Euler characteristic heuristic for approximating the distribution of the largest eigenvalue of an orthogonally invariant random matrix

Satoshi Kuriki [a] and Akimichi Takemura [b]

[a] *Institute of Statistical Mathematics*
4-6-7 Minami-azabu, Minato-ku, Tokyo 106-8569, Japan
Email: `kuriki@ism.ac.jp`

[b] *Graduate School of Information Science and Technology*
University of Tokyo
7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan
Email: `takemura@stat.t.u-tokyo.ac.jp`

Abstract

The Euler characteristic heuristic has been proposed as a method for approximating the upper tail probability of the maximum of a random field with smooth sample path. When the random field is Gaussian, this method is proved to be valid in the sense that the relative approximation error is exponentially smaller. However, very little is known about the validly of the method when the random field is non-Gaussian. In this paper, as a milestone to developing the general theory about the validity of the Euler characteristic heuristic, we examine the Euler characteristic heuristic for approximating the distribution of the largest eigenvalue of an orthogonally invariant non-Gaussian random matrix. In this particular example, if the probability density function of the random matrix converges to zero sufficiently fast at the the boundary of its support, the approximation error of the Euler characteristic heuristic is proved to be small and the approximation is valid. Moreover, for several standard orthogonally invariant random matrices, the approximation formula for the distribution of the largest eigenvalue and its asymptotic error are obtained explicitly. Our formulas are practical enough for the purpose of numerical calculations.

Key words: Random field, Morse's theorem, Tube method, Wishart distribution, Multivariate beta distribution, Inverse Wishart distribution

MAY 14, 2005

1 Introduction

Let $X(t)$ $(t \in S)$ be a random field on the index set $S \subset \mathbb{R}^n$ with smooth sample path. In this paper, we treat the approximation of the distribution of the maximum $\sup_{t\in S} X(t)$ of the random field. This distribution is required in testing hypothesis based on the maximum type statistic and multiple comparisons.

The set of indices t such that the value $X(t)$ exceeds a threshold x,

$$A_x = \{t \in S \mid X(t) \geq x\},$$

is called an *excursion set*. Note that A_x is a random set. By its definition,

$$P\Bigl(\sup_{t\in S} X(t) \geq x\Bigr) = E[1_{A_x}],$$

where 1 is the indicator function of a set defined by

$$1_{A_x} = \begin{cases} 1 & (A_x \neq \emptyset) \\ 0 & (A_x = \emptyset). \end{cases}$$

The Euler characteristic of the set A_x is denoted by $\chi(A_x)$. $\chi(A_x)$ is an integer-valued topological invariant. It holds that

$$\chi(A_x) = \begin{cases} 1 & (A_x \text{ is contractible}) \\ 0 & (A_x = \emptyset) \end{cases}$$

in particular. Note that a set is called contractible if the set can be transformed into a point set by a continuous map (i.e., the set is homotopy equivalent to a point set).

Suppose that there exists a unique $t = t^*$ which attains $\sup_{t\in S} X(t)$. Consider the case where x is large. If $x > X(t^*)$, then $A_x = \emptyset$. Even if $x \leq X(t^*)$, but if x is close to the maximum $X(t^*)$, then the excursion set A_x is expected to be a set homotopy equivalent to the point set $\{t^*\}$. Namely,

$$1_{A_x} \approx \chi(A_x) \quad \text{(for } x \text{ large).} \tag{1}$$

By taking the expectations of both sides, we have the approximation formula

$$P\Bigl(\sup_{t\in S} X(t) \geq x\Bigr) \approx E[\chi(A_x)] \quad \text{(for } x \text{ large).} \tag{2}$$

This approximation is called *Euler characteristic heuristic* (Adler (1981), Worsley (1994), Worsley (1995), Adler (2000), Taylor and Adler (2003)).

The reason why this approximation is proposed is that, in taking the expectations, $\chi(A_x)$ is easier to handle than 1_{A_x} in general. However, as we already saw, the approximation (1) is based on an intuitive consideration, and the validity of the approximation (2) is not obvious.

Actually, when the random field $X(t)$ is Gaussian with the mean 0 and constant variance, under suitable regularity conditions, the Euler characteristic heuristic is proved to be valid in the sense that the relative approximation error is exponentially smaller (Kuriki and Takemura (2001), Takemura and Kuriki (2002), Taylor et al. (2005)). However, very little is known about the validly of the method when the random field is non-Gaussian.

In this paper, as a milestone to developing the general theory about the validity of the Euler characteristic heuristic, we examine the approximation error of the Euler characteristic heuristic for approximating the distribution of the largest eigenvalue of an orthogonally invariant non-Gaussian random matrix. In Section 2, we formulate the largest eigenvalue of real symmetric random matrix as the maximum of a random field defined by the quadratic form, and give the approximation formula and its asymptotic error. It is shown that, if the probability density function of the random matrix converges to zero sufficiently fast at the the boundary of its support, the approximation error of the Euler characteristic heuristic is proved to be small, and the approximation method is valid. In Section 3, for several standard orthogonally invariant random matrices, the approximation formulas for the distribution of the largest eigenvalues and their asymptotic error are obtained explicitly. Usually the exact formulas for the distribution about the eigenvalues of random matrix are described in terms of special functions of matrix arguments (e.g., Muirhead (1982), Section 9.7). However, they are very complicated and not suitable for numerical calculations. The formulas obtained in this paper are practical enough for the purpose of numerical calculations.

2 Approximation of the distribution of the largest eigenvalue

2.1 Euler characteristic and its expectation

The set of $p \times p$ real symmetric matrices is denoted by $Sym(p)$, and the set of $p \times p$ orthogonal matrices is denoted by $O(p)$. By identifying the upper triangular elements of the matrix in $Sym(p)$ with the element of the Euclidean space $\mathbb{R}^{p(p+1)/2}$, we denote the Lebesgue measure of $Sym(p)$ at $W \in Sym(p)$ by $dW = \prod_{i \le j} dw_{ij}$. In this paper, we treat the case where the random matrix $W \in Sym(p)$ has a probability density function $f(W)$ with respect to the Lebesgue measure dW. Moreover, we assume that W is orthogonally invariant, that is, the distributions of W and QWQ' are the same for all $Q \in O(p)$. This assumption means that

$$f(W) = f(QWQ') \quad \text{for all } Q \in O(p).$$

Using the random matrix W, define a random field on the index set $\mathbb{S}^{p-1}$, the $p-1$ dimensional unit sphere, by the quadratic form

$$X(h) = h'Wh, \quad h \in \mathbb{S}^{p-1}.$$

Since the maximum of $X(h)$ is the largest eigenvalue of W,

$$\lambda_1(W) = \max_{h \in \mathbb{S}^{p-1}} X(h),$$

we can consider the approximation of the distribution of the largest eigenvalue of W by virtue of the Euler characteristic heuristic.

For an arbitrarily fixed $Q \in O(p)$, define a random field $Y(h) = X(Qh)$ on the index set $\mathbb{S}^{p-1}$. Because of the orthogonal invariance, the finite dimensional marginal distribution of $X(\cdot)$ is the same as that of $Y(\cdot)$. This means that $X(\cdot)$ is stationary with respect to the group action $O(p)$. Note that the marginal distribution of $X(h)$ does not depend on h.

In the following, we will derive an expression of $\chi(A_x)$ with the help of Morse's theorem. We begin with showing that $X(h)$ is a Morse function.

Let $t = (t^1, \ldots, t^{p-1})$ be a local coordinate of $\mathbb{S}^{p-1}$ so that $h = h(t)$. We use the abbreviations $h_i = \partial h/\partial t^i$, $h_{ij} = \partial^2 h/\partial t^i \partial t^j$, $X_i = \partial X(h)/\partial t^i$, *etc*. Note first that $h'h = 1$, $h_i'h = 0$, and $h_{ij}'h + h_i'h_j = 0$. Moreover

$$X_i = 2h_i'Wh, \qquad X_{ij} = 2h_i'Wh_j + 2h_{ij}'Wh.$$

Since $h, h_1, \ldots, h_{p-1}$ span $\mathbb{R}^p$, we can write

$$h_{ij} = c_{ij}h + d_{ij}^k h_k.$$

Here we adopt Einstein's convention of omitting summation symbols. Let $c_{ij} = h'h_{ij} = -h_i'h_j = -g_{ij}$ (say). Because $h_l'h_{ij} = d_{ij}^k g_{kl}$, it holds that $d_{ij}^k = (h_l'h_{ij})g^{kl} = \Gamma_{ij}^k$ (say), where g^{kl} is the element of the inverse of the matrix (g_{kl}). g_{ij} and Γ_{ij}^k are the metric tensor and the connection coefficient, respectively. Using these geometric quantities, the second derivatives are expressed as

$$\begin{aligned} h_{ij} &= -g_{ij}h + \Gamma_{ij}^k h_k, \\ X_{ij} &= 2h_i'Wh_j - 2g_{ij}h'Wh + 2\Gamma_{ij}^k h_k'Wh. \end{aligned}$$

The *critical point* $h = h^*$ is defined to be a point such that $X_i(h) = 0$ for all i. If the second derivative matrix $(X_{ij}(h))$ is not degenerate at each critical point h^*, the function $X(h)$ is said to be *non-degenerate* (or a *Morse function*).

Lemma 1 *The random field $X(h)$ is non-degenerate with probability one. The set of critical points is a finite set with probability one.*

Proof. By the Gram-Schmidt orthonormalization, let

$$(h_1, \ldots, h_{p-1}) = HT, \tag{3}$$

where H is a $p \times (p-1)$ matrix satisfying $H'H = I_{p-1}$, and T is a $(p-1)\times(p-1)$ upper triangular matrix with positive diagonal elements. Note that both H and T are defined as functions of h (i.e., $H = H(h)$, $T = T(h)$). The metric tensor is $(g_{ij}) = T'T$. Since T is non-singular, h is a critical point iff $h'WH = 0$. Moreover, because $h'H = 0$, h is a critical point iff h is an eigenvector of W.

Noting that $(X_{ij}(h)) = 2T'H'WHT - 2T'T(h'Wh)$ at a critical point h, we see that $X(h)$ is degenerate at a critical point h iff

$$\det((h'Wh)I_{p-1} - H'WH) = 0. \tag{4}$$

The ith largest eigenvalue of W is denoted by $\lambda_i(W)$. Suppose that a critical point h is an eigenvector of W with respect to the eigenvalue $\lambda_k(W)$. Then $H'WH$ has eigenvalues $\lambda_i(W)$ $(i \neq k)$, and (4) implies $\lambda_i(W) = \lambda_k(W)$ for some $i \neq k$. On the other hand, the random matrix W whose distribution is absolutely continuous with respect to the Lebesgue measure dW on $Sym(p)$ has distinct eigenvalues with probability one. (This follows from the facts that the discriminant of the eigenfunction of W is a polynomial of the elements of W, and that the Lebesgue measure of the zero-point set of a polynomial is 0. See Okamoto (1973).) Therefore, $X(h)$ is non-degenerate with probability one. Also, the set of critical points consists of p distinct points with probability one. $\square$

Lemma 2 *Let $h_{(k)}$ denote the eigenvector with respect to the kth largest eigenvalue of W. Then*

$$\chi(A_x) = \sum_{k=1}^{p} 1_{\{h'Wh \ge x\}} \operatorname{sgn} \det((h'Wh)I_{p-1} - H'WH)\Big|_{h=h_{(k)}} \tag{5}$$

$$= \sum_{k=1}^{p} (-1)^{k-1} 1_{\{\lambda_k(W) \ge x\}} \tag{6}$$

holds with probability one.

Proof. Obviously, $-X(h)$ is also non-degenerate with probability one. Noting that $-X(h)$ is a C^∞-function on a C^∞-manifold, and by applying Morse's theorem (Morse and Cairns (1969), Theorem 10.2; Worsley (1995), Theorem 1) to $-X(h)$, we have

$$\chi(A_x) = \sum_{h:\text{critical point}} 1_{\{X(h)\ge x\}} \operatorname{sgn}\det(-X_{ij}(h))$$

with probability one. The first equality now follows from the fact that the critical points of $X(h)$ are the eigenvectors $h_{(k)}$ $(1 \le k \le p)$ of W. Moreover, the right hand side of the first equality is rewritten as

$$\sum_{k=1}^{p} 1_{\{\lambda_k(W)\ge x\}} \operatorname{sgn} \prod_{i\ne k} (\lambda_k(W) - \lambda_i(W)),$$

from which the second equality follows. $\square$

Remark 3 *From (6) with probability one, $\chi(A_x)$ takes the value 0 or 1, and hence the expectation $E[\chi(A_x)]$ exists. Also $\chi(A_x) - 1_{A_x}$ takes the value 0 or -1, and hence $E[\chi(A_x)] \le P(\lambda_1(W) \ge x)$ holds for all x.*

Next we will derive the Jacobian in order to take the expectation of $\chi(A_x)$. Let $h \in \mathbb{S}^{p-1}$ be a p dimensional unit vector such that the first non-zero element is positive. Let $H = H(h)$ be a $p \times (p-1)$ matrix defined by (3). Consider the transform

$$W = (h, H) \begin{pmatrix} \tilde{w}_{11} & 0 \\ 0 & \tilde{W}_{22} \end{pmatrix} \begin{pmatrix} h' \\ H' \end{pmatrix}. \tag{7}$$

This transform represents that $\tilde{w}_{11}$ is an eigenvalue of W, and h is a corresponding eigenvector. By imposing a restriction that, for fixed k $(1 \le k \le p)$, $\tilde{w}_{11}$ is the kth largest eigenvalue of W (i.e., h is the kth eigenvector of W), the relation (7) becomes an one-to-one relation with probability one. Note that this restriction is rewritten as

$$\lambda_{k-1}(\tilde{W}_{22}) > \tilde{w}_{11} > \lambda_k(\tilde{W}_{22}). \tag{8}$$

By standard techniques of multivariate analysis, the Jacobian of (7) is shown to be

$$dW = d\tilde{w}_{11} d\tilde{W}_{22} |\det(\tilde{w}_{11} I_{p-1} - \tilde{W}_{22})| H' dh,$$

where $H'dh$ is a volume element of the unit sphere $\mathbb{S}^{p-1}$. Due to the restriction that the first non-zero element of h is positive, the total integral with respect to the volume element is a half of the volume Ω_p of $\mathbb{S}^{p-1}$, that is,

$$\int H' dh = \frac{\Omega_p}{2}, \qquad \Omega_p = \frac{2\pi^{\frac{p}{2}}}{\Gamma(\frac{p}{2})}.$$

Because $f(W) = f(\mathrm{diag}(\tilde{w}_{11}, \tilde{W}_{22}))$, the marginal distribution of $(\tilde{w}_{11}, \tilde{W}_{22})$ within the range (8) is

$$\frac{\Omega_p}{2} f(\mathrm{diag}(\tilde{w}_{11}, \tilde{W}_{22})) |\det(\tilde{w}_{11} I_{p-1} - \tilde{W}_{22})| d\tilde{w}_{11} d\tilde{W}_{22}. \tag{9}$$

By combining the density function (9) and the expression (5) of $\chi(A_x)$ in Lemma 2, we get the following.

Theorem 4 *The approximation formula $\widehat{P}(x)$ by the Euler characteristic heuristic is given by*

$$\begin{aligned}\widehat{P}(x) &= E[\chi(A_x)] \\ &= \frac{\Omega_p}{2} \iint\limits_{\mathbb{R}\times Sym(p-1)} dw_{11} dW_{22} 1_{\{w_{11} \geq x\}} \\ &\qquad \times \det(w_{11} I_{p-1} - W_{22}) f(\mathrm{diag}(w_{11}, W_{22})).\end{aligned}$$

2.2 Exact distribution and approximation error

Consider the transform (7) in the range that $\tilde{w}_{11}$ is the largest eigenvalue of W. The marginal distribution of $(\tilde{w}_{11}, \tilde{W}_{22})$ is

$$\frac{\Omega_p}{2} f(\mathrm{diag}(\tilde{w}_{11}, \tilde{W}_{22})) \det(\tilde{w}_{11} I_{p-1} - \tilde{W}_{22}) d\tilde{w}_{11} d\tilde{W}_{22} \quad (\tilde{w}_{11} I_{p-1} > \tilde{W}_{22}).$$

Here the inequality '>' of matrices means that the difference between the left hand side and the right hand side is positive definite. By taking the integral over the range $\tilde{w}_{11} \geq x$, we get the upper probability $P(x)$ of the largest eigenvalue $\lambda_1(W)$ and the approximation error $\Delta P(x) = \widehat{P}(x) - P(x)$ of the Euler characteristic heuristic as

$$\begin{aligned}P(x) &= P(\lambda_1(W) \geq x) \\ &= \frac{\Omega_p}{2} \iint\limits_{\mathbb{R}\times Sym(p-1)} dw_{11} dW_{22} 1_{\{w_{11} \geq x\}} 1_{\{w_{11} I_{p-1} > W_{22}\}} \\ &\qquad \times \det(w_{11} I_{p-1} - W_{22}) f(\mathrm{diag}(w_{11}, W_{22}))\end{aligned}$$

and

$$
\begin{aligned}
\Delta P(x) &= \widehat{P}(x) - P(x) \\
&= \frac{\Omega_p}{2} \iint_{\mathbb{R}\times Sym(p-1)} dw_{11} dW_{22} 1_{\{w_{11}\geq x\}} (1 - 1_{\{w_{11} I_{p-1} > W_{22}\}}) \\
&\qquad \times \det(w_{11} I_{p-1} - W_{22}) f(\mathrm{diag}(w_{11}, W_{22})),
\end{aligned}
$$

respectively. Note that $\Delta P(x) \leq 0$ for all x by Remark 3.

Let $l_1 = w_{11}$ and denote the eigenvalues of W_{22} by $l_2 > \cdots > l_p$. Let $L = \mathrm{diag}(l_i)_{1\leq i\leq p}$ be the diagonal matrix consisting of l_i's, and $dL = \prod_{i=1}^p dl_i$. By successive decomposition of W_{22} similar to (7), we get an expression in terms of the integral in eigenvalues

$$
P(x) = \prod_{i=1}^{p} \frac{\Omega_i}{2} \int_{\substack{l_1 \geq x \\ l_1 > l_2 > \cdots > l_p}} f(L) \prod_{i<j} (l_i - l_j) dL.
$$

Similarly,

$$
\widehat{P}(x) = \prod_{i=1}^{p} \frac{\Omega_i}{2} \int_{\substack{l_1 \geq x \\ l_2 > \cdots > l_p}} f(L) \prod_{i<j} (l_i - l_j) dL,
$$

$$
\Delta P(x) = \prod_{i=1}^{p} \frac{\Omega_i}{2} \int_{\substack{l_2 > l_1 \geq x \\ l_2 > \cdots > l_p}} f(L) \prod_{i<j} (l_i - l_j) dL.
$$

2.3 A class of orthogonally invariant distributions

Furthermore in this paper, we restrict our attention to the case where the density function is written as a product of non-negative functions in eigenvalues l_i of W,

$$
f(W) = f(L) = \prod_{i=1}^{p} g(l_i). \tag{10}
$$

The Wishart distribution, the multivariate beta distribution, the inverse Wishart distribution, and $p \times p$ multivariate symmetric normal distribution which has the density function

$$
f(A) = \frac{1}{e(p)} e^{-\frac{1}{2}\mathrm{tr}(A^2)} \quad (A \in Sym(p))
$$

with $e(p) = \pi^{p(p+1)/4} 2^{p/2}$ belong to the class (10). Also, the restriction of these distributions such that the eigenvalues are restricted to a particular region, for example, the conditional distribution of the $p \times p$ matrix A distributed as the multivariate symmetric normal distribution given the condition that A is positive definite is in the class (10).

Here we make an assumption on $g(\cdot)$.

Assumption 5 *Throughout the paper, let* $a = \inf\{l \mid g(l) > 0\}$ *and* $b = \sup\{l \mid g(l) > 0\}$. $g(\cdot)$ *is non-negative and piecewise continuous on* (a, b). *There exists* $c < b$ *such that* $g(l) > 0$ *for* $l \in (c, b)$.

The property that $f(\cdot)$ is a probability density function characterizes the function $g(\cdot)$.

Lemma 6 *The integral of* $f(\cdot)$ *in (10) over* $Sym(p)$ *is finite if and only if*

$$\int_a^b (|l|^{p-1} + 1) g(l) dl < \infty. \tag{11}$$

Proof. We begin with proving that if

$$\int_{b > l_1 > \cdots > l_p > a} \prod_i g(l_i) \prod_{i<j} (l_i - l_j) dl_1 \cdots dl_p < \infty, \tag{12}$$

then (11) holds. Let $b\ (= c_0) > c_1 > \cdots > c_n > a\ (= c_{n+1})$ be the discontinuities of $g(\cdot)$. We only have to check the convergence of (11) at $l = c_k$ $(0 \le k \le n+1)$.

It holds for $c \in (c_{k+1}, c_k)$ that

$$\int_c^{c_k} g(l_1) F(l_1) dl_1 < \infty$$

with

$$F(l) = \int_{l > l_2 > \cdots > l_p > c_{k+1}} \prod_{i \ge 2} g(l_i)(l - l_i) \prod_{2 \le i < j} (l_i - l_j) dl_2 \cdots dl_p.$$

Since $F(\cdot)$ is monotonically non-decreasing and $F(c) > 0$, it holds that

$$\int_c^{c_k} g(l) dl \le \frac{1}{F(c)} \int_c^{c_k} g(l) F(l) dl < \infty. \tag{13}$$

Also when $c_k = \infty$ (i.e., $k = 0$), it holds for $c > \max(c_1, 0)$ that

$$\int_c^\infty l_1^{p-1} g(l_1) F(l_1) dl_1 < \infty$$

with

$$F(l) = \int_{l > l_2 > \cdots > l_p > \max(c_1, 0)} \prod_{i \geq 2} g(l_i)\Big(1 - \frac{l_i}{l}\Big) \prod_{2 \leq i < j} (l_i - l_j) dl_2 \cdots dl_p.$$

Since $F(\cdot)$ is monotonically non-decreasing and $F(c) > 0$, it holds that

$$\int_c^\infty l^{p-1} g(l) dl \leq \frac{1}{F(c)} \int_c^\infty l^{p-1} g(l) F(l) dl < \infty. \tag{14}$$

From (13) and (14), in both cases $c_k = \infty$ and $c_k < \infty$, it holds that

$$\int_c^{c_k} (|l|^{p-1} + 1) g(l) dl < \infty \quad \text{for some } c \in (c_{k+1}, c_k).$$

We can prove similarly that

$$\int_{c_k}^{c} (|l|^{p-1} + 1) g(l) dl < \infty \quad \text{for some } c \in (c_k, c_{k-1})$$

and (11) follows.

Next we will prove that (11) implies (12). By expanding the linkage factor $\prod_{i<j}(l_i - l_j)$ in the integrand in (12), we see that (12) is a finite sum of the terms of the form

$$\int_{b > l_1 > \cdots > l_p > a} \prod_i g(l_i) l_i^{q_i} dl_1 \cdots dl_p \quad (0 \leq q_i \leq p-1). \tag{15}$$

Noting that $|l_i^{q_i}| \leq |l_i|^{p-1} + 1$, we see that the integral above is bounded above by

$$\int_{b > l_1 > \cdots > l_p > a} \prod_i g(l_i)(|l_i|^{p-1} + 1) dl_1 \cdots dl_p \leq \Big\{ \int_a^b g(l)(|l|^{p-1} + 1) dl \Big\}^p.$$

□

By the use of the cofactor expansion of the Vandermonde determinant,

$$\begin{aligned}\prod_{i<j}(l_i - l_j) &= \det(l_i^{p-j})_{1\le i,j\le p} \\ &= \sum_{j=0}^{p-1}(-1)^j l_1^{p-1-j}\mathrm{tr}_j(l_2,\ldots,l_p)\prod_{2\le i<j}(l_i - l_j) \qquad (16)\end{aligned}$$

or

$$\begin{aligned}\prod_{i<j}(l_i - l_j) &= \prod_{i<j}\{(b-l_j)-(b-l_i)\} \\ &= \sum_{j=0}^{p-1}(-1)^j (b-l_1)^j \mathrm{tr}_{p-1-j}(b-l_2,\ldots,b-l_p)\prod_{2\le i<j}(l_i - l_j) \qquad (17)\end{aligned}$$

($\mathrm{tr}_j(\cdot)$ is the jth elementary symmetric polynomial), we get expressions for $\widehat{P}(x)$.

Theorem 7 *When the density function is of the form (10), the approximation $\widehat{P}(x)$ by the Euler characteristic heuristic is written as follows.*
(i) The case $b=\infty$.

$$\begin{aligned}\widehat{P}(x) = \prod_{i=1}^{p}\frac{\Omega_i}{2}\sum_{j=0}^{p-1}(-1)^j\int_x^\infty l_1^{p-1-j}g(l_1)dl_1 \int_{\infty>l_2>\cdots>l_p>a}\mathrm{tr}_j(l_2,\ldots,l_p) \\ \times\prod_{2\le i<j}(l_i-l_j)g(l_2)\cdots g(l_p)dl_2\cdots dl_p. \qquad (18)\end{aligned}$$

(ii) The case $b<\infty$.

$$\begin{aligned}\widehat{P}(x) = \prod_{i=1}^{p}\frac{\Omega_i}{2}\sum_{j=0}^{p-1}(-1)^j\int_x^b (b-l_1)^j g(l_1)dl_1 \\ \times\int_{b>l_2>\cdots>l_p>a}\mathrm{tr}_{p-1-j}(b-l_2,\ldots,b-l_p) \\ \times\prod_{2\le i<j}(l_i-l_j)g(l_2)\cdots g(l_p)dl_2\cdots dl_p. \qquad (19)\end{aligned}$$

In addition, in each summation $\sum_{j=0}^{p-1}$ in (18) and (19), the term indexed by j is asymptotically smaller than that indexed by $j-1$ for $1\le j\le p-1$ as $x\uparrow b$. That is, the summation represents an expansion with terms ranked in descending order of asymptotic magnitude.

Recall that, for two functions $f(x)$ and $g(x)$, $f(x)$ is said to be *asymptotically smaller* than $g(x)$ if $f(x)=o(g(x))$ as $x\uparrow b$, or equivalently $\lim_{x\uparrow b} f(x)/g(x) =$

0. Before proving the theorem, we cite L'Hospital's rule (L'Hôpital's rule), which shall be used frequently in this paper (Rudin (1976), Theorem 5.13).

Theorem 8 *Let $f(x)$ and $g(x)$ be differential on an interval (c, b) $(c < b \leq \infty)$ for which $g'(x) \neq 0$. If both $\lim_{x\uparrow b} f(x)$ and $\lim_{x\uparrow b} g(x)$ are 0, or both are $\pm\infty$, and if the limit $\lim_{x\uparrow b} f'(x)/g'(x) = A$ $(-\infty \leq A \leq \infty)$ exists, then $\lim_{x\uparrow b} f(x)/g(x) = A$.*

Proof of Theorem 7. Confirm first that the definite integrals appearing in the expressions are finite. The finiteness of the integral with respect to l_1 in (18) follows immediately from (11), and the finiteness of the integral with respect to l_i $(i \geq 2)$ follows from the fact that the integral is a sum of the terms of the form (15). Finiteness of the integrals for (19) is established similarly.

In addition, by virtue of L'Hospital's rule (Theorem 8), for $0 \leq q, q' \leq p - 1$

$$\lim_{x\to\infty} \frac{\int_x^\infty l^q g(l)dl}{\int_x^\infty l^{q'} g(l)dl} = \lim_{x\to\infty} \frac{x^q g(x)}{x^{q'} g(x)} = \begin{cases} \infty & (q > q') \\ 1 & (q = q') \\ 0 & (q < q') \end{cases}$$

holds, from which (18) is shown to be an expansion of descending order. Here the conditions for L'Hospital's rule are assured from Assumption 5 and Lemma 6. Similarly (19) is shown to be an expansion of descending order. $\square$

2.4 Asymptotic evaluation of the approximation error

When the density function is assumed to be of the form (10), the approximation error $\Delta P(x)$ in terms of the integral in eigenvalues is expressed as

$$\Delta P(x) = \prod_{i=1}^{p} \frac{\Omega_i}{2} \int_{\substack{b>l_2>l_1\geq x \\ b>l_2>\cdots>l_p>a}} \prod_{i=1}^{p} g(l_i) \prod_{1\leq i<j\leq p} (l_i - l_j)dL.$$

In this subsection, we look at the asymptotic behavior of $\Delta P(x)$ as $x \uparrow b$. The results are summarized in the theorem below.

Theorem 9 *The approximation error $\Delta P(x)$ has the following asymptotic expressions.*
(i) The case $b = \infty$.

$$\Delta P(x) \sim -\prod_{i=1}^{p} \frac{\Omega_i}{2} \int_{\infty>l_2>l_1\geq x} (l_1 l_2)^{p-2}(l_2 - l_1)g(l_1)g(l_2)dl_1 dl_2$$

$$\times \int\limits_{\infty>l_3>\cdots>l_p>a} \prod_{i=3}^{p} g(l_i) \prod_{3\le i<j\le p} (l_i - l_j) \prod_{i=3}^{p} dl_i \quad (x \to \infty). \quad (20)$$

(ii) The case $b < \infty$.

$$\Delta P(x) \sim - \prod_{i=1}^{p} \frac{\Omega_i}{2} \int\limits_{b>l_2>l_1\ge x} g(l_1)g(l_2)(l_2 - l_1)dl_1 dl_2$$
$$\times \int\limits_{b>l_3>\cdots>l_p>a} \prod_{i=3}^{p} g(l_i)(b-l_i)^2 \prod_{3\le i<j\le p} (l_i - l_j) \prod_{i=3}^{p} dl_i$$
$$(x \uparrow b). \quad (21)$$

When $p = 2$, the integrals with respect to l_i $(i \ge 3)$ are assumed to be 1.

Here '$\sim$' means that the ratio of both sides converges to 1. In proving the theorem, we treat the cases $b = \infty$ and $b < \infty$ separately.

Proof. (i) The case $b = \infty$.
Let

$$q = \begin{pmatrix} p-1 & p-2 & \cdots & 0 \\ q_1 & q_2 & \cdots & q_p \end{pmatrix}$$

be a permutation of $\{p-1, p-2, \cdots, 0\}$. The set of all q such that

$$q_1 > q_2, \quad q_3 > \cdots > q_p$$

is denoted by Q_2. We can write an element of Q_2 as $q = (q_1, q_2)$ shortly without loss of information. Noting the Laplace expansion

$$\prod_{i<j}(l_i - l_j) = \det(l_i^{p-j})_{1\le i,j\le p}$$
$$= \sum_{q\in Q_2} \mathrm{sgn}(q) \det(l_i^{q_j})_{1\le i,j\le 2} \det(l_i^{q_j})_{3\le i,j\le p},$$

we have

$$\Delta P(x) = \prod_{i=1}^{p} \frac{\Omega_i}{2} \sum_{q\in Q_2} \mathrm{sgn}(q) \int\limits_{\infty>l_2>l_1\ge x} g(l_1)g(l_2) \det(l_i^{q_j})_{1\le i,j\le 2} F_q(l_2) dl_1 dl_2$$

with

$$F_q(l) = \int\limits_{l>l_3>\cdots>l_p>a} \prod_{i=3}^{p} g(l_i) \det(l_i^{q_j})_{3\le i,j\le p} \prod_{i=3}^{p} dl_i.$$

Note that $\det(l_i^{q_j})_{3\le i,j\le p}$ is an alternating polynomial with positive sign for $l_3 > \cdots > l_p$. Since $F_q(\cdot)$ is positive and monotonically increasing, and $F_q(\infty) < \infty$, it holds for $l_2 \ge x$ that

$$|F_q(\infty) - F_q(l_2)| \le |F_q(\infty) - F_q(x)| = o(1) \quad (x \to \infty),$$

and hence

$$F_q(l_2) = F_q(\infty) + o(1) \quad (x \to \infty).$$

Therefore,

$$\Delta P(x) \sim \prod_{i=1}^{p} \frac{\Omega_i}{2} \sum_{q \in Q_2} \mathrm{sgn}(q) H_q(x) F_q(\infty),$$

where

$$\begin{aligned} H_q(x) &= \int\limits_{\infty > l_2 > l_1 \ge x} g(l_1) g(l_2) \det(l_i^{q_j})_{1\le i,j\le 2} dl_1 dl_2 \\ &= \int_x^\infty l_1^{q_1} g(l_1) \Big[\int_{l_1}^\infty l_2^{q_2} g(l_2) dl_2\Big] dl_1 \\ &\quad -(\text{the term with } q_1 \text{ and } q_2 \text{ exchanged}). \end{aligned}$$

This includes the case $p = 2$ by letting $F_q(\infty) = 1$.

In the following, we evaluate the asymptotic behavior of $H_q(x)$ as $x \to \infty$. In the case $b = \infty$, we will show that $H_q(x)$ takes the maximum value asymptotically when $(q_1, q_2) = (p-1, p-2)$. Namely, we will show that for $q^0 = (p-1, p-2) \in Q_2$ and $q = (q_1, q_2) \in Q_2$ such that $q \ne q^0$,

$$\lim_{x\to\infty} \frac{H_q(x)}{H_{q^0}(x)} = 0. \tag{22}$$

Note that $q_2 \le p-3$ and $q_1 + q_2 \le 2p-4$.

Here we use L'Hospital's rule again. Because

$$H'_{q^0}(x) = x^{p-2} g(x) \int_x^\infty l^{p-2}(l-x) g(l) dl > 0$$

for sufficiently large x, it suffices to prove that

$$\frac{H'_q(x)}{H'_{q^0}(x)} = \frac{-x^{q_1-p+2}\int_x^\infty l^{q_2}g(l)dl + x^{q_2-p+2}\int_x^\infty l^{q_1}g(l)dl}{-x\int_x^\infty l^{p-2}g(l)dl + \int_x^\infty l^{p-1}g(l)dl} \tag{23}$$

converges to 0. Note that for $k \le p-1$ and $j+k \le p-1$,

$$x^j \int_x^\infty l^k g(l)dl = x^{j+k-p+1}\int_x^\infty x^{p-1-k}l^k g(l)dl \le x^{j+k-p+1}\int_x^\infty l^{p-1}g(l)dl \to 0.$$

In the following we treat the cases $q_1 = p-1$ and $q_1 \le p-2$ separately.

The case $q_1 = p-1$. The right hand side of (23) is rewritten as

$$x^{q_2-p+2} + \frac{-x\int_x^\infty l^{q_2}g(l)dl + x^{q_2-p+3}\int_x^\infty l^{p-2}g(l)dl}{-x\int_x^\infty l^{p-2}g(l)dl + \int_x^\infty l^{p-1}g(l)dl},$$

whose first term converges to 0. Therefore, we only have to check that the second term converges to 0. Since both numerator and denominator converge to 0, it suffices to show that

$$\begin{aligned}&\frac{-\int_x^\infty l^{q_2}g(l)dl + (q_2-p+3)x^{q_2-p+2}\int_x^\infty l^{p-2}g(l)dl}{-\int_x^\infty l^{p-2}g(l)dl}\\ &\quad = \frac{\int_x^\infty l^{q_2}g(l)dl}{\int_x^\infty l^{p-2}g(l)dl} - (q_2-p+3)x^{q_2-p+2}\end{aligned}$$

converges to 0 by L'Hospital's rule. Actually, it follows from $q_2 \le p-3$.

The case $q_1 \le p-2$. In the right hand side of (23), both numerator and denominator converge to 0. Therefore, by L'Hospital's rule, if

$$\lim_{x\to\infty} \frac{-(q_1-p+2)x^{q_1-p+1}\int_x^\infty l^{q_2}g(l)dl + (q_2-p+2)x^{q_2-p+1}\int_x^\infty l^{q_1}g(l)dl}{-\int_x^\infty l^{p-2}g(l)dl} \tag{24}$$

exists, then (23) has the same limit as (24). Noting that for $k \le p-2$ and $j+k \le p-3$,

$$\begin{aligned}\frac{x^j\int_x^\infty l^k g(l)dl}{\int_x^\infty l^{p-2}g(l)dl} &= \frac{x^{j+k-p+2}\int_x^\infty x^{p-2-k}l^k g(l)dl}{\int_x^\infty l^{p-2}g(l)dl}\\ &\le \frac{x^{j+k-p+2}\int_x^\infty l^{p-2}g(l)dl}{\int_x^\infty l^{p-2}g(l)dl} = x^{j+k-p+2} \to 0,\end{aligned}$$

we see that the limit in (24) exists and takes the value 0. This completes the proof of (22).

When $q = q^0$, it holds that $\mathrm{sgn}(q) = 1$,

$$\det(l_i^{q_j})_{1\le i,j\le 2} = (l_1 l_2)^{p-2}(l_1 - l_2), \qquad \det(l_i^{q_j})_{3\le i,j\le p} = \prod_{3\le i<j\le p}(l_i - l_j),$$

and hence (20) follows.

(ii) The case $b < \infty$.
Let

$$\bar{q} = \begin{pmatrix} 0 & 1 & \cdots & p-1 \\ \bar{q}_1 & \bar{q}_2 & \cdots & \bar{q}_p \end{pmatrix}$$

be a permutation of $\{0, 1, \ldots, p-1\}$. The set of all $\bar{q}$ such that

$$\bar{q}_1 < \bar{q}_2, \quad \bar{q}_3 < \cdots < \bar{q}_p$$

is denoted by $\bar{Q}_2$. We can write an element of $\bar{Q}_2$ as $\bar{q} = (\bar{q}_1, \bar{q}_2)$. For $1 \le i \le p$, let $\bar{l}_i = b - l_i$. Noting the Laplace expansion

$$\begin{aligned}\prod_{i<j}(l_i - l_j) &= (-1)^{\frac{p(p-1)}{2}}\prod_{i<j}(l_j - l_i) = (-1)^{\frac{p(p-1)}{2}}\prod_{i<j}(\bar{l}_i - \bar{l}_j) \\ &= \det(\bar{l}_i^{j-1})_{1\le i,j\le p} = \sum_{\bar{q}\in\bar{Q}_2}\mathrm{sgn}(\bar{q})\det(\bar{l}_i^{\bar{q}_j})_{1\le i,j\le 2}\det(\bar{l}_i^{\bar{q}_j})_{3\le i,j\le p},\end{aligned}$$

in the same manner as the case $b = \infty$, we see

$$\Delta P(x) \sim \prod_{i=1}^{p}\frac{\Omega_i}{2}\sum_{\bar{q}\in\bar{Q}_2}\mathrm{sgn}(\bar{q})\bar{H}_{\bar{q}}(x)\int\limits_{b>l_3>\cdots>l_p>a}\prod_{i=3}^{p}g(l_i)\det(l_i^{\bar{q}_j})_{3\le i,j\le p}\prod_{i=3}^{p}dl_i,$$

where

$$\begin{aligned}\bar{H}_{\bar{q}}(x) &= \int\limits_{b>l_2>l_1\ge x} g(l_1)g(l_2)\det(\bar{l}_i^{q_j})_{1\le i,j\le 2}dl_1 dl_2 \\ &= \int_x^b \bar{l}_1^{\bar{q}_1}g(l_1)\Big[\int_{l_1}^b \bar{l}_2^{\bar{q}_2}g(l_2)dl_2\Big]dl_1 \\ &\quad -(\text{the term with } \bar{q}_1 \text{ and } \bar{q}_2 \text{ exchanged}).\end{aligned}$$

In the following, as $x \uparrow b$, we will show that $\bar{H}_{\bar{q}}(x)$ takes the maximum value asymptotically when $(\bar{q}_1, \bar{q}_2) = (0, 1)$. Namely, we will show that for $\bar{q}^0 =$

$(0,1) \in \bar{Q}_2$ and $\bar{q} = (\bar{q}_1, \bar{q}_2) \in \bar{Q}_2$ such that $\bar{q} \neq \bar{q}^0$,

$$\lim_{x \uparrow b} \frac{\bar{H}_{\bar{q}}(x)}{\bar{H}_{\bar{q}^0}(x)} = 0. \tag{25}$$

Note that $\bar{q}_2 \geq 2$ and $\bar{q}_1 + \bar{q}_2 \geq 2$. Because

$$\bar{H}'_{\bar{q}^0}(x) = g(x) \int_x^b (l - x) g(l) dl > 0$$

for sufficiently large x, it suffices to prove that

$$\frac{\bar{H}'_{\bar{q}}(x)}{\bar{H}'_{\bar{q}^0}(x)} = \frac{-(b-x)^{\bar{q}_1} \int_x^b (b-l)^{\bar{q}_2} g(l) dl + (b-x)^{\bar{q}_2} \int_x^b (b-l)^{\bar{q}_1} g(l) dl}{-\int_x^b (b-l) g(l) dl + (b-x) \int_x^b g(l) dl} \tag{26}$$

converges to 0 by L'Hospital's rule. Note that for $k \geq 0$ and $j + k \geq 0$,

$$\begin{aligned} (b-x)^j \int_x^b (b-l)^k g(l) dl &= (b-x)^{j+k} \int_x^b (b-x)^{-k} (b-l)^k g(l) dl \\ &\leq (b-x)^{j+k} \int_x^b g(l) dl \to 0. \end{aligned}$$

In the right hand side of (26), both numerator and denominator converge to 0. Therefore, by L'Hospital's rule, if

$$\lim_{x \uparrow b} \frac{\bar{q}_1 (b-x)^{\bar{q}_1 - 1} \int_x^b (b-l)^{\bar{q}_2} g(l) dl - \bar{q}_2 (b-x)^{\bar{q}_2 - 1} \int_x^b (b-l)^{\bar{q}_1} g(l) dl}{\int_x^b g(l) dl} \tag{27}$$

exists, then (26) has the same limit as (27). Noting that for $k \geq 0$ and $j+k \geq 1$,

$$\begin{aligned} \frac{(b-x)^j \int_x^b (b-l)^k g(l) dl}{\int_x^b g(l) dl} &= \frac{(b-x)^{j+k} \int_x^b (b-x)^{-k} (b-l)^k g(l) dl}{\int_x^b g(l) dl} \\ &\leq \frac{(b-x)^{j+k} \int_x^b g(l) dl}{\int_x^b g(l) dl} = (b-x)^{j+k} \to 0, \end{aligned}$$

we see that (27) exists and takes the value 0. This completes the proof of (25).

When $\bar{q} = \bar{q}^0$, it holds that $\operatorname{sgn}(\bar{q}) = 1$,

$$\det(\bar{l}_i^{\bar{q}_j})_{1\le i,j\le 2} = \bar{l}_2 - \bar{l}_1 = l_1 - l_2,$$
$$\det(\bar{l}_i^{\bar{q}_j})_{3\le i,j\le p} = \prod_{i=3}^{p}(b-l_i)^2 \prod_{3\le i<j\le p}(l_i - l_j),$$

and hence (21) follows. $\square$

2.5 Validity of the Euler characteristic heuristic

In this subsection, we compare the approximation error $\Delta P(x)$ given in Theorem 9 with the expansion $\widehat{P}(x)$ given in Theorem 7.

The Euler characteristic heuristic is said to be *weakly valid* when $\Delta P(x)$ is asymptotically smaller than $\widehat{P}(x)$ as $x \uparrow b$, that is, $\Delta P(x) = o(\widehat{P}(x))$. The Euler characteristic heuristic is said to be *valid* when $\Delta P(x)$ is asymptotically smaller than each term of $\widehat{P}(x)$ as $x \uparrow b$. We treat the cases $b = \infty$ and $b < \infty$ separately. The conditions for L'Hospital's rule used below are fulfilled by Assumption 5 and Lemma 6.

(i) The case $b = \infty$.
Let $H_{q^0}(x)$ be the function defined in the previous subsection. We first consider the condition that the approximation error $\Delta P(x)$ in (20) is asymptotically smaller than the leading term of $\widehat{P}(x)$ in (18), or

$$\lim_{x\to\infty} \frac{H_{q^0}(x)}{\int_x^\infty l^{p-1} g(l)dl} = 0.$$

Actually, by L'Hospital's rule, the left hand side of the above is

$$\lim_{x\to\infty} \frac{H'_{q^0}(x)}{-x^{p-1}g(x)} = \lim_{x\to\infty}\Big\{\int_x^\infty l^{p-2}g(l)dl - x^{-1}\int_x^\infty l^{p-1}g(l)dl\Big\} = 0,$$

which means that $\Delta P(x)$ is always asymptotically smaller than the leading term of $\widehat{P}(x)$.

Next we consider the condition that the approximation error $\Delta P(x)$ is asymptotically smaller than each term of $\widehat{P}(x)$ expressed as (18), or

$$\lim_{x\to\infty} \frac{H_{q^0}(x)}{\int_x^\infty g(l)dl} = 0.$$

By L'Hospital's rule, it suffices to consider the condition for

$$\lim_{x\to\infty} \frac{H'_{q^0}(x)}{-g(x)} = \lim_{x\to\infty} \frac{x\int_x^\infty l^{p-2}g(l)dl - \int_x^\infty l^{p-1}g(l)dl}{x^{-(p-2)}} = 0,$$

which is, however, always true for $p = 2$. For $p \geq 3$, by using L'Hospital's rule again, we see that the above holds if

$$\lim_{x \to \infty} x^{p-1} \int_x^{\infty} l^{p-2} g(l) dl = 0. \tag{28}$$

Moreover, we have

$$\lim_{x \to \infty} x^{2p-2} g(x) = 0 \tag{29}$$

as a sufficient condition for (28).

(ii) The case $b < \infty$.

Let $\bar{H}_{\bar{q}^0}(x)$ be the function defined in the previous subsection. We first consider the condition that the approximation error $\Delta P(x)$ in (21) is asymptotically smaller than the leading term of $\widehat{P}(x)$ in (19), or

$$\lim_{x \uparrow b} \frac{\bar{H}_{\bar{q}^0}(x)}{\int_x^b g(l) dl} = 0.$$

Actually, by L'Hospital's rule, the left hand side of the above is

$$\lim_{x \uparrow b} \frac{\bar{H}'_{\bar{q}^0}(x)}{-g(x)} = \lim_{x \uparrow b} \left\{ \int_x^b (b-l) g(l) dl - (b-x) \int_x^b g(l) dl \right\} = 0,$$

which means that $\Delta P(x)$ is always asymptotically smaller than the leading term of $\widehat{P}(x)$.

Next we consider the condition that the approximation error $\Delta P(x)$ is asymptotically smaller than each term of $\widehat{P}(x)$ expressed as (19), or

$$\lim_{x \uparrow b} \frac{\bar{H}_{\bar{q}^0}(x)}{\int_x^b (b-l)^{p-1} g(l) dl} = 0.$$

By L'Hospital's rule, it suffices to consider the condition for

$$\lim_{x \uparrow b} \frac{\bar{H}'_{\bar{q}^0}(x)}{-(b-x)^{p-1} g(x)} = \lim_{x \uparrow b} \frac{\int_x^b (b-l) g(l) dl - (b-x) \int_x^b g(l) dl}{(b-x)^{p-1}} = 0,$$

which is, however, always true for $p = 2$. For $p \geq 3$, by using L'Hospital's rule again, we see that the above holds if

$$\lim_{x \uparrow b} (b-x)^{-(p-2)} \int_x^b g(l) dl = 0. \tag{30}$$

Moreover, we have

$$\lim_{x \uparrow b} (b-x)^{-(p-3)} g(x) = 0 \tag{31}$$

as a sufficient condition for (30).

The results so far are summarized as follows.

Theorem 10 *In both cases $b = \infty$ and $b < \infty$, the Euler characteristic heuristic is weakly valid. That is,*

$$\Delta P(x) = o(\widehat{P}(x)) \quad (x \uparrow b).$$

Moreover, if $p = 2$, or if $p \geq 3$ and

$$\lim_{x \to \infty} x^{2p-2} g(x) = 0 \quad (\textit{in the case } b = \infty),$$

$$\lim_{x \uparrow b} (b-x)^{-(p-3)} g(x) = 0 \quad (\textit{in the case } b < \infty),$$

then the Euler characteristic heuristic is valid. That is, $\Delta P(x)$ is asymptotically smaller than each term of $\widehat{P}(x)$ as $x \uparrow b$.

Remark 11 *The above results show that the faster $g(x)$ converges to 0 as $x \uparrow b$, the smaller is the approximation error $\Delta P(x)$.*

3 Examples

In the second half of the paper, we give the formulas of $\widehat{P}(x)$ and the asymptotic evaluations for $\Delta P(x)$ for the Wishart distribution, the multivariate beta distribution, and the inverse Wishart distribution. They are distributions of standard orthogonally invariant random matrices appearing in the multivariate analysis.

The $p \times p$ Wishart distribution with n ($n > p-1$) degrees of freedoms is denoted by $W_p(n)$. When $W \sim W_p(n)$, $X(h) = h'Wh$ for fixed $h \in \mathbb{S}^{p-1}$ is distributed as $\chi^2(n)$, the chi-squared distribution with n degrees of freedom. We can call $X(\cdot)$ a chi-squared field. In the following, let $\bar{G}_\nu(\cdot)$ denote the upper probability of the chi-squared distribution with ν degrees of freedom.

Theorem 12 *The approximation formula for the distribution of the largest eigenvalue of the Wishart matrix $W_p(n)$ by the Euler characteristic heuristic is given by*

$$\widehat{P}(x) = \frac{2^{p+n-2}\Gamma(\frac{p+1}{2})\Gamma(\frac{n+1}{2})}{\sqrt{\pi}} \sum_{j=0}^{p-1} \Big(-\frac{1}{2}\Big)^j \frac{\Gamma(\frac{n+p-1-2j}{2})}{\Gamma(p-j)\Gamma(n-j)j!} \bar{G}_{n+p-1-2j}(x). \tag{32}$$

The asymptotic error is evaluated as

$$\Delta P(x) \sim -\frac{1}{\Gamma(p-1)\Gamma(n-1)} x^{p+n-5} e^{-x} \quad (x \to \infty). \tag{33}$$

Remark 13 *Since $\bar{G}_\nu(x) = O(x^{\frac{\nu}{2}-1} e^{-\frac{x}{2}})$, the approximation error $\Delta P(x)$ is exponentially smaller than each term of $\widehat{P}(x)$.*

Remark 14 *We can propose an improved version of $\widehat{P}(x)$ by incorporating the right hand side of (33) as*

$$\widetilde{P}(x) = \widehat{P}(x) + \frac{1}{\Gamma(p-1)\Gamma(n-1)} x^{p+n-5} e^{-x}.$$

The error $\widetilde{P}(x) - P(x)$ is asymptotically smaller than $\Delta P(x)$.

Remark 15 *Kuriki and Takemura (2001) formulated the largest eigenvalue of the Wishart matrix as the maximum of a Gaussian random field, and derived the same formula (32) by virtue of the tube method, the Euler characteristic heuristic applied to the Gaussian random field. They showed that the approximation error is $O(x^{\frac{np}{2}-1} e^{-x})$ at most.*

For the Wishart $W_2(5)$, the exact upper probability of the largest eigenvalue $P(x)$, the approximation by the Euler characteristic heuristic $\widehat{P}(x)$, and its improvement $\widetilde{P}(x)$ are given as

$$P(x) = \frac{x^2}{3} e^{-\frac{x}{2}} + \Big(1 + x + \frac{x^2}{6}\Big) e^{-x},$$

$$\widehat{P}(x) = \frac{x^2}{3} e^{-\frac{x}{2}}, \quad \text{and} \quad \widetilde{P}(x) = \widehat{P}(x) + \frac{x^2}{6} e^{-x}.$$

These three are depicted in Fig. 1 as solid line, dotted line, and dashed line, respectively.

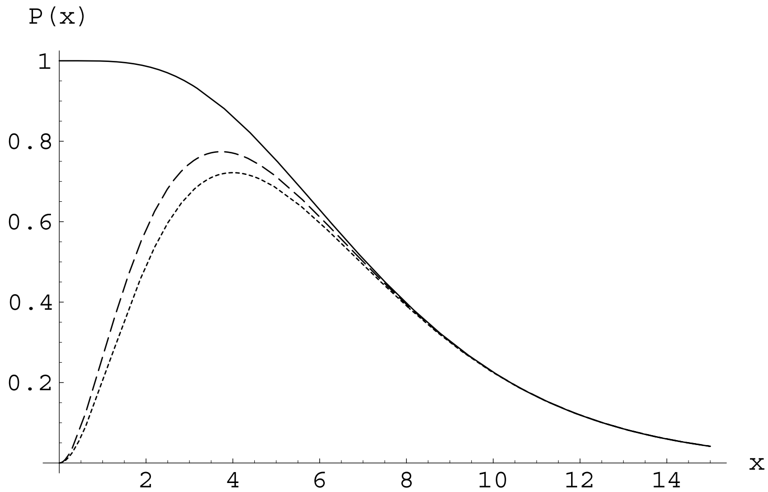

Fig. 1. $P(x)$ (—), $\widehat{P}(x)$ $(\cdots)$, and $\widetilde{P}(x)$ $(--)$ for the Wishart $W_2(5)$

From now on we prove Theorem 12. We will show (32) at first. The density function of the Wishart distribution is

$$f_{p,n}(W) = \frac{1}{c(p,n)} e^{-\frac{1}{2}\mathrm{tr}(W)} \det(W)^{\frac{1}{2}(n-p-1)} \quad (W > 0)$$

with

$$c(p,n) = 2^{\frac{pn}{2}} \Gamma_p\Big(\frac{n}{2}\Big), \quad \Gamma_p(a) = \pi^{\frac{p(p-1)}{4}} \prod_{i=1}^{p} \Gamma\Big(a - \frac{1}{2}(i-1)\Big). \tag{34}$$

For $w_{11} \in \mathbb{R}$ and $W_{22} \in Sym(p-1)$, it holds that

$$\begin{aligned} f_{p,n}(\mathrm{diag}(w_{11}, W_{22})) = & \frac{c(1, n-p+1) c(p-1, n-1)}{c(p,n)} \\ & \times f_{1,n-p+1}(w_{11}) f_{p-1,n-1}(W_{22}). \end{aligned}$$

Hence, if we suppose that $w_{11} \sim \chi^2(n-p+1)$ and $W_{22} \sim W_{p-1}(n-1)$ are independent random variables, then we have

$$\widehat{P}(x) = \frac{\Omega_p}{2} \frac{c(1, n-p+1) c(p-1, n-1)}{c(p,n)} E[1_{\{w_{11} \geq x\}} \det(w_{11} I_{p-1} - W_{22})].$$

The multiplicative constant in the expression above is

$$\frac{\Omega_p}{2}\frac{c(1,n-p+1)c(p-1,n-1)}{c(p,n)} = 2^{-(p-1)}\sqrt{\pi}\frac{\Gamma(\frac{n-p+1}{2})}{\Gamma(\frac{p}{2})\Gamma(\frac{n}{2})}.$$

Noting the expansion of the determinant

$$\det(w_{11}I_{p-1} - W_{22}) = \sum_{j=0}^{p-1}(-1)^j w_{11}^{p-1-j}\mathrm{tr}_j(W_{22}) \tag{35}$$

($\mathrm{tr}_j(\cdot)$ denotes the summation of all $j \times j$ principal minor determinants, see Muirhead (1982), Appendix A.7), we have

$$E[1_{\{w_{11}\ge x\}}\det(w_{11}I_{p-1} - W_{22})] = \sum_{j=0}^{p-1}(-1)^j E[1_{\{w_{11}\ge x\}}w_{11}^{p-1-j}]E[\mathrm{tr}_j(W_{22})].$$

The expectation with respect to w_{11} can be evaluated by reconsidering w_{11} as a chi-squared random variable with $n-p+1+2(p-1-j) = n+p-1-2j$ degrees of freedom. It follows that

$$E[1_{\{w_{11}\ge x\}}w_{11}^{p-1-j}] = 2^{p-1-j}\frac{\Gamma(\frac{n+p-1-2j}{2})}{\Gamma(\frac{n-p+1}{2})}\bar{G}_{n+p-1-2j}(x).$$

On the other hand, by letting χ^2_ν denote an independent chi-squared random variable with ν degrees of freedom, we can write

$$\begin{aligned} E[\mathrm{tr}_j(W_{22})] &= \binom{p-1}{j}E\Big[\prod_{i=1}^{j}\chi^2_{n-i}\Big] \\ &= \binom{p-1}{j}(n-1)(n-2)\cdots(n-j) \\ &= \frac{\Gamma(p)\Gamma(n)}{\Gamma(p-j)\Gamma(n-j)j!}. \end{aligned}$$

Summarizing the above, we get

$$\begin{aligned} \widehat{P}(x) = {} & 2^{-(p-1)}\sqrt{\pi}\frac{\Gamma(\frac{n-p+1}{2})}{\Gamma(\frac{p}{2})\Gamma(\frac{n}{2})} \\ & \times \sum_{j=0}^{p-1}(-1)^j 2^{p-1-j}\frac{\Gamma(\frac{n+p-1-2j}{2})}{\Gamma(\frac{n-p+1}{2})}\bar{G}_{n+p-1-2j}(x) \times \frac{\Gamma(p)\Gamma(n)}{\Gamma(p-j)\Gamma(n-j)j!}, \end{aligned}$$

which is reduced to (32). Here we used the duplication formula of the gamma function

$$\Gamma(z) = \frac{2^{z-1}}{\sqrt{\pi}} \Gamma\Big(\frac{z}{2}\Big)\Gamma\Big(\frac{z+1}{2}\Big). \tag{36}$$

Since the expansion (35) is equivalent to the expansion of the Vandermonde determinant (16), the resulting expansion of $\widehat{P}(x)$ corresponds to (18) in Theorem 7.

Next we will derive the asymptotic evaluation of the approximation error (33). Let l_i be the eigenvalue of W, and let

$$g(l) = l^{\frac{n-p-1}{2}} e^{-\frac{l}{2}}.$$

Then

$$f_{p,n}(W) = \frac{1}{c(p,n)} \prod_{i=1}^{p} g(l_i).$$

Recall that

$$\Delta P(x) \sim -\frac{1}{c(p,n)} \prod_{i=1}^{p} \frac{\Omega_i}{2} \int_{\infty > l_2 > l_1 \geq x} g(l_1)g(l_2)(l_1 l_2)^{p-2}(l_2 - l_1)dl_1 dl_2 \times F,$$

where

$$F = \int_{\infty > l_3 > \cdots > l_p > 0} \prod_{i=3}^{p} g(l_i) \prod_{3 \leq i < j \leq p} (l_i - l_j) \prod_{i=3}^{p} dl_i$$

for $p \geq 3$, and $F = 1$ for $p = 2$.

Assume that $p \geq 3$. By considering the Wishart $W_{p-2}(n-2)$, we have

$$c(p-2, n-2) = \prod_{i=1}^{p-2} \frac{\Omega_i}{2} \times F,$$

and hence

$$\begin{aligned} \Delta P(x) \sim & -\frac{c(p-2,n-2)}{c(p,n)} \frac{\Omega_p}{2} \frac{\Omega_{p-1}}{2} \\ & \times \int_{\infty > l_2 > l_1 \geq x} (l_1 l_2)^{\frac{n+p-5}{2}} (l_2 - l_1) e^{-\frac{1}{2}(l_1 + l_2)} dl_1 dl_2. \end{aligned}$$

The multiplicative constant contained above is

$$\begin{aligned}\frac{c(p-2,n-2)}{c(p,n)}\frac{\Omega_p}{2}\frac{\Omega_{p-1}}{2} &= \frac{2^{-(p+n-2)}\pi}{\Gamma(\frac{n}{2})\Gamma(\frac{n-1}{2})\Gamma(\frac{p}{2})\Gamma(\frac{p-1}{2})} \\ &= \frac{1}{4\Gamma(p-1)\Gamma(n-1)}.\end{aligned} \tag{37}$$

On the other hand, the multiplicative constant for $p = 2$ is

$$\frac{1}{c(2,n)}\frac{\Omega_2}{2}\frac{\Omega_1}{2} = \frac{1}{4\Gamma(n-1)}, \tag{38}$$

which is consistent with the constant for $p \geq 3$.

Write $k = \frac{n+p-5}{2}$ for simplicity. The integral with respect to l_1 and l_2 is evaluated by integration by parts as

$$\begin{aligned}&\int_x^\infty l_1^k e^{-\frac{l_1}{2}} dl_1 \int_{l_1}^\infty l_2^{k+1} e^{-\frac{l_2}{2}} dl_2 - \int_x^\infty l_2^k e^{-\frac{l_2}{2}} dl_2 \int_x^{l_2} l_1^{k+1} e^{-\frac{l_1}{2}} dl_1 \\ &\quad = 4\int_x^\infty l_1^{2k+1} e^{-l_1} dl_1 - 2x^k e^{-\frac{x}{2}} \int_x^\infty l_2^{k+1} e^{-\frac{l_2}{2}} dl_2 \\ &\quad \sim 4(2k+1)x^{2k}e^{-x} - 4kx^{x+1}e^{-\frac{x}{2}} \cdot 2x^{k-1}e^{-\frac{x}{2}} = 4x^{2k}e^{-x}.\end{aligned} \tag{39}$$

Multiplying (37) and (39) yields (33).

3.2 Multivariate beta distribution

The $p \times p$ multivariate beta distribution with parameters $(\frac{n_1}{2}, \frac{n_2}{2})$ $(n_1, n_2 > p-1)$ is denoted by $B_p(\frac{n_1}{2}, \frac{n_2}{2})$. When $B \sim B_p(\frac{n_1}{2}, \frac{n_2}{2})$, $X(h) = h'Bh$ for fixed $h \in \mathbb{S}^{p-1}$ is distributed as $B(\frac{n_1}{2}, \frac{n_2}{2})$, the beta distribution with parameters $(\frac{n_1}{2}, \frac{n_2}{2})$. We can call $X(\cdot)$ a beta random field. In the following, let $\bar{B}_{\frac{\nu_1}{2},\frac{\nu_2}{2}}(\cdot)$ denote the upper probability of the beta distribution with parameters $(\frac{\nu_1}{2}, \frac{\nu_2}{2})$.

Theorem 16 *The approximation formula for the distribution of the largest eigenvalue of the multivariate beta matrix $B_p(\frac{\nu_1}{2}, \frac{\nu_2}{2})$ by the Euler characteristic heuristic is given by*

$$\widehat{P}(x) = \frac{2^{-n_1+p}\Gamma(\frac{p+1}{2})\Gamma(\frac{n_2+1}{2})\Gamma(\frac{n_1-p+1}{2})}{\Gamma(\frac{n_1}{2})\Gamma(\frac{n_1+n_2}{2} - \frac{p}{2})}$$

$$\times \sum_{j=0}^{p-1} (-1)^j \frac{\Gamma(\frac{n_2-p+1+2j}{2})\Gamma(n_1+n_2-p+j)}{\Gamma(\frac{n_1+n_2}{2}-p+1+j)\Gamma(p-j)\Gamma(n_2-p+1+j)j!}$$
$$\times \bar{B}_{\frac{n_1-p+1}{2},\frac{n_2-p+1+2j}{2}}(x). \tag{40}$$

The asymptotic error is evaluated as

$$\Delta P(x) \sim -\frac{\Gamma(n_1+n_2-p+1)\Gamma(n_2+1)}{\Gamma(p-1)\Gamma(n_1-1)\Gamma(n_2-p+3)\Gamma(n_2-p+4)}$$
$$\times x^{n_1-p-1}(1-x)^{n_2-p+2} \quad (x \uparrow 1). \tag{41}$$

Remark 17 *Since* $\bar{B}_{\frac{\nu_1}{2},\frac{\nu_2}{2}}(x) = O((1-x)^{\frac{\nu_2}{2}})$, $\Delta P(x) = o(\widehat{P}(x))$. *In addition, since* $\nu_2 = n_2 + p - 1$ *in the last term of* $\widehat{P}(x)$, *the approximation error* $\Delta P(x)$ *is asymptotically smaller than each term of* $\widehat{P}(x)$ *whenever* $(n_2+p-1)/2 < n_2 - p + 2$, *i.e.,* $3p < n_2 + 5$.

Remark 18 *As in Remark 14, we can propose an improved version of* $\widehat{P}(x)$ *as*

$$\widetilde{P}(x) = \widehat{P}(x) + \frac{\Gamma(n_1+n_2-p+1)\Gamma(n_2+1)}{\Gamma(p-1)\Gamma(n_1-1)\Gamma(n_2-p+3)\Gamma(n_2-p+4)}$$
$$\times x^{n_1-p-1}(1-x)^{n_2-p+2}.$$

From now on we prove Theorem 16. We will show (40) at first. The density function of the multivariate beta distribution is

$$f_{p,n_1,n_2}(B) = \frac{1}{d(p,n_1,n_2)} \det(B)^{\frac{1}{2}(n_1-p-1)} \det(I_p - B)^{\frac{1}{2}(n_2-p-1)}$$
$$(0 < B < I_p)$$

with

$$d(p,n_1,n_2) = \frac{\Gamma_p(\frac{n_1}{2})\Gamma_p(\frac{n_2}{2})}{\Gamma_p(\frac{1}{2}(n_1+n_2))}.$$

For $b_{11} \in \mathbb{R}$ and $B_{22} \in Sym(p-1)$, it holds that

$$f_{p,n_1,n_2}(\mathrm{diag}(b_{11}, B_{22}))$$
$$= \frac{d(1,n_1-p+1,n_2-p+1)d(p-1,n_1-1,n_2-1)}{d(p,n_1,n_2)}$$
$$\times f_{1,n_1-p+1,n_2-p+1}(b_{11}) f_{p-1,n_1-1,n_2-1}(B_{22}).$$

Hence, if we suppose that

$$b_{11} \sim B\Big(\frac{n_1-p+1}{2}, \frac{n_2-p+1}{2}\Big) \quad \text{and} \quad B_{22} \sim B_{p-1}\Big(\frac{n_1-1}{2}, \frac{n_2-1}{2}\Big)$$

are independent random variables, then we have

$$\begin{aligned}\widehat{P}(x) = {} & \frac{\Omega_p}{2} \frac{d(1, n_1-p+1, n_2-p+1) d(p-1, n_1-1, n_2-1)}{d(p, n_1, n_2)} \\ & \times E[1_{\{b_{11} \ge x\}} \det(b_{11} I_{p-1} - B_{22})].\end{aligned}$$

The multiplicative constant in the expression above is

$$\begin{aligned}& \frac{\Omega_p}{2} \frac{d(1, n_1-p+1, n_2-p+1) d(p-1, n_1-1, n_2-1)}{d(p, n_1, n_2)} \\ & \quad = \frac{\sqrt{\pi}\Gamma(\frac{n_1+n_2}{2})\Gamma(\frac{n_1+n_2}{2}-\frac{1}{2})\Gamma(\frac{n_1-p+1}{2})\Gamma(\frac{n_2-p+1}{2})}{\Gamma(\frac{p}{2})\Gamma(\frac{n_1}{2})\Gamma(\frac{n_2}{2})\Gamma(\frac{n_1+n_2}{2}-\frac{p}{2})\Gamma(\frac{n_1+n_2}{2}-p+1)}.\end{aligned}$$

Noting the expansion of the determinant

$$\begin{aligned}\det(b_{11} I_{p-1} - B_{22}) &= \det(I_{p-1} - B_{22} - (1-b_{11}) I_{p-1}) \\ &= \sum_{j=0}^{p-1} (-1)^{p-1-j} (1-b_{11})^{p-1-j} \mathrm{tr}_j(I_{p-1} - B_{22}),\end{aligned} \tag{42}$$

we have

$$\begin{aligned}& E[1_{\{b_{11} \ge x\}} \det(b_{11} I_{p-1} - B_{22})] \\ & \quad = \sum_{j=0}^{p-1} (-1)^{p-1-j} E[1_{\{b_{11} \ge x\}} (1-b_{11})^{p-1-j}] E[\mathrm{tr}_j(I_{p-1} - B_{22})].\end{aligned}$$

The expectation with respect to w_{11} can be evaluated by reconsidering b_{11} as a beta random variable with the second parameter

$$\frac{n_2-p+1}{2} + (p-1-j) = \frac{n_2+p-1-2j}{2}.$$

It follows that

$$\begin{aligned}& E[1_{\{b_{11} \ge x\}} (1-b_{11})^{p-1-j}] \\ & \quad = \frac{\Gamma(\frac{n_1+n_2}{2}-p+1)\Gamma(\frac{n_2+p-1-2j}{2})}{\Gamma(\frac{n_2-p+1}{2})\Gamma(\frac{n_1+n_2}{2}-j)} \bar{B}_{\frac{n_1-p+1}{2}, \frac{n_2+p-1-2j}{2}}(x).\end{aligned}$$

On the other hand, noting that

$$I_{p-1} - B_{22} \sim B_{p-1}\Big(\frac{n_2-1}{2}, \frac{n_1-1}{2}\Big)$$

and from Theorem 3.3.3 of Muirhead (1982), we can write

$$\begin{aligned}
&E\Big[\mathrm{tr}_j(I_{p-1} - B_{22})\Big] \\
&\quad = \binom{p-1}{j} E\Big[\prod_{i=1}^{j} B_{\frac{n_2-i}{2},\frac{n_1-1}{2}}\Big] \\
&\quad = \binom{p-1}{j} \frac{(n_2-1)(n_2-2)\cdots(n_2-j)}{(n_1+n_2-2)(n_1+n_2-3)\cdots(n_1+n_2-1-j)} \\
&\quad = \frac{\Gamma(p)\Gamma(n_2)\Gamma(n_1+n_2-1-j)}{\Gamma(p-j)\Gamma(n_2-j)\Gamma(n_1+n_2-1)j!},
\end{aligned}$$

where $B_{\frac{\nu_1}{2},\frac{\nu_2}{2}}$ denotes an independent beta random variable with parameters $(\frac{\nu_1}{2}, \frac{\nu_2}{2})$. Summarizing the above, we get

$$\begin{aligned}
\widehat{P}(x) = {} & \frac{\sqrt{\pi}\Gamma(\frac{n_1+n_2}{2})\Gamma(\frac{n_1+n_2}{2}-\frac{1}{2})\Gamma(\frac{n_1-p+1}{2})\Gamma(\frac{n_2-p+1}{2})}{\Gamma(\frac{p}{2})\Gamma(\frac{n_1}{2})\Gamma(\frac{n_2}{2})\Gamma(\frac{n_1+n_2}{2}-\frac{p}{2})\Gamma(\frac{n_1+n_2}{2}-p+1)} \\
& \times \sum_{j=0}^{p}(-1)^{p-1-j}\frac{\Gamma(\frac{n_1+n_2}{2}-p+1)\Gamma(\frac{n_2+p-1-2j}{2})}{\Gamma(\frac{n_2-p+1}{2})\Gamma(\frac{n_1+n_2}{2}-j)}\bar{B}_{\frac{n_1-p+1}{2},\frac{n_2+p-1-2j}{2}}(x) \\
& \times \frac{\Gamma(p)\Gamma(n_2)\Gamma(n_1+n_2-1-j)}{\Gamma(p-j)\Gamma(n_2-j)\Gamma(n_1+n_2-1)j!},
\end{aligned}$$

which is reduced to (40). Here we used the duplication formula (36) again.

Since the expansion (42) is equivalent to the expansion of the Vandermonde determinant (17) with $b = 1$, the resulting expansion of $\widehat{P}(x)$ corresponds to (19) in Theorem 7.

Next we will derive the asymptotic evaluation of the approximation error (41). Let l_i be the eigenvalue of W, and let

$$g(l) = l^{\frac{n_1-p-1}{2}}(1-l)^{\frac{n_2-p-1}{2}}.$$

Then

$$f_{p,n_1,n_2}(B) = \frac{1}{d(p,n_1,n_2)}\prod_{i=1}^{p} g(l_i).$$

Remark 17 follows again from the fact that (31) with $b = 1$ holds when $3p < n_2 + 5$. Recall that

$$\Delta P(x) \sim -\frac{1}{d(p, n_1, n_2)} \prod_{i=1}^{p} \frac{\Omega_i}{2} \int_{1>l_2>l_1\geq x} g(l_1)g(l_2)(l_2 - l_1)dl_1dl_2 \times F,$$

where

$$F = \int_{1>l_3>\cdots>l_p>0} \prod_{i=3}^{p} g(l_i)(1-l_i)^2 \prod_{3\leq i<j\leq p} (l_i - l_j) \prod_{i=3}^{p} dl_i$$

for $p \geq 3$, and $F = 1$ for $p = 2$.

Assume that $p \geq 3$. Noting that

$$\frac{n_1 - p - 1}{2} = \frac{(n_1 - 2) - (p - 2) - 1}{2},$$
$$\frac{n_2 - p - 1}{2} + 2 = \frac{(n_2 + 2) - (p - 2) - 1}{2},$$

and considering the multivariate beta $B_{p-2}(\frac{n_1-2}{2}, \frac{n_2+2}{2})$, we have

$$d(p-2, n_1-2, n_2+2) = \prod_{i=1}^{p-2} \frac{\Omega_i}{2} \times F,$$

and hence

$$\begin{aligned}\Delta P(x) \sim &-\frac{d(p-2, n_1-2, n_2+2)}{d(p, n_1, n_2)} \frac{\Omega_p}{2} \frac{\Omega_{p-1}}{2} \\ &\times \int_{1>l_2>l_1\geq x} (l_1 l_2)^{\frac{n_1-p-1}{2}} \{(1-l_1)(1-l_2)\}^{\frac{n_2-p-1}{2}} (l_2 - l_1) dl_1 dl_2.\end{aligned}$$

The multiplicative constant contained above is

$$\begin{aligned}&\frac{d(p-2, n_1-2, n_2+2)}{d(p, n_1, n_2)} \frac{\Omega_p}{2} \frac{\Omega_{p-1}}{2} \\ &= \frac{\pi\Gamma(\frac{n_1+n_2-p+2}{2})\Gamma(\frac{n_1+n_2-p+1}{2})\Gamma(\frac{n_2+2}{2})\Gamma(\frac{n_2+1}{2})}{\Gamma(\frac{n_1}{2})\Gamma(\frac{n_1-1}{2})\Gamma(\frac{n_2-p+4}{2})\Gamma(\frac{n_2-p+3}{2})\Gamma(\frac{n_2-p+2}{2})\Gamma(\frac{n_2-p+1}{2})\Gamma(\frac{p}{2})\Gamma(\frac{p-1}{2})} \\ &= \frac{\Gamma(n_1 + n_2 - p + 1)\Gamma(n_2 + 1)}{4\Gamma(p-1)\Gamma(n_1 - 1)\Gamma(n_2 - p + 3)\Gamma(n_2 - p + 1)}.\end{aligned} \tag{43}$$

On the other hand, the multiplicative constant for $p = 2$ is

$$\frac{1}{d(2,n_1,n_2)}\frac{\Omega_2}{2}\frac{\Omega_1}{2} = \frac{\Gamma(n_1+n_2-1)}{4\Gamma(n_1-1)\Gamma(n_2-1)},$$

which is consistent with the constant for $p \geq 3$.

Write $k_1 = \frac{n_1-p-1}{2}$, $k_2 = \frac{n_2-p-1}{2}$ for simplicity. The integral with respect to l_1 and l_2 is evaluated by integration by parts as

$$\begin{aligned}
&-\int_x^1 l_1^{k_1}(1-l_1)^{k_2}dl_1 \int_{l_1}^1 l_2^{k_1}(1-l_2)^{k_2+1}dl_2 \\
&\qquad + \int_x^1 l_1^{k_1}(1-l_1)^{k_2+1}dl_1 \int_{l_1}^1 l_2^{k_1}(1-l_2)^{k_2}dl_2 \\
&\quad \sim -\int_x^1 l_1^{k_1}(1-l_1)^{k_2}dl_1 l_1^{k_1}\frac{1}{k_2+2}(1-l_1)^{k_2+2} \\
&\qquad + \int_x^1 l_1^{k_1}(1-l_1)^{k_2+1}dl_1 l_1^{k_1}\frac{1}{k_2+1}(1-l_1)^{k_2+1} \\
&\quad \sim \frac{1}{(k_2+1)(k_2+2)(2k_2+3)}x^{2k_1}(1-x)^{2k_2+3}.
\end{aligned} \tag{44}$$

Multiplying (43) and (44) yields (41).

3.3 Inverse Wishart distribution

The $p \times p$ inverse Wishart distribution with n $(n > p-1)$ degrees of freedom is denoted by $W_p^{-1}(n)$. When $V \sim W_p^{-1}(n)$, $X(h) = h'Vh$ for fixed $h \in \mathbb{S}^{p-1}$ is distributed as $1/\chi^2(n-p+1)$, the inverse chi-squared distribution with $n-p+1$ degrees of freedom. We can call $X(\cdot)$ an inverse chi-squared field. Recall that the inverse chi-squared distribution with ν degrees of freedom is the distribution of the reciprocal of the chi-squared random variable with ν degrees of freedom, and has the expectation $1/(\nu-2)$. In the following, let $G_\nu(\cdot)$ denote the cumulative distribution function of the chi-squared distribution with ν degrees of freedom.

Theorem 19 *The approximation formula for the distribution of the largest eigenvalue of the inverse Wishart matrix $W_p^{-1}(n)$ by the Euler characteristic heuristic is given by*

$$\widehat{P}(x) = \frac{2^{n-1}\Gamma(\frac{p+1}{2})\Gamma(\frac{n+1}{2})}{\sqrt{\pi}}$$

$$\times \sum_{j=0}^{p-1} (-2)^j \frac{\Gamma(\frac{n-p+1+2j}{2})}{\Gamma(p-j)\Gamma(n-p+j+1)j!} G_{n-p+1+2j}(x^{-1}). \tag{45}$$

The asymptotic error is evaluated as

$$\Delta P(x) \sim -\frac{\Gamma(n+1)}{\Gamma(p-1)\Gamma(n-p+4)\Gamma(n-p+3)} x^{-(n-p+2)} e^{-\frac{1}{x}} \quad (x \to \infty). \tag{46}$$

Remark 20 *Since $G_\nu(x^{-1}) = O(x^{-\frac{\nu}{2}})$, $\Delta P(x) = o(\widehat{P}(x))$. In addition, since $\nu = n+p-1$ in the last term of $\widehat{P}(x)$, the approximation error $\Delta P(x)$ is asymptotically smaller than each term of $\widehat{P}(x)$ whenever $(n+p-1)/2 < n-p+2$, i.e., $3p < n+5$.*

Remark 21 *As in Remark 14, we can propose an improved version of $\widehat{P}(x)$ as*

$$\widetilde{P}(x) = \widehat{P}(x) + \frac{\Gamma(n+1)}{\Gamma(p-1)\Gamma(n-p+4)\Gamma(n-p+3)} x^{-(n-p+2)} e^{-\frac{1}{x}}.$$

Remark 22 *Hanumara and Thompson (1968) proposed a procedure for approximating the distributions of the largest and smallest eigenvalues of the Wishart matrix. By some straightforward calculations, one can see that their procedure coincides with our formulas (32) and (45).*

For the inverse Wishart $W_2^{-1}(5)$, the exact upper probability of the largest eigenvalue $P(x)$, the approximation by the Euler characteristic heuristic $\widehat{P}(x)$, and its improvement $\widetilde{P}(x)$ are given as

$$P(x) = 1 - \Big(1 + \frac{1}{x} + \frac{1}{6x^2}\Big) e^{-\frac{1}{x}},$$

$$\widehat{P}(x) = \frac{1}{3x^2} e^{-\frac{1}{2x}}, \quad \text{and} \quad \widetilde{P}(x) = \widehat{P}(x) + \frac{1}{720x^5} e^{-\frac{1}{x}}.$$

These three are depicted in Fig. 2 as solid line, dotted line, and dashed line, respectively.

From now on we prove Theorem 19. We will show (45) at first. The density function of the inverse Wishart distribution is

$$f_{p,n}(V) = \frac{1}{c(p,n)} e^{-\frac{1}{2}\mathrm{tr}(V^{-1})} \det(V)^{-\frac{1}{2}(n+p+1)} \quad (V > 0).$$

The normalizing constant $c(p,n)$ is given in (34). For $v_{11} \in \mathbb{R}$ and $V_{22} \in Sym(p-1)$, it holds that

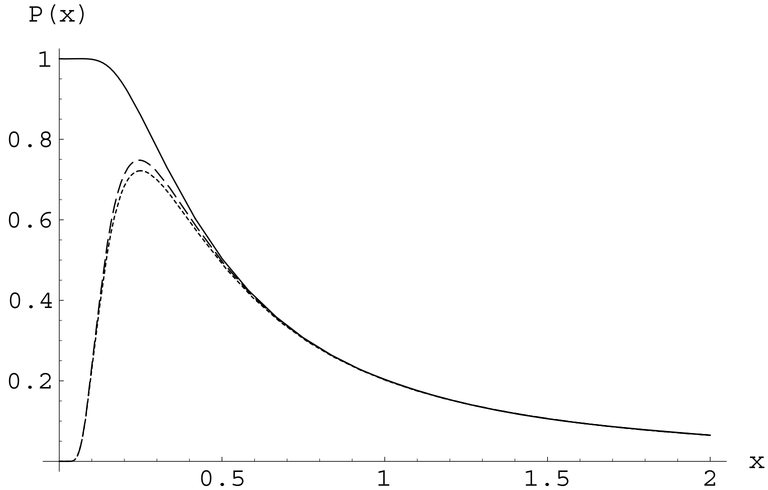

Fig. 2. $P(x)$ (—), $\widehat{P}(x)$ $(\cdots)$, and $\widetilde{P}(x)$ $(--)$ for the inverse Wishart $W_2^{-1}(5)$

$$f_{p,n}(\mathrm{diag}(v_{11}, V_{22})) = \frac{c(1, n+p-1)c(p-1, n+1)}{c(p,n)} \times f_{1,n+p-1}(v_{11}) f_{p-1,n+1}(V_{22}).$$

Hence, if we suppose that $v_{11} \sim 1/\chi^2(n+p-1)$ and $V_{22} \sim W_{p-1}^{-1}(n+1)$ are independent random variables, then we have

$$\widehat{P}(x) = \frac{\Omega_p}{2} \frac{c(1, n+p-1)c(p-1, n+1)}{c(p,n)} E[1_{\{v_{11} \geq x\}} \det(v_{11} I_{p-1} - V_{22})].$$

The multiplicative constant in the expression above is

$$\frac{\Omega_p}{2} \frac{c(1, n+p-1)c(p-1, n+1)}{c(p,n)} = 2^{p-1}\sqrt{\pi} \frac{\Gamma(\frac{n+p-1}{2})\Gamma(\frac{n+1}{2})}{\Gamma(\frac{p}{2})\Gamma(\frac{n-p+2}{2})\Gamma(\frac{n-p+1}{2})}.$$

According to the expansion of the determinant, we have

$$E[1_{\{v_{11} \geq x\}} \det(v_{11} I_{p-1} - V_{22})] = \sum_{j=0}^{p-1} (-1)^j E[1_{\{v_{11} \geq x\}} v_{11}^{p-1-j}] E[\mathrm{tr}_j(V_{22})].$$

The expectation with respect to v_{11} can be evaluated by reconsidering v_{11} as an inverse chi-squared random variable with $n+p-1-2(p-1-j) = n-p+1+2j$ degrees of freedom. It follows that

$$E[1_{\{v_{11} \geq x\}} v_{11}^{p-1-j}] = 2^{-(p-1-j)} \frac{\Gamma(\frac{n-p+1+2j}{2})}{\Gamma(\frac{n+p-1}{2})} G_{n-p+1+2j}(x^{-1}).$$

On the other hand, noting that $V_{22} \sim W_{p-1}^{-1}(n+1)$, and hence $j \times j$ principle minor of V_{22} is distributed as the inverse Wishart distribution with $(n+1)-$

$(p-1)+j = n-p+2+j$ degrees of freedom, we can write

$$
\begin{aligned}
E[\mathrm{tr}_j(V_{22})] &= \binom{p-1}{j} E\Big[\prod_{i=1}^{j} \frac{1}{\chi^2_{n-p+3+j-i}}\Big] \\
&= \binom{p-1}{j} \frac{1}{(n-p+j)(n-p+j-1)\cdots(n-p+1)} \\
&= \frac{\Gamma(p)\Gamma(n-p+1)}{\Gamma(p-j)\Gamma(n-p+j+1)j!},
\end{aligned}
$$

where χ^2_ν denotes an independent chi-squared random variable with ν degrees of freedom. Summarizing the above, we get

$$
\begin{aligned}
\widehat{P}(x) &= 2^{p-1}\sqrt{\pi} \frac{\Gamma(\frac{n+p-1}{2})\Gamma(\frac{n+1}{2})}{\Gamma(\frac{p}{2})\Gamma(\frac{n-p+2}{2})\Gamma(\frac{n-p+1}{2})} \\
&\quad \times \sum_{j=0}^{p-1} (-1)^j 2^{-(p-1-j)} \frac{\Gamma(\frac{n-p+1+2j}{2})}{\Gamma(\frac{n+p-1}{2})} G_{n-p+1+2j}(x^{-1}) \\
&\quad \times \frac{\Gamma(p)\Gamma(n-p+1)}{\Gamma(p-j)\Gamma(n-p+j+1)j!},
\end{aligned}
$$

which is reduced to (45). Here we used the duplication formula (36). The resulting expansion of $\widehat{P}(x)$ corresponds to the expansion (18) in Theorem 7.

Next we will derive the asymptotic evaluation of the approximation error (46). Let l_i be the eigenvalue of V, and let

$$
g(l) = l^{-\frac{n+p+1}{2}} e^{-\frac{1}{2l}}.
$$

Then

$$
f_{p,n}(V) = \frac{1}{c(p,n)} \prod_{i=1}^{p} g(l_i).
$$

Remark 20 follows again from the fact that (29) holds when $3p < n+5$. Recall that

$$
\Delta P(x) \sim -\frac{1}{c(p,n)} \prod_{i=1}^{p} \frac{\Omega_i}{2} \int_{\infty > l_2 > l_1 \ge x} g(l_1)g(l_2)(l_1 l_2)^{p-2}(l_2 - l_1)dl_1 dl_2 \times F,
$$

where

$$
F = \int_{\infty > l_3 > \cdots > l_p > 0} \prod_{i=3}^{p} g(l_i) \prod_{3 \le i < j \le p} (l_i - l_j) \prod_{i=3}^{p} dl_i
$$

for $p \geq 3$, and $F = 1$ for $p = 2$.

Assume that $p \geq 3$. By considering the inverse Wishart $W_{p-2}^{-1}(n+2)$, we have

$$c(p-2, n+2) = \prod_{i=1}^{p-2} \frac{\Omega_i}{2} \times F,$$

and hence

$$\Delta P(x) \sim -\frac{c(p-2,n+2)}{c(p,n)} \frac{\Omega_p}{2} \frac{\Omega_{p-1}}{2} \times \int\limits_{\infty > l_2 > l_1 \geq x} (l_1 l_2)^{-\frac{n-p+5}{2}} (l_2 - l_1) e^{-\frac{1}{2l_1} - \frac{1}{2l_2}} dl_1 dl_2.$$

The multiplicative constant contained above is

$$\begin{aligned} \frac{c(p-2,n+2)}{c(p,n)} \frac{\Omega_p}{2} \frac{\Omega_{p-1}}{2} &= \frac{2^{p-n-2} \pi \Gamma(\frac{n+2}{2}) \Gamma(\frac{n+1}{2})}{\Gamma(\frac{p}{2}) \Gamma(\frac{p-1}{2}) \Gamma(\frac{n-p+4}{2}) \Gamma(\frac{n-p+3}{2}) \Gamma(\frac{n-p+2}{2}) \Gamma(\frac{n-p+1}{2})} \\ &= \frac{\Gamma(n+1)}{4\Gamma(p-1)\Gamma(n-p+3)\Gamma(n-p+1)}, \end{aligned} \tag{47}$$

which is consistent with the constant for $p = 2$ in (38).

Write $k = \frac{n-p+5}{2}$ for simplicity. The integral with respect to l_1 and l_2 is evaluated by integration by parts as

$$\begin{aligned} &\int_x^\infty l_1^{-k} e^{-\frac{1}{2l_1}} dl_1 \int_{l_1}^\infty l_2^{-k+1} e^{-\frac{1}{2l_2}} dl_2 - \int_x^\infty l_1^{-k+1} e^{-\frac{1}{2l_1}} dl_1 \int_{l_1}^\infty l_2^{-k} e^{-\frac{1}{2l_2}} dl_2 \\ &\quad \sim \int_x^\infty l_1^{-k} e^{-\frac{1}{2l_1}} dl_1 \frac{1}{k-2} l_1^{-k+2} e^{-\frac{1}{2l_1}} - \int_x^\infty l_1^{-k+1} e^{-\frac{1}{2l_1}} dl_1 \frac{1}{k-1} l_1^{-k+1} e^{-\frac{1}{2l_1}} \\ &\quad \sim \frac{1}{(k-1)(k-2)(2k-3)} x^{-2k+3} e^{-\frac{1}{x}}. \end{aligned} \tag{48}$$

Multiplying (47) and (48) yields (46).

References

Adler, R. J., 1981. The geometry of random fields. Wiley, Chichester.

Adler, R. J., 2000. On excursion sets, tube formulas and maxima of random fields. The Annals of Applied Probability 10 (1), 1–74.

Hanumara, R. C., Thompson, Jr., W. A., 1968. Percentage points of the extreme roots of a Wishart matrix. Biometrika 55, 505–512.

Kuriki, S., Takemura, A., 2001. Tail probabilities of the maxima of multilinear forms and their applications. The Annals of Statistics 29 (2), 328–371.

Morse, M., Cairns, S. S., 1969. Critical point theory in global analysis and differential topology: An introduction. Pure and Applied Mathematics, Vol. 33. Academic Press, New York.

Muirhead, R. J., 1982. Aspects of multivariate statistical theory. Wiley, New York, wiley Series in Probability and Mathematical Statistics.

Okamoto, M., 1973. Distinctness of the eigenvalues of a quadratic form in a multivariate sample. The Annals of Statistics 1, 763–765.

Rudin, W., 1976. Principles of mathematical analysis, 3rd Edition. McGraw-Hill Book Co., New York, international Series in Pure and Applied Mathematics.

Takemura, A., Kuriki, S., 2002. On the equivalence of the tube and Euler characteristic methods for the distribution of the maximum of Gaussian fields over piecewise smooth domains. The Annals of Applied Probability 12 (2), 768–796.

Taylor, J. E., Adler, R. J., 2003. Euler characteristics for Gaussian fields on manifolds. The Annals of Probability 31 (2), 533–563.

Taylor, J. E., Takemura, A., Adler, R. J., 2005. Validity of the expected Euler characteristic heuristic. The Annals of Probability 33, in press.

Worsley, K. J., 1994. Local maxima and the expected Euler characteristic of excursion sets of χ^2, F and t fields. Advances in Applied Probability 26 (1), 13–42.

Worsley, K. J., 1995. Boundary corrections for the expected Euler characteristic of excursion sets of random fields, with an application to astrophysics. Advances in Applied Probability 27 (4), 943–959.