E l e c t r o n i c
J o u r n a l
o f
P r o b a b i l i t y

Stein's method and the multivariate CLT for traces of powers on the classical compact groups[*]

Christian Döbler[†]   Michael Stolz[‡]

Abstract

Let M_n be a random element of the unitary, special orthogonal, or unitary symplectic groups, distributed according to Haar measure. By a classical result of Diaconis and Shahshahani, for large matrix size n, the vector $(\mathrm{Tr}(M_n), \mathrm{Tr}(M_n^2), \ldots, \mathrm{Tr}(M_n^d))$ tends to a vector of independent (real or complex) Gaussian random variables. Recently, Jason Fulman has demonstrated that for a single power j (which may grow with n), a speed of convergence result may be obtained via Stein's method of exchangeable pairs. In this note, we extend Fulman's result to the multivariate central limit theorem for the full vector of traces of powers. .

Key words: random matrices, compact Lie groups, Haar measure, traces of powers, Stein's method, normal approximation, exchangeable pairs, heat kernel, power sum symmetric polynomials.

AMS 2010 Subject Classification: Primary 15B52 ; 60F05; Secondary: 60B15 ; 58J65.

[†]Ruhr-Universität Bochum, Fakultät für Mathematik, D-44780 Bochum, Germany. E-mail: christian.doebler@ruhr-uni-bochum.de

[‡]Ruhr-Universität Bochum, Fakultät für Mathematik, D-44780 Bochum, Germany. E-mail: michael.stolz@ruhr-uni-bochum.de

2375

One aspect of random matrix theory concerns random elements of compact Lie groups. A classical result, due to Diaconis and Shahshahani [DS94], is as follows: Let M_n be an element of U_n, O_n, or USp_{2n}, distributed according to Haar measure. Then, as $n \to \infty$, the vector

$$(\text{Tr}(M_n), \text{Tr}(M_n^2), \ldots, \text{Tr}(M_n^d))$$

converges weakly to a vector of independent, (real or complex) Gaussian random variables. The original proof deduced this from exact moment formulae, valid for n sufficiently large (see also [Sto05]). Different versions of the moment computations, also taking care of SO_n, have been proposed in [PV04] and [HR03].

Subsequently, the speed of convergence in the univariate version of this result was studied by Charles Stein [Ste95], who proved that in the orthogonal case the error decreases faster than any power of the dimension, and Kurt Johansson [Joh97], who obtained exponential convergence. While Johansson's approach had its roots in Szegö's limit theorem for Toeplitz determinants, Stein used the "exchangeable pairs" version of a set of techniques that he had been developing since the early 1970s (see [Ste72]) and that nowadays is referred to as "Stein's method". Recently, Jason Fulman [Ful10] has proposed an approach to the speed of convergence for a single power of M_n, based on combining Stein's method of exchangeable pairs with heat kernel techniques. While producing weaker results on the speed than Stein's and Johansson's, his theorems apply to the case that the power $M_n^{d(n)}$ grows with n. Furthermore, his techniques seem more likely to be useful in contexts beyond the standard representations of the classical groups.

In this note, we extend Fulman's approach and results to a multivariate setting, making use of recent extensions, due to Chatterjee, Meckes, Reinert, and Röllin [CM08, RR09, Mec09], of Stein's method of exchangeable pairs to cover multivariate normal approximations. This yields, to the best of our knowledge, the first rates of convergence result in the multivariate CLT for traces of powers on the classical compact groups, or, in any case, the first one that allows for the powers to grow with n. Our set-up contains Fulman's as a special case, and we recover, in Wasserstein distance, the rates that were proven by him in Kolmogorov distance (see Remark 1.2 below).

After stating the main result in Section 1, we will review Meckes' version of the multivariate exchangeable pairs method ([Mec09]) in Section 2, stating a complex version of her results that will be useful for the unitary case. This will lead to a slight improvement on Fulman's result even in the one dimensional (single power) case, in that now the convergence of the (complex) trace of a power of a random Haar unitary to a complex normal distribution can be quantified.

Fulman's approach consists in constructing useful exchangeable pairs from Brownian motion on the connected compact Lie groups $K_n = U_n, SO_n$, or USp_{2n}. This amounts to studying the heat kernel on K_n, and in particular the action of the Laplacian on power sum symmetric polynomials, which express products of traces of powers of group elements in terms of their eigenvalues. This is reviewed in Section 3. Explicit formulae are due to Rains [Rai97]. They are rooted in Schur-Weyl duality as explained in [Lév08].

In Sections 5, 6, and 7, we then prove the main result for the unitary, special orthogonal, and unitary symplectic groups.

1   The main result

Let $M = M_n$ be distributed according to Haar measure on $K_n \in \{U_n, SO_n, USp_{2n}\}$. For $d \in \mathbb{N}$, $r = 1, \ldots, d$, consider the r-dimensional (complex or real) random vector

$$W := W(d, r, n) := (f_{d-r+1}(M), f_{d-r+2}(M), \ldots, f_d(M))^T,$$

where $f_j(M) = \text{Tr}(M^j)$ in the unitary case,

$$f_j(M) = \begin{cases} \text{Tr}(M^j), & j \text{ odd}, \\ \text{Tr}(M^j) - 1, & j \text{ even} \end{cases}$$

in the orthogonal case and

$$f_j(M) = \begin{cases} \text{Tr}(M^j), & j \text{ odd}, \\ \text{Tr}(M^j) + 1, & j \text{ even} \end{cases}$$

in the symplectic case. In the orthogonal and symplectic cases, let $Z := (Z_{d-r+1} \ldots, Z_d)^T$ denote an r-dimensional real standard normal random vector. In the unitary case, Z is defined as standard *complex* normal, i.e., there are iid real random variables $X_{d-r+1}, \ldots, X_d$, $Y_{d-r+1}, \ldots, Y_d$ with distribution $N(0, 1/2)$ such that $Z_j = X_j + iY_j$ for $j = d - r + 1, \ldots, d$. In all cases, we take Σ to denote the diagonal matrix $\text{diag}(d - r + 1, d - r + 2, \ldots, d)$ and write $Z_\Sigma := \Sigma^{1/2} Z$.

THEOREM **1.1.** *If $n \geq 2d$ in the unitary and orthogonal cases, and $n \geq 4d + 1$ in the symplectic case, the Wasserstein distance between W and Z_Σ is*

$$d_{\mathscr{W}}(W, Z_\Sigma) = O\left(\frac{\max\left\{ \frac{r^{7/2}}{(d-r+1)^{3/2}}, (d-r)^{3/2}\sqrt{r} \right\}}{n} \right). \tag{1}$$

In particular, for $r = d$ we have

$$d_{\mathscr{W}}(W, Z_\Sigma) = O\left(\frac{d^{7/2}}{n} \right),$$

and for $r \equiv 1$

$$d_{\mathscr{W}}(W, Z_\Sigma) = O\left(\frac{d^{3/2}}{n} \right).$$

If $1 \leq r = \lfloor cd \rfloor$ for $0 < c < 1$, then

$$d_{\mathscr{W}}(W, Z_\Sigma) = O\left(\frac{d^2}{n} \right).$$

REMARK 1.2. The case $r \equiv 1$ means that one considers a (possibly shifted) single power $f_d(M)$. In his study [Ful10] of the univariate case, Fulman considers the random variable $\frac{f_d(M)}{\sqrt{d}}$ instead. By the scaling properties of the Wasserstein metric, the present result implies

$$d_{\mathscr{W}}\left(\frac{f_d(M)}{\sqrt{d}}, N(0, 1) \right) = O\left(\frac{d}{n} \right)$$

in the orthogonal and symplectic cases and an analogous result in the unitary case. So in the one-dimensional special case we recover the rate of convergence that was obtained by Fulman, albeit in Wasserstein rather than Kolmogorov distance.

2 Exchangeable pairs and multivariate normal approximation

The approach of univariate normal approximation by exchangeable pair couplings in Stein's method has a long history dating back to the monograph [Ste86] by Charles Stein in 1986. In order to show that a given real valued random variable W is approximately normally distributed, Stein proposes the construction of another random variable W' on the same probability space as W such that the pair (W, W') is exchangeable, i.e. satisfies the relation $(W, W') \overset{\mathscr{D}}{=} (W', W)$, and such that there exists $\lambda \in]0, 1[$ for which the linear regression property $\mathbb{E}[W - W'|W] = -\lambda W$ holds. In this situation, the distance between W and a standard normal Z can be bounded in various metrics, including Wasserstein's and Komogorov's.

The range of examples to which this technique could be applied was considerably extended in the work [RR97] of Rinott and Rotar who proved normal approximation theorems allowing the linear regression property to be satisfied only approximately. Specifically, they assumed the existence of a "small", random quantity R such that $\mathbb{E}[W' - W|W] = -\lambda W + R$.

In [CM08], Chatterjee and Meckes proposed a version of exchangeable pairs for multivariate normal aproximation. For a given random vector $W = (W_1, \ldots, W_d)^T$ they assume the existence of another random vector $W' = (W'_1, \ldots, W'_d)^T$ such that $W \overset{\mathscr{D}}{=} W'$ and of a constant $\lambda \in \mathbb{R}$ such that $\mathbb{E}[W' - W|W] = \lambda W$. In [RR09] Reinert and Röllin investigated the more general linear regression property $\mathbb{E}[W' - W|W] = -\Lambda W + R$, where now Λ is an invertible non-random $d \times d$ matrix and $R = (R_1, \ldots, R_d)^T$ is a small remainder term. However, in contrast to Chatterjee and Meckes, Reinert and Röllin need the full strength of the exchangeability of the vector (W, W'). Finally, in [Mec09], Elizabeth Meckes reconciled the two approaches, allowing for the more general linear regression property from [RR09] and using sharper coordinate-free bounds on the solution to the Stein equation suggested by those from [CM08].

Both [CM08] and [Mec09] contain an "infinitesimal version of Stein's method of exchangeable pairs", i.e., they provide error bounds for multivariate normal approximations in the case that for each $t > 0$ there is an exchangeable pair (W, W_t) and that some further limiting properties hold as $t \to 0$. It is such an infinitesimal version that will be applied in what follows.

The abstract multivariate normal approximation theorem provided by Meckes will be sufficient for the special orthogonal and symplectic cases of the present note, and the unitary case needs only a slight extension via basic linear algebra. So it is not really necessary to explain how these approximation theorems relate to the fundamentals of Stein's method (see [CS05] for a readable account). But it is crucial to understand that there are fundamental differences between the univariate and multivariate cases. So a few remarks are in order. The starting point of any version of Stein's method of normal approximation in the univariate case is the observation that a random variable W is standard normal if, and only if, for all f from a suitable class of piecewise C^1 functions one has

$$\mathbb{E}[f'(W) - W f(W)] = 0.$$

The quest for an analogous first order characterization of the multivariate normal distribution has proven unsuccessful, so Chatterjee and Meckes work with the following second-order substitute (see [Mec09, Lemma 1]): Let Σ be a positive semi-definite $d \times d$-matrix. A d-dimensional random vector X has the distribution $N(0, \Sigma)$ if, and only if, for all $f \in C_c^2(\mathbb{R}^d)$ the identity

$$\mathbb{E}\left[\langle \operatorname{Hess} f(X), \Sigma \rangle_{\mathrm{HS}} - \langle X, \nabla f(X) \rangle \right] = 0$$

holds (where HS stands for Hilbert-Schmidt, see below). Among the consequences of this is that the multivariate approximation theorems are phrased in Wasserstein distance (see below) rather than Kolmogorov's distance

$$\sup_{t \in \mathbb{R}} \left| \mu(] - \infty, t]) - \nu(] - \infty, t]) \right|$$

for probability measures μ, ν on the real line, i.e., the distance concept in which Fulman's univariate theorems are stated.

For a vector $x \in \mathbb{R}^d$ let $\|x\|_2$ denote its euclidean norm induced by the standard scalar product on $\mathbb{R}^d$ that will be denoted by $\langle \cdot, \cdot \rangle$. For $A, B \in \mathbb{R}^{d \times k}$ let $\langle A, B \rangle_{\mathrm{HS}} := \mathrm{Tr}(A^T B) = \mathrm{Tr}(B^T A) = \mathrm{Tr}(AB^T) = \sum_{i=1}^{d} \sum_{j=1}^{k} a_{ij} b_{ij}$ be the usual Hilbert-Schmidt scalar product on $\mathbb{R}^{d \times k}$ and denote by $\| \cdot \|_{\mathrm{HS}}$ the corresponding norm. For random matrices $M_n, M \in \mathbb{R}^{k \times d}$, defined on a common probability space $(\Omega, \mathscr{A}, \mathbb{P})$, we will say that M_n converges to M in $\mathrm{L}^1(\| \cdot \|_{\mathrm{HS}})$ if $\|M_n - M\|_{\mathrm{HS}}$ converges to 0 in $\mathrm{L}^1(\mathbb{P})$. For $A \in \mathbb{R}^{d \times d}$ let $\|A\|_{\mathrm{op}}$ denote the operator norm induced by the euclidean norm, i.e., $\|A\|_{\mathrm{op}} = \sup\{\|Ax\|_2 : \|x\|_2 = 1\}$. More generally, for a k-multilinear form $\psi : (\mathbb{R}^d)^k \to \mathbb{R}$ define the operator norm

$$\|\psi\|_{\mathrm{op}} := \sup \left\{ \psi(u_1, \ldots, u_k) : u_j \in \mathbb{R}^d, \|u_j\|_2 = 1, j = 1, \ldots, k \right\}.$$

For a function $h : \mathbb{R}^d \to \mathbb{R}$ define its minimum Lipschitz constant $M_1(h)$ by

$$M_1(h) := \sup_{x \neq y} \frac{|h(x) - h(y)|}{\|x - y\|_2} \in [0, \infty].$$

If h is differentiable, then $M_1(h) = \sup_{x \in \mathbb{R}^d} \|Dh(x)\|_{\mathrm{op}}$. More generally, for $k \geq 1$ and a $(k-1)$-times differentiable $h : \mathbb{R}^d \to \mathbb{R}$ let

$$M_k(h) := \sup_{x \neq y} \frac{\|D^{k-1}h(x) - D^{k-1}h(y)\|_{\mathrm{op}}}{\|x - y\|_2},$$

viewing the $(k-1)$-th derivative of h at any point as a $(k-1)$-multilinear form. Then, if h is actually k-times differentiable, we have $M_k(h) = \sup_{x \in \mathbb{R}^d} \|D^k h(x)\|_{\mathrm{op}}$. Having in mind this identity, we define $M_0(h) := \|h\|_\infty$.

Finally, recall the Wasserstein distance for probability distributions μ, ν on $(\mathbb{R}^d, \mathscr{B}^d)$. It is defined by

$$d_{\mathscr{W}}(\mu, \nu) := \sup \left\{ |\int h \, d\mu - \int h \, d\nu| : h : \mathbb{R}^d \to \mathbb{R} \text{ and } M_1(h) \leq 1 \right\}.$$

Now we are in a position to state the abstract multivariate normal approximation theorems that will be applied in this note, starting with the real version, taken from [Mec09, Thm. 4]. $Z = (Z_1, \ldots, Z_d)^T$ denotes a standard d-dimensional normal random vector, $\Sigma \in \mathbb{R}^{d \times d}$ a positive semi-definite matrix and $Z_\Sigma := \Sigma^{1/2} Z$ with distribution $\mathrm{N}(0, \Sigma)$.

PROPOSITION **2.1.** *Let W, W_t ($t > 0$) be $\mathbb{R}^d$-valued $\mathrm{L}^2(\mathbb{P})$ random vectors on the same probability space $(\Omega, \mathscr{A}, \mathbb{P})$ such that for any $t > 0$ the pair (W, W_t) is exchangeable. Suppose there exist an invertible non-random matrix Λ, a positive semi-definite matrix Σ, a random vector $R = (R_1, \ldots, R_d)^T$,*

a random $d \times d$-matrix S, a sub-σ-field $\mathscr{F}$ of $\mathscr{A}$ such that W is measurable w.r.t. $\mathscr{F}$ and a non-vanishing deterministic function $s :]0, \infty[\to \mathbb{R}$ such that the following three conditions are satisfied:

$$\frac{1}{s(t)} \mathbb{E}[W_t - W | \mathscr{F}] \xrightarrow{t \to 0} -\Lambda W + R \text{ in } L^1(\mathbb{P}). \tag{i}$$

$$\frac{1}{s(t)} \mathbb{E}[(W_t - W)(W_t - W)^T | \mathscr{F}] \xrightarrow{t \to 0} 2\Lambda\Sigma + S \text{ in } L^1(\|\cdot\|_{\mathrm{HS}}). \tag{ii}$$

$$\lim_{t \to 0} \frac{1}{s(t)} \mathbb{E}\left[\|W_t - W\|_2^2 \, \mathbb{1}_{\{\|W_t - W\|_2^2 > \epsilon\}} \right] = 0 \text{ for each } \epsilon > 0. \tag{iii}$$

Then

(a) *For each* $h \in C^2(\mathbb{R}^d)$,

$$|\mathbb{E}[h(W)] - \mathbb{E}[h(Z_\Sigma)]| \leq \|\Lambda^{-1}\|_{\mathrm{op}} \left(M_1(h)\mathbb{E}[\|R\|_2] + \frac{\sqrt{d}}{4} M_2(h) \, \mathbb{E}[\|S\|_{\mathrm{HS}}] \right).$$

(b) *If* Σ *is actually positive definite, then*

$$d_{\mathscr{W}}(W, Z_\Sigma) \leq \|\Lambda^{-1}\|_{\mathrm{op}} \left(\mathbb{E}[\|R\|_2] + \frac{1}{\sqrt{2\pi}} \|\Sigma^{-1/2}\|_{\mathrm{op}} \, \mathbb{E}[\|S\|_{\mathrm{HS}}] \right).$$

REMARK 2.2. In applications it is often easier to verify the following stronger condition in the place of (iii) of Proposition 2.1:

$$\lim_{t \to 0} \frac{1}{s(t)} \mathbb{E}\left[\|W_t - W\|_2^3 \right] = 0. \tag{iii$'$}$$

Now we turn to a version of Proposition 2.1 for complex random vectors, which will be needed for the case of the unitary group. $\|\cdot\|_{\mathrm{op}}$ and $\|\cdot\|_{\mathrm{HS}}$ extending in the obvious way, we now denote by $Z = (Z_1, \ldots, Z_d)$ a d-dimensional *complex* standard normal random vector, i.e., there are iid $N(0, 1/2)$ distributed real random variables $X_1, Y_1, \ldots, X_d, Y_d$ such that $Z_j = X_j + iY_j$ for all $j = 1, \ldots, d$.

PROPOSITION **2.3.** *Let* W, W_t $(t > 0)$ *be* $\mathbb{C}^d$*-valued* $L^2(\mathbb{P})$ *random vectors on the same probability space* $(\Omega, \mathscr{A}, \mathbb{P})$ *such that for any* $t > 0$ *the pair* (W, W_t) *is exchangeable. Suppose that there exist non-random matrices* $\Lambda, \Sigma \in \mathbb{C}^{d \times d}$, Λ *invertible,* Σ *positive semi-definite, a random vector* $R \in \mathbb{C}^d$, *random matrices* $S, T \in \mathbb{C}^{d \times d}$, *a sub-$\sigma$-field* $\mathscr{F}$ *of* $\mathscr{A}$ *such that* W *is measurable w.r.t.* $\mathscr{F}$, *and a non-vanishing deterministic function* $s :]0, \infty[\to \mathbb{R}$ *with the following properties:*

$$\frac{1}{s(t)} \mathbb{E}[W_t - W | \mathscr{F}] \xrightarrow{t \to 0} -\Lambda W + R \text{ in } L^1(\mathbb{P}). \tag{i}$$

$$\frac{1}{s(t)} \mathbb{E}\left[(W_t - W)(W_t - W)^* | \mathscr{F} \right] \xrightarrow{t \to 0} 2\Lambda\Sigma + S \text{ in } L^1(\|\cdot\|_{\mathrm{HS}}). \tag{ii}$$

$$\frac{1}{s(t)} \mathbb{E}\left[(W_t - W)(W_t - W)^T | \mathscr{F} \right] \xrightarrow{t \to 0} T \text{ in } L^1(\|\cdot\|_{\mathrm{HS}}). \tag{iii}$$

$$\lim_{t \to 0} \frac{1}{s(t)} \mathbb{E}\left[\|W_t - W\|_2^2 \, \mathbb{1}_{\{\|W_t - W\|_2^2 > \epsilon\}} \right] = 0 \text{ for each } \epsilon > 0. \tag{iv}$$

Then,

(a) *If $h \in C^2(\mathbb{R}^{2d})$, then*

$$|\mathbb{E}h(W) - \mathbb{E}h(\Sigma^{1/2}Z)| \leq \|\Lambda^{-1}\|_{\mathrm{op}}\left(M_1(h)\mathbb{E}\|R\|_2 + \frac{\sqrt{d}}{4}\mathbb{E}\left[\|S\|_{\mathrm{HS}} + \|T\|_{\mathrm{HS}}\right]\right).$$

(b) *If Σ is actually positive definite, then*

$$d_{\mathcal{W}}(W, \Sigma^{1/2}Z) \leq \|\Lambda^{-1}\|_{\mathrm{op}}\left(\mathbb{E}\|R\|_2 + \frac{1}{\sqrt{2\pi}}\|\Sigma^{-1/2}\|_{\mathrm{op}}\mathbb{E}\left[\|S\|_{\mathrm{HS}} + \|T\|_{\mathrm{HS}}\right]\right).$$

REMARK 2.4.

(i) Remark 2.2 above also applies to the present Condition (iv).

(ii) If Y is a centered d-dimensional complex normal random vector, then there exist $A \in \mathbb{C}^{d \times d}$ and a complex standard normal $Z \in \mathbb{C}^d$ such that $Y = AZ$. In this case, $\mathbb{E}[YY^T] = A\mathbb{E}[ZZ^T]A^T = 0$, since $\mathbb{E}[ZZ^T] = 0$. This heuristically motivates Condition (iii) above for W close to such a Y, if one thinks of T as a small remainder term. For an algebraic rationale see (3) below.

Proof of Proposition 2.3. The proof is by reduction to the real version in Proposition 2.1. To this end, we need some preparation. To each $z = (z_1, \ldots, z_d)^T \in \mathbb{C}^d$ assign the vector

$$z_{\mathbb{R}} := (\mathrm{Re}(z_1), \mathrm{Im}(z_1), \ldots, \mathrm{Re}(z_d), \mathrm{Im}(z_d))^T \in \mathbb{R}^{2d}.$$

Given $A = (a_{jk}) \in \mathbb{C}^{d \times d}$ define $A_{\mathbb{R}} \in \mathbb{R}^{2d \times 2d}$ uniquely by requiring that $(Az)_{\mathbb{R}} = A_{\mathbb{R}}z_{\mathbb{R}}$ for all $z \in \mathbb{C}$. In concrete terms, $A_{\mathbb{R}}$ consists of 2×2 blocks, where the block at position (j, k) is given as

$$\begin{pmatrix} \mathrm{Re}(a_{jk}) & -\mathrm{Im}(a_{jk}) \\ \mathrm{Im}(a_{jk}) & \mathrm{Re}(a_{jk}) \end{pmatrix}.$$

Observe that

$$(AB)_{\mathbb{R}} = A_{\mathbb{R}}B_{\mathbb{R}}, \ (A^*)_{\mathbb{R}} = (A_{\mathbb{R}})^T, \ \text{and} \ (AA^*)_{\mathbb{R}} = A_{\mathbb{R}}(A^*)_{\mathbb{R}} = A_{\mathbb{R}}(A_{\mathbb{R}})^T, \tag{2}$$

and that for $d = 1$ one has to specify whether a scalar is to be interpreted as a vector or as a matrix. Writing J for the Kronecker product $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \otimes I_d$, one easily verifies the following identity of real $2d \times 2d$ matrices:

$$z_{\mathbb{R}}w_{\mathbb{R}}^T = \frac{1}{2}(zw^*)_{\mathbb{R}} + \frac{1}{2}(zw^T)_{\mathbb{R}}J. \tag{3}$$

Now, given the data from Proposition 2.3, we will show that the pairs $(W_{\mathbb{R}}, (W_t)_{\mathbb{R}})$ satisfy the assumptions of Proposition 2.1 with $\Lambda_{\mathbb{R}}$ for Λ, $\Sigma' := \frac{1}{2}\Sigma_{\mathbb{R}}$ for Σ, $R_{\mathbb{R}}$ for R and $\frac{1}{2}(S_{\mathbb{R}} + T_{\mathbb{R}}J)$ for S. Clearly $(W_{\mathbb{R}}, (W_t)_{\mathbb{R}})$ is exchangeable and $\Lambda_{\mathbb{R}}$ is invertible. By (2), Σ' is symmetric and non-negative definite. To verify Assumption (i), observe that

$$\frac{1}{s(t)}\mathbb{E}[(W_t - W)_{\mathbb{R}}|\mathscr{F}] = \frac{1}{s(t)}\mathbb{E}[(W_t - W)|\mathscr{F}]_{\mathbb{R}}$$

$$\xrightarrow{t\to 0} (-\Lambda W + R)_{\mathbb{R}} = -\Lambda_{\mathbb{R}} W_{\mathbb{R}} + R_{\mathbb{R}}.$$

For (ii), by (3) above, we have

$$\frac{1}{s(t)}\mathbb{E}[(W_t - W)_{\mathbb{R}}(W_t - W)_{\mathbb{R}}^T | \mathscr{F}]$$

$$= \frac{1}{s(t)}\left(\frac{1}{2}\mathbb{E}[(W_t - W)(W_t - W)^* | \mathscr{F}]_{\mathbb{R}} + \frac{1}{2}\mathbb{E}[(W_t - W)(W_t - W)^T | \mathscr{F}]_{\mathbb{R}} J\right)$$

$$\xrightarrow{t\to 0} \frac{1}{2}(2\Lambda\Sigma + S)_{\mathbb{R}} + \frac{1}{2}T_{\mathbb{R}}J = \Lambda_{\mathbb{R}}\Sigma_{\mathbb{R}} + \frac{1}{2}(S_{\mathbb{R}} + T_{\mathbb{R}}J)$$

$$= 2\Lambda_{\mathbb{R}}\Sigma' + \frac{1}{2}(S_{\mathbb{R}} + T_{\mathbb{R}}J).$$

Since $\|(W_t - W)_{\mathbb{R}}\|_2 = \|W_t - W\|_2$, Condition (iii) is trivially satisfied. So we may apply Proposition 2.1 to the real version $W_{\mathbb{R}}$ of W. Note that if $\tilde{Z}$ is distributed according to $N(0, I_{2d})$, then $(\Sigma')^{1/2}\tilde{Z}$ is equal in distribution to $(\Sigma^{1/2}Z)_{\mathbb{R}}$. Observing that $\|\Lambda^{-1}\|_{\mathrm{op}} = \|\Lambda_{\mathbb{R}}^{-1}\|_{\mathrm{op}}, \|S_{\mathbb{R}}\|_{\mathrm{HS}} = \sqrt{2}\|S\|_{\mathrm{HS}}, \|T_{\mathbb{R}}J\|_{\mathrm{HS}} = \|T_{\mathbb{R}}\|_{\mathrm{HS}} = \sqrt{2}\|T\|_{\mathrm{HS}}, \mathbb{E}\|R\|_2 = \mathbb{E}\|R_{\mathbb{R}}\|_2$ and $\|(\Sigma')^{-1/2}\|_{\mathrm{op}} = \sqrt{2}\|\Sigma_{\mathbb{R}}^{-1}\|_{\mathrm{op}} = \sqrt{2}\|\Sigma^{-1}\|_{\mathrm{op}}$ yields the result. $\square$

Now we construct a family of exchangeable pairs to fit into Meckes' set-up. We do this along the lines of Fulman's univariate approach in [Ful10]. In a nutshell, let $(M_t)_{t\geq 0}$ be Brownian motion on K_n with Haar measure as initial distribution. Set $M := M_0$. Brownian motion being reversible w.r.t. Haar measure, (M, M_t) is an exchangeable pair for any $t > 0$. Suppose that the centered version W of the statistic we are interested in is given by $W = (f_1(M), \ldots, f_d(M))^T$ for suitable measurable functions $f_1, \ldots, f_d$. Defining $W_t = (f_1(M_t), \ldots, f_d(M_t))^T$ clearly yields an exchangeable pair (W, W_t).

To be more specific, let $\mathfrak{k}_n$ be the Lie algebra of K_n, endowed with the scalar product $\langle X, Y\rangle = \mathrm{Tr}(X^*Y)$. Denote by $\Delta = \Delta_{K_n}$ the Laplace–Beltrami operator of K_n, i.e., the diffential operator corresponding to the Casimir element of the enveloping algebra of $\mathfrak{k}_n$. Then $(M_t)_{t\geq 0}$ will be the diffusion on K_n with infinitesimal generator Δ and Haar measure as initial distribution. Reversibility then follows from general theory (see [Hel00, Section II.2.4], [IW89, Section V.4], for the relevant facts). Let $(T_t)_{t\geq 0}$, often symbolically written as $\left(e^{t\Delta}\right)_{t\geq 0}$, be the corresponding semigroup of transition operators on $C^2(K_n)$. From the Markov property of Brownian motion we obtain that

$$\mathbb{E}[f(M_t)|M] = (T_t f)(M) \quad \text{a.s.} \tag{4}$$

Note that $(t, g) \mapsto (T_t f)(g)$ satisfies the heat equation, so a Taylor expansion in t yields the following lemma, which is one of the cornerstones in Fulman's approach in that it shows that in first order in t a crucial quantity for the regression conditions of Propositions 2.1 and 2.3 is given by the action of the Laplacian.

LEMMA **2.5.** *Let $f : K_n \to \mathbb{C}$ be smooth. Then*

$$\mathbb{E}[f(M_t)|M] = f(M) + t(\Delta f)(M) + O(t^2).$$

REMARK 2.6. An elementary construction of the semigroup (T_t) via an eigenfunction expansion in irreducible characters of K_n can be found in [Ste70]. This construction immediately implies Lemma 2.5 for power sums, see Section 3, since they are characters of (reducible) tensor power representations.

REMARK 2.7. As K_n is compact, Lemma 2.5 implies that $\lim_{t\to 0}\mathbb{E}[f(M_t) - f(M)|M] = 0$ a.s. and in $L^2(\mathbb{P})$. Arguments of this type will occur frequently in what follows, usually without further notice.

3   Power sums and Laplacians

For indeterminates $X_1,\ldots,X_n$ and a finite family $\lambda = (\lambda_1,\ldots,\lambda_r)$ of positive integers, the power sum symmetric polynomial with index λ is given by

$$p_\lambda := \prod_{k=1}^{r}\sum_{j=1}^{n}X_j^{\lambda_k}.$$

For $A \in \mathbb{C}^{n\times n}$ with eigenvalues $c_1,\ldots,c_n$ (not necessarily distinct), we write $p_\lambda(A)$ in the place of $p_\lambda(c_1,\ldots,c_n)$. Then one has the identity

$$p_\lambda(A) = \prod_{k=1}^{r}\mathrm{Tr}(A^{\lambda_k}).$$

Using this formula, we will extend the definition of p_λ to integer indices by $p_0(A) = \mathrm{Tr}(I)$, $p_{-k}(A) = \mathrm{Tr}((A^{-1})^k)$ and $p_\lambda(A) = \prod_{j=1}^{r}p_j(A)$.

In particular, the p_λ may be viewed as functions on K_n, and the action of the Laplacian Δ_{K_n} on them, useful in view of Lemma 2.5, is available from [Rai97, Lév08]. We specialize their formulae to the cases that will be needed in what follows:

LEMMA **3.1.** *For the Laplacian Δ_{U_n} on U_n, one has*

$$\Delta_{U_n}p_j = -njp_j - j\sum_{l=1}^{j-1}p_{l,j-l}. \tag{i}$$

$$\Delta_{U_n}p_{j,k} = -n(j+k)p_{j,k} - 2jkp_{j+k} - jp_k\sum_{l=1}^{j-1}p_{l,j-l} - kp_j\sum_{l=1}^{k-1}p_{l,k-l}. \tag{ii}$$

$$\Delta_{U_n}(p_j\overline{p_k}) = 2jkp_{j-k} - n(j+k)p_j\overline{p_k} - j\overline{p_k}\sum_{l=1}^{j-1}p_{l,j-l} - kp_j\sum_{l=1}^{k-1}\overline{p_{l,k-l}}. \tag{iii}$$

LEMMA **3.2.** *For the Laplacian Δ_{SO_n} on SO_n,*

$$\Delta_{SO_n}p_j = -\frac{(n-1)}{2}jp_j - \frac{j}{2}\sum_{l=1}^{j-1}p_{l,j-l} + \frac{j}{2}\sum_{l=1}^{j-1}p_{2l-j}. \tag{i}$$

$$\begin{aligned}
\Delta_{SO_n}p_{j,k} = &-\frac{(n-1)(j+k)}{2}p_{j,k} - \frac{j}{2}p_k\sum_{l=1}^{j-1}p_{l,j-l} - \frac{k}{2}p_j\sum_{l=1}^{k-1}p_{l,k-l} \\
&- jk\,p_{j+k} + \frac{j}{2}p_k\sum_{l=1}^{j-1}p_{j-2l} + \frac{k}{2}p_j\sum_{l=1}^{k-1}p_{k-2l} + jk\,p_{j-k}.
\end{aligned} \tag{ii}$$

LEMMA **3.3.** *For the Laplacian* $\Delta_{\mathrm{USp}_{2n}}$ *on* USp_{2n},

$$\Delta_{\mathrm{USp}_{2n}} p_j = -\frac{(2n+1)}{2} jp_j - \frac{j}{2} \sum_{l=1}^{j-1} p_{l,j-l} - \frac{j}{2} \sum_{l=1}^{j-1} p_{2l-j}. \tag{i}$$

$$\Delta_{\mathrm{USp}_{2n}} p_{j,k} = -\frac{(2n+1)}{2} (j+k) \, p_{j,k} - jk \, p_{j+k} - \frac{j}{2} \, p_k \sum_{l=1}^{j-1} p_{l,j-l}$$

$$- \frac{k}{2} \, p_j \sum_{l=1}^{k-1} p_{l,k-l} - \frac{j}{2} \, p_k \sum_{l=1}^{j-1} p_{j-2l} - \frac{k}{2} \, p_j \sum_{l=1}^{k-1} p_{k-2l} + jk \, p_{j-k}. \tag{ii}$$

In what follows, we will need to integrate certain p_λ over the group K_n. Thus we will need special cases of the Diaconis-Shahshahani moment formulae that we now recall (see [DS94, HR03, PV04, Sto05] for proofs). Let $a = (a_1,\ldots,a_r)$, $b = (b_1,\ldots,b_q)$ be families of nonnegative integers and define

$$f_a(j) := \begin{cases} 1 & \text{if } a_j = 0, \\ 0 & \text{if } ja_j \text{ is odd}, a_j \geq 1, \\ j^{a_j/2}(a_j - 1)!! & \text{if } j \text{ is odd and } a_j \text{ is even}, a_j \geq 2, \\ 1 + \sum_{d=1}^{[a_j/2]} j^d \binom{a_j}{2d}(2d-1)!! & \text{if } j \text{ is even}, a_j \geq 1. \end{cases}$$

Here we have used the notation $(2m-1)!! = (2m-1)(2m-3)\cdot\ldots\cdot 3\cdot 1$. Further, we will write

$$k_a := \sum_{j=1}^{r} ja_j, \quad k_b := \sum_{j=1}^{q} jb_j, \quad \text{and} \quad \eta_j := \begin{cases} 1, & \text{if } j \text{ is even}, \\ 0, & \text{if } j \text{ is odd}. \end{cases}$$

LEMMA **3.4.** *Let* $M = M_n$ *be a Haar-distributed element of* U_n, $Z_1,\ldots,Z_r$ *iid complex standard normals. Then, if* $k_a \neq k_b$,

$$\alpha_{(a,b)} := \mathbb{E}\left(\prod_{j=1}^{r} (\mathrm{Tr}(M^j))^{a_j} \prod_{j=1}^{q} \overline{(\mathrm{Tr}(M^j))}^{b_j} \right) = 0.$$

If $k_a = k_b$ *and* $n \geq k_a$, *then*

$$\alpha_{(a,b)} = \delta_{a,b} \prod_{j=1}^{r} j^{a_j} a_j! = \mathbb{E}\left(\prod_{j=1}^{r} (\sqrt{j}Z_j)^{a_j} \prod_{j=1}^{q} \overline{(\sqrt{j}Z_j)}^{b_j} \right).$$

LEMMA **3.5.** *If* $M = M_n$ *is a Haar-distributed element of* SO_n, $n - 1 \geq k_a$, $Z_1,\ldots,Z_r$ *iid real standard normals, then*

$$\mathbb{E}\left(\prod_{j=1}^{r} (\mathrm{Tr}(M^j))^{a_j} \right) = \mathbb{E}\left(\prod_{j=1}^{r} (\sqrt{j}Z_j + \eta_j)^{a_j} \right) = \prod_{j=1}^{r} f_a(j). \tag{5}$$

LEMMA **3.6.** *If* $M = M_{2n}$ *is a Haar-distributed element of* USp_{2n}, $2n \geq k_a$, $Z_1,\ldots,Z_r$ *iid real standard normals, then*

$$\mathbb{E}\left(\prod_{j=1}^{r} (\mathrm{Tr}(M^j))^{a_j} \right) = \mathbb{E}\left(\prod_{j=1}^{r} (\sqrt{j}Z_j - \eta_j)^{a_j} \right) = \prod_{j=1}^{r} (-1)^{(j-1)a_j} f_a(j). \tag{6}$$

4 A useful lemma

In order to avoid repetitions in what follows, let us state the following lemma:

LEMMA **4.1.** *Let a, r, p be non-negative integers, $r > 0$. Then*

$$\sum_{k=a}^{a+r-1} k^p \le \frac{2^{p+1}}{p+1} \max\{ra^p, r^{p+1}\}.$$

Proof. Since $f(x) = x^p$ is non-decreasing on the non-negative real numbers,

$$\sum_{k=a}^{a+r-1} k^p \le \int_a^{a+r} x^p dx = \frac{(a+r)^{p+1} - a^{p+1}}{p+1}$$

$$= \frac{r}{p+1} \sum_{k=0}^{p} \binom{p+1}{k} a^k r^{p-k} \le \frac{r \max\{a^p, r^p\}}{p+1} \sum_{k=0}^{p} \binom{p+1}{k}$$

$$\le \frac{2^{p+1}}{p+1} \max\{ra^p, r^{p+1}\}.$$

$\square$

5 Proof of the main result in the unitary case

In order to prove Theorem 1 for the unitary group, we invoke the construction that was explained in Section 2 above, yielding a family $(W, W_t)_{t>0}$ of exchangeable pairs such that $W_t = W_t(d, r, n) = (\mathrm{Tr}(M_t^{d-r+1}), \mathrm{Tr}(M_t^{d-r+2}), \dots, \mathrm{Tr}(M_t^d))^T$. We have to check the conditions of Proposition 2.3. As to (i), from Lemma 2.5 and Lemma 3.1(i) we obtain

$$\mathbb{E}[p_j(M_t)|M] = p_j(M) + t(\Delta p_j)(M) + \mathrm{O}(t^2)$$

$$= p_j(M) - tnjp_j(M) - tj \sum_{l=1}^{j-1} p_{l,j-l}(M) + \mathrm{O}(t^2),$$

hence

$$\mathbb{E}[p_j(M_t) - p_j(M)|M] = -t\left(njp_j(M) + j \sum_{l=1}^{j-1} p_{l,j-l}(M) + \mathrm{O}(t) \right).$$

From this we see that

$$\frac{1}{t}\mathbb{E}[W_t - W|M] = \frac{1}{t}\left(\begin{array}{c} \mathbb{E}[p_{d-r+1}(M_t) - p_{d-r+1}(M)|M] \\ \mathbb{E}[p_{d-r+2}(M_t) - p_{d-r+2}(M)|M] \\ \vdots \\ \mathbb{E}[p_d(M_t) - p_d(M)|M] \end{array} \right)$$

tends, as $t \to 0$, a.s. and in $\mathrm{L}^1(\| \cdot \|_{\mathrm{HS}})$ to

$$
-\begin{pmatrix}
(d-r+1)\left[np_{d-r+1} + \sum_{l=1}^{d-r} p_{l,d-r+1-l} \right](M) \\
(d-r+2)\left[np_{d-r+2} + \sum_{l=1}^{d-r+1} p_{l,d-r+2-l} \right](M) \\
\vdots \\
d\left[np_d + \sum_{l=1}^{d-1} p_{l,d-l} \right](M)
\end{pmatrix} =: -\Lambda W + R,
$$

where $\Lambda = \mathrm{diag}(nj : j = d-r+1,\dots,d)$ and $R = (R_{d-r+1},\dots,R_d)^T$ with $R_j = -j \sum_{l=1}^{j-1} p_{l,j-l}(M)$. (See Remark 2.7 above for the type of convergence.)

To verify Conditions (ii) and (iii) of Proposition 2.3, we first prove the following lemma.

LEMMA **5.1.** *For $j,k = d-r+1,\dots,d$, one has*

$$
\mathbb{E}[(p_j(M_t) - p_j(M))(p_k(M_t) - p_k(M)) \mid M] = -2tjk\, p_{j+k}(M) + \mathrm{O}(t^2) \tag{i}
$$

$$
\mathbb{E}[(p_j(M_t) - p_j(M))(\overline{p_k}(M_t) - \overline{p_k}(M)) \mid M] = 2tjk\, p_{j-k}(M) + \mathrm{O}(t^2) \tag{ii}
$$

Proof. By well known properties of conditional expectation,

$$
\mathbb{E}[(p_j(M_t) - p_j(M))(p_k(M_t) - p_k(M))|M]
$$
$$
= \mathbb{E}[p_j(M_t)p_k(M_t)|M] - p_j(M)\mathbb{E}[p_k(M_t)|M] - p_k(M)\mathbb{E}[p_j(M_t)|M] + p_j(M)p_k(M).
$$

Applying Lemma 2.5 and 3.1(ii) to the first term yields

$$
\mathbb{E}[p_{j,k}(M_t)|M] = p_{j,k}(M) + t(\Delta p_{j,k})(M) + \mathrm{O}(t^2)
$$
$$
= p_{j,k}(M) + t\left(-n(j+k)p_{j,k}(M) - 2jkp_{j+k}(M) - jp_k(M)\sum_{l=1}^{j-1} p_{l,j-l}(M) \right.
$$
$$
\left. -kp_j(M)\sum_{l=1}^{k-1} p_{l,k-l}(M) \right) + \mathrm{O}(t^2).
$$

Analogously, for the second term,

$$
p_j(M)\mathbb{E}[p_k(M_t)|M] = p_j(M)(p_k(M) + t(\Delta p_k)(M) + \mathrm{O}(t^2))
$$
$$
= p_{j,k}(M) + tp_j(M)\left(-nkp_k(M) - k\sum_{l=1}^{k-1} p_{l,k-l}(M) \right) + \mathrm{O}(t^2)
$$
$$
= p_{j,k}(M) - tnkp_{j,k}(M) - tkp_j(M)\sum_{l=1}^{k-1} p_{l,k-l}(M) + \mathrm{O}(t^2),
$$

and by symmetry

$$
p_k(M)\mathbb{E}[p_j(M_t)|M] = p_{j,k}(M) - tnjp_{j,k}(M) - tjp_k(M)\sum_{l=1}^{j-1} p_{l,j-l}(M) + \mathrm{O}(t^2).
$$

Summing up, one obtains

$$\mathbb{E}[(p_j(M_t) - p_j(M))(p_k(M_t) - p_k(M))|M] = -2tjkp_{j+k}(M) + \mathrm{O}(t^2),$$

proving the first assertion. For the second, we compute analogously

$$
\begin{aligned}
\mathbb{E}[(p_j(M_t) - p_j(M))(\overline{p_k}(M_t) - \overline{p_k}(M)|M] \\
= \mathbb{E}[p_j\overline{p_k}(M_t)|M] - p_j(M)\mathbb{E}[\overline{p_k}(M_t)|M] - \overline{p_k}(M)\mathbb{E}[p_j(M_t)|M] + p_j\overline{p_k}(M),
\end{aligned}
$$

and we have by Lemma 3.1(*iii*)

$$
\begin{aligned}
\mathbb{E}[p_j\overline{p_k}(M_t)|M] &= p_j\overline{p_k}(M) + t\Delta(p_j\overline{p_k})(M) + \mathrm{O}(t^2) \\
&= p_j\overline{p_k}(M) + t\Bigg(2jkp_{j-k} - n(j+k)p_j\overline{p_k}(M) \\
&\qquad - j\overline{p_k}(M) \sum_{l=1}^{j-1} p_{l,j-l}(M) - kp_j(M) \sum_{l=1}^{k-1} \overline{p_{l,k-l}}(M) \Bigg) + \mathrm{O}(t^2)
\end{aligned}
$$

as well as

$$
\begin{aligned}
p_j(M)\mathbb{E}[\overline{p_k}(M_t)|M] &= p_j(M)\left(\overline{p_k}(M) + t(\Delta\overline{p_k})(M) + \mathrm{O}(t^2) \right) \\
&= p_j\overline{p_k}(M) + tp_j(M)\left(-nk\overline{p_k}(M) - k\sum_{l=1}^{k-1} \overline{p_{l,k-l}}(M) \right) + \mathrm{O}(t^2)
\end{aligned}
$$

and

$$
\begin{aligned}
\overline{p_k}(M)\mathbb{E}[p_j(M_t)|M] &= \overline{p_k}(M)\left(p_j(M) + t(\Delta p_j)(M) + \mathrm{O}(t^2) \right) \\
&= p_j\overline{p_k}(M) + t\,\overline{p_k}(M)\left(-njp_j(M) - j\sum_{l=1}^{j-1} p_{l,j-l}(M) \right) + \mathrm{O}(t^2).
\end{aligned}
$$

Summing up, one has

$$\mathbb{E}[(p_j(M_t) - p_j(M))(\overline{p_k}(M_t) - \overline{p_k}(M))|M] = 2tjk\,p_{j-k} + \mathrm{O}(t^2).$$

$\square$

Now we are in a position to identify the random matrices S, T of Proposition 2.3. By Lemma 5.1, $\frac{1}{t}\mathbb{E}[(W_t - W)(W_t - W)^T|M]$ converges almost surely, and in $\mathrm{L}^1(\|\cdot\|_{\mathrm{HS}})$, to $T = (t_{jk})_{j,k=d-r+1,\dots,d}$, where $t_{jk} = -2jkp_{j+k}(M)$ for $j, k = d - r + 1, \dots, d$. Observing that $\Lambda\Sigma = \mathrm{diag}(nj^2 : j = d - r + 1, \dots, d)$, one has that $\frac{1}{t}\mathbb{E}[(W_t - W)(W_t - W)^*|M]$ converges almost surely, and in $\mathrm{L}^1(\|\cdot\|_{\mathrm{HS}})$, to $2\Lambda\Sigma + S$, where $S = (s_{jk})_{j,k=d-r+1,\dots,d}$ is given by

$$
s_{jk} = \begin{cases} 0, & \text{if } j = k, \\ 2jk\,p_{j-k}(M), & \text{if } j \neq k. \end{cases}
$$

2387

Next we will verify Condition (iv), using Remark 2.4. Specifically, we will show that $\mathbb{E}[\|W_t - W\|_2^3] = \mathrm{O}(t^{3/2})$. Since

$$
\begin{aligned}
\mathbb{E}[\|W_t - W\|_2^3] &\leq \sum_{j,k,l=d-l+1}^{d} \mathbb{E}[|(W_{t,j} - W_j)(W_{t,k} - W_k)(W_{t,l} - W_l)|] \\
&\leq \sum_{j,k,l=d-r+1}^{d} \left(\mathbb{E}[|W_{t,j} - W_j|^3] \mathbb{E}[|W_{t,k} - W_k|^3] \mathbb{E}[|W_{t,l} - W_l|^3] \right)^{1/3},
\end{aligned}
$$

it suffices to prove that $\mathbb{E}[|W_{t,j} - W_j|^3] = \mathrm{O}(t^{3/2})$ for all $j = d - r + 1, \ldots, d$. This in turn follows from the next lemma, since

$$
\mathbb{E}[|W_{t,j} - W_j|^3] \leq \left(\mathbb{E}[|W_{t,j} - W_j|^4] \mathbb{E}[|W_{t,j} - W_j|^2] \right)^{1/2}.
$$

LEMMA **5.2.** *For $j = d - r + 1, \ldots, d$, $n \geq 2d$,*

$$
\mathbb{E}[|W_{t,j} - W_j|^2] = 2j^2 n t + \mathrm{O}(t^2), \tag{i}
$$
$$
\mathbb{E}[|W_{t,j} - W_j|^4] = \mathrm{O}(t^2). \tag{ii}
$$

Proof. By Lemma 5.1 (ii),

$$
\mathbb{E}[|W_{t,j} - W_j|^2] = \mathbb{E}[(p_j(M_t) - p_j(M))(\overline{p_j}(M_t) - \overline{p_j}(M))] = \mathbb{E}[2tj^2 n + t^2] = 2tj^2 n + \mathrm{O}(t^2),
$$

establishing (i). Turning to (ii), one calculates that

$$
\begin{aligned}
\mathbb{E}[|W_{t,j} - W_j|^4] &= \mathbb{E}\left[(W_{t,j} - W_j)^2 \cdot (\overline{W_{t,j}} - \overline{W_j})^2 \right] \\
&= \mathbb{E}\left[(p_j(M_t) - p_j(M))^2 (\overline{p_j}(M_t) - \overline{p_j}(M))^2 \right] \\
&= \mathbb{E}\left[(p_j^2(M_t) - 2p_j(M)p_j(M_t) + p_j^2(M))(\overline{p_j}^2(M_t) - 2\overline{p_j}(M)\overline{p_j}(M_t) + \overline{p_j}^2(M)) \right] \\
&= \mathbb{E}\left[p_j^2(M_t)\overline{p_j}^2(M_t) \right] - 2\mathbb{E}\left[p_j^2(M_t)\overline{p_j}(M_t)\overline{p_j}(M) \right] - 2\mathbb{E}\left[\overline{p_j}^2(M_t)p_j(M_t)p_j(M) \right] \\
&\quad + \mathbb{E}\left[p_j^2(M_t)\overline{p_j}^2(M) \right] + \mathbb{E}\left[\overline{p_j}^2(M_t)p_j^2(M) \right] + 4\mathbb{E}\left[p_j(M_t)\overline{p_j}(M_t)p_j(M)\overline{p_j}(M) \right] \\
&\quad - 2\mathbb{E}\left[p_j^2(M)\overline{p_j}(M)\overline{p_j}(M_t) \right] - 2\mathbb{E}\left[\overline{p_j}^2(M)p_j(M)p_j(M_t) \right] + \mathbb{E}\left[p_j^2(M)\overline{p_j}^2(M) \right] \\
&=: S_1 - 2S_2 - 2S_3 + S_4 + S_5 + 4S_6 - 2S_7 - 2S_8 + S_9.
\end{aligned}
$$

By exchangeability, we have $S_1 = S_9$, $S_3 = \overline{S_2}$, $S_4 = S_5$, $S_7 = S_2$, $S_8 = \overline{S_7} = S_3$.

Now, for $n \geq 2d$, i.e., large enough for the moment formulae of Lemma 3.4 to apply for all $j =$

$d - r + 1, \ldots, d$,

$$
\begin{aligned}
S_1 &= 2j^2, \\
S_8 &= \mathbb{E}\left[\overline{p_j}^2(M)p_j(M)p_j(M_t)\right] = \mathbb{E}\left[\overline{p_j}^2(M)p_j(M)\mathbb{E}\left[p_j(M_t)|M\right]\right] \\
&= \mathbb{E}\left[\overline{p_j}^2(M)p_j(M)\left(p_j(M) + t(\Delta p_j)(M) + O(t^2)\right)\right] \\
&= \mathbb{E}\left[\overline{p_j}^2(M)p_j^2(M)\right] + \mathbb{E}\left[\overline{p_j}^2(M)p_j(M)O(t^2)\right] \\
&\quad + t\,\mathbb{E}\left[\overline{p_j}^2(M)p_j(M)\left(-njp_j(M) - j\sum_{l=1}^{j-1}p_{l,j-l}(M)\right)\right] \\
&= 2j^2 + O(t^2) - tnj2j^2 - tj\sum_{l=1}^{j-1}\mathbb{E}\left[\overline{p_j}^2(M)p_j(M)p_{l,j-l}(M)\right] = 2j^2 - 2tnj^3 + O(t^2).
\end{aligned}
$$

Hence, $S_3 = S_2 = \overline{S_7} = S_8 = 2j^2 - 2tnj^3 + O(t^2)$. On the other hand,

$$
\begin{aligned}
S_4 &= \mathbb{E}\left[p_j^2(M_t)\overline{p_j}^2(M)\right] = \mathbb{E}\left[\overline{p_j}^2(M)\mathbb{E}\left[p_j^2(M_t)|M\right]\right] \\
&= \mathbb{E}\left[\overline{p_j}^2(M)\left(p_j^2(M) + t\,(\Delta p_j^2)(M) + O(t^2)\right)\right] = \mathbb{E}\left[\overline{p_j}^2(M)p_j^2(M)\right] + O(t^2) \\
&\quad + t\,\mathbb{E}\left[\overline{p_j}^2(M)\left(-2njp_j^2(M) - 2j^2p_{2j}(M) - 2jp_j(M)\sum_{l=1}^{j-1}p_{l,j-l}(M)\right)\right] \\
&= 2j^2 - 4ntj^3 - 2tj^2\mathbb{E}\left[\overline{p_j}^2(M)p_{2j}(M)\right] - 2tj\sum_{l=1}^{j-1}\mathbb{E}\left[\overline{p_j}^2(M)p_j(M)p_{l,j-l}(M)\right] + O(t^2) \\
&= 2j^2 - 4tnj^3 + O(t^2).
\end{aligned}
$$

Finally,

$$
\begin{aligned}
S_6 &= \mathbb{E}\left[p_j(M_t)\overline{p_j}(M_t)p_j(M)\overline{p_j}(M)\right] = \mathbb{E}\left[p_j(M)\overline{p_j}(M)\mathbb{E}\left[p_j(M_t)\overline{p_j}(M_t) \mid M\right]\right] \\
&= \mathbb{E}\left[p_j(M)\overline{p_j}(M)\left(p_j(M)\overline{p_j}(M) + t\,(\Delta p_j\overline{p_j})(M) + O(t^2)\right)\right] \\
&= \mathbb{E}\left[p_j^2(M)\overline{p_j}^2(M)\right] + O(t^2) + t\,\mathbb{E}\Bigg[p_j(M)\overline{p_j}(M) \\
&\qquad \left(2j^2n - 2njp_j(M)\overline{p_j}(M) - jp_j(M)\sum_{l=1}^{j-1}\overline{p_{l,j-l}}(M) - j\overline{p_j}(M)\sum_{l=1}^{j-1}p_{l,j-l}(M)\right)\Bigg] \\
&= 2j^2 + 2tj^2n\mathbb{E}\left[p_j(M)\overline{p_j}(M)\right] - 2ntj\mathbb{E}\left[p_j^2(M)\overline{p_j}^2(M)\right] \\
&\quad - tj\sum_{l=1}^{j-1}\mathbb{E}\left[p_j^2(M)\overline{p_j}(M)\overline{p_{l,j-l}}(M)\right] - tj\sum_{l=1}^{j-1}\mathbb{E}\left[p_j(M)p_{l,j-l}(M)\overline{p_j}^2(M)\right] + O(t^2) \\
&= 2j^2 + 2tnj^3 - 4ntj^3 + O(t^2) = 2j^2 - 2ntj^3 + O(t^2).
\end{aligned}
$$

Putting the pieces together,

$$
\begin{aligned}
\mathbb{E}\left[|W_{t,j} - W_j|^4\right] &= 2 \cdot j^2 - 8(2j^2 - 2tnj^3) + 2(2j^2 - 4tnj^3) + 4(2j^2 - 2ntj^3) + O(t^2) \\
&= j^2(4 - 16 + 4 + 8) + tnj^3(16 - 8 - 8) + O(t^2) = O(t^2).
\end{aligned}
$$

as asserted. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\quad\ \Box$

With the conditions of Theorem 2.3 in place, we have

$$d_{\mathcal{W}}(W, \Sigma^{1/2}Z) \leq \|\Lambda^{-1}\|_{\text{op}} \left(\mathbb{E}\|R\|_2 + \frac{1}{\sqrt{2\pi}} \|\Sigma^{-1/2}\|_{\text{op}} \mathbb{E}\left[\|S\|_{\text{HS}} + \|T\|_{\text{HS}} \right] \right). \tag{7}$$

To bound the quantities on the right hand side, we first observe that $\|\Lambda^{-1}\|_{\text{op}} = \frac{1}{n(d-r+1)}$ and $\|\Sigma^{-1/2}\|_{\text{op}} = \frac{1}{\sqrt{d-r+1}}$. Now

$$\mathbb{E}\|R\|_2^2 = \sum_{j=d-r+1}^{d} \mathbb{E}R_j\overline{R_j} = \sum_{j=d-r+1}^{d} j^2 \sum_{l,m=1}^{j-1} \mathbb{E}\left[p_{l,j-l}(M)\overline{p_{m,j-m}}(M) \right].$$

For $n \geq d$, Lemma 3.4 implies

$$\mathbb{E}\left[p_{l,j-l}(M)\overline{p_{m,j-m}}(M) \right] = \delta_{l,m} \mathbb{E}\left[p_{l,j-l}(M)\overline{p_{l,j-l}}(M) \right]$$

and

$$\mathbb{E}\left[p_{l,j-l}(M)\overline{p_{l,j-l}}(M) \right] = \left\{ \begin{array}{ll} 2l(j-l), & l = \frac{j}{2} \\ l(j-l), & \text{otherwise} \end{array} \right\} \leq 2l(j-l).$$

Hence, by using Lemma 4.1

$$\mathbb{E}\|R\|_2^2 \leq 2 \sum_{j=d-r+1}^{d} j^2 \sum_{l=1}^{j-1} l(j-l)$$

$$= 2 \sum_{j=d-r+1}^{d} j^2 \frac{j\,(j^2-1)}{6}$$

$$= \frac{1}{3} \sum_{j=d-r+1}^{d} (j^5 - j^3) \leq \sum_{j=d-r+1}^{d} j^5$$

$$\leq \frac{2^6}{3 \cdot 6} \max\{(r-1)(d-r+1)^5, (r-1)^6\}$$

$$= O(\max(r^6, r(d-r)^5)),$$

hence $\mathbb{E}\|R\|_2 \leq \sqrt{\mathbb{E}\|R\|_2^2} = O(\max(r^3, \sqrt{r}(d-r)^{5/2}))$. On the other hand, again by Lemma 4.1

$$\mathbb{E}\|S\|_{\text{HS}}^2 = \sum_{j,k=d-r+1}^{d} \mathbb{E}|s_{jk}|^2$$

$$= 8 \sum_{d-r+1 \leq k < j \leq d} j^2 k^2 \mathbb{E}\left[p_{j-k}(M)\overline{p_{j-k}}(M)\right] = 8 \sum_{d-r+1 \leq k < j \leq d} j^2 k^2 (j-k)$$

$$= 8 \sum_{d-r+1 \leq k < j \leq d} j^3 k^2 - 8 \sum_{d-r+1 \leq k < j \leq d} j^2 k^3$$

$$= 8 \sum_{d-r+2}^{d} j^3 \frac{(j-1)j(2j-1)}{6} - j^2 \frac{(j-1)^2 j^2}{4} = \frac{2}{3} \sum_{d-r+2}^{d} j^6 - j^4$$

$$\leq \frac{2}{3} \sum_{j=d-r+2}^{d} j^6$$

$$\leq \frac{2}{3} \cdot \frac{2^7}{7} \max\{(r-2)^7, (r-2)(d-r+2)^6\}$$

$$= O(\max(r^7, r(d-r)^6)).$$

So we have obtained that

$$\mathbb{E}\|S\|_{\text{HS}} \leq \sqrt{\mathbb{E}\|S\|_{\text{HS}}^2} = O(\max(r^{7/2}, \sqrt{r}(d-r)^3)).$$

As to $\|T\|_{\text{HS}}$, for $n \geq 2d$ we have, again using Lemma 4.1

$$\mathbb{E}[\|T\|_{\text{HS}}^2] = \sum_{j,k=d-r+1}^{d} \mathbb{E}[|t_{ik}|^2] = \sum_{j,k=d-r+1}^{d} 4j^2 k^2 \mathbb{E}[p_{j+k}(M)\overline{p_{j+k}}(M)]$$

$$= \sum_{j,k=d-r+1}^{d} 4j^2 k^2 (j+k) = 4 \sum_{j,k=d-r+1}^{d} j^3 k^2 + 4 \sum_{j,k=d-r+1}^{d} k^3 j^2$$

$$= 8 \left(\sum_{j=d-r+1}^{d} j^3\right) \left(\sum_{k=d-r+1}^{d} k^2\right)$$

$$\leq 8 \cdot \frac{2^4}{4} \max\{(r-1)^4, (r-1)(d-r+1)^3\} \cdot \frac{2^3}{3} \max\{(r-1)^3, (r-1)(d-r+1)^2\}$$

$$= O(\max(r^7, r^2(d-r)^5)).$$

hence

$$\mathbb{E}\|T\|_{\text{HS}} = O(\max(r^{7/2}, r(d-r)^{5/2})).$$

Plugging these bounds into (7), we obtain

$$d_{\mathcal{W}}(W, \Sigma^{1/2}Z) \leq \frac{1}{n(d-r+1)}\left(O\left(\max(r^3, \sqrt{r}(d-r)^{5/2})\right) \right.$$

$$\left. + \frac{O\left(\max(r^{7/2}, \sqrt{r}(d-r)^3) + \max(r^{7/2}, r(d-r)^{5/2})\right)}{\sqrt{d-r+1}}\right)$$

$$= O\left(\frac{\sqrt{d-r+1}\max(r^3,\sqrt{r}(d-r)^{5/2}) + \max(r^{7/2},\sqrt{r}(d-r)^3) + \max(r^{7/2},r(d-r)^{5/2})}{n(d-r+1)^{3/2}} \right). \quad (8)$$

In the case that $r \leq d - r$,

$$(8) \;=\; O\left(\frac{\sqrt{r(d-r+1)}(d-r)^{5/2}}{n(d-r+1)^{3/2}} \right) = O\left(\frac{\sqrt{r}(d-r)^{3/2}}{n} \right),$$

whereas in the case that $d - r < r$ one obtains

$$(8) = O\left(\frac{r^{7/2}}{n(d-r+1)^{3/2}} \right).$$

This yields the first claim, and the others follow easily. This proves Theorem 1 for the unitary group.

6 Proof of the main result in the orthogonal case

Again we invoke a family $(W, W_t)_{t>0}$ of exchangeable pairs such that $W_t := W_t(d, r, n) := (f_{d-r+1}(M_t), f_{d-r+2}(M_t), \ldots, f_d(M_t))^T$ for all $t > 0$. We will apply Proposition 2.1, so the first step is to verify its conditions. For Condition (i) we will need the following

LEMMA **6.1.** *For all $j = d - r + 1, \ldots, d$*

$$\mathbb{E}[W_{t,j} - W_j|M] = \mathbb{E}[f_j(M_t) - f_j(M)|M] = t \cdot \left(-\frac{(n-1)j}{2} f_j(M) + R_j + O(t) \right),$$

where

$$R_j \;=\; -\frac{j}{2}\sum_{l=1}^{j-1} p_{l,j-l}(M) + \frac{j}{2}\sum_{l=1}^{j-1} p_{2l-j}(M) \quad \text{if } j \text{ is odd,}$$

$$R_j \;=\; -\frac{(n-1)j}{2} - \frac{j}{2}\sum_{l=1}^{j-1} p_{l,j-l}(M) + \frac{j}{2}\sum_{l=1}^{j-1} p_{2l-j}(M) \quad \text{if } j \text{ is even.}$$

Proof. First observe that always $f_j(M_t) - f_j(M) = p_j(M_t) - p_j(M)$, no matter what the parity of j is. By Lemmas 2.5 and 3.2

$$\mathbb{E}[p_j(M_t) - p_j(M)|M] \;=\; t(\Delta p_j)(M) + O(t^2)$$

$$=\; t\left(-\frac{(n-1)j}{2} p_j(M) - \frac{j}{2}\sum_{l=1}^{j-1} p_{l,j-l}(M) + \frac{j}{2}\sum_{l=1}^{j-1} p_{2l-j}(M) \right) + O(t^2),$$

which is

$$t\left(-\frac{(n-1)j}{2} f_j(M) + \left(-\frac{j}{2}\sum_{l=1}^{j-1} p_{l,j-l}(M) + \frac{j}{2}\sum_{l=1}^{j-1} p_{2l-j}(M) \right) + O(t) \right)$$

if j is odd and which is

$$t\left(-\frac{(n-1)j}{2}f_j(M)+\left(-\frac{(n-1)j}{2}-\frac{j}{2}\sum_{l=1}^{j-1}p_{l,j-l}(M)+\frac{j}{2}\sum_{l=1}^{j-1}p_{2l-j}(M)\right)+\mathrm{O}(t)\right)$$

if j is even. This proves the lemma. $\qquad\square$

From Lemma 6.1 (see also Remark 2.7) we conclude

$$\frac{1}{t}\mathbb{E}[W_t-W|M]\xrightarrow{t\to0}-\Lambda W+R\ \text{almost surely and in } \mathrm{L}^1(\mathbb{P}),$$

where $\Lambda=\mathrm{diag}\left(\frac{(n-1)j}{2},j=d-r+1,\dots,d\right)$ and $R=(R_{d-r+1},\dots,R_d)^T$. Thus, Condition (i) of Proposition 2.1 is satisfied. In order to verify Condition (ii) we will first prove the following

LEMMA **6.2.** *For all* $j,k=d-r+1,\dots,d$

$$\mathbb{E}[(p_j(M_t)-p_j(M))(p_k(M_t)-p_k(M))|M]=t\left(jkp_{j-k}(M)-jkp_{j+k}(M)\right)+\mathrm{O}(t^2).$$

Proof. By well-known properties of conditional expectation

$$\mathbb{E}[(p_j(M_t)-p_j(M))(p_k(M_t)-p_k(M))|M]$$
$$=\mathbb{E}[p_{j,k}(M_t)|M]-p_j(M)\mathbb{E}[p_k(M_t)|M]-p_k(M)\mathbb{E}[p_j(M_t)|M]+p_{j,k}(M)\quad(9)$$

By Lemmas 2.5 and Claim (ii) of Lemma 3.2, we see that

$$\mathbb{E}[p_{j,k}(M_t)|M]=p_{j,k}(M)+t\left(\Delta p_{j,k}\right)(M)+\mathrm{O}(t^2)$$
$$=p_{j,k}(M)+t\left(-\frac{(n-1)(j+k)}{2}p_{j,k}(M)-\frac{j}{2}p_k(M)\sum_{l=1}^{j-1}p_{l,j-l}(M)\right.$$
$$-\frac{k}{2}p_j(M)\sum_{l=1}^{k-1}p_{l,k-l}(M)-kjp_{j+k}(M)+\frac{j}{2}p_k(M)\sum_{l=1}^{j-1}p_{j-2l}(M)$$
$$\left.+\frac{k}{2}p_j(M)\sum_{l=1}^{k-1}p_{k-2l}(M)+jkp_{j-k}(M)\right)+\mathrm{O}(t^2).$$

Also, by Lemma 2.5 and Claim (i) of Lemma 3.2

$$-p_j(M)\mathbb{E}[p_k(M_t)|M]=-p_j(M)\left(p_k(M)+t\left(\Delta p_k\right)(M)+\mathrm{O}(t^2)\right)$$
$$=-p_{j,k}(M)-tp_j(M)\left(-\frac{(n-1)k}{2}p_k(M)-\frac{k}{2}\sum_{l=1}^{k-1}p_{l,k-l}(M)+\frac{k}{2}\sum_{l=1}^{k-1}p_{2l-k}(M)\right)+\mathrm{O}(t^2)$$
$$=-p_{j,k}(M)+t\frac{(n-1)k}{2}p_{j,k}(M)+t\frac{k}{2}p_j(M)\sum_{l=1}^{k-1}p_{l,k-l}(M)$$
$$-t\frac{k}{2}p_j(M)\sum_{l=1}^{k-1}p_{2l-k}(M)+\mathrm{O}(t^2),$$

and for reasons of symmetry

$$-p_k(M)\mathbb{E}[p_j(M_t)|M] = -p_{j,k}(M) + t\frac{(n-1)j}{2}p_{j,k}(M) + t\frac{j}{2}p_k(M)\sum_{l=1}^{j-1}p_{l,j-l}(M)$$

$$- t\frac{j}{2}p_k(M)\sum_{l=1}^{j-1}p_{2l-j}(M) + \mathrm{O}(t^2).$$

Plugging these results into (9) above and noting that for $j \in \mathbb{N}$

$$p_{-j}(M) = \mathrm{Tr}(M^{-j}) = \mathrm{Tr}\left((M^T)^j\right) = \mathrm{Tr}\left((M^j)^T\right) = \mathrm{Tr}(M^j) = p_j(M),$$

we see that many terms cancel, and finally obtain

$$\mathbb{E}[(p_j(M_t) - p_j(M))(p_k(M_t) - p_k(M))|M] = t\left(jkp_{j-k}(M) - jkp_{j+k}(M)\right) + \mathrm{O}(t^2).$$

$$\square$$

Observing as above that, regardless of the parity of j, we have $f_j(M_t) - f_j(M) = p_j(M_t) - p_j(M)$, we can now easily compute

$$\frac{1}{t}\mathbb{E}[(W_{t,j} - W_j)(W_{t,k} - W_k)|M] = \frac{1}{t}\mathbb{E}[(f_j(M_t) - f_j(M))(f_k(M_t) - f_k(M))|M]$$

$$= \frac{1}{t}\mathbb{E}[(p_j(M_t) - p_j(M))(p_k(M_t) - p_k(M))|M] = jkp_{j-k}(M) - jkp_{j+k}(M) + \mathrm{O}(t^2)$$

$$\overset{t\to0}{\to} jkp_{j-k}(M) - jkp_{j+k}(M) \text{ a.s. and in } L^1(\mathbb{P}),$$

for all $j, k = 1, \ldots, d$. Noting that for $j = k$ the last expression is $j^2n - j^2p_{2j}(M)$ and that $2\Lambda\Sigma = \mathrm{diag}((n-1)j^2, j = 1, \ldots, d)$ we see that Condition (ii) of Proposition 2.1 is satisfied with the matrix $S = (S_{j,k})_{j,k=1,\ldots,d}$ given by

$$S_{j,k} = \begin{cases} j^2(1 - p_{2j}(M)), & j = k \\ jkp_{j-k}(M) - jkp_{j+k}(M), & j \neq k. \end{cases}$$

In order to show that Condition (iii) of Proposition 2.1 holds, we will need the following facts:

LEMMA **6.3.** *For all* $j = 1, \ldots, d$, $n \geq 4d + 1$,

$$\mathbb{E}\left[(W_{t,j} - W_j)^2\right] = tj^2(n-1) + \mathrm{O}(t^2). \tag{i}$$

$$\mathbb{E}\left[(W_{t,j} - W_j)^4\right] = \mathrm{O}(t^2). \tag{ii}$$

Proof. As for (i), by Lemma 6.2,

$$\mathbb{E}[(W_{t,j} - W_j)^2] = \mathbb{E}\left[\left(p_j(M_t) - p_j(M)\right)^2\right]$$

$$= \mathbb{E}\left[\mathbb{E}\left[\left(p_j(M_t) - p_j(M)\right)^2 \mid M\right]\right] = \mathbb{E}\left[tj^2\left(n - p_{2j}(M)\right) + \mathrm{O}(t^2)\right]$$

$$= tj^2(n-1) + \mathrm{O}(t^2),$$

since by Lemma 3.5 $\mathbb{E}[p_{2j}(M)] = 1$. For Claim (ii) we compute

$$\mathbb{E}\left[(W_{t,j} - W_j)^4\right] = \mathbb{E}\left[\left(p_j(M_t) - p_j(M)\right)^4\right]$$
$$= \mathbb{E}[p_j(M_t)^4] + \mathbb{E}[p_j(M)^4] - 4\mathbb{E}[p_j(M_t)^3 p_j(M)] - 4\mathbb{E}[p_j(M)^3 p_j(M_t)] + 6\mathbb{E}[p_j(M_t)^2 p_j(M)^2]$$
$$= 2\mathbb{E}[p_j(M)^4] - 8\mathbb{E}[p_j(M)^3 p_j(M_t)] + 6\mathbb{E}[p_{j,j}(M) p_{j,j}(M_t)],$$

where the last equality follows from exchangeability. By Lemma 2.5 and Claim (ii) of Lemma 3.2 for the case $k = j$

$$\mathbb{E}[p_{j,j}(M) p_{j,j}(M_t)] = \mathbb{E}\left[p_{j,j}(M) \mathbb{E}[p_{j,j}(M_t)|M]\right]$$
$$= \mathbb{E}\left[p_{j,j}(M)\left(p_{j,j}(M) + t\left(\Delta p_{j,j}\right)(M) + \mathrm{O}(t^2)\right)\right]$$
$$= \mathbb{E}[p_j(M)^4] + t\mathbb{E}[p_{j,j}(M) \cdot \Delta p_{j,j}(M)] + \mathrm{O}(t^2)$$
$$= \mathbb{E}[p_j(M)^4] + t\mathbb{E}\left[p_{j,j}(M)\left(-(n-1)jp_{j,j}(M) - jp_j(M)\sum_{l=1}^{j-1} p_{l,j-l}(M) - j^2 p_{2j}(M)\right.\right.$$
$$\left.\left. + jp_j(M)\sum_{l=1}^{j-1} p_{j-2l}(M) + j^2 n\right)\right] + \mathrm{O}(t^2).$$

Again by Lemma 2.5 and Claim (i) of Lemma 3.2,

$$\mathbb{E}[p_j(M)^3 p_j(M_t)] = \mathbb{E}\left[p_j(M)^3 \mathbb{E}[p_j(M_t)|M]\right]$$
$$= \mathbb{E}\left[p_j(M)^3\left((p_j(M) + t\left(\Delta p_j\right)(M) + \mathrm{O}(t^2))\right)\right]$$
$$= \mathbb{E}[p_j(M)^4] + t\mathbb{E}\left[p_j(M)^3\left(-\frac{(n-1)j}{2}p_j(M) - \frac{j}{2}\sum_{l=1}^{j-1} p_{l,j-l}(M) + \frac{j}{2}\sum_{l=1}^{j-1} p_{2l-j}(M)\right)\right] + \mathrm{O}(t^2).$$

Therefore,

$$\mathbb{E}\left[(W_{t,j} - W_j)^4\right]$$
$$= 2\mathbb{E}\left[p_j(M)^4\right] - 8\left(\mathbb{E}[p_j(M)^4] - t\frac{(n-1)j}{2}\mathbb{E}[p_j(M)^4]\right.$$
$$\left. - t\frac{j}{2}\sum_{l=1}^{j-1}\mathbb{E}\left[p_j(M)^3 p_{l,j-l}(M)\right] + t\frac{j}{2}\sum_{l=1}^{j-1}\mathbb{E}\left[p_j(M)^3 p_{2l-j}(M)\right]\right)$$
$$+ 6\left(\mathbb{E}\left[p_j(M)^4\right] - t(n-1)j\mathbb{E}\left[p_j(M)^4\right] - tj\sum_{l=1}^{j-1}\mathbb{E}\left[p_j(M)^3 p_{l,j-l}(M)\right]\right.$$
$$\left. - tj^2\mathbb{E}\left[p_{2j}(M)p_j(M)^2\right] + tj\sum_{l=1}^{j-1}\mathbb{E}\left[p_j(M)^3 p_{2l-j}(M)\right] + tj^2 n\mathbb{E}\left[p_j(M)^2\right]\right) + \mathrm{O}(t^2)$$
$$= -2t(n-1)j\mathbb{E}\left[p_j(M)^4\right] - 2tj\sum_{l=1}^{j-1}\mathbb{E}\left[p_j(M)^3 p_{l,j-l}(M)\right] + 2tj\sum_{l=1}^{j-1}\mathbb{E}\left[p_j(M)^3 p_{2l-j}(M)\right]$$
$$- 6tj^2\mathbb{E}\left[p_{2j}(M)p_j(M)^2\right] + 6tj^2 n\mathbb{E}\left[p_j(M)^2\right] + \mathrm{O}(t^2).$$

Case 1: j is odd. Then by Lemma 3.5

$$
\begin{aligned}
\mathbb{E}\left[(W_{t,j} - W_j)^4\right] &= -2t(n-1)j \cdot 3j^2 - 0 + 0 - 6tj^2 \cdot j + 6tj^2 n \cdot j + O(t^2) \\
&= O(t^2),
\end{aligned}
$$

as claimed.

Case 2: j is even. Then, again by Lemma 3.5

$$
\mathbb{E}\left[(W_{t,j} - W_j)^4\right] = -2t(n-1)j(1 + 6j + 3j^2) - 2tj(1+3j)\sum_{l=1}^{j-1}\mathbb{E}\left[p_{l,j-l}(M)\right]
$$

$$
+ 2tj(1+3j)\sum_{l=1}^{j-1}\mathbb{E}\left[p_{2l-j}(M)\right] - 6tj^2(1+j) + 6tj^2 n(1+j) + O(t^2).
$$

Consider the term $\sum_{l=1}^{j-1}\mathbb{E}\left[p_{l,j-l}(M)\right]$. Since $\mathbb{E}\left[p_{l,j-l}(M)\right] = 0$ whenever l is odd and $l \neq \frac{j}{2}$, we can write

$$
\begin{aligned}
\sum_{l=1}^{j-1}\mathbb{E}\left[p_{l,j-l}(M)\right] &= \mathbb{E}\left[p_{j/2,j/2}(M)\right] + \sum_{\substack{k=1 \\ k\neq j/4}}^{j/2-1}\mathbb{E}\left[p_{2k,j-2k}(M)\right] \\
&= \mathbb{E}\left[p_{j/2,j/2}(M)\right] + \sum_{\substack{k=1 \\ k\neq j/4}}^{j/2-1} 1,
\end{aligned}
$$

where the last equality follows again by Lemma 3.5. If $j/2$ is odd, then, clearly, $k \neq j/4$ is not really a restriction, and by Lemma 3.5

$$
\sum_{l=1}^{j-1}\mathbb{E}\left[p_{l,j-l}(M)\right] = \frac{j}{2} + \left(\frac{j}{2} - 1\right) = j - 1.
$$

If $j/2$ is even, then by Lemma 3.5

$$
\sum_{l=1}^{j-1}\mathbb{E}\left[p_{l,j-l}(M)\right] = \left(1 + \frac{j}{2}\right) + \left(\frac{j}{2} - 2\right) = j - 1.
$$

Hence, in either case we have $\sum_{l=1}^{j-1}\mathbb{E}\left[p_{l,j-l}(M)\right] = j - 1$.

Furthermore

$$
\sum_{l=1}^{j-1}\mathbb{E}\left[p_{2l-j}(M)\right] = 2\sum_{l=1}^{j/2-1}\mathbb{E}\left[p_{2l-j}(M)\right] + \mathbb{E}\left[\mathrm{Tr}(I_n)\right] = 2\left(\frac{j}{2} - 1\right) + n = j - 2 + n.
$$

Therefore, if j is even

$$
\begin{aligned}
\mathbb{E}\left[(W_{t,j} - W_j)^4\right] &= -2tj(n-1)(1 + 6j + 3j^2) - 2tj(1 + 3j)(j - 1) \\
&\quad + 2tj(1 + 3j)(j - 2 + n) - 6tj^2(1 + j) + 6tj^2n(1 + j) + \mathrm{O}(t^2) \\
&= tj\left(-12(n - 1) - 2 + 6 + 2 + 6(n - 2) - 6 + 6n\right) \\
&\quad + tj^3\left(-6(n - 1) - 6 + 6 - 6 + 6n\right) + \mathrm{O}(t^2) \\
&= \mathrm{O}(t^2),
\end{aligned}
$$

as asserted. $\qquad\square$

Now we are in a position to check Condition $(iii)'$ of Proposition 2.1. By Hölder's inequality

$$
\begin{aligned}
\mathbb{E}\left[\|W_t - W\|_2^3\right] &= \mathbb{E}\left[\left(\sum_{j=d-r+1}^{d}(W_{t,j} - W_j)^2\right)^{3/2}\right] \\
&= \mathbb{E}\left[\left(\sum_{j,k,l=d-r+1}^{d}(W_{t,j} - W_j)^2(W_{t,k} - W_k)^2(W_{t,l} - W_l)^2\right)^{1/2}\right] \\
&\leq \sum_{j,k,l=d-r+1}^{d}\mathbb{E}\left[\left|(W_{t,j} - W_j)(W_{t,k} - W_k)(W_{t,l} - W_l)\right|\right] \\
&\leq \sum_{j,k,l=d-r+1}^{d}\left(\mathbb{E}\left[|W_{t,j} - W_j|^3\right]\mathbb{E}\left[|W_{t,k} - W_k|^3\right]\mathbb{E}\left[|W_{t,l} - W_l|^3\right]\right)^{1/3}.
\end{aligned}
$$

Thus we have

$$
\frac{1}{t}\mathbb{E}\left[\|W_t - W\|_2^3\right] \leq \sum_{j,k,l=d-r+1}^{d}\left(\frac{1}{t}\mathbb{E}\left[|W_{t,j} - W_j|^3\right]\frac{1}{t}\mathbb{E}\left[|W_{t,k} - W_k|^3\right]\frac{1}{t}\mathbb{E}\left[|W_{t,l} - W_l|^3\right]\right)^{1/3},
$$

and it suffices to show that for all $j = d - r + 1, \dots, d$

$$
\lim_{t \to 0}\frac{1}{t}\mathbb{E}\left[|W_{t,j} - W_j|^3\right] = 0.
$$

But this follows from Lemma 6.3, since by the Cauchy-Schwarz inequality

$$
\begin{aligned}
\mathbb{E}\left[|W_{t,j} - W_j|^3\right] &\leq \sqrt{\mathbb{E}\left[|W_{t,j} - W_j|^2\right]\mathbb{E}\left[|W_{t,j} - W_j|^4\right]} = \sqrt{(tj^2(n-1) + \mathrm{O}(t^2))\cdot\mathrm{O}(t^2)} \\
&= \sqrt{\mathrm{O}(t^3)} = \mathrm{O}(t^{3/2})
\end{aligned}
$$

and hence $\frac{1}{t}\mathbb{E}\left[|W_{t,j} - W_j|^3\right] \overset{t \to 0}{\to} 0$.

By Proposition 2.1 we can now conclude that

$$d_{\mathscr{W}}(W, Z_\Sigma) \leq \|\Lambda^{-1}\|_{\mathrm{op}} \left(\mathbb{E}[\|R\|_2] + \frac{1}{\sqrt{2\pi}} \|\Sigma^{-1/2}\|_{\mathrm{op}} \mathbb{E}[\|S\|_{\mathrm{HS}}] \right).$$

Clearly, $\|\Lambda^{-1}\|_{\mathrm{op}} = \frac{2}{(n-1)(d-r+1)}$ and $\|\Sigma^{-1/2}\|_{\mathrm{op}} = \frac{1}{\sqrt{d-r+1}}$. In order to bound $\mathbb{E}[\|R\|_2]$ we will first prove the following lemma.

LEMMA **6.4.** *If* $n \geq 4d + 1$, *for all* $j = d - r + 1, \ldots, d$ *we have that* $\mathbb{E}[R_j^2] = O(j^5)$.

Proof. First suppose that j is odd. Then

$$\mathbb{E}[R_j^2] = \mathbb{E}\left[\left(-\frac{j}{2} \sum_{l=1}^{j-1} p_{l,j-l}(M) + \frac{j}{2} \sum_{l=1}^{j-1} p_{2l-j}(M) \right)^2 \right]$$

$$= \frac{j^2}{4} \left(\sum_{l,k=1}^{j-1} \mathbb{E}[p_{l,j-l}(M) p_{k,j-k}(M)] - 2 \sum_{l,k=1}^{j-1} \mathbb{E}[p_{l,j-l}(M) p_{2k-j}(M)] + \sum_{l,k=1}^{j-1} \mathbb{E}[p_{2l-j}(M) p_{2k-j}(M)] \right)$$

$$=: \frac{j^2}{4} \left(T_1 - 2T_2 + T_3 \right).$$

By Lemma 3.5 for j odd

$$\mathbb{E}[p_{l,j-l}(M) p_{k,j-k}(M)] = \begin{cases} l(j-l+1), & \text{if } k = l \text{ is odd}, \\ (l+1)(j-l), & \text{if } k = l \text{ is even}, \\ l(j-l+1), & \text{if } l \text{ is odd and } k = j-l, \\ 0, & \text{if } l \text{ is odd and } k \notin \{l, j-l\}, \\ (l+1)(j-l), & \text{if } l \text{ is even and } k = j-l, \\ 0, & \text{if } l \text{ is even and } k \notin \{l, j-l\}, \end{cases}$$

$$\mathbb{E}[p_{l,j-l}(M) p_{2k-j}(M)] = \begin{cases} l, & \text{if } l = 2k - j, \\ j - l, & \text{if } l = 2j - 2k, \\ l, & \text{if } l = j - 2k, \\ j - l, & \text{if } l = 2k, \\ 0, & \text{otherwise}, \end{cases}$$

$$\mathbb{E}[p_{2l-j}(M) p_{2k-j}(M)] = \begin{cases} |2l - j|, & \text{if } l = k \text{ or } l = j - k, \\ 0, & \text{otherwise}. \end{cases}$$

Therefore,

$$T_1 = 2\sum_{\substack{l=1 \\ l \text{ odd}}}^{j-1} l(j-l+1) + 2\sum_{\substack{l=1 \\ l \text{ even}}}^{j-1} (l+1)(j-l)$$

$$\leq 2j\sum_{l=1}^{j} l = j^2(j-1) = O(j^3),$$

$$T_2 = 2\sum_{\substack{l=1 \\ l \text{ odd}}}^{j-1} l + 2\sum_{\substack{l=1 \\ l \text{ even}}}^{j-1} (j-l) \leq 2\sum_{l=1}^{j} l = j(j-1) = O(j^2)$$

and

$$T_3 = 2\sum_{l=1}^{j-1} |2l-j| \leq 2\sum_{l=1}^{j} j = 2j^2 = O(j^2).$$

Hence for j odd

$$\mathbb{E}[R_j^2] = \frac{j^2}{4}\left(O(j^3) - 2O(j^2) + O(j^2)\right) = O(j^5).$$

The case that j is even can be treated similarly as can be seen by observing that in this case

$$R_j = \frac{j}{2}\left(1 - \sum_{l=1}^{j-1} p_{l,j-l}(M) + \sum_{\substack{l=1 \\ l \neq j/2}}^{j-1} p_{2l-j}\right).$$

But as in Case 2 of the proof of Lemma 6.3 (ii), one has to distinguish between the cases that $j \equiv 0 \bmod 4$ or else that $j \equiv 2 \bmod 4$, and the explicit formulae for $\mathbb{E}[R_j^2]$ are more complicated. □

By Lemma 6.4 there is a constant $C > 0$ neither depending on j nor on n such that $\mathbb{E}[R_j^2] \leq Cj^5$. Thus by Jensen's inequality and Lemma 4.1 we obtain

$$\mathbb{E}[\|R\|_2] \leq \sqrt{\sum_{j=d-r+1}^{d} \mathbb{E}[R_j^2]} \leq \sqrt{C}\sqrt{\sum_{j=d-r+1}^{d} j^5}$$

$$\leq \sqrt{C}\sqrt{\frac{2^6}{6}\max\{(r-1)^6, (r-1)(d-r+1)^5\}}$$

$$= O\left(\max\{r^3, (d-r)^{5/2}\sqrt{r}\}\right). \tag{10}$$

Next, we turn to bounding $\mathbb{E}[\|S\|_{\text{HS}}]$. By Lemma 3.5 we have

2399

$$\mathbb{E}[\|S\|_{\text{HS}}^2] = \sum_{j,k=d-r+1}^{d} \mathbb{E}[S_{j,k}^2] = \sum_{j=d-r+1}^{d} j^4 \mathbb{E}\left[1 - 2p_{2j}(M) + p_{2j}(M)^2\right]$$

$$+ \sum_{\substack{j,k=d-r+1 \\ j \neq k}}^{d} j^2 k^2 \mathbb{E}\left[p_{j-k}(M)^2 - 2p_{j-k}(M)p_{j+k}(M) + p_{j+k}(M)^2\right]$$

$$= 2 \sum_{j=d-r+1}^{d} j^5 + 2 \sum_{d-r+1 \leq k < j \leq d} j^2 k^2 \left(\mathbb{E}\left[p_{j-k}(M)^2\right] - 2\mathbb{E}\left[p_{j-k}(M)p_{j+k}(M)\right] + \mathbb{E}\left[p_{j+k}(M)^2\right]\right).$$

Again, by Lemma 3.5

$$\mathbb{E}\left[p_{j-k}(M)^2\right] = \begin{cases} 1 + j - k, & \text{if } j + k \text{ even,} \\ j - k, & \text{if } j + k \text{ odd,} \end{cases}$$

$$\mathbb{E}\left[p_{j-k}(M)p_{j+k}(M)\right] = \begin{cases} 1, & \text{if } j + k \text{ even,} \\ 0, & \text{if } j + k \text{ odd,} \end{cases}$$

$$\mathbb{E}\left[p_{j+k}(M)^2\right] = \begin{cases} 1 + j + k, & \text{if } j + k \text{ even} \\ j + k, & \text{if } j + k \text{ odd.} \end{cases}$$

Hence,

$$\mathbb{E}[\|S\|_{\text{HS}}^2] = 2 \sum_{j=d-r+1}^{d} j^5 + 2 \sum_{\substack{d-r+1 \leq k < j \leq d \\ k+j \text{ even}}} j^2 k^2 \left(1 + j - k - 2 + 1 + j + k\right)$$

$$+ 2 \sum_{\substack{d-r+1 \leq k < j \leq d \\ k+j \text{ odd}}} j^2 k^2 \left(j - k - 2 \cdot 0 + j + k\right) = 2 \sum_{j=d-r+1}^{d} j^5 + 4 \sum_{d-r+1 \leq k < j \leq d} k^2 j^3.$$

From (10) we know that

$$\sum_{j=d-r+1}^{d} j^5 = O\left(\max\{r^6, r(d-r)^5\}\right)$$

and furthermore we can bound

$$\sum_{d-r+1\leq k<j\leq d} k^2 j^3 \leq \sum_{j=d-r+2}^{d} j^2(j-(d-r+1))j^3$$

$$= \sum_{j=2}^{r}(j+d-r)^2(j+d-r-(d-r+1))(j+d-r)^3 = \sum_{j=2}^{r}(j+d-r)^5(j-1)$$

$$\leq r\sum_{j=1}^{r}(j+d-r)^5 \leq 32r\max\{r^6,(d-r)^5 r\} = 32\max\{r^7,(d-r)^5 r^2\}.$$

Thus, we obtain

$$\mathbb{E}[\|S\|_{HS}^2] = O\left(\max\{r^6,(d-r)^5 r\}\right) + 128\max\{r^7,(d-r)^5 r^2\}$$
$$= \qquad\qquad\qquad\qquad\qquad\qquad\qquad O\left(\max\{r^7,(d-r)^5 r^2\}\right),$$

and hence, by Jensen's inequality

$$\mathbb{E}[\|S\|_{HS}] \leq \sqrt{\mathbb{E}[\|S\|_{HS}^2]} = O\left(\max\{r^{7/2},(d-r)^{5/2} r\}\right),$$

Collecting terms, we see that

$$d_{\mathcal{W}}(W,Z_\Sigma) \leq \|\Lambda^{-1}\|_{op}\mathbb{E}[\|R\|_2] + \frac{1}{\sqrt{2\pi}}\|\Lambda^{-1}\|_{op}\|\Sigma^{-1/2}\|_{op}\mathbb{E}[\|S\|_{HS}]$$

$$= \frac{2}{(n-1)(d-r+1)}O\left(\max\{r^3,(d-r)^{5/2}\sqrt{r}\}\right)$$

$$+ \frac{1}{\sqrt{2\pi}}\frac{2}{(n-1)(d-r+1)^{3/2}}O\left(\max\{r^{7/2},(d-r)^{5/2} r\}\right)$$

$$= O\left(\frac{\max\left\{\frac{r^3}{d-r+1},\frac{r^{7/2}}{(d-r+1)^{3/2}},\frac{(d-r)^{5/2}\sqrt{r}}{d-r+1},\frac{(d-r)^{5/2} r}{(d-r+1)^{3/2}}\right\}}{n}\right).$$

Proceeding as in the unitary case and treating the cases that $r \leq d-r$ and $d-r < r$ separately, one obtains that the last expression is of order

$$\frac{1}{n}\max\left\{\frac{r^{7/2}}{(d-r+1)^{3/2}},(d-r)^{3/2}\sqrt{r}\right\}.$$

This concludes the proof of the orthogonal case.

7 Proof of the main result in the symplectic case

Again we use exchangeable pairs $(W,W_t)_{t>0}$ such that

$$W_t = (f_{d-r+1}(M_t),f_{d-r+2}(M_t),\ldots,f_d(M_t))^T,$$

and apply Proposition 2.1. The verification of Condition (i) involves the following analog of Lemma 6.1 above.

LEMMA **7.1.** *For all $j = d - r + 1, \ldots, d$,*

$$\mathbb{E}[W_{t,j} - W_j | M] = \mathbb{E}[f_j(M_t) - f_j(M) | M] = t \cdot \left(-\frac{(2n+1)j}{2} f_j(M) + R_j + O(t) \right),$$

where

$$R_j = -\frac{j}{2} \sum_{l=1}^{j-1} p_{l,j-l}(M) - \frac{j}{2} \sum_{l=1}^{j-1} p_{2l-j}(M) \text{ if } j \text{ is odd,}$$

$$R_j = \frac{(2n+1)j}{2} - \frac{j}{2} \sum_{l=1}^{j-1} p_{l,j-l}(M) - \frac{j}{2} \sum_{l=1}^{j-1} p_{2l-j}(M) \text{ if } j \text{ is even.}$$

Proof. First observe that always $f_j(M_t) - f_j(M) = p_j(M_t) - p_j(M)$, no matter what the parity of j is. By Lemmas 2.5 and 3.3,

$$\mathbb{E}[p_j(M_t) - p_j(M) | M] = t(\Delta p_j)(M) + O(t^2)$$

$$= t \left(-\frac{(2n+1)j}{2} p_j(M) - \frac{j}{2} \sum_{l=1}^{j-1} p_{l,j-l}(M) - \frac{j}{2} \sum_{l=1}^{j-1} p_{2l-j}(M) \right) + O(t^2),$$

which is

$$t \left(-\frac{(2n+1)j}{2} f_j(M) + \left(-\frac{j}{2} \sum_{l=1}^{j-1} p_{l,j-l}(M) - \frac{j}{2} \sum_{l=1}^{j-1} p_{2l-j}(M) \right) + O(t) \right)$$

if j is odd and which is

$$t \left(-\frac{(2n+1)j}{2} f_j(M) + \left(\frac{(2n+1)j}{2} - \frac{j}{2} \sum_{l=1}^{j-1} p_{l,j-l}(M) - \frac{j}{2} \sum_{l=1}^{j-1} p_{2l-j}(M) \right) + O(t) \right)$$

if j is even. This proves the lemma. $\qquad\square$

From Lemma 7.1 we conclude that

$$\frac{1}{t} \mathbb{E}[W_t - W | M] \xrightarrow{t \to 0} -\Lambda W + R \text{ a.s. and in } L^1(\mathbb{P}),$$

where $\Lambda = \text{diag}\left(\frac{(2n+1)j}{2}, j = d - r + 1, \ldots, d \right)$ and $R = (R_{d-r+1}, \ldots, R_d)^T$. Thus, Condition (i) of Proposition 2.1 is satisfied. The validity of Condition (ii) follows from the next lemma, whose proof only differs from that of Lemma 6.2 in that it makes use of Lemma 3.3 in the place of Lemma 3.2.

LEMMA **7.2.** *For all $j, k = d - r + 1, \ldots, d$,*

$$\mathbb{E}[(p_j(M_t) - p_j(M))(p_k(M_t) - p_k(M)) | M] = t \left(jk p_{j-k}(M) - jk p_{j+k}(M) \right) + O(t^2).$$

Observing that for all $j = d - r + 1, \ldots, d$ we have $f_j(M_t) - f_j(M) = p_j(M_t) - p_j(M)$ and using Lemma 7.2 we can now easily compute

$$\frac{1}{t}\mathbb{E}[(W_{t,j} - W_j)(W_{t,k} - W_k)|M] = \frac{1}{t}\mathbb{E}[(f_j(M_t) - f_j(M))(f_k(M_t) - f_k(M))|M]$$

$$= \frac{1}{t}\mathbb{E}[(p_j(M_t) - p_j(M))(p_k(M_t) - p_k(M))|M] = jkp_{j-k}(M) - jkp_{j+k}(M) + \mathrm{O}(t^2)$$

$$\xrightarrow{t \to 0} jkp_{j-k}(M) - jkp_{j+k}(M) \text{ a.s. and in } \mathrm{L}^1(\mathbb{P}),$$

for all $d - r + 1 \leq j, k \leq d$. Noting that for $j = k$ the last expression is $2j^2 n - j^2 p_{2j}(M)$ and that $2\Lambda\Sigma = \mathrm{diag}((2n+1)j^2, j = d-r+1, \ldots, d)$, we see that Condition (ii) of Proposition 2.1 is satisfied with the matrix $S = (S_{j,k})_{j,k=d-r+1,\ldots,d}$ given by

$$S_{j,k} = \begin{cases} -j^2(1 + p_{2j}(M)), & j = k \\ jkp_{j-k}(M) - jkp_{j+k}(M), & j \neq k. \end{cases}$$

The validity of Condition $(iii)'$ of Proposition 2.1 is based on the following lemma, which can be proven in the same way as its analog in Section 6.

LEMMA **7.3.** *If $n \geq 2d$, for all $j = 1, \ldots, d$ there holds*

$$\mathbb{E}\left[(W_{t,j} - W_j)^2\right] = tj^2(2n+1) + \mathrm{O}(t^2). \tag{i}$$

$$\mathbb{E}\left[(W_{t,j} - W_j)^4\right] = \mathrm{O}(t^2). \tag{ii}$$

So we obtain from Proposition 2.1 that

$$d_{\mathscr{W}}(W, Z_\Sigma) \leq \|\Lambda^{-1}\|_{\mathrm{op}}\left(\mathbb{E}[\|R\|_2] + \frac{1}{\sqrt{2\pi}}\|\Sigma^{-1/2}\|_{\mathrm{op}}\mathbb{E}[\|S\|_{\mathrm{HS}}]\right).$$

Again, it is easy to see that $\|\Lambda^{-1}\|_{\mathrm{op}} = \frac{2}{(2n+1)(d-r+1)}$ and $\|\Sigma^{-1/2}\|_{\mathrm{op}} = \frac{1}{\sqrt{d-r+1}}$. Since we can show in a similar way as in Section 6 that $\mathbb{E}[\|R\|_2] = O\left(\max\{r^3, (d-r)^{5/2}\sqrt{r}\}\right)$ and $\mathbb{E}[\|S\|_{\mathrm{HS}}] = O\left(\max\{r^{7/2}, (d-r)^{5/2}r\}\right)$, we may conclude the proof of the symplectic case as before.

Acknowledgements

References

[CM08]   S. Chatterjee and E. Meckes. Multivariate normal approximation using exchangeable pairs. *ALEA Lat. Am. J. Probab. Math. Stat.*, 4:257–283, 2008. MR2453473

[CS05]   L. H. Y. Chen and Q.-M. Shao. Stein's method for normal approximation. In *An introduction to Stein's method*, volume 4 of *Lect. Notes Ser. Inst. Math. Sci. Natl. Univ. Singap.*, pages 1–59. Singapore Univ. Press, Singapore, 2005. MR2235448

[DS94]   P. Diaconis and M. Shahshahani. On the eigenvalues of random matrices. *J. Appl. Probab.*, 31A:49–62, 1994. MR1274717

[Ful10] J. Fulman. Stein's method, heat kernel, and traces of powers of elements of compact Lie groups, 2010. available on arXiv.org:1005.1306.

[Hel00] S. Helgason. *Groups and geometric analysis*, volume 83 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 2000. Integral geometry, invariant differential operators, and spherical functions, Corrected reprint of the 1984 original. MR1790156

[HR03]   C. P. Hughes and Z. Rudnick. Mock-Gaussian behaviour for linear statistics of classical compact groups. *J. Phys. A*, 36(12):2919–2932, 2003. Random matrix theory. MR1986399

[IW89]   N. Ikeda and S. Watanabe. *Stochastic differential equations and diffusion processes*, volume 24 of *North-Holland Mathematical Library*. North-Holland Publishing Co., Amsterdam, second edition, 1989. MR1011252

[Joh97] K. Johansson. On random matrices from the compact classical groups. *Ann. of Math. (2)*, 145(3):519–545, 1997. MR1454702

[Lév08] T. Lévy. Schur-Weyl duality and the heat kernel measure on the unitary group. *Adv. Math.*, 218(2):537–575, 2008. MR2407946

[Mec09] E. Meckes. On Stein's method for multivariate normal approximation. In *High dimensional probability V: The Luminy volume*, pages 153–158. IMS, 2009. MR2797946

[PV04]   L. Pastur and V. Vasilchuk. On the moments of traces of matrices of classical groups. *Comm. Math. Phys.*, 252(1-3):149–166, 2004. MR2104877

[Rai97] E. M. Rains. Combinatorial properties of Brownian motion on the compact classical groups. *J. Theoret. Probab.*, 10(3):659–679, 1997. MR1468398

[RR97]   Y. Rinott and V. Rotar. On coupling constructions and rates in the CLT for dependent summands with applications to the antivoter model and weighted *U*-statistics. *Ann. Appl. Probab.*, 7(4):1080–1105, 1997. MR1484798

[RR09]   G. Reinert and A. Röllin. Multivariate normal approximation with Stein's method of exchangeable pairs under a general linearity condition. *Ann. Probab.*, 37(6):2150–2173, 2009. MR2573554

[Ste70] E. M. Stein. *Topics in harmonic analysis related to the Littlewood-Paley theory*. Annals of Mathematics Studies, No. 63. Princeton University Press, Princeton, N.J., 1970. MR0252961

[Ste72] C. Stein, A bound for the error in the normal approximation to the distribution of a sum of dependent random variables. Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability, Vol. 2, 583–602. MR0402873

[Ste86] C. Stein. *Approximate computation of expectations*. Institute of Mathematical Statistics Lecture Notes—Monograph Series, 7. Institute of Mathematical Statistics, Hayward, CA, 1986. MR0882007

[Ste95] C. Stein, The accuracy of the normal approximation to the distribution of the traces of powers of random orthogonal matrices. Department of Statistics, Stanford University, Technical Report No. 470, 1995

[Sto05] M. Stolz. On the Diaconis-Shahshahani method in random matrix theory. *J. Algebraic Combin.*, 22(4):471–491, 2005. MR2191648