^2\leq z_2(b)$ where

$$z_2(b)=2\alpha_1(b)(1-\alpha_1(b))^2b^2+1-\alpha_1(b),$$

and the opposite inequality holds if $|z_0|^2>z_2(b)$, as desired.

Notice that one can compute numerically that starting from $b>b_1\approx 1.11$ we get $z_2(b)<b$, and so the interval $\max(1,b)\leq|z_0|^2\leq z_2(b)$ is empty. □

3.3 Integral estimates

Now we proceed to the integral estimates. Consider first the domain Ω_1 where the $*$-saddle point dominates (see Lemmas 4-7). In a standard way the integration domain in (3.1) can be restricted as follows

$$\mathsf{f}_2(z_1,z_2)=\frac{2n^5}{\pi}\int\limits_{\Sigma_r}\triangle^2(T^2)t_1t_2\times e^{nf(T,U,V,v)}d\mu(U)d\mu(V)dTd\bar{v}dv+O(e^{-nr/2}),$$

where

$$\Sigma_r=\{(T,U,V,v)\in\mathcal{D}\mid \|T\|+|v|\le r\}$$

and $\mathcal{D}$ is defined in (3.2).

The next step is to restrict the integration domain by $\frac{\log n}{\sqrt{n}}$-neighborhood of the $*$-saddle point. To this end we need to expand f near the $*$-saddle point (t_*I,x_*):

Lemma 8. *Let $\tilde{T}$ be a 2×2 diagonal matrix such that $\|\tilde{T}\|\le\log n$ and $\tilde{v}=\tilde{x}+i\tilde{y}$ be a complex number with $|\tilde{v}|\le\log n$. Then uniformly in U and V*

$$\begin{aligned}f(t_*I+\tfrac{1}{\sqrt{n}}\tilde{T},U,V,x_*+\tfrac{1}{\sqrt{n}}\tilde{v})=f_{0*}+n^{-1/2}\frac{|z_0|^2+t_*^2}{h_*}\operatorname{tr}(\bar{z}_0\mathcal{Z}+z_0\mathcal{Z}^*)\\+n^{-1}\xi(\tilde{T},U,V,\tilde{v})+O\big(n^{-3/2}\log^3 n\big)\end{aligned}\tag{3.56}$$

where

$$\begin{aligned}\xi(\tilde{T},U,V,\tilde{v})=&\frac{1}{2h_*}\operatorname{tr}\Big[-4(t_*^2+bx_*)\tilde{T}^2-4t_*\tilde{T}P_1-P_1^2+2(|z_0|^2+t_*^2)\mathcal{Z}_U\mathcal{Z}_V^*\Big]\\&+\frac{1}{2h_*}\Big[(2t_*\operatorname{tr}\tilde{T}+\operatorname{tr}P_1)^2+4bt_*\tilde{x}\operatorname{tr}\tilde{T}+2bx_*(\operatorname{tr}\tilde{T})^2+2b^2|\tilde{v}|^2\Big]\\&-\frac{1}{2h_*^2}\Big[2h_*(t_*\operatorname{tr}\tilde{T}+x_*\tilde{x})+(|z_0|^2+t_*^2)\operatorname{tr}P_1\Big]^2-|\tilde{v}|^2;\end{aligned}\tag{3.57}$$

$$\mathcal{Z}=\begin{pmatrix}\zeta_1&0\\0&\zeta_2\end{pmatrix},\quad \mathcal{Z}_B=B^*\mathcal{Z}B;\tag{3.58}$$

$$P_1=\bar{z}_0\mathcal{Z}_U+z_0\mathcal{Z}_V^*,$$

and $f_{0*}=f_0(t_*I,x_*)$.

Proof. If $Q=U(t_*I+n^{-1/2}\tilde{T})V^*$ then the matrix A (2.4) has the form

$$A=\begin{pmatrix}U&0\\0&V\end{pmatrix}\left(A_0+\frac{1}{\sqrt{n}}A_1\right)\begin{pmatrix}U^*&0\\0&V^*\end{pmatrix},$$

where

$$A_0=\begin{pmatrix}-z_0I&t_*I\\-t_*I&-\bar{z}_0I\end{pmatrix},\quad A_1=\begin{pmatrix}-\mathcal{Z}_U&\tilde{T}\\-\tilde{T}&-\mathcal{Z}_V^*\end{pmatrix}.$$

One gets

$$\begin{aligned}\log\det A&=\log\det A_0+\log\det A_0^{-1}A=\log\det A_0+\operatorname{tr}\log(1+n^{-1/2}A_0^{-1}A_1)\\&=\log\det A_0+\frac{1}{\sqrt{n}}\operatorname{tr}A_0^{-1}A_1-\frac{1}{2n}\operatorname{tr}(A_0^{-1}A_1)^2+O\left(\frac{\log^3 n}{\sqrt{n^3}}\right)\end{aligned}\tag{3.59}$$

uniformly in U and V. Moreover,

$$
\begin{aligned}
A_0^{-1} &= \frac{1}{|z_0|^2+t_*^2}\begin{pmatrix} -\bar z_0 I & -t_* I \\ t_* I & -z_0 I\end{pmatrix}, \\
A_0^{-1}A_1 &= \frac{1}{|z_0|^2+t_*^2}\begin{pmatrix} \bar z_0 \mathcal{Z}_U + t_*\tilde T & -\bar z_0 \tilde T + t_* \mathcal{Z}_V^* \\ -t_*\mathcal{Z}_U + z_0\tilde T & t_*\tilde T + z_0 \mathcal{Z}_V^*\end{pmatrix}, \\
\operatorname{tr} A_0^{-1}A_1 &= \frac{1}{|z_0|^2+t_*^2}\operatorname{tr}\left[2t_*\tilde T + P_1\right], \\
\operatorname{tr}(A_0^{-1}A_1)^2 &= \frac{1}{(|z_0|^2+t_*^2)^2}\operatorname{tr}\Big[2(t_*^2-|z_0|^2)\tilde T^2 + 4t_*P_1\tilde T \\
&\quad + \bar z_0^2\mathcal{Z}_U^2 + z_0^2(\mathcal{Z}_V^*)^2 - 2t_*^2\mathcal{Z}_U\mathcal{Z}_V^*\Big]
\end{aligned}
$$

where $P_1 = \bar z_0\mathcal{Z}_U + z_0\mathcal{Z}_V^*$. (3.59) implies

$$
\begin{aligned}
\det A = (|z_0|^2+t_*^2)^2 \Bigg(1 + \frac{1}{\sqrt{n}}\operatorname{tr} A_0^{-1}A_1 \\
+\frac{1}{2n}\left(\operatorname{tr} A_0^{-1}A_1)^2 - \operatorname{tr}(A_0^{-1}A_1)^2\right)\Bigg) + O\left(\frac{\log^3 n}{\sqrt{n^3}}\right).
\end{aligned}
\tag{3.60}
$$

Further,

$$
\begin{aligned}
xt_1t_2 &= x_*t_*^2 + \frac{1}{\sqrt{n}}(t_*^2\tilde x + x_*t_*\operatorname{tr}\tilde T) \\
&\quad + \frac{1}{2n}\left(2t_*\tilde x\operatorname{tr}\tilde T + x_*(\operatorname{tr}\tilde T)^2 - x_*\operatorname{tr}\tilde T^2\right) + O\left(\frac{\log^3 n}{\sqrt{n^3}}\right).
\end{aligned}
\tag{3.61}
$$

Equations (3.60) and (3.61) yield for $Q = U(t_*I + n^{-1/2}\tilde T)V^*$, $v = x_* + n^{-1/2}\tilde v$

$$
\begin{aligned}
h(Q,v) &= h_* + \frac{1}{\sqrt{n}}\left[(|z_0|^2+t_*^2)a_1 + 2bt_*^2\tilde x + 2bx_*t_*\operatorname{tr}\tilde T + 2b^2x_*\tilde x\right] \\
&+ \frac{1}{2n}\left[a_1^2 - a_2 + 4bt_*\tilde x\operatorname{tr}\tilde T + 2bx_*(\operatorname{tr}\tilde T)^2 - 2bx_*\operatorname{tr}\tilde T^2 + 2b\tilde x^2 + 2b\tilde y^2\right] + O\left(\frac{\log^3 n}{\sqrt{n^3}}\right),
\end{aligned}
\tag{3.62}
$$

where h is defined in (3.7), h_* is defined in (3.27), and

$$
\begin{aligned}
a_1 &= (|z_0|^2+t_*^2)\operatorname{tr} A_0^{-1}A_1 = \operatorname{tr}\left[2t_*\tilde T + P_1\right], \\
a_2 &= (|z_0|^2+t_*^2)^2\operatorname{tr}(A_0^{-1}A_1)^2 \\
&= \operatorname{tr}\Big[2(t_*^2-|z_0|^2)\tilde T^2 + 4t_*P_1\tilde T + \bar z_0^2\mathcal{Z}_U^2 + z_0^2(\mathcal{Z}_V^*)^2 - 2t_*^2\mathcal{Z}_U\mathcal{Z}_V^*\Big].
\end{aligned}
\tag{3.63}
$$

Using (3.27) and (3.15), the equation (3.62) can be transformed to

$$
\begin{aligned}
h(Q,v) &= h_* + \frac{1}{\sqrt{n}}\left[2h_*t_*\operatorname{tr}\tilde T + 2h_*x_*\tilde x + (|z_0|^2+t_*^2)\operatorname{tr}P_1\right] \\
&+ \frac{1}{2n}\left[a_1^2 - a_2 + 4bt_*\tilde x\operatorname{tr}\tilde T + 2bx_*(\operatorname{tr}\tilde T)^2 - 2bx_*\operatorname{tr}\tilde T^2 + 2b\tilde x^2 + 2b\tilde y^2\right] + O\left(\frac{\log^3 n}{\sqrt{n^3}}\right).
\end{aligned}
\tag{3.64}
$$

Substituting (3.64) into (3.3) and expanding the logarithm, we obtain

$$f\left(t_*I + \tfrac{1}{\sqrt{n}}\tilde{T}, U, V, x_* + \tfrac{1}{\sqrt{n}}\tilde{v}\right) = f_{0*} + n^{-1/2}\frac{|z_0|^2 + t_*^2}{h_*}\operatorname{tr}(\bar{z}_0\mathcal{Z} + z_0\mathcal{Z}^*)$$
$$+ n^{-1}\Big\{-\operatorname{tr}\tilde{T}^2 - \tilde{x}^2 - \tilde{y}^2 - \frac{1}{2h_*^2}\left[2h_*t_*\operatorname{tr}\tilde{T} + 2h_*x_*\tilde{x} + (|z_0|^2 + t_*^2)\operatorname{tr}P_1\right]^2$$
$$+ \frac{1}{2h_*}\left[a_1^2 - a_2 + 4bt_*\tilde{x}\operatorname{tr}\tilde{T} + 2bx_*(\operatorname{tr}\tilde{T})^2 - 2bx_*\operatorname{tr}\tilde{T}^2 + 2b\tilde{x}^2 + 2b\tilde{y}^2\right]\Big\} + O\Big(\tfrac{\log^3 n}{n^{3/2}}\Big).$$

The last expansion, (3.63), and (3.27) imply (3.56). □

We also need

Lemma 9. *Let the $*$-saddle-point $(t_*, t_*, x_*, 0)$ defined by Proposition 2 be a unique global maximum point of the function $f_0(T, v) = f_0(t_1, t_2, x, y)$. Set*

$$\tilde{f}(T, U, V, v) = f(T, U, V, v) - f_*,$$

where $f_* = f(t_*I, I, I, x_*)$ with f of (3.3). Then for sufficiently large n and $c > 0$ there exists $C > 0$

$$\max_{\frac{\log n}{\sqrt{n}} \le \|T - t_*I\| + |v - x_*| \le c} \tilde{f}(T, U, V, v) \le -C\frac{\log^2 n}{n}$$

uniformly in U and V.

Proof. First let us check that the first derivatives of f_r are bounded in the δ-neighborhood of the manifold (t_*I, U, V, x_*), $U, V \in U(2)$, where f_r is defined in (3.6) and δ is n-independent. Indeed, since h and h_0 are polynomials,

$$\left|\frac{1}{\sqrt{n}}\frac{\partial f_r}{\partial s}\right| = \left|\frac{\partial(f - f_0)}{\partial s}\right| = \left|\frac{\partial(\log h - \log h_0)}{\partial s}\right| \le \left|\frac{1}{h_0}\cdot\frac{\partial h_0}{\partial s} - \frac{1}{h}\cdot\frac{\partial h}{\partial s}\right| \le \frac{C}{\sqrt{n}},$$

where s is either t_1, t_2 x or y. Let T_E be a real diagonal 2×2 matrix of unit norm and let $v_E \in \mathbb{C}$ be a number on the unit circle. Then for any T_E and v_E and for $\frac{\log n}{\sqrt{n}} \le \sigma \le \delta$ we have

$$\frac{d}{d\sigma}\tilde{f}(t_*I + \sigma T_E, U, V, x_* + \sigma v_E) = \langle \nabla_{T,x,y} f_0(t_*I + \sigma T_E, x_* + \sigma v_E), \boldsymbol{u}_E\rangle$$
$$+ n^{-1/2}\langle \nabla_{T,x,y} f_r(t_*I + \sigma T_E, U, V, x_* + \sigma v_E), \boldsymbol{u}_E\rangle$$
$$= \langle \nabla_{T,x,y} f_0(t_*I + \sigma T_E, x_* + \sigma v_E), \boldsymbol{u}_E\rangle + O(n^{-1/2}),$$

where $\boldsymbol{u}_E$ denotes a vector $(t_{E1}, t_{E2}, x_E, y_E)$ and $\langle\cdot,\cdot\rangle$ is a standard real scalar product. Expanding the scalar product by Taylor formula and considering that $\nabla_{T,x,y} f_0(t_*I, x_*) = 0$, we obtain

$$\frac{d}{d\sigma}\tilde{f}(t_*I + \sigma T_E, U, V, x_* + \sigma v_E) = \sigma\langle f_0''(t_*I, x_*)\boldsymbol{u}_E, \boldsymbol{u}_E\rangle + r_1 + O(n^{-1/2}),$$

where f_0'' is a matrix of second order derivatives of f_0 w.r.t. T, x and y and $|r_1| \le C\sigma^2$. $f_0''(t_*I, x_*)$ is non-negative definite, because (t_*I, x_*) is a global maximum. Putting $\mathcal{Z} = 0$ into (3.56) and (3.57) one obtains

$$f_0''(t_*I, x_*) = -2\begin{pmatrix} 2t_*^2 + \frac{bx_*}{h_*} & 2t_*^2 - \frac{2t_*^2 + bx_*}{h_*} & t_*\left(2x_* - \frac{b}{h_*}\right) & 0 \\ 2t_*^2 - \frac{2t_*^2 + bx_*}{h_*} & 2t_*^2 + \frac{bx_*}{h_*} & t_*\left(2x_* - \frac{b}{h_*}\right) & 0 \\ t_*\left(2x_* - \frac{b}{h_*}\right) & t_*\left(2x_* - \frac{b}{h_*}\right) & 1 + 2x_*^2 - \frac{b^2}{h_*} & 0 \\ 0 & 0 & 0 & 1 - \frac{b^2}{h_*} \end{pmatrix}.$$

A straightforward check shows that $\det f_0''(t_*I, x_*) > 0$. Hence $f_0''(t_*I, x_*)$ is negative definite and $\frac{d}{d\sigma}(t_*I + \sigma T_E, U, V, x_* + \sigma v_E)$ is negative and

$$\begin{aligned}\max_{\frac{\log n}{\sqrt{n}} \le \|T - t_*I\| + |v - x_*| \le \delta} \tilde{f}(T, U, V, v) &= \max_{\|T - t_*I\| + |v - x_*| = \frac{\log n}{\sqrt{n}}} \tilde{f}(T, U, V, v) \\ &\le f(t_*I, U, V, x_*) - C\frac{\log^2 n}{n} - f_*.\end{aligned} \tag{3.65}$$

Notice that f_r is bounded from above uniformly in n and $(t_*, t_*, x_*, 0)$ is a point of global maximum of the function f_0. These facts imply that δ in (3.65) can be replaced by c

$$\max_{\frac{\log n}{\sqrt{n}} \le \|T - t_*I\| + |v - x_*| \le c} \tilde{f}(T, U, V, v) \le f(t_*I, U, V, x_*) - f_* - C\frac{\log^2 n}{n}.$$

It remains to deduce from Lemma 8 that $f(t_*I, U, V, x_*) - f_* = O(n^{-1})$ uniformly in U and V. □

Lemma 9 yields

$$\mathsf{f}_2(z_1, z_2) = \frac{2n^5 e^{nf_*}}{\pi}\left(\int\limits_{\Omega_n(t_*I, x_*)} \triangle^2(T^2) t_1 t_2 e^{n\tilde{f}(T,U,V,v)} d\mu(U) d\mu(V) dT d\bar{v} dv + O(e^{-C_1 \log^2 n})\right),$$

where by $\Omega_n(\hat{T}, \hat{v})$ we denote a $\frac{\log n}{\sqrt{n}}$-neighborhood of the point $(\hat{T}, \hat{v})$, i.e.

$$\Omega_n(\hat{T}, \hat{v}) = \left\{(T, U, V, v) \in \mathcal{D} \mid \left\|T - \hat{T}\right\| \le \frac{\log n}{\sqrt{n}}, |v - \hat{v}| \le \frac{\log n}{\sqrt{n}}\right\}, \tag{3.66}$$

Changing the variables $T = t_*I + \frac{1}{\sqrt{n}}\tilde{T}$ and $v = x_* + \frac{1}{\sqrt{n}}\tilde{v}$, and expanding $\triangle^2(T^2)t_1t_2$ and the function f according to Lemma 8, we obtain

$$\mathsf{f}_2(z_1, z_2) = \frac{8t_*^4}{\pi}\mathsf{k}_n \int\limits_{\sqrt{n}\Omega_n(0)} \triangle^2(\tilde{T}) \exp\left\{\xi(\tilde{T}, U, V, \tilde{v})\right\} d\mu(U) d\mu(V) d\tilde{T} d\bar{\tilde{v}} d\tilde{v} (1 + o(1)) \tag{3.67}$$

with ξ of (3.57) and

$$\mathsf{k}_n = n^2 \exp\left\{n f_{0*} + \sqrt{n}\frac{|z_0|^2 + t_*^2}{h_*} \operatorname{tr}(\bar{z}_0 \mathcal{Z} + z_0 \mathcal{Z}^*)\right\}. \tag{3.68}$$

Let us change the variables $V = WU$. Taking into account that the Haar measure is invariant w.r.t. shifts, we get

$$\mathsf{f}_2(z_1, z_2) = \frac{8t_*^4}{\pi}\mathsf{k}_n \int\limits_{\mathbb{R}^4} d\tilde{T} d\tilde{x} d\tilde{y} \int\limits_{U(2)} d\mu(U) \int\limits_{U(2)} d\mu(W) \triangle^2(\tilde{T}) \exp\left\{\xi_1(\tilde{T}, U, W, \tilde{v})\right\}(1 + o(1)),$$

where

$$\begin{aligned}\xi_1(\tilde{T}, U, W, \tilde{v}) = &\frac{1}{2h_*} \operatorname{tr}\left[-4(t_*^2 + bx_*)U\tilde{T}^2U^* - 4t_*U\tilde{T}U^*P - P^2 + 2(|z_0|^2 + t_*^2)\mathcal{Z}\mathcal{Z}_W^*\right] \\ &+ \frac{1}{2h_*}\left[(2t_* \operatorname{tr}\tilde{T} + \operatorname{tr}P)^2 + 4bt_*\tilde{x}\operatorname{tr}\tilde{T} + 2bx_*(\operatorname{tr}\tilde{T})^2 + 2b^2|\tilde{v}|^2\right] - |\tilde{v}|^2 \\ &- \frac{1}{2h_*^2}\left[2t_*h_* \operatorname{tr}\tilde{T} + 2x_*h_*\tilde{x} + (|z_0|^2 + t_*^2)\operatorname{tr}P\right]^2\end{aligned}$$

with

$$P = \bar{z}_0 \mathcal{Z} + z_0 \mathcal{Z}_W^*.$$

The next step is to change the variables $(\tilde{T}, U) \to H$ such that $H = U\tilde{T}U^*$. The Jacobian is $\frac{2}{\pi}\triangle^{-2}(\tilde{T})$ (see e.g. [16]). Thus

$$\mathsf{f}_2(z_0 + \tfrac{\zeta_1}{\sqrt{n}}, z_0 + \tfrac{\zeta_2}{\sqrt{n}}) = \frac{16t_*^4}{\pi^2}\mathsf{k}_n \int\limits_{\mathcal{H}_2} dH \int\limits_{\mathbb{R}^2} d\tilde{x} d\tilde{y} \int\limits_{U(2)} d\mu(W) \exp\left\{\xi_2(H, W, \tilde{v})\right\}(1 + o(1)),$$

where $\mathcal{H}_2$ is a space of hermitian 2×2 matrices and

$$\begin{aligned} dH &= d(H)_{11} d(H)_{22} d\Re(H)_{12} d\Im(H)_{12}, \\ \xi_2(H, W, \tilde{v}) &= \frac{1}{2h_*}\left[-\operatorname{tr}\left(2\sqrt{t_*^2 + bx_*}H + \frac{t_*}{\sqrt{t_*^2 + bx_*}}P\right)^2 + 2(|z_0|^2 + t_*^2)\operatorname{tr}\mathcal{Z}\mathcal{Z}_W^*\right] \\ &+ \frac{1}{2h_*}\left[(2t_*\operatorname{tr}H + \operatorname{tr}P)^2 + 4bt_*\tilde{x}\operatorname{tr}H + 2bx_*(\operatorname{tr}H)^2 + 2b^2|\tilde{v}|^2 - \frac{bx_*}{t_*^2 + bx_*}\operatorname{tr}P^2\right] \\ &- \frac{1}{2h_*^2}\left[2t_*h_*\operatorname{tr}H + 2x_*h_*\tilde{x} + (|z_0|^2 + t_*^2)\operatorname{tr}P\right]^2 - |\tilde{v}|^2. \end{aligned}$$

Shifting the variables $H \to H - \frac{t_*}{2(t_*^2 + bx_*)}P$ and moving integration back to the real axis, one has

$$\mathsf{f}_2(z_0 + \tfrac{\zeta_1}{\sqrt{n}}, z_0 + \tfrac{\zeta_2}{\sqrt{n}}) = \frac{16t_*^4}{\pi^2}\mathsf{k}_n \int\limits_{\mathcal{H}_2} dH \int\limits_{\mathbb{R}^2} d\tilde{x} d\tilde{y} \int\limits_{U(2)} d\mu(W) \exp\left\{\xi_3(H, W, \tilde{v})\right\}(1 + o(1)),$$

where

$$\begin{aligned} \xi_3(H, W, \tilde{v}) = &\frac{1}{2h_*}\operatorname{tr}\left[-\frac{bx_*}{t_*^2 + bx_*}P^2 + 2(|z_0|^2 + t_*^2)\mathcal{Z}\mathcal{Z}_W^*\right] \\ &- \frac{4(t_*^2 + bx_*)}{h_*}|(H)_{12}|^2 - \left(1 - \frac{b^2}{h_*}\right)\tilde{y}^2 - \frac{1}{2}\langle B\boldsymbol{u}, \boldsymbol{u}\rangle - \langle \boldsymbol{q}, \boldsymbol{u}\rangle \\ &+ \frac{bx_*(h_*t_*^2 + 2(h_* - |z_0|^4)bx_*)}{4h_*^2(t_*^2 + bx_*)^2}(\operatorname{tr}P)^2 \end{aligned}$$

with

$$B = 2\begin{pmatrix} 2t_*^2 + \frac{bx_*}{h_*} & 2t_*^2 - \frac{2t_*^2 + bx_*}{h_*} & t_*\left(2x_* - \frac{b}{h_*}\right) \\ 2t_*^2 - \frac{2t_*^2 + bx_*}{h_*} & 2t_*^2 + \frac{bx_*}{h_*} & t_*\left(2x_* - \frac{b}{h_*}\right) \\ t_*\left(2x_* - \frac{b}{h_*}\right) & t_*\left(2x_* - \frac{b}{h_*}\right) & 1 + 2x_*^2 - \frac{b^2}{h_*} \end{pmatrix},$$

$$\boldsymbol{u} = \begin{pmatrix} (H)_{11} \\ (H)_{22} \\ \tilde{x} \end{pmatrix}, \qquad \boldsymbol{q} = \frac{b\operatorname{tr}P}{h_*(t_*^2 + bx_*)}\begin{pmatrix} x_*t_*(2|z_0|^2 - 1) \\ x_*t_*(2|z_0|^2 - 1) \\ t_*^2 + 2|z_0|^2 x_*^2 \end{pmatrix}.$$

The Gaussian integration over H and $\tilde{v}$ implies

$$\mathsf{f}_2(z_0+\tfrac{\zeta_1}{\sqrt{n}}, z_0+\tfrac{\zeta_2}{\sqrt{n}}) = C(\zeta_1,\zeta_2)\mathsf{k}_n \int\limits_{U(2)} d\mu(W) \tag{3.69}$$
$$\times \exp\left\{\frac{1}{h_*}\left[-\frac{bx_*|z_0|^2}{t_*^2+bx_*}+(|z_0|^2+t_*^2)\right]\operatorname{tr}\mathcal{Z}W^*\mathcal{Z}^*W\right\}(1+o(1))$$
$$= C(\zeta_1,\zeta_2)\mathsf{k}_n \int\limits_{U(2)} d\mu(W)\exp\left\{\left[1-\frac{b^2}{h_*}\right]\operatorname{tr}\mathcal{Z}W^*\mathcal{Z}^*W\right\}(1+o(1)),$$

where

$$C(\zeta_1,\zeta_2) = \frac{4\pi t_*^4 h_*}{t_*^2+bx_*}\sqrt{\frac{8h_*}{(h_*-b^2)\det B}}\exp\left\{-\frac{bx_*}{2h_*(t_*^2+bx_*)}\operatorname{tr}(\bar{z}_0^2\mathcal{Z}^2+z_0^2(\mathcal{Z}^*)^2)\right\} \tag{3.70}$$
$$\times\exp\left\{\frac{1}{2}\left\langle B^{-1}\boldsymbol{q},\boldsymbol{q}\right\rangle+\frac{bx_*(h_*t_*^2+2(h_*-|z_0|^4)bx_*)}{4h_*^2(t_*^2+bx_*)^2}(\operatorname{tr}P)^2\right\}.$$

Straightforwardly substituting the expressions (3.22)–(3.23), (3.27) for t_*^2, h_*, and then for $|z_0|^2$ into (3.70) one can get

$$\exp\left\{\frac{1}{2}\left\langle B^{-1}\boldsymbol{q},\boldsymbol{q}\right\rangle+\frac{bx_*(h_*t_*^2+2(h_*-|z_0|^4)bx_*)}{4h_*^2(t_*^2+bx_*)^2}(\operatorname{tr}P)^2\right\} = \exp\left\{\gamma(\operatorname{tr}P)^2/4\right\}.$$

where

$$\gamma = \frac{2b^2(1-(1-4\alpha+2\alpha^2)b^2)}{(1-(1-4\alpha+3\alpha^2)b^2)(1+(2\alpha-2\alpha^2)b^2)} \tag{3.71}$$

and α is as in Lemma 3.

For computing the integral over the unitary group, we use the well-known Harish Chandra/Itsykson–Zuber formula

Proposition 3 (see, e.g., [18], A5)**.** *Let A and B be normal $d\times d$ matrices with distinct eigenvalues $\{a_j\}_{j=1}^d$ and $\{b_j\}_{j=1}^d$ respectively. Then*

$$\int\limits_{U(d)}\exp\{t\operatorname{tr}AU^*BU\}d\mu(U) = \left(\prod_{j=1}^{d-1}j!\right)\frac{\det\{\exp(ta_jb_k)\}_{j,k=1}^d}{t^{(d^2-d)/2}\triangle(A)\triangle(B)},$$

where t is some constant, μ is a Haar measure, and $\triangle(A)=\prod\limits_{j>k}(a_j-a_k)$.

An application of the formula to (3.69) gives

$$\mathsf{f}_2(z_0+\tfrac{\zeta_1}{\sqrt{n}}, z_0+\tfrac{\zeta_2}{\sqrt{n}}) = C(\zeta_1,\zeta_2)\mathsf{k}_n\frac{\det\left\{\exp\left[\left(1-\frac{b^2}{h_*}\right)\zeta_j\bar{\zeta}_k\right]\right\}_{j,k=1}^2}{\left(1-\frac{b^2}{h_*}\right)|\triangle(\mathcal{Z})|^2}(1+o(1)), \tag{3.72}$$

where

$$C(\zeta_1,\zeta_2) = \frac{4\pi t_*^4 h_*}{t_*^2+bx_*}\sqrt{\frac{8h_*}{(h_*-b^2)\det B}}\exp\left\{-\frac{b^2}{2h_*^2}\operatorname{tr}(\bar{z}_0^2\mathcal{Z}^2+z_0^2(\mathcal{Z}^*)^2)+\gamma(\operatorname{tr}P)^2/4\right\} \tag{3.73}$$

with γ of (3.71). If we denote $\beta = 1 - \frac{b^2}{h_*}$ and

$$C_1 = f_{0*}, \tag{3.74}$$

$$C_2(\zeta_1, \zeta_2) = \frac{|z_0|^2 + t_*^2}{h_*} \operatorname{tr}(\bar{z}_0 \mathcal{Z} + z_0 \mathcal{Z}^*), \tag{3.75}$$

then (3.72) implies (1.14). The equation on β follows from

$$\beta = 1 - \frac{b^2(1-\alpha)}{|z_0|^2} \Longrightarrow \alpha = -\frac{(1-\beta)|z_0|^2}{b^2} + 1$$

and (3.23). Here we also used (3.27). Now using

$$\operatorname{tr} P = z_0(\bar{\zeta}_1 + \bar{\zeta}_2) + \bar{z}_0(\zeta_1 + \zeta_2),$$
$$(z_0(\bar{\zeta}_1 + \bar{\zeta}_2) + \bar{z}_0(\zeta_1 + \zeta_2))^2 - 2(\bar{z}_0\zeta_1 + z_0\bar{\zeta}_1)^2 - 2(\bar{z}_0\zeta_2 + z_0\bar{\zeta}_2)^2 = -4(\Re(\bar{z}_0(\zeta_1 - \zeta_2)))^2,$$

we get (1.11).

Suppose now $(b, |z_0|^2) \in \Omega_2$ where the main contribution is given by v-saddle point. Writing

$$v_0 = r_0 e^{i\varphi}, \quad r_0 = \sqrt{1 - |z_0|^4 / b^2}, \quad \tilde{v} = \tilde{r} e^{i\varphi},$$

changing the variables $T = \frac{1}{\sqrt{n}}\tilde{T}$ and $v = (r_0 + \frac{1}{\sqrt{n}}\tilde{r})e^{i\varphi}$ and proceeding similarly to Lemmas 8-9, we obtain

$$\mathsf{f}_2(z_1, z_2) = \frac{2r_0 \mathsf{k}_n^{II}}{\pi} \int\limits_{\sqrt{n}\Omega_n(0)} \triangle^2(\tilde{T}^2)\tilde{t}_1\tilde{t}_2 \exp\left\{\xi^{II}(\tilde{T}, U, V, \tilde{v})\right\} d\mu(U)d\mu(V)d\tilde{T}d\tilde{r}d\varphi(1 + o(1)), \tag{3.76}$$

where

$$\begin{aligned}\xi^{II}(\tilde{T}, U, V, \tilde{r}e^{i\varphi}) &= \frac{1}{2b^2}\operatorname{tr}\left[-2(br_0\cos\varphi + b^2 - |z_0|^2)\tilde{T}^2 - \bar{z}_0^2\mathcal{Z}^2 - z_0^2(\mathcal{Z}^*)^2\right] \\ &+ \frac{1}{2b^2}\left[(\operatorname{tr} P_1)^2 + 2br_0\cos\varphi(\operatorname{tr}\tilde{T})^2\right] - \frac{1}{2b^4}\left[2b^2 r_0\tilde{r} + |z_0|^2\operatorname{tr} P_1\right]^2,\end{aligned}$$

$$\mathsf{k}_n^{II} = n^{1/2}\exp\left\{n\left(-1 + \frac{|z_0|^4}{b^2} + \log b^2\right) + \sqrt{n}\frac{|z_0|^2}{b^2}\operatorname{tr}(\bar{z}_0\mathcal{Z} + z_0\mathcal{Z}^*)\right\}. \tag{3.77}$$

Taking the integral w.r.t. $\tilde{r}$ one gets

$$\mathsf{f}_2(z_1, z_2) = C^{II}\mathsf{k}_n^{II}\exp\left\{\frac{1}{2b^2}\left[(\operatorname{tr}\bar{z}_0\mathcal{Z} + \operatorname{tr} z_0\mathcal{Z}^*)^2 - \bar{z}_0^2\operatorname{tr}\mathcal{Z}^2 - z_0^2\operatorname{tr}(\mathcal{Z}^*)^2\right]\right\}(1 + o(1)),$$

where

$$\begin{aligned}C^{II} &= \sqrt{\frac{2}{\pi}}\int\limits_0^{+\infty} d\tilde{t}_1 \int\limits_0^{+\infty} d\tilde{t}_2 \int\limits_0^{2\pi} d\varphi\, \triangle^2(\tilde{T}^2)\tilde{t}_1\tilde{t}_2(1 + o(1)) \\ &\times \exp\left\{\frac{1}{b^2}\left[-(br_0\cos\varphi + b^2 - |z_0|^2)\operatorname{tr}\tilde{T}^2 + br_0\cos\varphi(\operatorname{tr}\tilde{T})^2\right]\right\}.\end{aligned} \tag{3.78}$$

Since $1/2b^2 = p/4$ according to (1.9), this implies (1.12) and (1.15).

Similarly, if $(b, |z_0|^2) \in \Omega_3$ where zero saddle point is dominant, changing the variables $T = \frac{1}{\sqrt{n}}\tilde{T}$ and $v = \frac{\tilde{v}}{\sqrt{n}}$, one obtains

$$\mathsf{f}_2(z_1, z_2) = \frac{2\mathsf{k}_n^{III}}{\pi} \int\limits_{\sqrt{n}\Omega_n(0)} \triangle^2(\tilde{T}^2)\tilde{t}_1\tilde{t}_2 \exp\left\{\xi^{III}(\tilde{T}, U, V, \tilde{v})\right\} d\mu(U)d\mu(V)d\tilde{T}d\bar{\tilde{v}}d\tilde{v}(1 + o(1)),$$

where

$$\xi^{III}(\tilde{T}, U, V, \tilde{v}) = \frac{1}{2\,|z_0|^4} \operatorname{tr}\left[-2(|z_0|^4 - |z_0|^2)\tilde{T}^2 - \bar{z}_0^2\mathcal{Z}^2 - z_0^2(\mathcal{Z}^*)^2\right] + \frac{b^2}{|z_0|^4}\,|\tilde{v}|^2 - |\tilde{v}|^2\,,$$

and

$$\mathsf{k}_n^{III} = \exp\left\{n\log|z_0|^4 + \frac{\sqrt{n}}{|z_0|^2}\operatorname{tr}\left(\bar{z}_0\mathcal{Z} + z_0\mathcal{Z}^*\right)\right\}. \tag{3.79}$$

Taking integration with respect to $\tilde{t}_1, \tilde{t}_2 > 0$ and with respect to $\tilde{v}$ we get

$$\mathsf{f}_2(z_0 + \tfrac{\zeta_1}{\sqrt{n}}, z_0 + \tfrac{\zeta_2}{\sqrt{n}}) = \frac{\mathsf{k}_n^{III}(1 + o(1))}{\left(1 - \frac{1}{|z_0|^2}\right)^4\left(1 - \frac{b^2}{|z_0|^4}\right)} \exp\left\{-\frac{1}{2\,|z_0|^4}\left[\bar{z}_0^2\operatorname{tr}\mathcal{Z}^2 + z_0^2\operatorname{tr}(\mathcal{Z}^*)^2\right]\right\},$$

which implies (1.13) and (1.16).

Theorem 2 follows from the consideration above for $b = 0$ (see (1.10)) with minor modifications.

Statements and Declarations.

Funding: I.A. partially supported by the Grant EFDS-FL2-08 of the found The European Federation of Academies of Sciences and Humanities (ALLEA). T.S. is supported in part by Alfred P. Sloan Foundation grant FG-2022-18916 and the National Science Foundation grant DMS-2346379.

Data Availability: Data sharing is not applicable to this article as no datasets were generated or analysed during the current study.

References

[1] Afanasiev, I.: On the correlation functions of the characteristic polynomials of the sparse hermitian random matrices. J. Stat. Phys, **163**, 324 – 356 (2016)

[2] Afanasiev, I. :On the correlation functions of the characteristic polynomials of non-Hermitian random matrices with independent entries, J.Stat.Phys. 176, 1561 - 1582 (2019)

[3] Afanasiev, I. On the correlation functions of the characteristic polynomials of random matrices with Independent entries: interpolation between complex and real cases. Journal of Mathematical Physics, Analysis, Geometry, 18:2, p. 159–181 (2022)

[4] Afanasiev, I., Shcherbina, M., Shcherbina, T. Universality of the second correlation function of the deformed Ginibre ensemble, arXiv:2405.00617 (2024)

[5] Akemann, G. ,Phillips, M.J. and Sommers, H.-J.: Characteristic polynomials in real Ginibre ensembles, J. Phys. A: Math. Theor. 42 (2009)

[6] Akemann, G. , Vernizzi, G., Characteristic Polynomials of Complex Random Matrix Models, Nucl. Phys. B, 3:600, p. 532-556 (2003)

[7] Basak, A and Rudelson, M.: The circular law for sparse non-Hermitian matrices, Ann. Probab. 47, no. 4, 2359 -2416 (2019)

[8] Berezin, F. A. Introduction to superanalysis, Math. Phys. Appl. Math. (9), D. Reidel Publishing Co., Dordrecht, 1987. Edited and with a foreword by A. A. Kirillov. With an appendix by V. I. Ogievetsky. Translated from the Russian by J. Niederle and R. Kotecký. Translation edited by Dimitri Leĭtes.

[9] Brézin, E., Hikami, S.: Characteristic polynomials of random matrices. Commun. Math. Phys. 214, p. 111 – 135 (2000)

[10] Brézin, E., Hikami, S.: Characteristic polynomials of real symmetric random matrices. Commun. Math. Phys., vol. 223, p. 363 – 382 (2001)

[11] Cipolloni, G., Erdős, L., Schröder, D.: Edge universality for non-Hermitian random matrices, Prob.Theory and Related Fields **179**, 1– 28 (2021)

[12] Dubova, S. and Yang, K. Bulk universality for complex eigenvalues of real non-symmetric random matrices with iid entries,arXiv:2402.10197 (2024)

[13] Efetov, K.: Supersymmetry in disorder and chaos. Cambridge university press, New York (1997)

[14] Götze, F. , Tikhomirov, A.: The circular law for random matrices, Ann. Probab. 38 , no. 4, 1444–1491 (2010)

[15] He, Y.: Edge universality of sparse Erdős-Rényi digraphs, arXiv:2304.04723v2 (2023)

[16] Hua, L. K. Harmonic Analysis of Functions of Several Complex Variables in the Classical Domains, American Mathematical Society, Providence, RI, 1963

[17] Maltsev, A., Osman, M. Bulk universality for complex non-hermitian matrices with independent and identically distributed entries, arXiv:2310.11429v4 (2023)

[18] Mehta, M.L. Random Matrices. Academic Press Inc., Boston, second edition, 1991

[19] Mirlin, A. D.: Statistics of energy levels. New Directions in Quantum Chaos, (Proceedings of the International School of Physics Enrico Fermi, Course CXLIII), ed. by G.Casati, I.Guarneri, U.Smilansky, IOS Press, Amsterdam, 223-298 (2000)

[20] Osman, M. Bulk universality for real matrices with independent and identically distributed entries,arXiv:2402.04071 (2024)

[21] Rudelson, M. and Tikhomirov, K.: The sparse circular law under minimal assumptions, Geom. Funct. Anal. 29 , no. 2, 561– 637 (2019)

[22] Sah, A., Sahasrabudhe, J. and Sawhney, M.,The limiting spectral law for sparse iid matrices, https://arxiv.org/abs/2310.17635 (2023)

[23] Shcherbina, M., Shcherbina, T.: Universality for 1 d random band matrices, Commun. Math. Phys. 385, 667 – 716 (2021)

[24] Shcherbina, T.: On the correlation function of the characteristic polynomials of the Hermitian Wigner ensemble. Commun. Math. Phys., vol. 308, p. 1 – 21 (2011),

[25] Shcherbina, T.: On the correlation functions of the characteristic polynomials of the hermitian sample covariance ensemble, Probab. Theory Relat. Fields, vol. 156, p. 449 – 482 (2013)

[26] Shcherbina, T. : On the second mixed moment of the characteristic polynomials of the 1D band matrices. Commun. Math. Phys., vol. 328, p. 45 – 82 (2014)

[27] Strahov and Y. V. Fyodorov, Y.V., Strahov, E. Universal results for correlations of characteristic polynomials: Riemann-Hilbert approach Comm. Math. Phys., 241:2-3, p. 343-382 (2003)

[28] Tao,T., Vu, V.: Random matrices: the circular law. Commun. Contemp. Math., 10(2):261–307 (2008)

[29] Tao, T., Vu, V.: Random matrices: universality of local spectral statistics of non- Hermitian matrices, Ann. Probab. 43, 782–874 (2015)

[30] Wood, P. M.. Universality and the circular law for sparse random matrices. Ann. Appl. Probab., 22(3):1266–1300 (2012)