Bernoulli **6**(1), 2000, 113–167

Random matrix approximation of spectra of integral operators

VLADIMIR KOLTCHINSKII[1] and EVARIST GINÉ[2]

[1]*Department of Mathematics and Statistics, University of New Mexico, Albuquerque NM 87131-1141, USA. E-mail: vlad@math.unm.edu*
[2]*Department of Mathematics and Department of Statistics, University of Connecticut, Storrs CT 06269-3009, USA. E-mail: gine@uconnvm.uconn.edu*

Let $H: L_2(S, \mathscr{S}, P) \mapsto L_2(S, \mathscr{S}, P)$ be a compact integral operator with a symmetric kernel h. Let X_i, $i \in \mathbb{N}$, be independent S-valued random variables with common probability law P. Consider the $n \times n$ matrix $\tilde{H}_n$ with entries $n^{-1}h(X_i, X_j)$, $1 \leq i, j \leq n$ (this is the matrix of an empirical version of the operator H with P replaced by the empirical measure P_n), and let H_n denote the modification of $\tilde{H}_n$, obtained by deleting its diagonal. It is proved that the ℓ_2 distance between the ordered spectrum of H_n and the ordered spectrum of H tends to zero a.s. if and only if H is Hilbert–Schmidt. Rates of convergence and distributional limit theorems for the difference between the ordered spectra of the operators H_n (or $\tilde{H}_n$) and H are also obtained under somewhat stronger conditions. These results apply in particular to the kernels of certain functions $H = \varphi(L)$ of partial differential operators L (heat kernels, Green functions).

This paper is dedicated to Richard M. Dudley on his sixtieth birthday.

Keywords: eigenvalues; heat kernels; integral operators; limit theorems; random matrices

1. Introduction

Let $(S, \mathscr{S}, P)$ be a probability space and let $h: S^2 \to \mathbb{R}$ be a symmetric kernel, that is, a measurable function symmetric in its two entries. Suppose the formula

$$Hg(x) = \int_S h(x, y)g(y)\,\mathrm{d}P(y), \qquad x \in S, \qquad g \in L_2(S, \mathscr{S}, P), \tag{1.1}$$

defines a compact operator on $L_2(P) := L_2(S, \mathscr{S}, P)$. For instance, if $\int_{S^2} h^2\,\mathrm{d}(P \times P) < \infty$, then H is Hilbert–Schmidt (i.e. $\mathbb{E}h^2(X_1, Y) < \infty$), and conversely. Let $\{X, Y, X_i : i \in \mathbb{N}\}$ be S-valued random variables, independently and identically P-distributed, and let $P_n = \frac{1}{n}\sum_{i=1}^n \delta_{X_i}$, $n \in \mathbb{N}$, be the corresponding empirical measures. Then, a candidate for the empirical counterpart to the operator H is, for each n, the random linear operator

$$\tilde{H}_n : L_2(P_n) \to L_2(P_n)$$

defined by

$$\tilde{H}_n g(x) = \int_S h(x, y)g(y)\,\mathrm{d}P_n(y), \qquad x \in S, \quad g \in L_2(S, \mathscr{S}, P_n).$$

The map $g \to (\frac{1}{\sqrt{n}}(g(X_1(\omega)), \ldots, g(X_n(\omega)))$ defines, for each $\omega \in \Omega$, an isometry of $L_2(P_n(\omega))$ onto a subspace of $\mathbb{R}^n$ (onto the whole of $\mathbb{R}^n$ if the values $X_i(\omega)$, $i \leqslant n$, are all different). By means of this isometry $\tilde{H}_n$ is identified to the operator of $\mathbb{R}^n$ whose matrix in the canonical basis is

$$\tilde{H}_n = \frac{1}{n}(h(X_i, X_j) : 1 \leqslant i, j \leqslant n). \tag{1.2}$$

Actually, as will be made clear in the next section, $\tilde{H}_n$ is not the best random operator that can be used in order to recover properties of the spectrum of H (unless H satisfies certain additional integrability conditions). It is rather its modification

$$H_n = \frac{1}{n}((1 - \delta_{ij})h(X_i, X_j) : 1 \leqslant i, j \leqslant n), \tag{1.3}$$

obtained by deleting the diagonal in the matrix $\tilde{H}_n$, that should be used to this effect. Clearly, this should be so because, if P is continuous, the operator H is not altered by modifications on the diagonal of the function $h(x, y)$.

The object of this paper is to show that the (usually infinite) spectrum of H can be approximated by the (finite) spectrum of H_n or the spectrum of $\tilde{H}_n$. Our results include a law of large numbers (LLN), convergence rates, central limit theorems (CLTs) and some examples. There is an extensive literature on estimation of spectra of random matrices. Most often one encounters limit theorems for spectra of sample covariance matrices (see, for example, Hsu 1939; Anderson 1948; James 1954; Geman 1980; Dauxois *et al.* 1982; Silverstein 1985; Girko 1990; Eaton and Tyler 1991; Bai 1993b), or for spectra of Wigner matrices, that is, matrices with independent entries (Wigner 1955; Dyson 1962a; 1962b; 1962c; Grenander 1963; Pastur 1973; Mehta 1991; Girko 1990; Voiculescu 1991; Bai 1993a; and references therein), or for certain random operators of interest in physics (see Pastur 1973; Cycon *et al.* 1987). However, the problems considered here, which somehow originate in U-statistics, do not seem to have been treated before.

Although we do not develop applications in this paper, we believe the results are potentially useful in several respects. As a first observation in this regard, our law of large numbers should provide a dimension-robust Monte Carlo method for estimating the spectrum of an integral Hilbert–Schmidt operator, which should at least be useful in higher dimensions. The results may also be of interest in nonparametric indirect estimation problems (for an interesting practical example of such a problem, see Vardi *et al.* 1985): suppose, for example, that the goal is to estimate an unknown function g by the observations of its integral transform Hg at random points X_i, $i = 1, \ldots, n$. Suppose also that the kernel h of H is unknown, but its observations at the sample X_i, $i = 1, \ldots, n$, are available. In such cases, preliminary estimation of the spectrum of H (as well as its eigenfunctions) is crucial for the construction of asymptotically optimal estimators of the unknown function g.

Another possible application may be found in the theory of U-statistics. Let U_n be a U-statistic with a P-completely degenerate Hilbert–Schmidt kernel h,

$$U_n := \frac{2}{n(n-1)} \sum_{i<j\leqslant n} h(X_i, X_j).$$

The limit distribution of nU_n is that of the random variable $\sum_{r=1}^{\infty} \lambda_r(g_r^2 - 1)$, where λ_r, $r \geqslant 1$, are the eigenvalues of h and g_r, $r \geqslant 1$, are independent standard normal random variables (see, for example, Serfling 1980). The eigenvalues of the random matrix H_n can then be used to approximate the limit distribution of U-statistics. Giné and Zhang (1996) formulated one of the conditions for the law of the iterated logarithm for degenerate U-statistics in terms of stochastic boundness of the maximal eigenvalue of the random matrix

$$\left(n^{-1} h(X_i, X_j) I_{h^2(X_i,X_j)\leqslant(i\vee j)^2 \log\log(i\vee j)} : 1 \leqslant i, j \leqslant n\right),$$

which is a truncated version of the matrix $\tilde{H}_n$. It is plausible that refinements of the present results may help in dealing with this not too practical condition. Dehling and Mikosch (1994) used the asymptotic properties of the matrices H_n and related quadratic forms in their study of the bootstrap for U-statistics. They proved a version of the law of large numbers, for the spectra of H_n, which follows from our Theorem 3.1.

It is also worth noting that the operators $\tilde{H}_n$, H_n provide a very simple model for asymptotically small random perturbations of an operator in a Hilbert space. Thus, the study of the asymptotic behaviour of their spectra might be related to other problems on fluctuations of the spectra of randomly perturbed operators. A physically meaningful example of this sort is the so-called 'crushed ice' problem (see Simon 1979), where one studies the asymptotics of the spectrum of the Laplacian with a Dirichlet boundary condition on a bounded open region G with random holes, as the number of holes tends to infinity. Ozawa (1987; 1993) obtained rates of convergence and distributional limit theorems for such random spectra (in some special cases) by considering the Green function of the random operator as a small random perturbation of the Green function of the Laplace operator in G. Some of his asymptotics look very similar to the asymptotics in our problem.

In Section 2 we describe a few known facts on perturbation of spectra of operators, namely, some very useful inequalities due to Lidskii and to Wielandt. Section 3 contains the LLN: we show that the ℓ_2 distance between the ordered spectra of H_n and H tends to zero almost surely if and only if the operator H is Hilbert–Schmidt. We also obtain rates for this LLN in Section 4. Under strong enough conditions these rates become of the order of $\sqrt{n}$, suggesting the possiblity of obtaining CLTs. These are considered in Sections 5 and 6. The limiting distribution is that of a suitable functional of a Gaussian process. It is worth noting that we do not assume the spectrum of H to be simple, and therefore application of the delta method must be circumvented. The conditions on H for rates of convergence or for the CLT are not always easy to verify. In Section 7 we show, in part via Tauberian type arguments, how bounds on the heat kernels of certain elliptic partial differential operators (Davies 1989) imply our conditions and provide estimators for their eigenvalues.

The main non-probabilitic tools throughout are the classical inequalities of Lidskii (1950) on perturbation of spectra and a less known inequality of Wielandt (1967) (see also Eaton and Tyler 1991). This last inequality allows us to treat multiple eigenvalues in the CLT.

Several *ad hoc* lemmas on perturbation of operators are also required, and we develop them as needed (Sections 5 and 6).

2. Some notation and several facts on operator theory

In this paper we consider mostly compact symmetric operators in a Hilbert space $\mathscr{H}$, mainly, $\mathscr{H} = L_2(S, \mathscr{S}, P)$. Given an operator $T : \mathscr{H} \mapsto \mathscr{H}$, we use the notation $\|T\|$ for the operator norm and $\|T\|_{\mathrm{HS}}$ for the Hilbert–Schmidt norm of T (the latter, defined as $\|T\|_{\mathrm{HS}} = (\sum_\alpha |Te_\alpha|^2)^{1/2}$ for any complete orthonormal system $\{e_\alpha\}$).

Let J be a countable set (more specifically, $J = \mathbb{N}$ or $J = \mathbb{Z}$). As usual $\ell_2 := \ell_2(J)$ will denote the Hilbert space of all sequences $\mu := \{\mu_n\}_{n\in J}$, such that $\sum_{n\in J}|\mu_n|^2 < +\infty$ with inner product $\langle \mu, \nu\rangle_{\ell_2} := \sum_{n\in J}\mu_n\nu_n$ and norm $\|\cdot\|_{\ell_2}^2 := \langle\cdot,\cdot\rangle_{\ell_2}$. Let $\|\mu\|_{c_0} := \sup_{n\in J}|\mu_n|$ denote the sup-norm of the sequence μ. Then, $c_0(\mathbb{N})$ will denote the Banach space of all sequences $\mu := \{\mu_n\}_{n\in\mathbb{N}}$ such that $\mu_n \to 0$ as $n \to +\infty$, equipped with the norm $\|\cdot\|_{c_0}$. Similarly, $c_0(\mathbb{Z})$ will denote the Banach space of all sequences $\mu := \{\mu_n\}_{n\in\mathbb{Z}}$, such that $\mu_n \to 0$ as $|n| \to \infty$, also with the norm $\|\cdot\|_{c_0}$. We will use the same norms in finite-dimensional spaces $\mathbb{R}^n$, considering them as subspaces of c_0 or ℓ_2, respectively.

Let $\mathscr{P} = \mathscr{P}(J)$ be the set of all bijections on the set J, and let $\mathscr{P}$ also denote the equivalence relation on $\mathbb{R}^J$ given by $\{x_i : i \in J\}\mathscr{P}\{y_i : i \in J\}$ if and only if the sequence $\{y_i\}$ is a rearrangement of the sequence $\{x_i\}$, that is, $y_i = x_{\sigma(i)}$, $i \in J$, for some $\sigma \in \mathscr{P}$. Then, the spectrum $\lambda(T)$ of a compact, symmetric operator T on $\mathscr{H}$, which is the set of all its eigenvalues, each counted with its multiplicity, is a point in $c_0/\mathscr{P}$, and if moreover T is Hilbert–Schmidt, then $\lambda(T) \in \ell_2/\mathscr{P}$, as is well known (see, for example, Dunford and Schwartz 1963, Vol. II, Chapters X and XI). If A is a symmetric operator on $\mathbb{R}^n$, $n < \infty$, then we make the convention of adding to its spectrum an infinite number of zeros, so as to make of it also a point in $\ell_2/\mathscr{P}$ (or, more formally, we denote by $\lambda(A)$ the spectrum of the operator on ℓ_2 that coincides with A on $\mathbb{R}^n$ and is zero on the orthocomplement of $\mathbb{R}^n$ in ℓ_2). These formalities are introduced because we will measure closeness of spectra of operators by the following distance δ_2 defined on $\ell_2/\mathscr{P}$, which we call the ℓ_2 *rearrangement* distance: if $x = \{x_i : i \in J\}$ and $y = \{y_i : i \in J\}$ are in ℓ_2, and if $[x]$, $[y]$ are the corresponding equivalence classes, then $\delta_2([x], [y])$, or, with some abuse of notation, $\delta_2(x, y)$, is defined as

$$\delta_2(x, y) = \inf_{\sigma\in\mathscr{P}}\left[\sum(x_i - y_{\sigma(i)})^2\right]^{\frac{1}{2}}.$$

To see that δ_2 is a distance on $\ell_2/\mathscr{P}$, given $x, y, z \in \ell_2$ and $\varepsilon > 0$, let $\tau \in \mathscr{P}$ be such that $\delta_2(x, z) \geqslant [\sum(x_i - z_{\tau(i)})^2]^{\frac{1}{2}} - \varepsilon/2$ and let $\sigma \in \mathscr{P}$ be such that $\delta_2(y, z) \geqslant [\sum(z_{\tau(i)} - y_{\sigma(i)})^2]^{\frac{1}{2}} - \varepsilon/2$. Then,

$$\delta_2(x, y) \leqslant \left[\sum(x_i - y_{\sigma(i)})^2\right]^{\frac{1}{2}} \leqslant \delta_2(x, z) + \delta_2(z, y) - \varepsilon$$

by the triangle inequality for the ℓ_2 norm. So, δ_2 is a well-defined distance between spectra of Hilbert–Schmidt operators and/or spectra of operators in $\mathbb{R}^n$, $n<\infty$.

Given two vectors $u := (u_1, \ldots, u_n) \in \mathbb{R}^n$ and $v := (v_1, \ldots, v_m) \in \mathbb{R}^m$, we set

$$u \oplus v := (u_1, \ldots, u_n, v_1, \ldots, v_m) \in \mathbb{R}^{n+m},$$

with obvious extensions to the cases of finite sums, infinite sums, infinite-dimensional vectors, etc.

Given a vector μ in $\mathbb{R}^d$ we will denote by $\mu^\downarrow$ ($\mu^\uparrow$) its non-increasing (non-decreasing) rearrangement, that is, the vector of $\mathbb{R}^d$ with the same coordinates as μ, arranged in non-increasing (non-decreasing) order.

For $x = (x_i) \in c_0$, we set $\tilde{x}_+ = (x_i \vee 0)$ and $\tilde{x}_- = x - \tilde{x}_+$. Then, for any $x, y \in \ell_2$, $\delta_2^2(x, y) = \delta_2^2(\tilde{x}_+, \tilde{y}_+) + \delta_2^2(\tilde{x}_-, \tilde{y}_-)$, and if we let $\tilde{x}_+^\downarrow$ denote the point in ℓ_2 with the same coordinates of $\tilde{x}_+$, but arranged in non-increasing order (with the zero coordinates suppressed if $\tilde{x}_+$ has an infinite number of non-zero coordinates), we have $\delta_2(\tilde{x}_+, \tilde{y}_+) = \|\tilde{x}_+^\downarrow - \tilde{y}_+^\downarrow\|_{\ell_2}$. Likewise, $\delta_2(\tilde{x}_-, \tilde{y}_-) = \|\tilde{x}_-^\uparrow - \tilde{y}_-^\uparrow\|_{\ell_2}$, where the superscript $\uparrow$ indicates non-decreasing rearrangement. (For $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$, $\tilde{x}_+ = (x_1 \vee 0, \ldots, x_n \vee 0, 0, \ldots) \in \ell_2$, and the seemingly superfluous tilde is introduced in order to distinguish this point from the vector in a Euclidean space of lower dimension obtained by deleting all the zeros from $\tilde{x}_+$, which will be denoted by x_+ in Section 5.)

If $\mu \in c_0(\mathbb{Z})$, we set $\mu^{\uparrow\downarrow} = \tilde{\mu}_-^\uparrow \oplus \tilde{\mu}_+^\downarrow$. (For instance, if $\mu = (-2, -3, 1, 2)$ then $\mu^\downarrow = (2, 1, -2, -3)$ and $\mu^{\uparrow\downarrow} = (-3, -2, 0, \ldots) \oplus (2, 1, 0, \ldots) \in c_0(\mathbb{Z})$.) Clearly, for any two sequences $\mu, \nu \in \ell_2(\mathbb{Z}) = \ell_2 \oplus \ell_2$,

$$\delta_2(\mu, \nu) = \|\mu^{\uparrow\downarrow} - \nu^{\uparrow\downarrow}\|_{\ell_2}.$$

We will make frequent use of the following inequality of Lidskii (1950) (see also Kato 1982).

Theorem 2.1 (Lidskii's inequality). *Let A, B be symmetric operators on* $\mathbb{R}^d$. *Then, for any convex function* φ *on* $\mathbb{R}$,

$$\sum_{j=1}^{d} \varphi(\lambda_j^\downarrow(B) - \lambda_j^\downarrow(A)) \leqslant \sum_{j=1}^{d} \varphi(\lambda_j^\downarrow(B - A)). \tag{2.1}$$

In particular, for all $p \geqslant 1$,

$$\left(\sum_{j=1}^{d} |\lambda_j^\downarrow(B) - \lambda_j^\downarrow(A)|^p \right)^{1/p} \leqslant \left(\sum_{j=1}^{d} |\lambda_j^\downarrow(B - A)|^p \right)^{1/p}. \tag{2.2}$$

For $p = +\infty$,

$$\max_{1 \leqslant j \leqslant d} |\lambda_j^\downarrow(B) - \lambda_j^\downarrow(A)| \leqslant \max_{1 \leqslant j \leqslant d} |\lambda_j^\downarrow(B - A)|. \tag{2.3}$$

We are especially interested in the case $p = 2$. In this case Lidskii's inequality was extended by Hoffman and Wielandt (1953) to the more general case of normal operators A

and B. We formulate this result here in a form covenient for our goals (using the δ_2 distance).

Theorem 2.2 (Hoffman–Wielandt Inequality). *If A and B are normal operators of $\mathbb{R}^d$, in particular if they are symmetric, then*

$$\delta_2(\lambda(A), \lambda(B)) \leqslant \|A - B\|_{\mathrm{HS}}. \tag{2.4}$$

The δ_2 distance is slightly different from the distance used, for example, in Hoffman and Wielandt (1953); in fact it only coincides with it on pairs of operators that have both non-negative (or non-positive) spectra. However, as is easy to see, their inequality also applies to δ_2. In what follows we will refer to inequality (2.4), which we only use on symmetric operators, as the Lidskii–Hoffman–Wielandt inequality.

Lidskii's inequality (2.3) can be improved for special types of perturbations: in some cases, the bound on the right-hand side of (2.3) can be replaced by a constant times $\|B - A\|^2$ (as opposed to $\|B - A\|$). The following result is due to Wielandt (1967) (see, for example, Eaton and Tyler 1991):

Theorem 2.3 (Wielandt's inequalities). *Let A be a symmetric operator on $\mathbb{R}^n$ such that $\lambda_d^{\downarrow}(A) - \lambda_{d+1}^{\downarrow}(A) > 0$, for some $1 \leqslant d < n$. Let P^d denote the orthogonal projector of $\mathbb{R}^n$ onto the subspace generated by the eigenvectors corresponding to the largest d eigenvalues of A, and let P_d denote the orthogonal projector onto the subspace generated by the remaining eigenvectors of A. If B is a symmetric operator such that $P_d B P_d = 0$ and $P^d B P^d = 0$, then*

$$0 \leqslant \lambda_j^{\downarrow}(A + B) - \lambda_j^{\downarrow}(A) \leqslant \frac{\|B\|^2}{\lambda_j^{\downarrow}(A) - \lambda_{d+1}^{\downarrow}(A)}, \qquad j = 1, \ldots, d, \tag{2.5}$$

and

$$0 \leqslant \lambda_j^{\downarrow}(A) - \lambda_j^{\downarrow}(A + B) \leqslant \frac{\|B\|^2}{\lambda_d^{\downarrow}(A) - \lambda_j^{\downarrow}(A)}, \qquad j = d + 1, \ldots, n. \tag{2.6}$$

Wielandt's inequalities will be used at a crucial step in the derivation of the limiting distribution of the not necessarily simple eigenvalues of the random matrices considered in this paper. They have been put to a somewhat similar use in a different situation by Eaton and Tyler (1991).

3. The law of large numbers

In the setup of Section 1 and with these definitions, the LLN for spectra of Hilbert–Schmidt operators is as follows:

Theorem 3.1. *If $\mathbb{E}h^2(X, Y) < \infty$, then*

$$\delta_2(\lambda(H_n), \lambda(H)) \to 0 \text{ a.s.} \tag{3.1}$$

Proof. h being symmetric and $\mathbb{E}h^2(X, Y) < \infty$, the operator H that corresponds to the kernel h by equation (1.1) is Hilbert–Schmidt. Then, by the spectral theorem (see e.g. Theorem 4, Chapter X, and Section XI.6 in Dunford and Schwartz, Part II, 1964), there exists an orthonormal set of $L_2(P)$, $\{\phi_i : i \in J\}$, where $J = \mathbb{N}$ or $J = \{1, \ldots, R\}$ for some $R < \infty$, and a sequence of real numbers, $\{\lambda_i : i \in J\}$, with non-increasing absolute values and satisfying $\sum_{i\in J}\lambda_i^2 < \infty$, such that

$$h(x, y) = \sum_{i\in J} \lambda_i \phi_i(x)\phi_i(y) \tag{3.2}$$

in the $L_2(P)$ sense. The set $\lambda(H) := \{\lambda_i : i \in J\}$ (that we increase with a sequence of zeros when necessary) is the spectrum of H.

We first prove the theorem in the case where the identity in (3.2) holds pointwise and $J = \{1, \ldots, R\}$ for some $R < \infty$ (this will turn out to be the basic case). If we define the random vectors

$$\mathbf{\Phi}_r^n = \left(\frac{\phi_r(X_1)}{n^{\frac{1}{2}}}, \ldots, \frac{\phi_r(X_n)}{n^{\frac{1}{2}}}\right), \qquad r \leqslant R,$$

then the operator $\tilde{H}_n$ (see (1.2)) becomes

$$\tilde{H}_n \mathbf{u} = \sum_{r=1}^{R} \lambda_r \langle \mathbf{\Phi}_r^n, \mathbf{u}\rangle \mathbf{\Phi}_r^n, \qquad \mathbf{u} \in \mathbb{R}^n. \tag{3.3}$$

By the LLN and orthogonality of the functions ϕ_r in $L_2(P)$, we have

$$\langle \mathbf{\Phi}_r^n, \mathbf{\Phi}_s^n\rangle \to \delta_{r,s} \text{ a.s.} \tag{3.4}$$

for all $r, s \leqslant R$. The limit (3.4) indicates that the restriction of $\tilde{H}_n$ to the linear span of the vectors $\mathbf{\Phi}_1^n, \ldots, \mathbf{\Phi}_R^n$ is an asymptotically small perturbation of an operator with spectrum $\lambda_1, \ldots, \lambda_R$, and therefore, by general principles (the Lidskii–Hoffman–Wielandt inequality), its spectrum must approach $\{\lambda_1, \ldots, \lambda_R\}$. Indeed, the proof that follows makes this observation precise and shows as well that the spectra of $\tilde{H}_n$ and H_n are also asymptotically close.

Let E_n, $n \in \mathbb{N}$, denote the random operators on $\mathbb{R}^R$ whose matrices in the canonical basis are given by the relations

$$I + E_n = (\langle \mathbf{\Phi}_r^n, \mathbf{\Phi}_s^n\rangle : 1 \leqslant r, s \leqslant R), \qquad n \in \mathbb{N}, \tag{3.5}$$

$I := I_R$ denoting the identity operator on $\mathbb{R}^R$. Clearly,

$$\sum_{1\leqslant r,s\leqslant R} \langle \mathbf{\Phi}_r^n, \mathbf{\Phi}_s^n\rangle u_r u_s = \left\langle \sum_{r=1}^{R} u_r \mathbf{\Phi}_r^n, \sum_{r=1}^{R} u_r \mathbf{\Phi}_r^n \right\rangle \geqslant 0$$

for all $\mathbf{u} = (u_1, \ldots, u_R) \in \mathbb{R}^R$. It follows that the operator $I + E_n$ is positive definite and, therefore, it has a square root, say $A_n := (I + E_n)^{1/2}$. Let $\mathbf{e}_1, \ldots, \mathbf{e}_R$ denote the canonical basis of $\mathbb{R}^R$ and set $\hat{\mathbf{e}}_r = A_n \mathbf{e}_r$, $r = 1, \ldots, R$. Then, for all $1 \leqslant r, s \leqslant R$,

$$\langle \hat{\mathbf{e}}_r, \hat{\mathbf{e}}_s \rangle = \langle A_n \mathbf{e}_r, A_n \mathbf{e}_s \rangle = \langle A_n^2 \mathbf{e}_r, \mathbf{e}_s \rangle = \langle \mathbf{\Phi}_r^n, \mathbf{\Phi}_s^n \rangle,$$

that is, the correspondence $\hat{\mathbf{e}}_r \leftrightarrow \mathbf{\Phi}_r^n$, $r = 1, \ldots, R$, defines an isometry between the linear spans of the sets of vectors $\{\hat{\mathbf{e}}_r : 1 \leqslant r \leqslant R\}$ and $\{\mathbf{\Phi}_r^n : 1 \leqslant r \leqslant R\}$. This isometry implies that the random operator $\hat{H}_n$ of $\mathbb{R}^R$, given by

$$\hat{H}_n \mathbf{u} = \sum_{r=1}^{R} \lambda_r \langle \hat{\mathbf{e}}_r, \mathbf{u} \rangle \hat{\mathbf{e}}_r, \qquad \mathbf{u} \in \mathbb{R}^R, \tag{3.6}$$

has the same spectrum as $\tilde{H}_n$,

$$\lambda(\tilde{H}_n) = \lambda(\hat{H}_n).$$

Finally, let K be the operator defined by

$$K\mathbf{u} = \sum_{r=1}^{R} \lambda_r \langle \mathbf{e}_r, \mathbf{u} \rangle \mathbf{e}_r, \qquad \mathbf{u} \in \mathbb{R}^R. \tag{3.7}$$

Obviously,

$$\lambda(K) = \{\lambda_i : 1 \leqslant i \leqslant R\} = \lambda(H).$$

It is also easy to see, using the symmetry of A_n, that

$$\hat{H}_n = A_n K A_n. \tag{3.8}$$

We then have

$$\begin{aligned} \delta_2(\lambda(H), \lambda(H_n)) &\leqslant \delta_2(\lambda(H), \lambda(\tilde{H}_n)) + \delta_2(\lambda(\tilde{H}_n), \lambda(H_n)) \\ &\leqslant \delta_2(\lambda(K), \lambda(\hat{H}_n)) + \delta_2(\lambda(\tilde{H}_n), \lambda(H_n)). \end{aligned} \tag{3.9}$$

Then, using (2.4) and the Marcinkiewicz law of large numbers ($\mathbb{E}\phi_r^2(X) < \infty$ for all r), we obtain

$$\begin{aligned} \delta_2^2(\lambda(\tilde{H}_n), \lambda(H_n)) &\leqslant \sum_{i=1}^{n} \left(\frac{1}{n} \sum_{r=1}^{R} \lambda_r \phi_r^2(X_i) \right)^2 \\ &= \sum_{1 \leqslant r,s \leqslant R} \lambda_r \lambda_s \left[\frac{1}{n^2} \sum_{i=1}^{n} \phi_r^2(X_i) \phi_s^2(X_i) \right] \to 0 \text{ a.s.} \end{aligned} \tag{3.10}$$

In order to estimate the first summand on the right of (3.9), we first observe that, by (3.8) and Lidskii–Hoffman–Wielandt,

$$\delta_2(\lambda(K), \lambda(\hat{H}_n)) \leqslant \|A_n K A_n - K\|_{\mathrm{HS}}. \tag{3.11}$$

Since $I + E_n$ is positive definite, we have $\lambda(E_n) \subset [-1, +\infty)$. Let

$$\alpha(\lambda) := (1 + \lambda)^{1/2} - 1, \qquad \lambda \geqslant -1.$$

Then, for all $\lambda \geqslant -1$,

$$|\alpha(\lambda)| = \left| \frac{\lambda}{(1+\lambda)^{1/2}+1} \right| \leqslant |\lambda|.$$

This observation implies, for example by diagonalization, that

$$\|A_n - I\|_{\mathrm{HS}} = \|(I+E_n)^{1/2} - I\|_{\mathrm{HS}} = \|\alpha(E_n)\|_{\mathrm{HS}} \leqslant \|E_n\|_{\mathrm{HS}}. \tag{3.12}$$

Since $\|E_n\|_{\mathrm{HS}}^2 = \sum_{1\leqslant r,s\leqslant R}(\langle \mathbf{\Phi}_r^n, \mathbf{\Phi}_s^n\rangle - \delta_{r,s})^2 \to 0$ a.s. by (3.4), we have

$$\|A_n - I\|_{\mathrm{HS}} \to 0 \text{ a.s.}$$

Hence, by continuity of operator composition with respect to the Hilbert–Schmidt topology ($\|AB\|_{\mathrm{HS}} \leqslant \|A\|_{\mathrm{HS}}\|B\|$ and $\|AB\|_{\mathrm{HS}} \leqslant \|A\|\,\|B\|_{\mathrm{HS}}$), it follows that

$$\|A_n K A_n - K\|_{\mathrm{HS}} \to 0 \text{ a.s.}$$

and, by (3.11), that

$$\delta_2(\lambda(K), \lambda(\hat{H}_n)) \to 0 \text{ a.s.} \tag{3.13}$$

Finally, (3.9), (3.10) and (3.13) give

$$\delta_2(\lambda(H_n), \lambda(H)) \to 0 \text{ a.s.}$$

in the case $h(x, y) = \sum_{r=1}^{R}\lambda_r\phi_r(x)\phi_r(y)$ pointwise and $R < \infty$.

In the general case, h has the representation (3.2) with $J = \mathbb{N}$. For such h, let $h_R(x, y) := \sum_{r=1}^{R}\lambda_r\phi_r(x)\phi_r(y)$, let H_R be the integral operator with kernel h_R, and let $H_{R,n}$ be the operator of $\mathbb{R}^n$ whose matrix in the canonical basis is given by

$$H_{R,n} = \frac{1}{n}((1-\delta_{ij})h_R(X_i, X_j) : 1 \leqslant i, j \leqslant n).$$

Then we have

$$\lim_{R\to\infty} \delta_2(\lambda(H), \lambda(H_R)) = \lim_{R\to\infty}\left(\sum_{r=R+1}^{\infty}\lambda_r^2\right)^{\frac{1}{2}} = 0$$

since H is Hilbert–Schmidt, and

$$\lim_{n\to\infty} \delta_2(\lambda(H_R), \lambda(H_{R,n})) = 0 \text{ a.s.}$$

for all $R < \infty$ by the first part of this proof. Next we note that if f is a symmetric square-integrable kernel, then the LLN for U-statistics (see, for example, Serfling 1980, Theorem 5.4.A) gives

$$\frac{1}{n^2}\sum_{1\leqslant i\neq j\leqslant n} f^2(X_i, X_j) \to \mathrm{E}f^2(X, Y) \text{ a.s.} \tag{3.14}$$

Then the Lidskii–Hoffman–Wielandt inequality, together with the limit (3.14) for $f(x, y) = h(x, y) - h_R(x, y)$, implies

$$\lim_{R\to\infty}\lim_{n\to\infty}\delta_2(\lambda(H_{R,n}),\lambda(H_n))\leqslant\lim_{R\to\infty}\lim_{n\to\infty}\|H_{R,n}-H_n\|_{\mathrm{HS}}$$

$$=\lim_{R\to\infty}\lim_{n\to\infty}\left[\frac{1}{n^2}\sum_{1\leqslant i\neq j\leqslant n}(h-h_R)^2(X_i,X_j)\right]$$

$$=\lim_{R\to\infty}\left(\sum_{r=R+1}^{\infty}\lambda_r^2\right)^{\frac{1}{2}}=0\text{ a.s.}$$

Collecting together the previous three limits, we obtain

$$\lim_{n\to\infty}\delta_2(\lambda(H_n),\lambda(H))$$

$$\leqslant\lim_{R\to\infty}\limsup_{n\to\infty}[\delta_2(\lambda(H),\lambda(H_R))+\delta_2(\lambda(H_R),\lambda(H_{R,n}))+\delta_2(\lambda(H_{R,n}),\lambda(H_n))]=0\text{ a.s.},$$

proving the theorem. □

There is some interest in noticing that, if $\mathrm{E}h^2(X,Y)<\infty$, then there is *convergence of all the moments of order* $2-\delta$, $0<\delta<2$, in the law of large numbers (3.1). By the triangle inequality, $\delta_2(\lambda(H_n),\lambda(H))$ is dominated by $\delta_2(\lambda(H),0)+\delta_2(\lambda(H_n),0)$, with the first summand bounded by $(\mathbb{E}h^2)^{1/2}$ and the second satisfying

$$\mathbb{E}[\delta_2(\lambda(H_n),0)]^2=\frac{1}{n^2}\mathbb{E}\left[\sum_{1\leqslant i\neq j\leqslant n}h^2(X_i,X_j)\right]=\frac{n-1}{n}\mathbb{E}h^2(X,Y),$$

so that the assertion follows by uniform integrability.

Theorem 3.1 has a converse. In fact, the following stronger statement holds:

Theorem 3.2. *Let h be a symmetric, measurable real function on $(S^2,\mathscr{S}^2)$, let P be a probability measure on $(S,\mathscr{S})$, and let H_n be the random symmetric operators on $\mathbb{R}^n$ corresponding to h and to a sample from P via equation (1.3). If the sequence of spectra $\{\lambda(H_n)\}$ is stochastically δ_2-bounded, then $\mathbb{E}h^2(X,Y)<\infty$.*

Proof. Since $\delta_2^2(\lambda(H_n),0)=\sum_{1\leqslant i\neq j\leqslant n}h^2(X_i,X_j)/n^2$, the hypothesis implies, by the decoupling inequality in Giné and Zinn (1994, Theorem 2.b), that the sequence $\{\sum_{1\leqslant i,j\leqslant n}h^2(X_i,X_j')/n^2\}_{n=1}^{\infty}$ is stochastically bounded, where $\{X_i'\}$ is an independent copy of the sequence $\{X_i\}$. Hence we have

$$\lim_{K\to\infty}\sup_{n\in\mathbb{N}}\Pr\left\{\frac{1}{n^2}\sum_{1\leqslant i,j\leqslant n}h^2(X_i,X_j')>K\right\}=0. \tag{3.15}$$

This limit will allow us to apply the extension to U-statistics of Hoffmann-Jørgensen's inequality given by Giné and Zinn (1992) in a way similar to Giné and Zhang's (1996) inequality: as mentioned in the latter reference, Hoffmann-Jørgensen's inequality for U-

statistics also holds for non-negative kernels, and, applied to the kernels $h^2 \wedge n$, it gives that there exist universal constants c_1 and c_2 such that, for all $n \in \mathbb{N}$,

$$\mathbb{E}\left(\frac{1}{n^2}\sum_{1\leqslant i,j\leqslant n}(h^2(X_i, X'_j)\wedge n)\right)^2 \leqslant c_1\left[\mathbb{E}\left(\max_{1\leqslant i\leqslant n}\sum_{j=1}^{n}(h^2(X_i, X'_j)\wedge n)\right)^2 + t_{0,n}^2\right], \qquad (3.16)$$

where $t_{0,n}$ are any numbers satisfying

$$\Pr\left\{\frac{1}{n^2}\sum_{1\leqslant i,j\leqslant n}(h^2(X_i, X'_j)\wedge n)^2 > t_{0,n}\right\} \leqslant c_2.$$

Now, by positivity, (3.15) implies that the numbers $t_{0,n}$ can be chosen so that $t_0 := \sup_n t_{0,n} < \infty$; this observation and (3.16) then give

$$\sup_{n\in\mathbb{N}}(\mathbb{E}[h^2(X, Y)\wedge n])^2 \leqslant c_1(1 + t_0^2) < \infty,$$

proving the theorem. □

Let δ denote Hausdorff distance between subsets of $\mathbb{R}$, that is,

$$\delta(A, B) = [\sup_{a\in A}\inf_{b\in B}|b-a|] \vee [\sup_{b\in B}\inf_{a\in A}|b-a|], \qquad A, B \subset \mathbb{R}.$$

Closeness of spectra of operators is sometimes also measured by the Hausdorff distance between the corresponding sets (although this distance is less appropriate than the δ_2 distance because it does not detect multiplicities). Let us also momentarily denote by $\lambda(T)$ the set of different eigenvalues of a compact operator T. It is then obvious that $\delta(\lambda(T), \lambda(U)) \leqslant \delta_2(\lambda(T), \lambda(U))$ for any Hilbert–Schmidt operators T and U. Therefore, Theorem 3.1 immediately gives:

Corollary 3.3. *If* $\mathbb{E}h^2(X, Y) < \infty$, *then*

$$\delta(\lambda(H_n), \lambda(H)) \to 0 \text{ a.s.}$$

Next we make two observations regarding the previous results: one, quite simple, to show that we cannot replace H_n by $\tilde{H}_n$ in Theorem 3.1 and Corollary 3.3; and the other, which requires some more work, to indicate that the condition $\mathbb{E}h^2 < \infty$ in Corollary 3.3 is sharp.

Of course, the proof of Theorem 3.1 does not work for $\tilde{H}_n$: the limit (3.14) does not hold without extra assumptions if the sum includes the terms $i = j$ (the diagonal). Moreover, it is easy to construct an example of a square-integrable kernel h for which $\limsup_{n\to\infty}\delta(\lambda(\tilde{H}_n), \lambda(H)) = \infty$: take P continuous and h such that $h(x, y) = 0$ for $x \neq y$ and $\mathbb{E}|h(X, X)| = \infty$; in this case, the spectrum of H is $\{0\}$, whereas the spectrum of $\tilde{H}_n$ is $\{h(X_i, X_i)/n : i = 1, \ldots, n\}$, so that we have

$$\limsup_{n\to\infty}\delta(\lambda(\tilde{H}_n), \lambda(H)) = \limsup_{n\to\infty}\frac{1}{n}\max_{1\leqslant i\leqslant n}|h(X_i, X_i)| = \infty \text{ a.s.}$$

In order to prove that Corollary 3.3 is sharp, we will consider a generalization of the example introduced in Giné and Zhang (1996) in connection with the law of the iterated logarithm. We take $(S, \mathscr{S}, P) = ([0, 1], \mathscr{B}, \lambda)$, λ denoting Lebesgue measure, and denote by I_r the indicator function of the interval $(1 - 2^{-r+1}, 1 - 2^{-r}]$. Then, given a sequence $\{\lambda_r\}$ of real numbers, we set

$$h(x, y) = \sum_{r=1}^{\infty} \lambda_r 2^r I_r(x) I_r(y), \tag{3.17}$$

a kernel whose eigenfunctions are $2^{r/2} I_r$, $r \in \mathbb{N}$, each with eigenvalue λ_r (as well as $I_{\{0\}}$, which has eigenvalue zero). The discussion that follows applies also to the degenerate kernel

$$h((x, \alpha), (y, \beta)) = \sum_{r=1}^{\infty} \lambda_r 2^r I_r(x) I_r(y) \alpha\beta, \tag{3.17$'$}$$

where now $S = [0, 1] \times \{-1, 1\}$ and $P = \lambda \times \frac{1}{2}(\delta_{-1} + \delta_1)$. (This kernel corresponds to replacing, for each r, the eigenvector $2^{r/2} I_r$ by the Haar function with support I_r.) The following proposition subsumes a previous weaker result and was obtained in collaboration with J. Zinn.

Proposition 3.4. *Let h be the kernel defined by (3.17) or by (3.17′), and let $\{X_i\}_{i=1}^{\infty}$ be an independently and identically distributed sequence of random variables uniform on* [0, 1]. *Then, a necessary and sufficient condition for*

$$\limsup_{n \to \infty} \max_{1 \leqslant i \neq j \leqslant n} \frac{1}{n} |h(X_i, X_j)| < \infty \text{ a.s.} \tag{3.18}$$

is $\sum_{i=1}^{\infty} \lambda_i^2 < \infty$ (i.e., $\mathbb{E} h^2(X, Y) < \infty$ or, equivalently, that h is the kernel of a Hilbert–Schmidt operator on $L_2([0, 1], \mathscr{B}, \lambda)$).

Proof. If $\sum \lambda_i^2 < \infty$, that is, if $\mathbb{E} h^2(X, Y) < \infty$, then (3.18) is a consequence of the LLN for the U-statistics applied to the kernel h^2. Let us now assume that (3.18) holds. Then, by Borel–Cantelli, there exists $K < \infty$ such that

$$\sum_{m=1}^{\infty} \Pr\left\{ \frac{1}{2^m} \max_{1 \leqslant i \neq j \leqslant 2^m} |h(X_i, X_j)| \geqslant K \right\} < \infty. \tag{3.19}$$

Since at most one summand in (3.17) or (3.17′) is different from zero for any given x, y, we can replace λ_r by $|\lambda_r| \wedge c$ in the definition of h and continue having (a fortiori) convergence of the series in (3.19). In other words, we can assume the eigenvalues λ_r to be non-negative and bounded. We can also take $K = 1$ in (3.19). Let $X_{m,1}$, $X_{m,2}$ be the first two order statistics of the sequence $X_1, \ldots, X_{2^m}$ (with any of the usual conventions for ties). Since

$$\max_{1 \leqslant i \neq j \leqslant 2^m} |h(X_i, X_j)| \geqslant 2^r \lambda_r I_r(X_{m,1}) I_r(X_{m,2}),$$

and since the events $\{X_{m,1} \in I_r, X_{m,2} \in I_r\}$, $r \in \mathbb{N}$, are disjoint, we have

$$\Pr\left\{\frac{1}{2^m}\max_{1\leqslant i\neq j\leqslant 2^m}|h(X_i, X_j)|\geqslant 1\right\}\geqslant\sum_{r:2^r\lambda_r\geqslant 2^m}\Pr\{\mathrm{X}_{m,1}\in I_r, X_{m,2}\in I_r\}. \tag{3.20}$$

Since $2^m/2^r\leqslant c$ in the domain of summation of this last series, there exists $C>0$ such that, for $r\geqslant 2$,

$$\begin{aligned}\Pr\{X_{m,1}\in I_r, X_{m,2}\in I_r\}&=\sum_{k=0}^{2^m-2}\binom{2^m}{k}\left(1-\frac{1}{2^{r-1}}\right)^k\left(\frac{1}{2^r}\right)^{2^m-k}\\&\geqslant\binom{2^m}{2}\frac{1}{2^{2r}}\left(1-\frac{1}{2^{r-1}}\right)^{2^m-2}\\&\geqslant C\frac{2^{2m}}{2^{2r}},\end{aligned} \tag{3.21}$$

and, for $r=1$, $\Pr\{X_{m,1}\in I_1, X_{m,2}\in I_1\}=2^{-2m}\geqslant C2^{2m}/2^2$ since $m\leqslant\log_2(2c)$. Now, (3.19)–(3.21) give

$$\begin{aligned}\infty&>\sum_{m=1}^{\infty}\sum_{r:2^r\lambda_r\geqslant 2^m}\frac{2^{2m}}{2^{2r}}\\&=\sum_{r=1}^{\infty}\sum_{m:2^r\lambda_r\geqslant 2^m}\frac{2^{2m}}{2^{2r}}\\&\geqslant\sum_{r=1}^{\infty}\lambda_r^2 I_{2^r\lambda_r\geqslant 2}.\end{aligned}$$

Since $\sum_{r=1}^{\infty}\lambda_r^2 I_{2^r\lambda_r<2}<\infty$, it follows that $\sum(|\lambda_r|\wedge c)^2<\infty$ and, hence, that $\sum\lambda_r^2<\infty$. □

Let $\lambda_{1,n}$ be the eigenvalue of the matrix H_n of largest absolute value. Then

$$|\lambda_{1,n}|=\frac{1}{n}\sup_{|\mathbf{u}|\leqslant 1}\left|\sum_{1\leqslant i\neq j\leqslant n}h(X_i, X_j)u_iu_j\right|, \tag{3.22}$$

where $|\mathbf{u}|$ denotes the Euclidean norm of the vector $\mathbf{u}=(u_1, \ldots, u_n)\in\mathbb{R}^n$. If the maximum of $|h(X_i, X_j)|$ on the set $1\leqslant i\neq j\leqslant n$ is attained at $i=r$, $j=s$, taking $u_r=u_s=1/\sqrt{2}$ and $u_i=0$ for $i\neq r$, $i\neq s$, gives

$$|\lambda_{1,n}|\geqslant\frac{1}{n}\max_{1\leqslant i\neq j\leqslant n}|h(X_i, X_j)|. \tag{3.23}$$

It is also clear that $\lambda_{1,n}^2\leqslant n^{-2}\sum_{1\leqslant i\neq j\leqslant n}h^2(X_i, X_j)$. It then follows from these two inequalities that Proposition 3.4 has the following corollary to the effect that the LLN in Corollary 3.3 is the best possible on the class of kernels (3.17), (3.17′):

Corollary 3.5. *Let H_n be the random operators corresponding, via equation (1.3), to the* *kernel h defined in (3.17) or (3.17′), and to a sample from the uniform distribution on* [0, 1]. *Let* $\lambda_{1,n}$ *denote the eigenvalue of the operator* H_n *with largest absolute value,* $n \in \mathbb{N}$. *Then*

$$\limsup_{n\to\infty} |\lambda_{1,n}| < \infty \text{ a.s.} \tag{3.24}$$

if and only if $\sum_{r=1}^{\infty} \lambda_r^2 < \infty$.

Recently, Latala (1998) solved a problem which we raised in the first version of this paper. He proved that the converse of Corollary 3.3 holds in general, that is, that the square-integrability of an arbitrary kernel h is necessary for the limit (3.24) to hold.

4. Rates of convergence

In this section $\{\phi_r : r \in \mathbb{N}\}$ and $\{\lambda_r : r \in \mathbb{N}\}$ will be as in the proof of Theorem 3.1, that is, $\{\phi_r : r \in \mathbb{N}\}$ is an orthonormal sequence in $L_2(P)$ and $\{\lambda_r : r \in \mathbb{N}\}$ is a sequence of real numbers, non-increasing in absolute value, such that $\sum \lambda_r^2 < \infty$. We will also assume some extra integrability for the functions ϕ_r, namely,

$$\int \phi_r^4 \mathrm{d}P < \infty, \qquad r \in \mathbb{N}. \tag{4.1}$$

Our goal is to obtain rates of convergence in the law of large numbers (3.1) for the function

$$h(x, y) = \sum_{r=1}^{\infty} \lambda_r \phi_r(x) \phi_r(y) \tag{4.2}$$

in terms of the asymptotic behaviour of the eigenvalues λ_r and the eigenfunctions ϕ_r, under assumption (4.1).

The main part of the proof of our result will be to obtain a bound for kernels that consist of a finite sum of terms of the form $\lambda_r \phi_r(x) \phi_r(y)$. We present it separately in the following lemma. For this, it is convenient to introduce the following notation:

$$\zeta^2(R) := \sum_{1 \leqslant r,s \leqslant R} (\lambda_r^2 + \lambda_s^2) \int \phi_r^2 \phi_s^2 \, \mathrm{d}P. \tag{4.3}$$

Lemma 4.1. *Let* $h(x, y) = \sum_{r=1}^{R} \lambda_i \phi_r(x) \phi_r(y)$ *for all* $x, y \in S$ *and for some* $R < \infty$, *with* ϕ_r *and* λ_r *as just described. Then,*

$$\mathbb{E}\delta_2^2(\lambda(\tilde{H}_n), \lambda(H)) \leqslant \frac{\zeta^2(R)}{n} - \frac{2\sum_{r=1}^{R} \lambda_r^2}{n}. \tag{4.4}$$

Proof. As in the first part of the proof of Theorem 3.1, the Lidskii–Hoffman–Wielandt inequality implies

$$\begin{aligned}\delta_2^2(\lambda(\tilde{H}_n), \lambda(H)) &= \delta_2^2(\lambda(\hat{H}_n), \lambda(K)) \\ &\leqslant \|\hat{H}_n - K\|_{\mathrm{HS}}^2 \\ &= \|A_n K A_n - K\|_{\mathrm{HS}}^2, \end{aligned} \tag{4.5}$$

with the notation as in the proof of Theorem 2.1. Let $g_1, \ldots, g_R$ be a (random) orthonormal basis of $\mathbb{R}^R$ of eigenvectors of the (random) symmetric operator E_n, and let $\mu_1, \ldots, \mu_R$ be the corresponding (random) eigenvalues. In this basis, the operator $A_n = (I + E_n)^{1/2}$ has a diagonal matrix with entries $(1 + \mu_j)^{1/2}$, $j = 1, \ldots, R$. Then, letting

$$k_{rs} = \langle K g_r, g_s \rangle, \; r, s = 1, \ldots, R,$$

we have, from (4.5),

$$\begin{aligned}\delta_2^2(\lambda(\tilde{H}_n), \lambda(H)) &\leqslant \|A_n K A_n - K\|_{\mathrm{HS}}^2 \\ &= \sum_{1 \leqslant r,s \leqslant R} k_{rs}^2 [(1+\mu_r)^{\frac{1}{2}}(1+\mu_s)^{\frac{1}{2}} - 1]^2. \end{aligned} \tag{4.6}$$

To bound this expression we use the following elementary inequality:

$$[(1+a)^{\frac{1}{2}}(1+b)^{\frac{1}{2}} - 1]^2 \leqslant a^2 \vee b^2, \; a, b \geqslant -1, \tag{4.7}$$

which follows because either

$$0 \leqslant (1+a)^{\frac{1}{2}}(1+b)^{\frac{1}{2}} - 1 \leqslant (1 + a \vee b)^{\frac{1}{2}}(1 + a \vee b)^{\frac{1}{2}} - 1 = a \vee b$$

or

$$0 \leqslant 1 - (1+a)^{\frac{1}{2}}(1+b)^{\frac{1}{2}} \leqslant 1 - (1 + a \wedge b)^{\frac{1}{2}}(1 + a \wedge b)^{\frac{1}{2}} = -(a \wedge b).$$

Then, using (4.7), we obtain from (4.6) that

$$\begin{aligned}\delta_2^2(\lambda(\tilde{H}_n), \lambda(H)) &\leqslant \sum_{1 \leqslant r,s \leqslant R} k_{rs}^2 [(1+\mu_r)^{\frac{1}{2}}(1+\mu_s)^{\frac{1}{2}} - 1]^2 \\ &\leqslant \sum_{1 \leqslant r,s \leqslant R} k_{rs}^2 (\mu_r^2 \vee \mu_s^2) \\ &\leqslant \sum_{1 \leqslant r,s \leqslant R} k_{rs}^2 (\mu_r^2 + \mu_s^2) \\ &= \|E_n K\|_{\mathrm{HS}}^2 + \|K E_n\|_{\mathrm{HS}}^2 \\ &= 2\|E_n K\|_{\mathrm{HS}}^2, \end{aligned} \tag{4.8}$$

where the last identity follows from the symmetry of K and E_n. Using the matrix representation of E_n and K in the canonical basis, it is clear that

$$\|E_nK\|_{\mathrm{HS}}^2 = \sum_{1\leq r,s\leq R} \lambda_s^2\left[\int \phi_r\phi_s\,\mathrm{d}(P_n-P)\right]^2.$$

Therefore, since $\int \phi_r^2\,\mathrm{d}P = 1$ for all r,

$$\begin{aligned}\mathbb{E}\|E_nK\|_{\mathrm{HS}}^2 &= \sum_{1\leq r,s\leq R} \lambda_s^2\mathbb{E}\left[\int \phi_r\phi_s\,\mathrm{d}(P_n-P)\right]^2 \\ &= \frac{1}{n}\sum_{1\leq r,s\leq R}\lambda_s^2\int(\phi_r\phi_s-\delta_{rs})^2\,\mathrm{d}P \\ &= \frac{\zeta^2(R)}{2n} - \frac{\sum_{r=1}^R \lambda_r^2}{n}.\end{aligned}$$

The lemma follows from this identity and inequality (4.8). □

Theorem 4.2. *If h is as in (4.2), with the eigenfunctions ϕ_r, $r<\infty$, satisfying (4.1), then, for all R, $n\in\mathbb{N}$,*

$$\mathbb{E}\delta_2^2(\lambda(H_n),\lambda(H)) \leq \frac{4}{n}\sum_{1\leq r,s\leq R}(\lambda_r^2+\lambda_s^2+\lambda_r\lambda_s)\int\phi_r^2\phi_s^2\,\mathrm{d}P + 8\sum_{r=R+1}^{\infty}\lambda_r^2. \tag{4.9}$$

In particular,

$$\mathbb{E}\delta_2^2(\lambda(H_n),\lambda(H)) \leq \frac{6}{n}\zeta^2(R) + 8\sum_{r=R+1}^{\infty}\lambda_r^2. \tag{4.9$'$}$$

Proof. Let us consider the following decomposition of $\delta_2^2(\lambda(H_n),\lambda(H))$:

$$\begin{aligned}\tfrac14\delta_2^2(\lambda(H_n),\lambda(H)) \leq{}& \delta_2^2(\lambda(H),\lambda(H_R)) + \delta_2^2(\lambda(H_R),\lambda(\tilde H_{R,n})) \\ &+ \delta_2^2(\lambda(\tilde H_{R,n}),\lambda(H_{R,n})) + \delta_2^2(\lambda(H_{R,n}),\lambda(H_n)),\end{aligned} \tag{4.10}$$

where $\tilde H_{R,n}$ is defined as in (1.2) with h replaced by h_R. Lemma 4.1 already gives

$$\mathbb{E}\delta_2^2(\lambda(H_R),\lambda(\tilde H_{R,n})) \leq \frac{\zeta^2(R)}{n}.$$

By (3.10),

$$\mathbb{E}\delta_2^2(\lambda(\tilde H_{R,n}),\lambda(H_{R,n})) \leq \frac{1}{n}\sum_{1\leq r,s\leq R}\lambda_r\lambda_s\int\phi_r^2\phi_s^2\,\mathrm{d}P.$$

The Lidskii–Hoffman–Wielandt inequality gives, as in the proof of Theorem 3.1, that

$$\mathbb{E}\delta_2^2(\lambda(H_{R,n}), \lambda(H_n)) \leqslant \mathbb{E}\|H_{R,n} - H_n\|_{\mathrm{HS}}^2$$
$$= \mathbb{E}\left[\frac{1}{n^2}\sum_{1\leqslant i\neq j\leqslant n}(h - h_R)^2(X_i, X_j)\right]$$
$$= \frac{n-1}{n}\sum_{r=R+1}^{\infty}\lambda_r^2.$$

Finally, it is clear from the definition and Lidskii–Hoffman–Wielandt that

$$\mathbb{E}\delta_2^2(\lambda(H), \lambda(H_R)) \leqslant \sum_{r=R+1}^{\infty}\lambda_r^2.$$

Plugging these four estimates into inequality (4.10) proves the theorem. □

Next we single out two particular instances of Theorem 4.2 to be used in Section 7.

Corollary 4.3. *Suppose that*

$$\sum_{r,s=1}^{\infty}(\lambda_r^2 + \lambda_s^2)\int_S \phi_r^2\phi_s^2\,\mathrm{d}P < +\infty. \tag{4.11}$$

Then

$$\mathbb{E}\delta_2^2(\lambda(H_n), \lambda(H)) \leqslant \frac{6}{n}\sum_{r,s=1}^{\infty}(\lambda_r^2 + \lambda_s^2)\int_S \phi_r^2\phi_s^2\,\mathrm{d}P. \tag{4.12}$$

Proof. Under condition (4.11) we can let $R \to \infty$ in (4.9′) and obtain (4.12). □

Corollary 4.4. *Suppose that, for some* $\alpha > 0$,

$$\zeta^2(R) = O(R^\alpha) \tag{4.13}$$

and that, for some $\beta > 0$,

$$\lambda_r^2 = O(r^{-1-\beta}). \tag{4.14}$$

Then

$$\mathbb{E}\delta_2^2(\lambda(H_n), \lambda(H)) = O(n^{-\frac{\beta}{\alpha+\beta}}) \tag{4.15}$$

Proof. Under conditions (4.13) and (4.14), the bound in (4.9′) is dominated by $C[R^\alpha/n + R^{-\beta}]$, for some constant $C < \infty$, and this expression is minimized at $R = {}^{1/(\alpha+\beta)}$. □

Given $\alpha > 0$, $\beta > 0$ and $C > 0$, let $\mathscr{H}(\alpha, \beta, C)$ denote the class of symmetric kernels

$$h(x, y) = \sum_{r \in I} \lambda_r \phi_r(x) \phi_r(y)$$

with

$$|\lambda_r|^2 \leqslant C r^{-1-\beta}, \qquad r \in I,$$

and

$$\sum_{r,s=1}^{R} (\lambda_r^2 + \lambda_s^2) \int_S \phi_r^2 \phi_s^2 \, \mathrm{d}P \leqslant C R^{\alpha}, \qquad R \geqslant 1.$$

It follows from the proof of Theorem 4.2 that in fact

$$\sup_{h \in \mathscr{H}(\alpha,\beta,\mathscr{C})} \mathbb{E} \delta_2^2(\lambda(H_n), \lambda(H)) = O(n^{-\beta/(\alpha+\beta)}).$$

5. Asymptotic normality of spectra I: finite-dimensional convergence

In this section we show that the limiting distribution of the spectrum of $\tilde{H}_n$ (and also of H_n) can be expressed in terms of the generalized Brownian bridge associated with P. This is the centred Gaussian process G_P indexed by functions $f \in L_2(P)$ whose covariance is that of P, that is,

$$\mathbb{E} G_P(f) G_P(g) = \int_S fg \, \mathrm{d}P - \int_S f \, \mathrm{d}P \int_S g \, \mathrm{d}P, \qquad f, g \in L_2(P).$$

It will be convenient in this section to index eigenvalues and eigenfunctions of the operator H with kernel h by all integers as follows:

$$h(x, y) = \sum_{r \in \mathbb{Z}} \lambda_r \phi_r(x) \phi_r(y), \tag{5.1}$$

with convergence taking place in $L_2(P^2)$, where

$$\lambda_0 \geqslant \lambda_1 \geqslant \lambda_2 \geqslant \ldots > 0 > \ldots \geqslant \lambda_{-2} \geqslant \lambda_{-1}$$

(we disregard the possible eigenvalue 0 of H, that is, we consider only the restriction of H to the Hilbert space spanned by $\{\phi_r : r \in \mathbb{Z}\}$). We denote by $\{\mu_r\}$ the ordered set of distinct eigenvalues of H other than 0,

$$\mu_0 > \mu_1 > \mu_2 > \ldots > 0 > \ldots > \mu_{-2} > \mu_{-1},$$

and let m_r be the multiplicity of the eigenvalue μ_r, $r \in \mathbb{Z}$. (So, for example, there are m_1 eigenvalues λ_i equal to μ_1, etc.) A blanket assumption that will be in force throughout this section is that the eigenfunctions of H are in $L_4(P)$, that is,

$$\int_S \phi_r^4 \, \mathrm{d}P < +\infty, \qquad r \in \mathbb{Z}.$$

With the above notation, the object of this section becomes that of proving the following theorem about weak convergence of the finite-dimensional distributions of the spectra of H_n and $\tilde{H}_n$ to the corresponding finite-dimensional distributions of a random vector associated with the process G_P.

Theorem 5.1. *Let h be a symmetric kernel in $L_2(P^2)$ with the property that there exists a sequence $R_n \to \infty$ satisfying*

$$\sum_{|r| > R_n} \lambda_r^2 = o(n^{-1}) \tag{5.2}$$

and

$$\sum_{|r| \leq R_n, |s| \leq R_n} \int_S \phi_r^2 \phi_s^2 \, \mathrm{d}P \quad \sum_{|r| \leq R_n, |s| \leq R_n} (\lambda_r^2 + \lambda_s^2) \int_S \phi_r^2 \phi_s^2 \, \mathrm{d}P = o(n). \tag{5.3}$$

Suppose that, moreover,

$$\sum_{r \in \mathbb{Z}} |\lambda_r| \phi_r^2 \in L_2(P). \tag{5.4}$$

Let Δ_r be the set of indices $i \in \mathbb{Z}$ such that $H\phi_i = \mu_r \phi_i$. Let G_P be the generalized Brownian bridge associated with P. Let Γ_r be the Gaussian matrix

$$\Gamma_r := \mu_r(\mathrm{G}_P(\phi_i \phi_j) \colon i, j \in \Delta_r), \qquad r \in \mathbb{Z}. \tag{5.5}$$

Then the finite-dimensional distributions of the sequence

$$n^{1/2}(\lambda^{\uparrow\downarrow}(H_n) - \lambda^{\uparrow\downarrow}(H))$$

converge weakly to the corresponding finite-dimensional distributions of

$$(\oplus_{r<0} \lambda^{\uparrow}(\Gamma_r)) \oplus (\oplus_{r \geq 0} \lambda^{\downarrow}(\Gamma_r)).$$

If condition (5.4) is replaced by

$$\sum_{r \in \mathbb{Z}} \lambda_r \phi_r^2(x) = h(x, x) \in L_2(P), \tag{5.4$'$}$$

then the statement holds with H_n replaced by $\tilde{H}_n$.

Dauxois *et al.* (1982) observed a somewhat similar type of limit behaviour for spectra of empirical covariance operators.

In fact, we will prove a slightly stronger statement (Theorem 5.6). We begin by proving three lemmas about approximation of operators.

Lemma 5.2. *Let A be a symmetric linear operator on $\mathbb{R}^d$ with eigenvalues $\mu_1, \ldots, \mu_k$ of respective multiplicities $m_1, \ldots, m_k$ and eigenspaces $W_1, \ldots, W_k$. Let $P_j(A)$ be the* *orthogonal projector of* $\mathbb{R}^d$ *onto* W_j, $1 \leqslant j \leqslant k$. *For* $r \leqslant k$, *let*

$$\delta_r = \tfrac{1}{2}\min[\min\{|\mu_i - \mu_j| : 1 \leqslant i < j \leqslant r\}, \min\{|\mu_i - \mu_j| : 1 \leqslant i \leqslant r, r+1 \leqslant j \leqslant k\}].$$

Let $0 < \varepsilon \leqslant \delta_r/2$ *and let* B *be another symmetric linear operator with* $\|B\|_{\mathrm{HS}} < \varepsilon$. *Then:*

(i) *the set of eigenvalues of* $A + B$ *partitions into* $r+1$ *subsets* $\Lambda_j(A+B)$, $j = 1, \ldots, r$, *and* R_r *such that*

$$\Lambda_j(A+B) \subset B(\mu_j, \varepsilon) \quad \text{and} \quad \mathrm{dist}(\mathrm{R}_r, \{\mu_1, \ldots, \mu_r\}) > 2\delta_r - \varepsilon,$$

where $B(\mu_j, \varepsilon)$ *denotes the open ball with centre* μ_j *and radius* ε;

(ii) *if* $P_j(A+B)$ *denotes the orthogonal projection onto the direct sum of the eigenspaces of* $A+B$ *with eigenvalues in the cluster* $\Lambda_j(A+B)$, *then*

$$\mathrm{tr}(\mathrm{P}_j(A+B)) = \mathrm{tr}(\mathrm{P}_j(A)), \qquad j = 1, \ldots, r; \tag{5.6}$$

(iii) *if* γ_j *denotes the (positively oriented) circle of radius* δ_r *about* μ_j *(in* $\mathbb{C}$*) and* $R_A(\zeta) := (A - \zeta I)^{-1}$ *is the resolvent of* A, *then*

$$P_j(A+B) = P_j(A) + \frac{1}{2\pi \mathrm{i}} \oint_{\gamma_j} R_A(\zeta) B R_A(\zeta)\, \mathrm{d}\zeta + S_j, \qquad j = 1, \ldots, r, \tag{5.7}$$

where

$$\|S_j\|_{\mathrm{HS}} \leqslant 2\frac{\|B\|_{\mathrm{HS}}^2}{\delta_r^2}, \; j = 1, \ldots, r. \tag{5.8}$$

Proof. (i) and (ii) are direct consequences of the Lidskii–Hoffmann–Wielandt inequality and only (iii) requires proof. γ_j does not intersect the spectrum $\Lambda(A+B)$ of $A+B$ and the intersection of the corresponding disk $B(\mu_j, \delta_r)$ with $\Lambda(A+B)$ is just $\Lambda_j(A+B)$, $j = 1, \ldots, r$. Therefore (Kato 1982, p. 39),

$$P_j(A+B) = -\frac{1}{2\pi \mathrm{i}} \oint_{\gamma_j} R_{A+B}(\zeta)\, \mathrm{d}\zeta, \tag{5.9}$$

where R_{A+B} is the resolvent of $A+B$. For $\zeta \in \gamma_j$,

$$\|R_A(\zeta)\| = \frac{1}{\mathrm{dist}(\zeta, \Lambda(A))} < \frac{1}{\delta_r}$$

(Kato 1982, p. 60), so that

$$\|R_{A+B}(\zeta)B\| < \tfrac{1}{2}$$

and the following representation holds with uniform (in $\zeta \in \gamma_j$) and absolute convergence of the series on the right-hand side:

$$
\begin{aligned}
R_{A+B}(\zeta) &= (A + B - \zeta I)^{-1} = [(\mathrm{A} - \zeta I)(I + R_A(\zeta)B)]^{-1} \\
&= (I + R_A(\zeta)B)^{-1} R_A(\zeta) \\
&= \sum_{n=0}^{\infty} (-1)^n (R_A(\zeta)B)^n R_A(\zeta).
\end{aligned}
$$

Hence,

$$
\begin{aligned}
\|R_{A+B}(\zeta) - R_A(\zeta) + R_A(\zeta) B R_A(\zeta)\|_{\mathrm{HS}} &\leqslant \sum_{n=2}^{\infty} \|R_A(\zeta)\| \, \|R_A(\zeta)B\|_{\mathrm{HS}}^n \\
&\leqslant \sum_{n=2}^{\infty} \|R_A(\zeta)\|^{n+1} \|B\|_{\mathrm{HS}}^n
\end{aligned}
$$

which is dominated by

$$
\sum_{n=2}^{\infty} \frac{\|B\|_{\mathrm{HS}}^n}{\delta_r^{n+1}} = \left(\frac{\|B\|_{\mathrm{HS}}}{\delta_r} \right)^2 (\delta_r - \|B\|_{\mathrm{HS}})^{-1} \leqslant 2 \frac{\|B\|_{\mathrm{HS}}^2}{\delta_r^3}.
$$

The lemma follows upon setting $C(\zeta) = R_{A+B}(\zeta) - R_A(\zeta) + R_A(\zeta) B R_A(\zeta)$ and $S_j = \frac{1}{2\pi \mathrm{i}} \oint_{\gamma_j} C(\zeta) \, \mathrm{d}\zeta$. □

Given an operator C on $\mathbb{R}^d$ and $\varepsilon > 0$, we say that a subset Λ of the spectrum $\lambda(C)$ of C is an ε-cluster of (the spectrum of) C if Λ has diameter smaller than $\varepsilon > 0$ and is at a distance larger than ε from $\lambda(C) \Delta \Lambda$. Then, parts (i) and (ii) of Lemma 5.2 assert that if A is perturbed by B with $\|B\|_{\mathrm{HS}} < \varepsilon$ and if the eigenvalues of A are separated by more than 4ε, then to each eigenvalue μ_r of A there corresponds an ε-cluster of $A + B$ about μ_r such that the dimension of the sum of the eigenspaces of $A + B$ corresponding to the eigenvalues in such a cluster equals the multiplicity m_r of μ_r.

Part (iii) will now be used to 'expand' the restriction of $A + B$ to the orthogonal sum of the eigenspaces corresponding to any ε-cluster of eigenvalues $\Lambda_r(A + B)$ as the sum of the restriction of A to W_r, a 'linear' term in B and 'lower-order' terms.

Lemma 5.3. *Let A and B be symmetric operators on $\mathbb{R}^d$ satisfying the conditions of Lemma 5.2 with $\delta_r \leqslant 1$. Let $e_1, \ldots, e_d$ be an orthonormal basis of $\mathbb{R}^d$ consisting of eigenvectors of A and let Δ_r denote the set of indices i such that $Ae_i = \mu_r e_i$. Let $L_r(B)$ be the symmetric operator of $\mathbb{R}^d$ defined by the equations*

$$
\langle L_r(B) e_i, e_j \rangle = \begin{cases} \langle B e_i, e_j \rangle, & \text{if } i, j \in \Delta_r \\ -\dfrac{\mu_r}{\lambda_j - \mu_r} \langle B e_i, e_j \rangle, & \text{if } i \in \Delta_r \text{ and } j \notin \Delta_r \\ -\dfrac{\mu_r}{\lambda_i - \mu_r} \langle B e_i, e_j \rangle, & \text{if } i \notin \Delta_r \text{ and } j \in \Delta_r \\ 0, & \text{otherwise,} \end{cases} \tag{5.10}
$$

where, for all $1 \leqslant j \leqslant d$, $\lambda_j = \mu_\ell$ *if* $Ae_j = \mu_\ell e_j$. *Then there exists a symmetric operator* C_r *on* $\mathbb{R}^d$ *such that*

$$\|C_r\|_{\mathrm{HS}} \leqslant 8\frac{\|B\|_{\mathrm{HS}}^2}{\delta_r^2} \tag{5.11}$$

and

$$P_r(A+B)(A+B)P_r(A+B) = P_r(A)AP_r(A) + L_r(B) + C_r. \tag{5.12}$$

Proof. Equation (5.7) gives

$$\begin{aligned} P_r(A+B)(A+B)P_r(A+B) = {} & P_r(A)AP_r(A) + P_r(A)BP_r(A) \\ & + \left(\frac{1}{2\pi \mathrm{i}}\oint_{\gamma_r} R_A(\zeta)BR_A(\zeta)\,\mathrm{d}\zeta\right)AP_r(A) \\ & + P_r(A)A\left(\frac{1}{2\pi \mathrm{i}}\oint_{\gamma_r} R_A(\zeta)BR_A(\zeta)\,\mathrm{d}\zeta\right) \\ & + P_r(A)B\left(\frac{1}{2\pi \mathrm{i}}\oint_{\gamma_r} R_A(\zeta)BR_A(\zeta)\,\mathrm{d}\zeta\right) \\ & + \left(\frac{1}{2\pi \mathrm{i}}\oint_{\gamma_r} R_A(\zeta)BR_A(\zeta)\,\mathrm{d}\zeta\right)\mathrm{BP}_r(A) \\ & + \left(\frac{1}{2\pi \mathrm{i}}\oint_{\gamma_r} R_A(\zeta)BR_A(\zeta)\,\mathrm{d}\zeta\right)B\left(\frac{1}{2\pi \mathrm{i}}\oint_{\gamma_r} R_A(\zeta)BR_A(\zeta)\,\mathrm{d}\zeta\right) \\ & + \text{`multiples' of } S_r. \end{aligned} \tag{5.13}$$

Define $L_r(B)$ as the sum of the operators in the second, third and fourth terms of the right-hand side of (5.13) and C_r as the sum of the operators in subsequent terms. With these definitions, equation (5.12) is just (5.13) and we only need to show that $L_r(A)$ can be written as in (5.10), and that C_r satisfies the bound (5.11). In order to compute L_r, we first observe that, by definition and by symmetry of $P_r(A)$,

$$\begin{aligned} \langle P_r(A)BP_r(A)e_i, e_j\rangle &= \langle BP_r(A)e_i, P_r(A)e_j\rangle \\ &= \begin{cases} \langle Be_i, e_j\rangle, & \textit{if } i, j \in \Delta_r \\ 0, & \textit{otherwise}. \end{cases} \end{aligned}$$

Next, if $i \notin \Delta_r$ then, obviously,

$$\left\langle \frac{1}{2\pi i}\oint_{\gamma_r} R_A(\zeta)BR_A(\zeta)AP_r(A)\,\mathrm{d}\zeta AP_r(A)e_i,\, e_j \right\rangle = 0.$$

Now, if $i \in \Delta_r$ and $j \notin \Delta_r$, since

$$\begin{aligned}\langle R_A(\zeta)BR_A(\zeta)AP_r(A)e_i,\, e_j\rangle &= \langle \mathrm{B}R_A(\zeta)AP_r(A)e_i,\, R_A(\overline{\zeta})e_j\rangle \\ &= \frac{\mu_r}{(\mu_r-\zeta)(\lambda_j-\zeta)}\langle Be_i,\, e_j\rangle,\end{aligned}$$

with $\lambda_j \neq \mu_r$ (hence, λ_j lies outside the disk bounded by γ_r), we have

$$\begin{aligned}\left\langle \frac{1}{2\pi i}\oint_{\gamma_r} R_A(\zeta)BR_A(\zeta)AP_r(A)\,\mathrm{d}\zeta AP_r(A)e_i,\, e_j \right\rangle &= \frac{\mu_r\langle Be_i,\, e_j\rangle}{2\pi i}\oint_{\gamma_r}\frac{\mathrm{d}\zeta}{(\mu_r-\zeta)(\lambda_j-\zeta)} \\ &= -\frac{\mu_r}{\lambda_j-\mu_r}\langle Be_i,\, e_j\rangle.\end{aligned}$$

Thus,

$$\left\langle \frac{1}{2\pi i}\oint_{\gamma_r} R_A(\zeta)BR_A(\zeta)AP_r(A)\,\mathrm{d}\zeta AP_r(A)e_i,\, e_j \right\rangle = \begin{cases} 0, & \text{if } i \notin \Delta_r, \\ -\dfrac{\mu_r}{\lambda_j-\mu_r}\langle Be_i,\, e_j\rangle, & \text{if } i \in \Delta_r,\, j \notin \Delta_r. \end{cases}$$

The third component of $L_r(B)$ can be handled in a similar way. The main difference occurs for $i, j \in \Delta_r$. In this case we just observe

$$\begin{aligned}\langle P_r(A)AR_A(\zeta)BR_A(\zeta)e_i,\, e_j\rangle &= \mu_r\langle R_A(\zeta)BR_A(\zeta)e_i,\, e_j\rangle \\ &= \mu_r\langle BR_A(\zeta)e_i,\, R_A(\overline{\zeta})e_j\rangle \\ &= -\mu_r\left\langle \frac{1}{\mu_r-\zeta}Be_i,\, \frac{1}{\mu_r-\overline{\zeta}}e_j \right\rangle \\ &= -\frac{\mu_r}{(\mu_r-\zeta)^2}\langle Be_i,\, e_j\rangle,\end{aligned}$$

whose integral over γ_r is zero. So (5.10) is proved.

Since, as observed above, $\|R_A(\zeta)\| = 1/\delta_r$ for $\zeta \in \gamma_r$, it follows that

$$\left\| \frac{1}{2\pi i}\oint_{\gamma_r} R_A(\zeta)BR_A(\zeta)\,\mathrm{d}\zeta \right\|_{\mathrm{HS}} \leqslant \frac{\|B\|_{\mathrm{HS}}}{\delta_r},$$

showing that the Hilbert–Schmidt norm of each of the operators in the fifth, sixth and seventh terms in (5.13) is dominated by $\|B\|_{\mathrm{HS}}^2/\delta_r^2$ (note that both $\delta_r \leqslant 1$ and $\|B\|_{\mathrm{HS}} \leqslant 1$). The bound (5.8) for S_r gives the same bound for the ten operators in (5.13) containing S_r as a factor. Inequality (5.11) follows. □

We will apply Lemma 5.3 to

$$A = K_{R_n} \text{ and } B = \hat{H}_{R_n,n} - K_{R_n} = (I + E_{R_n,n})^{1/2} K_{R_n} (I + E_{R_n,n})^{1/2} - K_{R_n},$$

in the notation of Section 4. The following lemma will allow us to simplify the linear term in equation (5.12).

Lemma 5.4. *Let* $K, E: \mathbb{R}^d \mapsto \mathbb{R}^d$ *be symmetric operators. Suppose that* $I + E$ *is non-negative definite. Then*

$$\left\| (I+E)^{1/2} K (I+E)^{1/2} - \left(K + \frac{EK + KE}{2} \right) \right\|_{\mathrm{HS}} \leqslant 3\|E\|_{\mathrm{HS}} [\|\mathrm{EK}\|_{\mathrm{HS}} + \|KE\|_{\mathrm{HS}}]. \tag{5.14}$$

Proof. Let

$$R := (I+E)^{1/2} - I - \frac{E}{2}.$$

Then we have

$$(I+E)^{1/2} K (I+E)^{1/2} = K + \frac{EK + KE}{2} + \frac{EKE}{4} + RK + KR + \frac{RKE + EKR}{2} + RKR.$$

It then follows that

$$\left\| (I+E)^{1/2} K (I+E)^{1/2} - \left(K + \frac{EK + KE}{2} \right) \right\|_{\mathrm{HS}} \leqslant \frac{\|EKE\|_{\mathrm{HS}}}{4} + \|RK\|_{\mathrm{HS}} + \|KR\|_{\mathrm{HS}} + \frac{\|RKE\|_{\mathrm{HS}} + \|EKR\|_{\mathrm{HS}}}{2} + \|RKR\|_{\mathrm{HS}}. \tag{5.15}$$

We now prove that

$$\|RK\|_{\mathrm{HS}} \leqslant \|E^2 K\|_{\mathrm{HS}}. \tag{5.16}$$

Indeed, let μ_i, $i = 1, \ldots, d$, be eigenvalues of the operator E, and let g_i, $i = 1, \ldots, d$, be the corresponding orthonormal eigenvectors (recall that E is symmetric and $I + E$ is non-negative definite). Then the eigenvalues of R are

$$(1 + \mu_i)^{1/2} - 1 - \frac{\mu_i}{2}, \qquad i = 1, \ldots, d.$$

Let k_{ij}, $1 \leqslant i, j \leqslant m$, denote the entries of the matrix of the operator K in the basis g_i, $i = 1, \ldots, d$. Then we have

$$\|RK\|_{\mathrm{HS}}^2 \leqslant \sum_{i,j=1}^{d} \left((1 + \mu_i)^{1/2} - 1 - \frac{\mu_i}{2} \right)^2 k_{ij}^2. \tag{5.17}$$

Expression (5.16) will follow from the following elementary inequality:

$$0 \leqslant 1 + \frac{\mu}{2} - (1 + \mu)^{1/2} \leqslant |\mu| \wedge \mu^2, \qquad \mu \geqslant -1. \tag{5.18}$$

To prove (5.18), we note that $(1 + \mu)^{1/2} \leqslant 1 + \mu/2$ for all $\mu \geqslant -1$. For $|\mu| \leqslant \frac{1}{2}$ the inequality

$$1+\frac{\mu}{2}-(1+\mu)^{1/2}\leqslant\frac{\mu^2}{8}\leqslant\frac{|\mu|}{8}$$

follows from the Taylor expansion of the square root. For $\mu \geqslant \frac{1}{2}$, we have

$$1+\frac{\mu}{2}-(1+\mu)^{1/2}\leqslant\frac{\mu}{2}\leqslant\mu^2.$$

Finally, for $-1 \leqslant \mu \leqslant -\frac{1}{2}$,

$$1+\frac{\mu}{2}-(1+\mu)^{1/2}\leqslant 1+\frac{\mu}{2}-(1+\mu)=\frac{|\mu|}{2}\leqslant\mu^2.$$

Using (5.17), (5.18) gives

$$\|RK\|_{\mathrm{HS}}^2\leqslant\sum_{i,j=1}^{d}\mu_i^4k_{ij}^2=\|E^2K\|_{\mathrm{HS}}^2,$$

thus proving inequality (5.16).

Quite similarly, it can be shown, using (5.18), that

$$\|KR\|_{\mathrm{HS}}\leqslant\|KE^2\|_{\mathrm{HS}}, \tag{5.19}$$

$$\|RKE\|_{\mathrm{HS}}\leqslant\|EKE\|_{\mathrm{HS}}, \tag{5.20}$$

$$\|EKR\|_{\mathrm{HS}}\leqslant\|EKE\|_{\mathrm{HS}}, \tag{5.21}$$

$$\|RKR\|_{\mathrm{HS}}\leqslant\|EKE\|_{\mathrm{HS}}, \tag{5.22}$$

and we obtain from (5.15), (5.16), (5.19)–(5.22) that

$$\left\|(I+E)^{1/2}K(I+E)^{1/2}-\left(K+\frac{EK+KE}{2}\right)\right\|_{\mathrm{HS}}\leqslant\frac{9}{4}\|EKE\|_{\mathrm{HS}}+\|KE^2\|_{\mathrm{HS}}+\|E^2K\|_{\mathrm{HS}}. \tag{5.23}$$

Since

$$\|EKE\|_{\mathrm{HS}}\leqslant\frac{1}{2}\|E\|_{\mathrm{HS}}[\|EK\|_{\mathrm{HS}}+\|KE\|_{\mathrm{HS}}],$$

$$\|KE^2\|_{\mathrm{HS}}\leqslant\|E\|_{\mathrm{HS}}\|KE\|_{\mathrm{HS}},$$

and

$$\|E^2K\|_{\mathrm{HS}}\leqslant\|E\|_{\mathrm{HS}}\|EK\|_{\mathrm{HS}},$$

(5.23) implies (5.14). □

We now come back to the kernel h defined by equation (5.1). We assume, without loss of generality, that $\delta_r \leqslant 1$ for all r, where the numbers δ_r are associated with the eigenvalues μ_r of h essentially as in Lemma 5.2, as follows:

$$\delta_r = \begin{cases} \frac{1}{2}\min\{\mu_i - \mu_{i+1} : 0 \leqslant i \leqslant r\} \text{ for } r > 0 \\ \frac{1}{2}\min\{\mu_{i-1} - \mu_i : r \leqslant i \leqslant -2\} \text{ for } r < 0. \end{cases}$$

Slightly modifying some of the notation of Sections 3 and 4, we set

$$h_R(x, y) := \sum_{|r|\leqslant R} \lambda_r \phi_r(x)\phi_r(y),$$

and denote by H_R the integral operator with kernel h_R, and by $H_{R,n}$, $\tilde{H}_{R,n}$ the operators of $\mathbb{R}^n$ whose matrices in the canonical basis are respectively

$$H_{R,n} := \frac{1}{n}((1-\delta_{ij})h_R(X_i, X_j) : 1 \leqslant i, j \leqslant n) \quad \text{and} \quad \tilde{H}_{R,n} := \frac{1}{n}(h_R(X_i, X_j) : 1 \leqslant i, j \leqslant n).$$

The following lemma gives conditions under which the spectrum of H_n and/or the spectrum of $\tilde{H}_n$ can be approximated by that of $H_{R_n,n}$ and/or that of $\tilde{H}_{R_n,n}$, for suitable $R_n \to \infty$.

Lemma 5.5. *Let $R_n \to \infty$ be a sequence satisfying (5.2). We then have:*

(i)

$$\mathbb{E}\|\lambda^{\uparrow\downarrow}(H_n) - \lambda^{\uparrow\downarrow}(H_{R_n,n})\|_{\ell_2}^2 = o(n^{-1});$$

(ii) *if, moreover, $\sum_{r\in\mathbb{Z}}|\lambda_r|\phi_r^2 \in L_2(P)$, then*

$$\mathbb{E}\sup_{R\geqslant 1}\|\lambda^{\uparrow\downarrow}(\tilde{H}_{R,n}) - \lambda^{\uparrow\downarrow}(H_{R,n})\|_{c_0}^2 = o(n^{-1});$$

(iii) *and if, moreover, $\sum_{r\in\mathbb{Z}}\lambda_r\phi_r^2$ converges in $L_2(P)$ to the function $h(x, x)$, then*

$$\mathbb{E}\|\lambda^{\uparrow\downarrow}(\tilde{H}_n) - \lambda^{\uparrow\downarrow}(\tilde{H}_{R_n,n})\|_{\ell_2}^2 = o(n^{-1}).$$

Proof. The three statements are direct consequences of the Lidskii and Lidskii–Hoffman–Wielandt inequalities, the simplest being statement (i). We only prove (ii) and (iii). Lidskii's inequality for $p = \infty$ gives that, for all $R \geqslant 1$,

$$\|\lambda^{\uparrow\downarrow}(\tilde{H}_{R,n}) - \lambda^{\uparrow\downarrow}(H_{R,n})\|_{c_0} \leqslant \|\tilde{H}_{R,n} - H_{R,n}\|$$
$$= n^{-1}\max_{1\leqslant i\leqslant n}|h_R(X_i, X_i)|.$$

Since

$$|h_R(x, x)| = \left|\sum_{|r|\leqslant R}\lambda_r\phi_r^2(x)\right| \leqslant \psi(x),$$

where

$$\psi := \sum_{r\in\mathbb{Z}}|\lambda_r|\phi_r^2 \in L_2(P),$$

we have

$$n^{1/2}\sup_{R\geqslant 1}\|\lambda^{\uparrow\downarrow}(\tilde{H}_{R,n})-\lambda^{\uparrow\downarrow}(H_{R,n})\|_{c_0}\leqslant\frac{\max_{1\leqslant i\leqslant n}\psi(X_i)}{n^{1/2}}.$$

Now, the inequality in (ii) follows from this and the fact that

$$\mathbb{E}\max_{1\leqslant i\leqslant n}\psi^2(X_i)=o(n)$$

because $\mathbb{E}\psi^2(X)<+\infty$.

To prove (iii) we first note that, by the Lidskii–Hoffman–Wielandt inequality,

$$\|\lambda^{\uparrow\downarrow}(\tilde{H}_n)-\lambda^{\uparrow\downarrow}(\tilde{H}_{R_n,n})\|^2_{\ell_2}\leqslant\|\tilde{H}_n-\tilde{H}_{R_n,n}\|^2_{\mathrm{HS}}$$
$$=n^{-2}\sum_{1\leqslant i\neq j\leqslant n}(h-h_{R_n})^2(X_i,X_j)+n^{-2}\sum_{i=1}^{n}(h-h_{R_n})^2(X_i,X_i).$$

Therefore

$$\mathbb{E}\|\lambda^{\downarrow}(\tilde{H}_n)-\lambda^{\downarrow}(\tilde{H}_{R_n,n})\|^2_{\ell_2}\leqslant\frac{n(n-1)}{n^2}\mathbb{E}(h-h_{R_n})^2(X,Y)+n^{-1}\mathbb{E}(h-h_{R_n})^2(X,X)$$
$$=\frac{n(n-1)}{n^2}\sum_{|r|>R_n}\lambda_r^2+n^{-1}\sum_{|r|>R_n,|s|>R_n}\lambda_r\lambda_s\int_S\phi_r^2\phi_s^2\,\mathrm{d}P.$$

Since $\sum_{r\in\mathbb{Z}}\lambda_r\phi_r^2$ converges in $L_2(P)$ to the function $h(x,x)$,

$$\sum_{|r|>R_n,|s|>R_n}\lambda_r\lambda_s\int_S\phi_r^2\phi_s^2\,\mathrm{d}P=\Bigg\|\sum_{|r|>R_n}\lambda_r\phi_r^2\Bigg\|^2_{L_2P}\to 0,$$

showing that the second summand on the right of the last inequality is $o(n^{-1})$. Since, by condition (5.2), the first summand is also $o(n^{-1})$, the proof of (iii) is completed. □

From this point on, we change the definition of the ε-cluster of an operator C in two ways: if Λ is an ε-cluster of C in the sense indicated just before Lemma 5.3, and Λ consists of positive (negative) eigenvalues, then we redefine Λ as a vector whose coordinates are the elements of the set Λ each counted with its multiplicity, arranged in non-increasing (non-decreasing) order. Only strictly positive and strictly negative clusters will be considered.

The statement of the next theorem requires numbering of the clusters of H_n and $\tilde{H}_n$. For $r\geqslant 0$ ($r<0$), an ε-cluster of an operator C is the rth ε-cluster of C, $\Lambda_r^\varepsilon(C)$, if there are exactly r ($-r-1$) ε-clusters of C whose components are larger (smaller) than those of $\Lambda_r^\varepsilon(C)$. $\Lambda_r^\varepsilon(C)$ need not exist.

Theorem 5.6. *Let h be a symmetric kernel in $L_2(P^2)$ with the property that there exists a sequence $R_n\to\infty$ satisfying (5.2) and (5.3). If, moreover, (5.4) holds, then so do the following statements:*

(i) *For all $r\in\mathbb{Z}$, and $\varepsilon_n=cn^{-1/4}$, $c>0$, the probability that the ε_n-cluster $\Lambda_r^{\varepsilon_n}(H_n)$ of*

H_n exists, that it is within ε_n from μ_r and that $\mathrm{tr}(P_r(H_n)) = \mathrm{tr}(P_r(H))$ *tends to 1 as $n \to \infty$.*

(ii) *Let Δ_r be the set of indices $i \in \mathbb{Z}$ such that $H\phi_i = \mu_r \phi_i$. Let G_p be the generalized Brownian bridge associated with P. Let Γ_r be the Gaussian matrix defined in (5.5). Let $M_r := (\mu_r, \ldots, \mu_r)$, where the u_r term occurs d_r times, $r \in \mathbb{Z}$. Then, for all $0 \leqslant R_1 \leqslant R_2 < \infty$,*

$$n^{1/2}\Big(\oplus_{r=R_1}^{R_2}\Lambda_r^{\varepsilon_n}(H_n) - \oplus_{r=R}^{R_2} M_r\Big) \to_{\mathscr{L}} \oplus_{r=R_1}^{R_2}\lambda^{\downarrow}(\Gamma_r) \tag{5.24}$$

and, for all $-\infty < R_1 \leqslant R_2 \leqslant -1$,

$$n^{1/2}\Big(\oplus_{r=R_1}^{R_2}\Lambda_r^{\varepsilon_n}(H_n) - \oplus_{r=R_1}^{R_2} M_r\Big) \to_{\mathscr{L}} \oplus_{r=R_1}^{R_2}\lambda^{\uparrow}(\Gamma_r) \tag{5.24$'$}$$

as random vectors in ℓ_2.

If condition (5.4) is replaced by

$$\sum_{r\in\mathbb{Z}} \lambda_r \phi_r^2(x) = h(x, x) \in L_2(P), \tag{5.4$'$}$$

then statements (i) and (ii) hold with H_n replaced by $\tilde{H}_n$.

Proof. By Lemma 5.5, we can replace H_n and $\tilde{H}_n$ by $\tilde{H}_{R_n,n}$ in all the conclusions above. As in the proof of Theorem 3.1, we set

$$E_n := ((P_n - P)(\phi_r\phi_s) : 0 \leqslant |r|, |s| \leqslant R_n),$$

observe that $I_{2R_n+1} + E_n$ is non-negative definite and also define

$$A_n := (I_{2R_n+1} + E_n)^{1/2}.$$

Letting K_{R_n} be the diagonal matrix with numbers $\lambda_r : |r| \leqslant R_n$, on the diagonal, we further define

$$\hat{H}_{R_n,n} = A_n K_{R_n} A_n.$$

By the same argument as in Section 3 (see the proof of Theorem 3.1) we have

$$\lambda(K_{R_n}) = \lambda(H_{R_n}), \quad \lambda(\hat{H}_{R_n,n}) = \lambda(\tilde{H}_{R_n,n}), \qquad n \in \mathbb{N}. \tag{5.25}$$

As previously mentioned, we will apply Lemma 5.3 with $A = K_{R_n}$ and $B = \hat{H}_{R_n,n} - K_{R_n}$. Thus, we begin by estimating the size of B. The following estimate is contained in the derivation of the bounds (4.6) and (4.8):

$$\|\hat{H}_{R_n,n} - K_{R_n}\|_{\mathrm{HS}}^2 = \|A_n K_{R_n} A_n - K_{R_n}\|_{\mathrm{HS}}^2 \leqslant \|E_n K_{R_n}\|_{\mathrm{HS}}^2 + \|K_{R_n} E_n\|_{\mathrm{HS}}^2. \tag{5.26}$$

Next we observe that

$$\mathbb{E}\|E_n\|_{\mathrm{HS}}^2 = \sum_{|r|\leqslant R_n, |s|\leqslant R_n} \mathbb{E}((P_n - P)(\phi_r\phi_s))^2$$

$$= n^{-1} \sum_{|r|\leqslant R_n, |s|\leqslant R_n} (P(\phi_r^2\phi_s^2) - \delta_{rs}) \tag{5.27}$$

and that, similarly,

$$\mathbb{E}\|E_n K_{R_n}\|_{\mathrm{HS}}^2 = n^{-1} \sum_{|r|\leqslant R_n, |s|\leqslant R_n} \lambda_s^2 (P(\phi_r^2 \phi_s^2) - \delta_{rs}) \tag{5.28}$$

and

$$\mathbb{E}\|K_{R_n} E_n\|_{\mathrm{HS}}^2 = n^{-1} \sum_{|r|\leqslant R_n, |s|\leqslant R_n} \lambda_r^2 (P(\phi_r^2 \phi_s^2) - \delta_{rs}). \tag{5.28$'$}$$

Combining these estimates with (5.26), we obtain, by conditions (5.2) and (5.3),

$$\begin{aligned}
\mathbb{E}\|\hat{H}_{R_n,n} - K_{R_n}\|_{\mathrm{HS}}^2 &\leqslant n^{-1} \sum_{|r|\leqslant R_n, |s|\leqslant R_n} (\lambda_r^2 + \lambda_s^2) \int_S \phi_r^2 \phi_s^2 \, \mathrm{d}P \\
&\leqslant 2 \frac{\sup_{r\in\mathbb{Z}} |\lambda_r|}{n} \left(\sum_{\substack{|r|\leqslant R_n \\ |s|\leqslant R_n}} \int_S \phi_r^2 \phi_s^2 \, \mathrm{d}P \right)^{1/2} \left(\sum_{\substack{|r|\leqslant R_n \\ |s|\leqslant R_n}} (\lambda_r^2 + \lambda_s^2) \int_S \phi_r^2 \phi_s^2 \, \mathrm{d}P \right)^{1/2} \\
&= o(n^{-1/2}).
\end{aligned} \tag{5.29}$$

Then (5.29), Lidskii–Hoffman–Wielandt and the definition of ε_n imply that, for each $r \in \mathbb{Z}$, with probability tending to 1, $\Lambda_r^{\varepsilon_n}(\hat{H}_{R_n,n})$ is well defined and $\mathrm{tr}(P_r(\hat{H}_{R_n,n})) = \mathrm{tr}(P_r(K_{R_n}))$. In particular, conclusion (i) holds for $\hat{H}_{R_n,n}$ under the stated hypotheses. Also, we can apply Lemma 5.3 'in probability' and obtain, by (5.29), that, for each $r \in \mathbb{Z}$,

$$P_r(\hat{H}_{R_n,n}) \hat{H}_{R_n,n} P_r(\hat{H}_{R_n,n}) = P_r(K_{R_n}) K_{R_n} P_r(K_{R_n}) + L_{r,n} + o_{Pr}(n^{-1/2}) \tag{5.30}$$

in the sense of Hilbert–Schmidt norms, where $L_{r,n}$ is given by (5.10) with B replaced by $\hat{H}_{R_n,n} - K_{R_n}$.

Next, we make use of Lemma 5.4 to replace $B = \hat{H}_{R_n,n} - K_{R_n}$ in the definition of $L_{r,n}$ by its linearization $(EK + KE)/2$ (with the corresponding subindices). To this end we just observe that, by (5.27) and (5.28), (5.28$'$) the expected value of the Hilbert–Schmidt norm of the remainder term in this substitution is bounded by

$$\begin{aligned}
&3\mathbb{E}[\|E_n\|_{\mathrm{HS}}(\|E_n K_{R_n}\|_{\mathrm{HS}} + \|K_{R_n} E_n\|_{\mathrm{HS}})] \\
&\quad \leqslant 3\sqrt{2} n^{-1} \left(\sum_{\substack{|r|\leqslant R_n \\ |s|\leqslant R_n}} \int_S \phi_r^2 \phi_s^2 \, \mathrm{d}P \right)^{1/2} \left(\sum_{\substack{|r|\leqslant R_n \\ |s|\leqslant R_n}} (\lambda_r^2 + \lambda_s^2) \int_S \phi_r^2 \phi_s^2 \, \mathrm{d}P \right)^{1/2} = o(n^{-1/2}).
\end{aligned} \tag{5.31}$$

Hence, if $\bar{L}_{r,n}$ is defined via equation (5.10) with B replaced by $(E_n K_{R_n} + K_{R_n} E_n)/2$, since also $\mu_r/(\lambda_j - \mu_r) \leqslant \mu_r/\delta_r$, we have

$$\|\bar{L}_{r,n} - L_{r,n}\|_{\mathrm{HS}} = o_{Pr}(n^{-1/2}). \tag{5.32}$$

Then (5.30)–(5.32) give, by Lidskii–Hoffman–Wielandt, that

$$n^{1/2}(\Lambda_r^{\varepsilon_n}(\hat{H}_{R_n,n}) - \lambda^{\downarrow}(P_r(K_{R_n})K_{R_n}P_r(K_{R_n}) + \bar{L}_{r,n})) \to 0 \text{ in probability.} \tag{5.33}$$

for $r \geqslant 0$, and the same limit but with $\lambda^{\downarrow}$ replaced by $\lambda^{\uparrow}$ for $r < 0$. We show next that we can discard from the matrix $\bar{L}_n$ in (5.33) the entries not in $\Delta_r \times \Delta_r$, by applying Wielandt's inequalities (2.5) and (2.6). Note that the matrix representation of $(E_n K_{R_n} + K_{R_n} E_n)/2$ in the basis $\{e_i\}$ is

$$\frac{1}{2}(E_n K_{R_n} + K_{R_n} E_n) = \left(\frac{\lambda_i + \lambda_j}{2}(P_n - P)(\phi_i \phi_j) : |i|,\, |j| \leqslant R_n\right),$$

and therefore the matrix entries of the operator

$$\overline{\overline{L}}_{r,n} := P_r(K_{R_n}) \frac{E_n K_{R_n} + K_{R_n} E_n}{2} P_r(K_{R_n})$$

are

$$\langle \overline{\overline{L}}_{r,n} e_i,\, e_j\rangle = \begin{cases} \mu_r(P_n - P)(\phi_i\phi_j), & \text{for } i,\, j \in \Delta_r, \\ 0, & \text{otherwise.} \end{cases} \tag{5.34}$$

Let us set

$$A_n := P_r(K_{R_n})K_{R_n}P_r(K_{R_n}) + \overline{\overline{L}}_{r,n} \quad \text{and} \quad B_n := \overline{\overline{L}}_{r,n} - \overline{L}_{r,n}.$$

A_n has a block diagonal matrix such that, by (5.34) and Lidskii–Hoffman–Wieland, all the eigenvalues of its $\Delta_r \times \Delta_r$ block converge in probability to $\mu_r \neq 0$, whereas the $\Delta_r^c \times \Delta_r^c$ block of A_n is 0. Also

$$\|B_n\|_{\mathrm{HS}}^2 = \|\overline{\overline{L}}_{r,n} - \bar{L}_{r,n}\|_{\mathrm{HS}}^2 = 2\sum_{i\in\Delta_r}\sum_{j\notin\Delta_r, |j|\leqslant R_n} \left(\frac{\mu_r + \lambda_j}{2}\right)^2 \frac{\mu_r^2}{(\lambda_j - \mu_r)^2}((P_n - P)(\phi_i\phi_j))^2$$

so that, by (5.29),

$$\begin{aligned} \mathbb{E}\|B_n\|_{\mathrm{HS}}^2 &= \mathbb{E}\|\overline{\overline{L}}_{r,n} - \bar{L}_{r,n}\|_{\mathrm{HS}}^2 \\ &\leqslant \frac{2\mu_r^2}{\delta_r^2}\frac{\sup_{r\in\mathbb{Z}}|\lambda_r|}{n}\left(\sum_{\substack{|r|\leqslant R_n \\ |s|\leqslant R_n}} \int_S \phi_r^2\phi_s^2\, \mathrm{d}P\right)^{1/2}\left(\sum_{\substack{|r|\leqslant R_n \\ |s|\leqslant R_n}} (\lambda_r^2 + \lambda_s^2)\int_S \phi_r^2\phi_s^2\, \mathrm{d}P\right)^{1/2} \\ &= o(n^{-1/2}). \end{aligned}$$

Then, we can apply Wielandt's inequalities (2.5) and (2.6) in probability to $A = A_n$ and $B = B_n$ and obtain

$$\lambda^{\downarrow}(A_n + B_n) - \lambda^{\downarrow}(A_n) = o_{Pr}(n^{-1/2}).$$

This estimate, combined with (5.33), gives

$$n^{1/2}(\Lambda_r^{\varepsilon_n}(\hat{H}_{R_n,n}) - \lambda^{\downarrow}(P_r(K_{R_n})K_{R_n}P_r(K_{R_n}) + \overline{\overline{L}}_{r,n})) \to 0 \text{ in probability} \tag{5.35}$$

for $r \geqslant 0$, and the same limit with $\lambda^{\downarrow}$ replaced by $\lambda^{\uparrow}$ for $r < 0$. But $P_r(K_{R_n})K_{R_n}P_r(K_{R_n})$ being μ_r times the identity on the linear span of e_i, $i \in \Delta_r$, we have

$$\lambda^{\downarrow}(P_r(K_{R_n})K_{R_n}P_r(K_{R_n}) + \overline{\overline{L}}_{r,n}) = M_r + \lambda^{\downarrow}\left(\overline{\overline{L}}_{r,n}\right) \text{ for } r \geqslant 0$$

$$\lambda^{\uparrow}(P_r(K_{R_n})K_{R_n}P_r(K_{R_n}) + \overline{\overline{L}}_{r,n}) = M_r + \lambda^{\uparrow}\left(\overline{\overline{L}}_{r,n}\right) \text{ for } r < 0.$$

Now, the central limit theorem in finite dimensions applied to the vectors $(\langle \overline{\overline{L}}_{r,n} e_i, e_j \rangle :$ $i, j \in \Delta_r)$ (see (5.34)), together with continuity of the mappings $\lambda^{\uparrow}$, $\lambda^{\downarrow}$, shows that

$$n^{1/2}\lambda^{\downarrow}\left(\overline{\overline{L}}_{r,n}\right) \to_{\mathscr{L}} \lambda^{\downarrow}(\Gamma_r)$$

and

$$n^{1/2}\lambda^{\uparrow}\left(\overline{\overline{L}}_{r,n}\right) \to_{\mathscr{L}} \lambda^{\uparrow}(\Gamma_r)$$

as random vectors in ℓ_2. We then conclude from this and (5.35) that

$$n^{1/2}(\Lambda_r^{\varepsilon_n}(\hat{H}_{R_n,n}) - M_r) \to_{\mathscr{L}} \begin{cases} \lambda^{\downarrow}(\Gamma_r), & \text{for } r \geqslant 0 \\ \lambda^{\uparrow}(\Gamma_r), & \text{for } r < 0, \end{cases} \tag{5.36}$$

this being just the limit (5.24) or (5.24′) in the case $R_1 = R_2 = r$. Now (5.24) and (5.24′) with arbitrary $R_1 \leqslant R_2$ (of the same sign) follow by continuity of the direct sum $\oplus$. □

For a compact symmetric operator A, let $\lambda_r(A)$ denote the rth eigenvalue of A in the ordering given below (5.1). With this notation we have:

Corollary 5.7. *If all the eigenvalues of H are simple and hypotheses (5.2), (5.3), and (5.4) of Theorem 5.6 hold, then, with probability tending to one, the first r eigenvalues of H_n are also simple for $|r| < \infty$ and conclusions (5.24), (5.24′) become*

$$n^{1/2}[(\lambda_{R_1}(H_n), \lambda_{R_1+1}(H_n), \ldots, \lambda_{R_2}(H_n)) - (\lambda_{R_1}(H), \lambda_{R_1+1}(H) \ldots, \lambda_{R_2}(H))$$
$$\to_{\mathscr{L}} (\lambda_{R_1}(H)G_P(\phi_{R_1}^2), \lambda_{R_1+1}(H)G_P(\phi_{R_1+1}^2), \ldots, \lambda_{R_2}(H)G_P(\phi_{R_2}^2)).$$

If (5.4′) holds instead of (5.4) and we consider $\tilde{H}_n$ rather than H_n, then the analogous conclusion for $\tilde{H}_n$ holds.

Corollary 5.8. *Suppose that, for some $R \geqslant 1$,*

$$h(x, y) := \sum_{r=1}^{R} \lambda_r \phi_r(x)\phi_r(y),$$

where

$$\lambda_1 \geqslant \lambda_2 \geqslant \ldots \geqslant \lambda_R$$

and

$$\int_S \phi_r^4 \,\mathrm{d}P < +\infty, \qquad r = 1, \ldots, R.$$

Let $\mu_1 > \ldots > \mu_r$ be the ordered sequence of distinct eigenvalues λ_j. Then the sequence

$$\{n^{1/2}(\lambda^{\downarrow}(\tilde{H}_n) - \lambda^{\downarrow}(H))\}_{n=1}^{\infty}$$

converges weakly in $\mathbb{R}^R$ to

$$\lambda^{\downarrow}(\Gamma_1) \oplus \ldots \oplus \lambda^{\downarrow}(\Gamma_r),$$

which, if the eigenvalues λ_j are all distinct, becomes

$$(\lambda_1 G_P(\phi_1^2), \ldots, \lambda_R G_P(\phi_R^2)).$$

Remark 5.9. It is worth mentioning that the condition

$$\int_S \phi_r^4 \,\mathrm{d}P < +\infty, \qquad r = 1, \ldots, R,$$

is sharp, in fact, that it is necessary for the CLT whenever the eigenfunctions $\phi_r : 1 \leqslant r \leqslant R$, have disjoint supports. In this case the matrix E_n is diagonal with eigenvalues $(P_n - P)(\phi_r^2)$, $r = 1, \ldots, R$. It follows that $\hat{H}_n$ is also diagonal with eigenvalues $\lambda_r P_n(\phi_r^2)$, $r = 1, \ldots, R$. Since, by the LLN, $P_n(\phi_r^2) \to P(\phi_r^2) = 1$ as $n \to \infty$ almost everywhere, assuming for simplicity that

$$\lambda_1 > \lambda_2 > \ldots > \lambda_R,$$

we have that, with probability tending to 1,

$$\lambda^{\downarrow}(\tilde{H}_n) = \lambda^{\downarrow}(\hat{H}_n) = (\lambda_r P_n(\phi_r^2) : 1 \leqslant r \leqslant R).$$

It follows that, with probability tending to 1,

$$n^{1/2}(\lambda^{\downarrow}(\tilde{H}_n) - \lambda^{\downarrow}(H)) = (n^{1/2}(P_n - P)(\phi_r^2) : 1 \leqslant r \leqslant R),$$

and, if this sequence converges in distribution to a Gaussian random vector in $\mathbb{R}^R$, then, by the converse CLT, $\int \phi_r^4 \,\mathrm{d}P < +\infty$ for all $r = 1, \ldots, R$.

6. Asymptotic normality of spectra II: tightness in $\ell_2(\mathbb{Z})$ and $c_0(\mathbb{Z})$

The ordered spectrum $\lambda^{\uparrow\downarrow}(H)$ of a Hilbert–Schmidt operator H is a vector in $\ell_2(\mathbb{Z})$ and therefore also in $c_0(\mathbb{Z})$, and so are the ordered spectra $\lambda^{\uparrow\downarrow}(H_n)$ and $\lambda^{\uparrow\downarrow}(\tilde{H}_n)$ (see Section 2). The object of this section is to strengthen the finite-dimensional convergence in distribution proved in Theorem 5.1 to weak convergence of the corresponding spectra as random vectors in $\ell_2(\mathbb{Z})$ and $c_0(\mathbb{Z})$. Concretely, we prove the following two theorems.

Theorem 6.1. *Suppose that*

$$\sum_{r,s\in\mathbb{Z}}(\lambda_r^2+\lambda_s^2)\int_S \phi_r^2\phi_s^2\,\mathrm{d}P<+\infty, \tag{6.1}$$

and that, moreover, there exists a sequence $R_n\to\infty$ such that both

$$\sum_{|r|\leq R_n,|s|\leq R_n}\int_S \phi_r^2\phi_s^2\,\mathrm{d}P=o(n) \tag{6.2}$$

and condition (5.2) hold. Suppose, in addition, that P-a.e.

$$h(x,x)=\sum_{r\in\mathbb{Z}}\lambda_r\phi_r^2(x). \tag{6.3}$$

Then the sequence

$$\{n^{1/2}(\lambda^{\uparrow\downarrow}(\tilde{H}_n)-\lambda^{\uparrow\downarrow}(H))\}_{n=1}^{\infty} \tag{6.4}$$

converges weakly in $\ell_2(\mathbb{Z})$ to the random vector

$$(\oplus_{r<0}\lambda^{\uparrow}(\Gamma_r))\oplus(\oplus_{r\geq 0}\lambda^{\downarrow}(\Gamma_r)), \tag{6.5}$$

where Γ_r, $r\in\mathbb{Z}$, denote the Gaussian matrices defined by (5.5).

Theorem 6.2. *If conditions (5.2), (6.1) and (6.2) hold, then the sequence of random vectors*

$$\{n^{1/2}(\lambda^{\uparrow\downarrow}(H_n)-\lambda^{\uparrow\downarrow}(H))\}_{n=1}^{\infty}$$

converges weakly in $c_0(\mathbb{Z})$ to the random vector given by (6.5).

The following corollary is immediate:

Corollary 6.3. *(a) If conditions (5.2), (6.1)–(6.3) hold, then the sequence of random variables*

$$\{n\delta_2^2(\lambda(\tilde{H}_n),\lambda(H))\}_{n=1}^{\infty}$$

converges weakly to the square of the ℓ_2 norm of the random vector given by (6.5).

(b) If conditions (5.2), (6.1) and (6.2) hold, then the sequence of random variables

$$\{n^{1/2}\|\lambda^{\uparrow\downarrow}(H_n)-\lambda^{\uparrow\downarrow}(H)\|_{c_0}\}_{n=1}^{\infty}$$

converges weakly to the c_0 norm of the random vector (6.5).

Remark 6.4. If the spectrum of H consists of only simple eigenvalues, then the limiting random vector in the previous theorems becomes $(\lambda_r G_P(\phi_r^2))_{r\in\mathbb{Z}}$ (see Corollaries 5.7 and 5.8), and the same change occurs in the corollary.

The proofs are based on the results from Section 5 and the following two lemmas. To formulate them, we introduce some more notation.

If A is a linear operator on $\mathbb{R}^n$ and B is a linear operator on $\mathbb{R}^m$, we let $A \oplus B$ denote the linear operator on $\mathbb{R}^{n+m}$ such that

$$(A \oplus B)(u \oplus v) = (Au) \oplus (Bv).$$

For $u = (u_1, \ldots, u_n) \in \mathbb{R}^n$, if d^+ is the number of non-negative coordinates among $u_1, \ldots, u_n$ and $d^- = n - d^+$, we denote by u_+ the vector of $\mathbb{R}^{d+}$ obtained from u by deleting its negative coordinates, and by u_- the vector in $\mathbb{R}^{d-}$ obtained from u by deleting its non-negative coordinates. Finally, 0_n will denote 0 in $\mathbb{R}^n$.

Lemma 6.5. *Let K_1 and K_2 be symmetric operators respectively on $\mathbb{R}^{m_1}$ and $\mathbb{R}^{m_2}$. Suppose that*

$$\delta := \min\{|\mu| : \mu \in \lambda(K_1)\} - \max\{|\mu| : \mu \in \lambda(K_2)\} > 0. \tag{6.6}$$

Then, for any symmetric operators C_1 on $\mathbb{R}^{m_1}$ and C_2 on $\mathbb{R}^{m_2}$ such that

$$\|C_1 - K_1\|_{\mathrm{HS}} \vee \|C_2 - K_2\|_{\mathrm{HS}} < \frac{\delta}{2}, \tag{6.7}$$

we have

$$\|\lambda^{\uparrow\downarrow}(C_1 \oplus C_2) - \lambda^{\uparrow\downarrow}(K_1 \oplus K_2) - (\lambda^{\uparrow\downarrow}(C_1) - \lambda^{\uparrow\downarrow}(K_1))\|_{\ell_2} \leqslant \|C_2 - K_2\|_{\mathrm{HS}}. \tag{6.8}$$

Proof. Proving (6.8) is obviously equivalent to showing

$$\|\lambda^{\downarrow}(C_1 \oplus C_2) - \lambda^{\downarrow}(K_1 \oplus K_2) - (\lambda^{\downarrow}_+(C_1) - \lambda^{\downarrow}_+(K_1)) \oplus 0_{m_2} \oplus (\lambda^{\downarrow}_-(C_1) - \lambda^{\downarrow}_-(K_1))\|_{\ell_2} \leqslant \|C_2 - K_2\|_{\mathrm{HS}}. \tag{6.8$'$}$$

To prove (6.8′), first we note that, under condition (6.6),

$$\lambda^{\downarrow}(K_1 \oplus K_2) = \lambda^{\downarrow}_+(K_1) \oplus \lambda^{\downarrow}(K_2) \oplus \lambda^{\downarrow}_-(K_1). \tag{6.9}$$

Condition (6.7) and the Lidskii–Hoffman–Wielandt inequality imply that

$$\|\lambda^{\downarrow}(C_1) - \lambda^{\downarrow}(K_1)\|_{\ell_2} \vee \|\lambda^{\downarrow}(C_2) - \lambda^{\downarrow}(K_2)\|_{\ell_2} < \frac{\delta}{2}. \tag{6.10}$$

It follows from (6.9) and (6.10) that

$$d^+(\lambda^{\downarrow}(C_1)) = d^+(\lambda^{\downarrow}(K_1)), \qquad d^-(\lambda^{\downarrow}(C_1)) = d^-(\lambda^{\downarrow}(K_1)).$$

This and (6.10) imply

$$\|\lambda^{\downarrow}_+(C_1) - \lambda^{\downarrow}_+(K_1)\|_{\ell_2} \vee \|\lambda^{\downarrow}_-(C_1) - \lambda^{\downarrow}_-(K_1)\|_{\ell_2} < \frac{\delta}{2},$$

so that, by (6.9),

$$\lambda^{\downarrow}(C_1 \oplus C_2) = \lambda^{\downarrow}_+(C_1) \oplus \lambda^{\downarrow}(C_2) \oplus \lambda^{\downarrow}_-(C_1).$$

Therefore

$$\lambda^{\downarrow}(C_1 \oplus C_2) - \lambda^{\downarrow}(K_1 \oplus K_2) = (\lambda_+^{\downarrow}(C_1) - \lambda_+^{\downarrow}(K_1)) \oplus (\lambda^{\downarrow}(C_2) - \lambda^{\downarrow}(K_2)) \oplus (\lambda_-^{\downarrow}(C_1) - \lambda_-^{\downarrow}(K_1)),$$

and we have

$$\begin{aligned}\lambda^{\downarrow}(C_1 \oplus C_2) - \lambda^{\downarrow}(K_1 \oplus K_2) - (\lambda_+^{\downarrow}(C_1) - \lambda_+^{\downarrow}(K_1)) \oplus 0_{m_2} \oplus (\lambda_-^{\downarrow}(C_1) - \lambda_-^{\downarrow}(K_1))& \\ = 0_{m_1^+} \oplus (\lambda^{\downarrow}(C_2) - \lambda^{\downarrow}(K_2)) \oplus 0_{m_1^-}&,\end{aligned}$$

where

$$m_1^+ := d^+(\lambda^{\downarrow}(K_1)), \qquad m_1^- := d^-(\lambda^{\downarrow}(K_1)).$$

This implies (6.8′) by yet another application of Lidskii–Hoffman–Wielandt inequality:

$$\begin{aligned}\|\lambda^{\downarrow}(C_1 \oplus C_2) - \lambda^{\downarrow}(K_1 \oplus K_2) - (\lambda_+^{\downarrow}(C_1) - \lambda_+^{\downarrow}(K_1)) \oplus 0_{m_2} \oplus (\lambda_-^{\downarrow}(C_1) - \lambda_-^{\downarrow}(K_1))\|_{\ell_2}& \\ = \|0_{m_1^+} \oplus (\lambda^{\downarrow}(C_2) - \lambda^{\downarrow}(K_2)) \oplus 0_{m_1^-}\|_{\ell_2}& \\ = \|\lambda^{\downarrow}(C_2) - \lambda^{\downarrow}(K_2)\|_{\ell_2}& \\ \leqslant \|C_2 - K_2\|_{\mathrm{HS}}.&\end{aligned}$$

□

Let $K_1 : \mathbb{R}^{m_1} \mapsto \mathbb{R}^{m_1}$, $K_2 : \mathbb{R}^{m_2} \mapsto \mathbb{R}^{m_2}$ be symmetric operators and let $K := K_1 \oplus K_2$. We denote by P_1, P_2 the orthogonal projectors of $\mathbb{R}^{m_1+m_2}$ onto $\mathbb{R}^{m_1}$, $\mathbb{R}^{m_2}$, respectively; and by I, I_1, I_2 the identity operators on $\mathbb{R}^{m_1}$, $\mathbb{R}^{m_2}$, $\mathbb{R}^{m_1+m_2}$, respectively. Let $E : \mathbb{R}^{m_1+m_2} \mapsto \mathbb{R}^{m_1+m_2}$ be a symmetric operator, and let us set $E_1 := P_1 E P_1$ and $E_2 := P_2 E P_2$.

Lemma 6.6. *Assuming $I + E$ to be non-negative definite and setting*

$$C := (I+E)^{1/2} K (I+E)^{1/2},$$

$$C_1 := (I_1 + E_1)^{1/2} K_1 (I_1 + E_1)^{1/2}, \qquad C_2 := (I_2 + E_2)^{1/2} K_2 (I_2 + E_2)^{1/2},$$

we have

$$\begin{aligned}\|C - (C_1 \oplus C_2)\|_{\mathrm{HS}} \leqslant \tfrac{1}{2}[\|(P_1 E P_2 + P_2 E P_1) K\|_{\mathrm{HS}} + \|K(P_1 E P_2 + P_2 E P_1)\|_{\mathrm{HS}}]& \\ + \ 6\|E\|_{\mathrm{HS}}[\|EK\|_{\mathrm{HS}} + \|KE\|_{\mathrm{HS}}].&\end{aligned}$$

Proof. Setting

$$\tilde{E} := E_1 \oplus E_2,$$

we have

$$C_1 \oplus C_2 = (I + \tilde{E})^{1/2} K (I + \tilde{E})^{1/2}. \tag{6.11}$$

By Lemma 5.4,

$$\left\|(I+E)^{1/2} K (I+E)^{1/2} - \left(\mathrm{K} + \frac{EK + KE}{2}\right)\right\|_{\mathrm{HS}} \leqslant 3\|E\|_{\mathrm{HS}}[\|EK\|_{\mathrm{HS}} + \|KE\|_{\mathrm{HS}}]. \tag{6.12}$$

Similarly,

$$\left\|(I+\tilde{E})^{1/2}K(I+\tilde{E})^{1/2}-\left(K+\frac{\tilde{E}K+K\tilde{E}}{2}\right)\right\|_{\mathrm{HS}} \leqslant 3\|\tilde{E}\|_{\mathrm{HS}}[\|\tilde{E}K\|_{\mathrm{HS}}+\|K\tilde{E}\|_{\mathrm{HS}}], \quad (6.13)$$

and, since

$$\|\tilde{E}\|_{\mathrm{HS}} \leqslant \|E\|_{\mathrm{HS}}, \qquad \|K\tilde{E}\|_{\mathrm{HS}} \leqslant \|KE\|_{\mathrm{HS}}, \qquad \|\tilde{E}K\|_{\mathrm{HS}} \leqslant \|EK\|_{\mathrm{HS}},$$

(6.13) implies

$$\left\|(I+\tilde{E})^{1/2}K(I+\tilde{E})^{1/2}-\left(K+\frac{\tilde{E}K+K\tilde{E}}{2}\right)\right\|_{\mathrm{HS}} \leqslant 3\|E\|_{\mathrm{HS}}[\|EK\|_{\mathrm{HS}}+\|KE\|_{\mathrm{HS}}]. \quad (6.14)$$

It follows from (6.12) and (6.14) that

$$\|C-(C_1\oplus C_2)\|_{\mathrm{HS}} = \|(I+E)^{1/2}K(I+E)^{1/2}-(I+\tilde{E})^{1/2}K(I+\tilde{E})^{1/2}\|_{\mathrm{HS}}$$
$$\leqslant \frac{1}{2}[\|(E-\tilde{E})K\|_{\mathrm{HS}}+\|K(E-\tilde{E})\|_{\mathrm{HS}}]+6\|E\|_{\mathrm{HS}}[\|EK\|_{\mathrm{HS}}+\|KE\|_{\mathrm{HS}}],$$

which is the bound in the lemma. □

Proof of Theorem 6.1. The proof of the theorem for general kernels h is completely analogous to the proof for non-negative definite kernels, but the notation it requires is more complicated. So, in order to simplify notation, we assume in what follows that $h(x, y)=\sum_{r=1}^{\infty}\lambda_r\phi_r(x)\phi_r(y)$, $\lambda_1 \geqslant \lambda_2 \geqslant \ldots \geqslant 0$ (instead of $r\in\mathbb{Z}$). We continue using the notation introduced before Lemma 5.5 for the operators H_R, $H_{R,n}$, $\tilde{H}_{R,n}$, as well as the notation in the proof of Theorem 5.6 for the operators K_R and $\hat{H}_{R,n}$, applied now to the non-negative definite case, and recall that

$$\lambda(K_R)=\lambda(H_R),\ \lambda(\hat{H}_{R,n})=\lambda(\tilde{H}_{R,n}), \qquad R\geqslant 1,\ n\in\mathbb{N} \quad (6.15)$$

(see the proof of Theorem 3.1 and (5.25)). We also set

$$E_{R,n} := ((P_n-P)(\phi_r\phi_s) : 1\leqslant r,\, s\leqslant R), \qquad A_{R,n} := (I_R+E_{R,n})^{1/2}$$

for all $R>0$, although will keep the notation E_n and A_n for $E_{R_n,n}$ and $A_{R_n,n}$, as in the proof of Theorem 5.6. And we introduce the following notation: for $1\leqslant R_1\leqslant R_2$ and $n\in\mathbb{N}$, we set

$$E_{R_1,R_2,n} := ((P_n-P)(\phi_r\phi_s) : R_1<r,\, s\leqslant R_2),$$
$$A_{R_1,R_2,n} := (I_{R_2-R_1}+E_{R_1,R_2,n})^{1/2}$$

(note that $I_{R_2-R_1}+E_{R_1,R_2,n}$ is a non-negative definite matrix) and

$$\hat{H}_{R_1,R_2,n} := A_{R_1,R_2,n}K_{R_1,R_2}A_{R_1,R_2,n},$$

where K_{R_1,R_2} is the diagonal matrix with entries λ_r, $R_1<r\leqslant R_2$, on its diagonal.

For the proof, first we observe that the hypotheses imply that we can replace H and $\tilde{H}_n$ by H_{R_n} and $\tilde{H}_{R_n,n}$: conditions (5.2), (6.1) and (6.3) imply, by Lemma 5.5(iii), that

$$\mathbb{E}\|\lambda^{\downarrow}(\tilde{H}_n)-\lambda^{\downarrow}(\tilde{H}_{R_n,n})\|_{\ell_2}^2 = o(n^{-1}), \quad (6.16)$$

and the Lidskii–Hoffman–Wielandt inequality together with (5.2) gives

$$\|\lambda^{\downarrow}(H) - \lambda^{\downarrow}(H_{R_n})\|_{\ell_2}^2 \leqslant \|H - H_{R_n}\|_{\mathrm{HS}}^2 = \sum_{r > R_n} \lambda_r^2 = o(n^{-1}). \tag{6.17}$$

The theorem will follow from the finite-dimensional result in Section 5 after 'decoupling' the convergence to infinity of n and R in the sequence $n^{1/2}(\lambda^{\downarrow}(\tilde{H}_{R_n,n}) - \lambda^{\downarrow}(H_{R_n}))$.

Our first objective is now to show that

$$\lim_{R\to\infty} \limsup_{n\to\infty} n^{1/2}\mathbb{E}\|\hat{H}_{R_n,n} - (\hat{H}_{R,n} \oplus \hat{H}_{R,R_n,n})\|_{\mathrm{HS}} = 0. \tag{6.18}$$

To this end, given $R \geqslant 1$ and $R_n > R$, we define P_R and P_{R,R_n} as the orthogonal projectors of the space $\mathbb{R}^{R_n}$ onto $\mathbb{R}^R$ and onto $\mathbb{R}^{R_n}$ given respectively by the equations

$$P_R(u_r : r \leqslant R_n) = (u_r : r \leqslant R)$$

and

$$P_{R,R_n}(u_r : r \leqslant R_n) = (u_r : R < r \leqslant R_n).$$

It follows from Lemma 6.6 that for all n large enough, so that $R_n > R$,

$$\begin{aligned}\|\hat{H}_{R_n,n} - (\hat{H}_{R,n} \oplus \hat{H}_{R,R_n,n})\|_{\mathrm{HS}} \leqslant\ & \tfrac{1}{2}[\|(P_R E_n P_{R,R_n} + P_{R,R_n} E_n P_R)K_{R_n}\|_{\mathrm{HS}} \\ & + \|K_{R_n}(P_R E_n P_{R,R_n} + P_{R,R_n} E_n P_R)\|_{\mathrm{HS}}] \\ & + 6\|E_n\|_{\mathrm{HS}}[\|E_n K_{R_n}\|_{\mathrm{HS}} + \|K_{R_n} E_n\|_{\mathrm{HS}}].\end{aligned} \tag{6.19}$$

Then, regarding the first summand in (6.19), we have

$$\begin{aligned}\mathbb{E}\|(P_R E_n P_{R,R_n} + P_{R,R_n} E_n P_R)K_{R_n}\|_{\mathrm{HS}}^2 &= \mathbb{E}\left(\sum_{r\leqslant R, R < s\leqslant R_n} + \sum_{R < r\leqslant R_n, s\leqslant R}\right)((P_n - P)(\phi_r\phi_s))^2\lambda_s^2 \\ &= n^{-1}\left(\sum_{r\leqslant R R < s\leqslant R_n} + \sum_{R < r\leqslant R_n, s\leqslant R}\right)\lambda_s^2(P(\phi_r^2\phi_s^2) - \delta_{rs}).\end{aligned}$$

Similarly, we have, for the second, that

$$\begin{aligned}&\mathbb{E}\|K_{R_n}(P_R E_n P_{R,R_n} + P_{R,R_n} E_n P_R)\|_{\mathrm{HS}}^2 \\ &\qquad = n^{-1}\left(\sum_{r\leqslant R, R < s\leqslant R_n} + \sum_{R < r\leqslant R_n, s\leqslant R}\right)\lambda_r^2(P(\phi_r^2\phi_s^2) - \delta_{rs}).\end{aligned}$$

Therefore, by Hölder and condition (6.1), we have

$$\lim_{R\to\infty}\limsup_{n\to\infty} n^{1/2}\mathbb{E}\tfrac{1}{2}[\|(P_R E_n P_{R,R_n} + P_{R,R_n} E_n P_R)K_{R_n}\|_{\mathrm{HS}} + \|K_{R_n}(P_R E_n P_{R,R_n} + P_{R,R_n} E_n P_R)\|_{\mathrm{HS}}]$$

$$\leqslant \lim_{R\to\infty}\limsup_{n\to\infty} n^{-1/2}\left(\frac{1}{2}\left(\sum_{r\leqslant R, R<s\leqslant R_n} + \sum_{R<r\leqslant R_n, s\leqslant R}\right)(\lambda_r^2+\lambda_s^2)(P(\phi_r^2\phi_s^2)-\delta_{rs})\right)^{1/2}$$

$$= 0. \tag{6.20}$$

Under conditions (6.1) and (6.2), it follows from (5.27)–(5.28′) that

$$\mathbb{E}\|E_n\|_{\mathrm{HS}}^2 = o(1), \qquad \mathbb{E}\|E_n K_{R_n}\|_{\mathrm{HS}}^2 = O(n^{-1}) \quad \text{and} \quad \mathbb{E}\|K_{R_n}E_n\|_{\mathrm{HS}}^2 = O(n^{-1}),$$

and therefore,

$$\mathbb{E}(\|\mathrm{E}_n\|_{\mathrm{HS}}[\|E_n K_{R_n}\|_{\mathrm{HS}} + \|K_{R_n}E_n\|_{\mathrm{HS}}]) = o(n^{-1/2}). \tag{6.21}$$

Combining (6.19), (6.20) and (6.21) gives (6.18).

Equation (6.18) and the Lidskii–Hoffman–Wielandt inequality yield

$$\lim_{R\to\infty}\limsup_{n\to\infty} n^{1/2}\mathbb{E}\|\lambda^{\downarrow}(\hat{H},\ _{R_n,n}) - \lambda^{\downarrow}(\hat{H}_{R,n}\oplus\hat{H}_{R,R_n,n})\|_{\ell_2} = 0. \tag{6.22}$$

The last step in the proof will then consist of showing that, for all $\varepsilon > 0$,

$$\lim_{\substack{R\to\infty\\ R\in\mathscr{R}}}\limsup_{n\to\infty} \Pr\{n^{1/2}\|\lambda^{\downarrow}(\hat{H}_{R,n}\oplus\hat{H}_{R,R_n,n}) - \lambda^{\downarrow}(K_{R_n}) - (\lambda^{\downarrow}(\hat{H}_{R,n}) - \lambda^{\downarrow}(K_R))\|_{\ell_2} \geqslant \varepsilon\} = 0, \tag{6.23}$$

where $\mathscr{R} = \{r : \lambda_r \neq \lambda_{r+1}\}$. For this, we apply Lemma 6.5 to the operators

$$K_1 := K_R, \qquad K_2 := K_{R,R_n}, \qquad K := K_{R_n} = K_R \oplus K_{R,R_n},$$

$$C_1 := \hat{H}_{R,n}, \qquad C_2 = \hat{H}_{R,R_n,n}.$$

By the definition of $\mathscr{R}$, for every $R \in \mathscr{R}$ we have

$$\delta_R := \min\{|\mu| : \mu \in \lambda(K_R)\} - \max\{|\mu| : | \in \lambda(K_{R,R_n})\} > 0,$$

so that condition (6.6) holds. Using bounds similar to (4.6) and (4.8), we obtain

$$\|\hat{H}_{R,n} - K_R\|_{\mathrm{HS}}^2 \leqslant 2\|E_{R,n}K_R\|_{\mathrm{HS}}^2,$$

which, by (5.5), implies (see (5.28))

$$\mathbb{E}\|\hat{H}_{R,n} - K_R\|_{\mathrm{HS}}^2 \leqslant \sum_{r,s\leqslant R}(\lambda_r^2+\lambda_s^2)\mathbb{E}((P_n-P)(\phi_r\phi_s))^2$$

$$= n^{-1}\sum_{r,s\leqslant R}(\lambda_r^2+\lambda_s^2)(P(\phi_r^2\phi_s^2)-\delta_{rs})^2 = O(n^{-1}). \tag{6.24}$$

Similarly,

$$\mathbb{E}\|\hat{H}_{R,R_n,n} - K_{R,R_n}\|_{\mathrm{HS}}^2 \leqslant \sum_{R<r\leqslant R_n, R<s\leqslant R_n} (\lambda_r^2 + \lambda_s^2)\mathbb{E}((P_n - P)(\phi_r\phi_s))^2$$
$$= n^{-1} \sum_{R<r\leqslant R_n, R<s\leqslant R_n} (\lambda_r^2 + \lambda_s^2)(P(\phi_r^2\phi_s^2) - \delta_{rs})^2$$
$$= O(n^{-1}). \tag{6.25}$$

In particular, this implies that, for all $R \geqslant 1$,

$$\lim_{n\to\infty} \Pr\{\|\hat{H}_{R,n} - K_R\|_{\mathrm{HS}} \geqslant \delta\} = 0 \tag{6.26}$$

and

$$\lim_{n\to\infty} \Pr\{\|\hat{H}_{R,R_n,n} - K_{R,R_n}\|_{\mathrm{HS}} \geqslant \delta\} = 0. \tag{6.27}$$

Equation (6.25) also gives that, for all $\varepsilon > 0$,

$$\lim_{R\to\infty} \limsup_{n\to\infty} \Pr\{n^{1/2}\|\hat{H}_{R,R_n,n} - K_{R,R_n}\|_{\mathrm{HS}} \geqslant \varepsilon\} = 0. \tag{6.28}$$

By Lemma 6.5, on the event

$$\{\|\hat{H}_{R,n} - K_{R,n}\|_{\mathrm{HS}} \vee \|\hat{H}_{R,R_n,n} - K_{R,R_n}\|_{\mathrm{HS}} < \delta_R\},$$

we have

$$n^{1/2}\|\lambda^{\downarrow}(\hat{H}_{R,n} \oplus \hat{H}_{R,R_n,n}) - \lambda^{\downarrow}(K_{R_n}) - (\lambda^{\downarrow}(\hat{H}_{R,n}) - \lambda^{\downarrow}(K_R))\|_{\ell_2} \leqslant n^{1/2}\|\hat{H}_{R,R_n,n} - K_{R,R_n}\|_{\mathrm{HS}}. \tag{6.29}$$

Now (6.23) easily follows from (6.29), (6.28), (6.26) and (6.27).

Combining (6.22) and (6.23), we obtain that, for all $\varepsilon > 0$,

$$\lim_{\substack{R\to\infty\\ R\in\mathscr{R}}} \limsup_{n\to\infty} \Pr\{n^{1/2}\|\lambda^{\downarrow}(\hat{H}_{R_n,n}) - \lambda^{\downarrow}(K_{R_n}) - (\lambda^{\downarrow}(\hat{H}_{R,n}) - \lambda^{\downarrow}(K_R))\|_{\ell_2} \geqslant \varepsilon\} = 0,$$

which, in view of (6.15), implies

$$\lim_{\substack{R\to\infty\\ R\in\mathscr{R}}} \limsup_{n\to\infty} \Pr\{n^{1/2}\|\lambda^{\downarrow}(\tilde{H}_{R_n,n}) - \lambda^{\downarrow}(H_{R_n}) - (\lambda^{\downarrow}(\tilde{H}_{R,n}) - \lambda^{\downarrow}(H_R))\|_{\ell_2} \geqslant \varepsilon\} = 0. \tag{6.30}$$

Since, for all $R < \infty$, $R \in \mathscr{R}$, the sequence $\{n^{1/2}(\lambda^{\downarrow}(\tilde{H}_{R,n}) - \lambda^{\downarrow}(H_R))\}_{n=1}^{\infty}$ converges in law as $n \to \infty$ to $\oplus_{\ell\leqslant S(R)}\lambda^{\downarrow}(\Gamma_\ell)$ for an appropriate $S(R) < \infty$ by Theorem 5.1 (note that (6.1) implies (5.4)), and since these random vectors (in ℓ_2) converge a.s. to $\oplus_{r\in\mathbb{N}}\lambda^{\downarrow}(\Gamma_r)$ as $R \to \infty$ (as $S(R) \to \infty$ with R), it follows from (6.30) that the sequence

$$\{n^{1/2}(\lambda^{\downarrow}(\tilde{H}_{R_n,n}) - \lambda^{\downarrow}(H_{R_n}))\}_{n=1}^{\infty}$$

converges weakly in ℓ_2 to $\oplus_{r\in\mathbb{N}}\lambda^{\downarrow}(\Gamma_r)$. Together with (6.15) and (6.16), this completes the proof. □

Proof of Theorem 6.2. As in the previous proof, we assume that h is a non-negative defnite kernel since the proof in the general case is the same up to formal, obvious changes. By Lemma 5.5(ii),

$$\sup_{R\geq 1}\|\lambda^{\downarrow}(\tilde{H}_{R,n})-\lambda^{\downarrow}(H_{R,n})\|_{c_0}=o_p(n^{-1/2}). \tag{6.31}$$

It then follows that

$$\|\lambda^{\downarrow}(\tilde{H}_{R_n,n})-\lambda^{\downarrow}(H_{R_n,n})\|_{c_0}=o_p(n^{-1/2}). \tag{6.32}$$

Now (6.31) and (6.32) along with (6.30) (whose proof above does not make use of condition (6.3)) imply

$$\lim_{\substack{R\to\infty\\ R\in\mathscr{R}}}\limsup_{n\to\infty}\Pr\{n^{1/2}\|\lambda^{\downarrow}(H_{R_n,n})-\lambda^{\downarrow}(H_{R_n})-(\lambda^{\downarrow}(H_{R,n})-\lambda^{\downarrow}(H_R))\|_{c_0}\geq\varepsilon\}=0. \tag{6.33}$$

Now, the result follows as in the argument following (6.30) in the previous proof, but invoking the part of Theorem 5.1 corresponding to the operators H_n. □

7. Some examples

Whereas it is easy to decide whether the LLN in Section 3 applies for a given kernel (one just checks if $\mathbb{E}h^2(X, Y)<\infty$), it is not necessarily straightforward to verify the hypotheses ensuring the validity of the bounds on the speed of convergence in the LLN (Section 4) or of the distributional limit theorems for the spectrum of H_n in Sections 5 and 6. In this section we outline two techniques, one based on integrablity and smoothness of the kernel and the other more specific to kernels arising from partial differential operators.

7.1. Kernel integrability and smoothness

It is known that smoothness of the kernel $h(x, y)$ has a direct bearing on the rate of decrease of the eigenvalues: for instance (Gohberg and Krein, 1968, §10, Propositions 3 and 4), if h defined on $[0, 1]\times[0, 1]$ is non-negative definite, Hilbert–Schmidt and ℓ times differentiable in square mean (see the reference cited for the definition: this is weaker than h being pointwise differentiable with continuous partial derivatives of order ℓ), with ℓth derivative being Hilbert–Schmidt, then the eigenvalues of h for Lebesgue measure decrease at least as fast as $n^{-\ell-1/2}$. So, in the following proposition we will concentrate on an easy estimate for the function $\zeta^2(R)$ appearing in Theorem 4.2. See the beginning of Section 4 for notation and conventions.

Proposition 7.1. *Let S be a locally compact space with Borel probability measure P. Suppose that the mapping*

$$S\times S\times S\times S\ni(x, y; u, v)\mapsto\int_S h(z, x)h(z, y)h(z, u)h(z, v)\,\mathrm{d}P(z)$$

is continuous, that

$$\int_S\int_S\int_S h^2(z, x)h^2(z, y)\,\mathrm{d}P(x)\,\mathrm{d}P(y)\,\mathrm{d}P(z) < +\infty, \tag{7.1}$$

and that, for some $\beta > 0$,

$$\lambda_r^2 \asymp r^{-1-\beta}. \tag{7.2}$$

Then

$$\mathbb{E}\delta_2^2(\lambda(H_n), \lambda(H)) = O(n^{-\beta/(2\beta+1)}). \tag{7.3}$$

Proof. Let K be the integral operator on $L_2(P^2) := L_2(S \times S, \mathscr{S} \otimes \mathscr{S}, P \times P)$ defined as

$$Kg(x, y) = \int_S\int_S k(x, y; u, v)g(u, v)\,\mathrm{d}P(u)\,\mathrm{d}P(v), \qquad x, y \in S,\ g \in L_2(P^2),$$

where

$$k(x, y; u, v) := \int_S h(z, x)h(z, y)h(z, u)h(z, v)\,\mathrm{d}P(z), \qquad x, y, u, v \in S.$$

It is easy to check that K is non-negative definite. Then, condition (7.1) and continuity of the kernel k imply that K is a nuclear operator with trace

$$\mathrm{tr}(K) = \int_S\int_S k(x, y; x, y)\,\mathrm{d}P(x)\,\mathrm{d}P(y) = \int_S\int_S\int_S h^2(z, x)h^2(z, y)\,\mathrm{d}P(x)\,\mathrm{d}P(y)\,\mathrm{d}P(z).$$

In order to estimate $\zeta^2(R)$ (recall its definition from (4.3)), we note

$$\begin{aligned}
&\lambda_r^2\lambda_s^2\phi_r^2(z)\phi_s^2(z)\\
&\quad = (H\phi_r(z))^2(H\phi_s(z))^2\\
&\quad = \int_S h(z, x)\phi_r(x)\,\mathrm{d}P(x)\int_S h(z, u)\phi_r(u)\,\mathrm{d}P(u)\int_S h(z, y)\phi_s(y)\,\mathrm{d}P(u)\int_S h(z, v)\phi_s(v)\,\mathrm{d}P(v)\\
&\quad = \int_S\int_S\int_S\int_S h(z, x)h(z, y)h(z, u)h(z, v)\phi_r(x)\phi_s(y)\phi_r(u)\phi_s(v)\,\mathrm{d}P(x)\,\mathrm{d}P(y)\,\mathrm{d}P(u)\,\mathrm{d}P(v).
\end{aligned}$$

Therefore

$$\begin{aligned}
&\lambda_r^2\lambda_s^2\int_S \phi_r^2(z)\phi_s^2(z)\,\mathrm{d}P(z)\\
&\quad = \int_S\int_S\int_S\int_S k(x, y; u, v)(\phi_r \otimes \phi_s)(x, y)(\phi_r \otimes \phi_s)(u, v)\,\mathrm{d}P(x)\,\mathrm{d}P(y)\,\mathrm{d}P(u)\,\mathrm{d}P(v)\\
&\quad = \langle K(\phi_r \otimes \phi_s), \phi_r \otimes \phi_s\rangle_{L_2(P^2)},
\end{aligned}$$

which, since $\phi_r \otimes \phi_s$, $r, s \in \mathbb{N}$, is an orthonormal system in $L_2(P^2)$, gives

$$\zeta^2(R) = 2 \sum_{1 \leq r,s \leq R} \lambda_r^2 \int_S \phi_r^2 \phi_s^2 \, \mathrm{d}P$$

$$= 2 \sum_{1 \leq r,s \leq R} \lambda_s^{-2} \langle K(\phi_r \otimes \phi_s), \phi_r \otimes \phi_s \rangle_{L_2(P^2)}$$

$$\leq 2 \Big(\min_{1 \leq s \leq R} |\lambda_s| \Big)^{-2} \mathrm{tr}(K).$$

Hence, it follows from conditions (7.1) and (7.2) that

$$\zeta^2(R) = O(R^{(1+\beta)}).$$

This and (7.2) give the estimate (7.3) by Corollary 4.4 applied with $\alpha = 1 + \beta$. □

For example, if $h(x, y) = (1 + xy - |x - y|)/2$, $-1 \leq x, y \leq 1$, which is the Green function for the Dirichlet problem $-u'' = f$, $u(1) = u(-1) = 0$, then $\lambda_r = 4\pi^{-2} r^{-2}$ and the previous proposition gives $\mathbb{E}\delta_2^2(\lambda(H_n), \lambda(H)) = O(n^{-3/7})$.

7.2. Heat kernels and the estimation of spectra of differential operators

Our aim here is to show that heat kernels and, less often, also Green functions of certain strictly elliptic differential operators on bounded domains, the Laplacian on a compact Riemannian manifold and Schrödinger operators satisfy the conditions of some of the theorems in Sections 4 and 6 so that, if these kernels are known, Monte Carlo techniques should be of value in the approximate evaluation of the spectra of these operators. (The fact that the Green function or the heat kernel have to be known does indeed hinder the usefulness of the results; on the other hand, one only has to be able to simulate $h(X_i, X_j)$ at random points X_i, and this may conceivably require less than full knowledge of h; at any rate, we only show a way to check the conditions of our theorems and do not claim to have answers to any numerical questions regarding solutions of partial differential equations.) Even though our immediate interest lies in the concrete type of operators just mentioned, we place ourselves in a more abstract setting. This whole subsection is inspired by Davies' (1989) book on heat kernels.

Let $\{\phi_r : r \geq 1\}$ be an orthonormal system in $L_2(P)$ and let $0 < \mu_1 \leq \mu_2 \leq \ldots$ be a sequence of real numbers such that, for all $t > 0$,

$$\sum_{r=1}^{\infty} \mathrm{e}^{-\mu_r t} < +\infty. \tag{7.4}$$

The map $\phi_r \mapsto L\phi_r := \mu_r \phi_r$, $r \geq 1$, extends to a linear operator L, on the linear span $\mathscr{D}_L$ of the functions ϕ_r, $r \geq 1$ (in general, L is unbounded). Define

$$h_t(x, y) := \sum_{r=1}^{\infty} \mathrm{e}^{-\mu_r t} \phi_r(x) \phi_r(y). \tag{7.5}$$

Under condition (7.4) the series (7.5) converges in $L_2(P^2)$ for all $t > 0$. We also assume that the kernel h_t is defined by (7.5) on the diagonal, that is,

$$h_t(x, x) := \sum_{r=1}^{\infty} \mathrm{e}^{-\mu_r t} \phi_r^2(x).$$

Then

$$\sum_{r=1}^{\infty} \mathrm{e}^{-\mu_r t} |\phi_r(x)\phi_r(y)| \leqslant \left(\sum_{r=1}^{\infty} \mathrm{e}^{-\mu_r t} |\phi_r(x)|^2 \right)^{1/2} \left(\sum_{r=1}^{\infty} \mathrm{e}^{-\mu_r t} |\phi_r(y)|^2 \right)^{1/2}$$
$$= h_t(x, x)^{1/2} h_t(y, y)^{1/2},$$

which means that the series (7.5) converges for all $x, y \in S$ such that $h_t(x, x) < +\infty$ and $h_t(y, y) < +\infty$.

The kernel h_t defines a bounded operator H_t from $L_2(P)$ into $L_2(P)$. Moreover, $\{H_t : t > 0\}$ is a semigroup of bounded operators in $L_2(P)$ (since $H_{t+s} = H_t H_s$ for all $t > 0$, $s > 0$) and

$$\tau(t) := \mathrm{tr}(H_t) = \int_S h_t(x, x) P(\mathrm{d}x) = \sum_{r=1}^{\infty} \mathrm{e}^{-\mu_r t} < +\infty, \tag{7.6}$$

$$\|H_t\|_{\mathrm{HS}} = \left(\int_{S \times S} h_t^2(x, y) P(\mathrm{d}x) P(\mathrm{d}y) \right)^{1/2} = \left(\sum_{r=1}^{\infty} \mathrm{e}^{-2\mu_r t} \right)^{1/2} < +\infty. \tag{7.7}$$

In fact, H_t can be viewed as the operator e^{-Lt}. We call the function $(t, x, y) \mapsto h_t(x, y)$ from $(0, \infty) \times S \times S$ into $\mathbb{R}^1$ the *heat kernel* of the operator L (since this is an abstract version of the heat kernel of a differential operator). Most of the information we will need on regularity of heat kernels can be found in Davies (1989).

A warning on notation: whereas H_t or H_s will denote a 'heat operator' as just defined, we will still use the notation H_n for the random operator defined by (1.3); no confusion should arise from this slight inconsistency.

Let ν be a measure on $[0, +\infty)$ such that $\nu([0, t]) < +\infty$ for all $t > 0$ and $\nu(\mathbb{R}) \neq 0$. Assume that the Laplace transform of ν,

$$\varphi(\mu) := \int_0^{+\infty} \mathrm{e}^{-\mu t} \nu(\mathrm{d}t) \tag{7.8}$$

exists for all $\mu \geqslant \mu_1$. Let $\lambda_r := \varphi(\mu_r)$. Clearly, $\lambda_r > 0$ for all $r \geqslant 1$. The condition

$$\int_0^{+\infty} \int_0^{+\infty} \tau(t + s)\, \mathrm{d}\nu(t)\, \mathrm{d}\nu(s) < +\infty \tag{7.9}$$

implies that

$$\sum_{r=1}^{\infty}\lambda_r^2 = \sum_{r=1}^{\infty}\int_0^{+\infty}\int_0^{+\infty} \mathrm{e}^{-\mu_r(t+s)}\,\mathrm{d}\nu(t)\,\mathrm{d}\nu(s)$$

$$= \int_0^{+\infty}\int_0^{+\infty} \tau(t+s)\,\mathrm{d}\nu(t)\,\mathrm{d}\nu(s) < +\infty, \tag{7.10}$$

which means that the operator H from $L_2(P)$ into $L_2(P)$ with kernel

$$h(x,\ y) := \sum_{r=1}^{\infty}\lambda_r\phi_r(x)\phi_r(y)$$

is Hilbert–Schmidt. This operator can be and will be viewed as $H = \varphi(L)$. The kernel h is well defined as a function in $L_2(P\times P)$. In what follows, we assume for convenience that

$$h(x,\ x) := \sum_{r=1}^{\infty}\lambda_r\phi_r^2(x), \qquad x\in S.$$

Then, $h(x,\ y)$ is also well defined at least for all $x,\ y\in S$ such that $h(x,\ x) < +\infty$ and $h(y,\ y) < +\infty$, since

$$\sum_{r=1}^{\infty}\lambda_r|\phi_r(x)||\phi_r(y)| \leqslant \left(\sum_{r=1}^{\infty}\lambda_r\phi_r^2(x)\right)^{1/2}\left(\sum_{r=1}^{\infty}\lambda_r\phi_r^2(y)\right)^{1/2} = h^{1/2}(x,\ x)h^{1/2}(y,\ y) < +\infty.$$

Suppose now that

$$\int_0^{\infty}\|H_t\|_{\mathrm{HS}}\,\mathrm{d}\nu(t) < +\infty. \tag{7.11}$$

Then

$$H = \int_0^{+\infty} H_t\,\mathrm{d}\nu(t) \tag{7.12}$$

and its kernel h has the representation

$$h(x,\ y) = \int_0^{+\infty} h_t(x,\ y)\,\mathrm{d}\nu(t). \tag{7.13}$$

Note that condition (7.11) implies (7.8) (since $\|H_t\|_{\mathrm{HS}} \geqslant \mathrm{e}^{-\mu_1 t}$ by (7.7)) and (7.10) (since $\mathrm{tr}(H_{t+s}) = \mathrm{tr}(H_t H_s) \leqslant \|H_t\|_{\mathrm{HS}}\|H_s\|_{\mathrm{HS}}$).

Define

$$\gamma(t) := \int_0^{+\infty}\int_0^{+\infty}\int_S h_t(x;\ x)h_{t+u+v}(x;\ x)P(\mathrm{d}x)\,\mathrm{d}\nu(u)\,\mathrm{d}\nu(v), \qquad t>0. \tag{7.14}$$

Theorem 7.2. *(a) If $\gamma(0+) < \infty$, then*

$$\mathbb{E}\delta_2^2(\lambda(H_n),\ \lambda(H)) = O(n^{-1}). \tag{7.15}$$

(b) Suppose that, for some $\rho > 0$,

$$\gamma(t) = O(t^{-\rho}) \qquad \text{as } t \to 0; \tag{7.16}$$

for some $a > 0$,

$$\tau(t) \asymp t^{-a} \qquad \text{as } t \to 0; \tag{7.17}$$

and, for some $b > a$,

$$\varphi(\mu) = O(\mu^{-b/2}) \qquad \text{as } \mu \to \infty. \tag{7.18}$$

Then

$$\mathbb{E}\delta_2^2(\lambda(H_n), \lambda(H)) = O(n^{-(b-a)/(b-a+\rho)}). \tag{7.19}$$

To prove the theorem we need the following lemma (which is in fact a weak version of known Tauberian theorems).

Lemma 7.3. *Let* F *be a non-decreasing function on* $[0, +\infty)$ *with* $F(0) = 0$. *Suppose that, for all* $t > 0$,

$$f(t) := \int_0^{+\infty} \mathrm{e}^{-tu}\, \mathrm{d}F(u) < +\infty.$$

Then,

(*i*) $$f(0+) < +\infty \Leftrightarrow F(+\infty) < +\infty.$$

For any $\rho > 0$,

(*ii*) $$F(u) = O(u^{\rho}) \text{ as } u \to \infty \Leftrightarrow f(t) = O(t^{-\rho}) \text{ as } t \to 0,$$

and

(*iii*) $$F(u) \asymp u^{\rho} \text{ as } u \to \infty \Leftrightarrow f(t) \asymp t^{-\rho} \text{ as } t \to 0.$$

Proof (sketch). Indeed, we have, for all $u > 0$,

$$f(t) \geqslant \int_0^u \mathrm{e}^{-vt}\, \mathrm{d}F(v) \geqslant \mathrm{e}^{-ut} F(u),$$

which implies (with $t = 1/u$)

$$F(u) \leqslant \mathrm{e} f(1/u), \qquad u > 0. \tag{7.20}$$

On the other hand, if $F(u) = o(\mathrm{e}^{tu})$ as $u \to \infty$ for all $t > 0$, then, by integration by parts,

$$f(t) = \int_0^{+\infty} F(u) t \mathrm{e}^{-ut}\, \mathrm{d}u = \int_0^{+\infty} F(v/t) \mathrm{e}^{-v}\, \mathrm{d}v. \tag{7.21}$$

The lemma follows rather trivially from (7.20) and (7.21). We just check (iii). If $F(u) \asymp u^{\rho}$, then (7.21) implies $f(t) = O(t^{-\rho})$ and (7.20) implies $f(t) \geqslant ct^{-\rho}$ for some constant $c > 0$, so $f(t) \asymp t^{-\rho}$ holds. If, on the other hand, we have $f(t) \asymp t^{-\rho}$, then, by (7.20), $F(u) = O(u^{\rho})$. Using (7.21), we obtain that, for some constants $c > 0$, $C > 0$ and for all $R > 0$,

$$ct^{-\rho} \leqslant f(t) \leqslant F(R/t) + \int_R^{+\infty} F(v/t)\mathrm{e}^{-v}\,\mathrm{d}v \leqslant F(R/t) + Ct^{-\rho}\int_R^{+\infty} v^{\rho}\mathrm{e}^{-v}\,\mathrm{d}v.$$

Choosing R large enough, so that

$$\int_R^{+\infty} v^{\rho}\mathrm{e}^{-v}\mathrm{d}v < \frac{c}{2C},$$

the previous inequality gives that, for all $t > 0$,

$$F(R/t) \geqslant \frac{c}{2}t^{-\rho}.$$

It follows that $F(u) \asymp u^{\rho}$, and (iii) is proved. □

Proof of Theorem 7.2. To prove part (a), we define

$$G(u) := \frac{1}{2}\sum_{r,s:\mu_r+\mu_s\leqslant u} (\lambda_r^2 + \lambda_s^2)P(\phi_r^2\phi_s^2), \qquad u > 0,$$

and observe

$$\begin{aligned}
\gamma(t) &= \int_0^{+\infty}\int_0^{+\infty}\int_S \sum_{r=1}^{\infty} \mathrm{e}^{-\mu_r t}\phi_r^2(x) \sum_{s=1}^{\infty} \mathrm{e}^{-\mu_s(t+u+v)}\phi_s^2(x)\,\mathrm{d}P(x)\,\mathrm{d}\nu(u)\,\mathrm{d}\nu(v) \\
&= \sum_{r=1}^{\infty}\sum_{s=1}^{\infty}\int_S \mathrm{e}^{-(\mu_r+\mu_s)t}\phi_r^2(x)\phi_s^2(x)P(\mathrm{d}x)\int_0^{+\infty}\int_0^{+\infty} \mathrm{e}^{-\mu_s u}\mathrm{e}^{-\mu_s v}\,\mathrm{d}\nu(u)\,\mathrm{d}\nu(v) \\
&= \sum_{r=1}^{\infty}\sum_{s=1}^{\infty} \lambda_s^2 \mathrm{e}^{-(\mu_r+\mu_s)t}\int_S \phi_r^2(x)\phi_s^2(x)\,\mathrm{d}P(x) \\
&= \frac{1}{2}\sum_{r=1}^{\infty}\sum_{s=1}^{\infty}(\lambda_r^2 + \lambda_s^2)\mathrm{e}^{-(\mu_r+\mu_s)t}P(\phi_r^2\phi_s^2) \\
&= \int_0^{+\infty} \mathrm{e}^{-tu}\,\mathrm{d}G(u). \qquad (7.22)
\end{aligned}$$

Then the condition $\gamma(0+) < \infty$ and the representation (7.22) imply, by Lemma 7.3(i), that

$$\sum_{r,s=1}^{\infty}(\lambda_r^2 + \lambda_s^2)P(\phi_r^2\phi_s^2) = G(+\infty) < +\infty,$$

which gives (7.15) by Corollary 4.3.

To prove part (b), we will apply Corollary 4.4 and, once more, Lemma 7.3. By the representation (7.22) for γ, condition (7.16) and Lemma 7.3 imply

$$G(u) = O(u^{\rho}) \qquad \text{as } u \to \infty. \qquad (7.23)$$

Setting

$$F(u) := \sum_{r=1}^{\infty} I_{\{\mu_r \leq u\}} \qquad \text{as } u > 0,$$

we have, by (7.6),

$$\tau(t) = \mathrm{tr}(\mathrm{H}_t) = \int_0^{+\infty} \mathrm{e}^{-ut}\, \mathrm{d}F(u).$$

Then, condition (7.17) implies, by Lemma 7.3(iii), that

$$F(u) \asymp u^a \qquad \text{as } u \to +\infty.$$

$\{\mu_r\}$ being a non-decreasing sequence, it follows easily that

$$\mu_r \asymp r^{1/a} \qquad \text{as } r \to \infty. \tag{7.24}$$

Setting $\alpha := \rho a^{-1}$, it follows from (7.23) and (7.24), that

$$\sum_{1 \leq r,s \leq R} (\lambda_r^2 + \lambda_s^2) P(\phi_r^2 \phi_s^2) \leq \sum_{r,s:\mu_r+\mu_s \leq CR^{1/a}} (\lambda_r^2 + \lambda_s^2) P(\phi_r^2 \phi_s^2) = O(R^\alpha), \tag{7.25}$$

for some constant $C > 0$. On the other hand, by (7.18) and (7.24)

$$\lambda_r^2 = \varphi^2(\mu_r) = O(r^{-b/a}) = O(r^{-1-\beta}), \tag{7.26}$$

where $\beta := ba^{-1} - 1$. Expressions (7.25) and (7.26) are all that is needed in order to conclude, by Corollary 4.4, that

$$\mathbb{E}\delta_2^2(\lambda(H_n), \lambda(H)) = O(n^{-\beta/(\alpha+\beta)}) = O(n^{-(b-a)/(b-a+\rho)}),$$

proving (7.19). □

Next we give an example of a CLT for the spectrum of H. More refined results can be obtained, but the one whose proof we present suffices for our purposes. In order to prove the CLT for the spectrum of H, it is convenient to introduce a new function:

$$\psi(t) := \int_S h_t^2(x, x)\, \mathrm{d}P(x), \qquad t > 0.$$

Theorem 7.4. *Let H be defined by (7.12) and let H_n and $\tilde{H}_n$ be the random operators corresponding to H given respectively by equations (1.3) and (1.2). Suppose $\psi(t) < \infty$ for all $t > 0$,*

$$\int_0^{\infty} \psi^{1/2}(t)\, \mathrm{d}\nu(t) < \infty \tag{7.27}$$

and

$$\varphi(\mu) \left(= \int_0^{\infty} \mathrm{e}^{-t\mu}\, \mathrm{d}\nu(t) \right) = O(\mathrm{e}^{-b\mu}) \qquad \text{as } \mu \to +\infty. \tag{7.28}$$

Then the finite-dimensional distributions of the sequences

$$\{n^{1/2}(\lambda^{\downarrow}(H_n) - \lambda^{\downarrow}(H)\}_{n=1}^{\infty} \quad and \quad \{n^{1/2}(\lambda^{\downarrow}(\tilde{H}_n) - \lambda^{\downarrow}(H)\}_{n=1}^{\infty} \tag{7.29}$$

converge in law to the corresponding finite-dimensional distributions of the random vector $\oplus_{r=1}^{\infty}\lambda^{\downarrow}(\Gamma_r)$*, defined by equation (5.5) for the operator H. If, moreover,* $\gamma(0+) < \infty$ *then the sequences (7.29) converge weakly in* ℓ_2.

Proof. Condition (7.27) gives, by the generalized Minkowski inequality, that

$$h(x, x) = \sum_{r=1}^{\infty} \lambda_r \phi_r^2(x) \in L_2(P)$$

which is conditions (5.4) and (5.4′) in Theorem 5.1. The finiteness of $\varphi(t)$ implies that, for all $\varepsilon > 0$,

$$\begin{aligned}\sum_{r,s=1}^{R} P(\phi_r^2\phi_s^2) &\leqslant e^{2\varepsilon\mu_R} \sum_{r,s=1}^{R} P(e^{-\varepsilon\mu_r}\phi_r^2 e^{-\varepsilon\mu_s}\phi_s^2) \\ &\leqslant e^{2\varepsilon\mu_R} P\left(\sum_{r=1}^{\infty} e^{-\varepsilon\mu_r}\phi_r^2 \sum_{s=1}^{\infty} e^{-\varepsilon\mu_s}\phi_s^2\right) \\ &= e^{2\varepsilon\mu_R} \int_S h_\varepsilon^2(x, x) P(dx) \\ &= O\big(e^{2\varepsilon\mu_R}\big) \qquad \text{as } R \to \infty. \end{aligned} \tag{7.30}$$

Since the eigenvalues λ_r are uniformly bounded, it follows from (7.30) that

$$\sum_{1 \leqslant r,s \leqslant R} (\lambda_r^2 + \lambda_s^2) P(\phi_r^2\phi_s^2) \sum_{1 \leqslant r,s \leqslant R} P(\phi_r^2\phi_s^2) = O(e^{4\varepsilon\mu_R}) \qquad \text{as } R \to \infty. \tag{7.31}$$

This will allow us to verify condition (5.3) in Theorem 5.1, and we prepare next for condition (5.2). To this effect we begin by noting that, for all $r \geqslant 1$,

$$r e^{-\varepsilon\mu_r} \leqslant \sum_{k=1}^{\infty} e^{-\varepsilon\mu_k} = \int_S h_\varepsilon(x, x) P(dx) \leqslant \psi^{1/2}(\varepsilon),$$

which implies

$$\mu_r \geqslant \frac{\log r}{\varepsilon} - \frac{0.5 \log \psi(\varepsilon)}{\varepsilon} \tag{7.32}$$

for all $r \geqslant 1$. Condition (7.28) implies

$$\sum_{r \geqslant R+1} \lambda_r^2 = \sum_{r \geqslant R+1} \lambda_r^2 \leqslant C \sum_{r \geqslant R+1} e^{-2b\mu_r},$$

for some constant $C > 0$ and for all $R \geqslant 1$. Combining this with (7.32) gives that, for all $\varepsilon > b$,

$$\sum_{r\geqslant R+1}\lambda_r^2 \leqslant C\mathrm{e}^{-b\mu_{R+1}}\sum_{r\geqslant 1}\mathrm{e}^{-b\mu_r} \leqslant C\exp\{0.5\log\psi(\varepsilon)/\varepsilon\}\mathrm{e}^{-b\mu_{R+1}}\sum_{r\geqslant 1} r^{-b/\varepsilon} = C(\varepsilon)\mathrm{e}^{-b\mu_{R+1}}. \tag{7.33}$$

Taking $R = R_n$ such that

$$\exp\{\mu_{R_n}b/2\} \leqslant n < \exp\{\mu_{R_n+1}b/2\},$$

in (7.30) (where we can take $\varepsilon < b/8$) and in (7.33) yields conditions (5.3) and (5.2) in Theorem 5.1. Now Theorem 5.1 gives finite-dimensional convergence of the sequences (7.29). If $\gamma(0+) < \infty$ then the proof of Theorem 7.2(a), shows that condition (6.1) in Theorem 6.1 holds; on the other hand, condition (6.2) follows from (7.30) with the same R_n as above. Therefore, convergence in ℓ_2 of the sequences (7.29) follows from Theorem 6.1. □

With some extra care, the following can be proved using Lemma 7.3.

Theorem 7.5. *Suppose that condition (7.27) holds, that*

$$\gamma(t)\psi(t) = O(t^{-\rho}) \qquad \text{as } t \to 0$$

for some $\rho > 0$, *that condition (7.17) holds for some* $a > 0$ *and that condition (7.18) holds for some* $b > \rho + a$. *Then the finite-dimensional distributions of the sequences (7.29) converge in law to the corresponding finite-dimensional distributions of the random vector* $\oplus_{r=1}^{\infty}\lambda^{\downarrow}(\Gamma_r)$, *defined by equation (5.5) for the operator H. If, moreover,* $\gamma(0+) < \infty$ *then the sequences (7.29) converge weakly in* ℓ_2.

Next we show how the above theorems allow us to estimate the eigenvalues of the operator L by means of random matrices associated to its heat kernel. Consider the family of sequences of random matrices

$$\left\{\tilde{H}_n^s := \left(\frac{h_s(X_i, X_j)}{n}\right)_{i,j=1}^{n} : n \geqslant 1\right\} \tag{7.34}$$

for some $s > 0$, and let H_n^s denote the matrix obtained from $\tilde{H}_n^s$ by deleting the diagonal. It is easily seen that $\tilde{H}_n^s$ is a non-negative definite matrix, so that $\lambda_r^{\downarrow}(\tilde{H}_n^s) \geqslant 0$ for all $r \geqslant 1$. Define

$$\mu_{r,n}^s := -\frac{1}{s}\log\lambda_r^{\downarrow}(\tilde{H}_n^s), \qquad r \geqslant 1. \tag{7.35}$$

The next corollary shows that, for all $s > 0$, $\mu_{r,n}^s$ provides an asymptotic estimator of the eigenvalue μ_r of the operator L for all $r \geqslant 1$.

Corollary 7.6. *Suppose that* $\psi(t) < +\infty$ *for all* $t > 0$. *Then, for all* $s > 0$, *the finite dimensional distributions of the sequence of random vectors*

$$\{(n^{1/2}(\mu_{r,n}^s - \mu_r) : r \geqslant 1)\}_{n=1}^{\infty}$$

converge weakly to the corresponding finite-dimensional distributions of the random vector $s^{-1} \oplus_{\ell=1}^{\infty} \lambda^{\downarrow}(\Gamma_\ell)$, *where* Γ_ℓ *is the matrix* $(G_P(\phi_i \phi_j) : i, j \in \Delta_\ell)$ *and* Δ_ℓ *is the set of indices corresponding to the* ℓ*th largest eigenvalue of* H_s.

Proof. Take $\nu = \delta_s$ for some $s > 0$, so that $\varphi(\mu) = \mathrm{e}^{-\mu s}$ and $h = h_s$. Theorem 7.4 then implies that the sequence

$$\{n^{1/2}(\lambda_r^{\downarrow}(\tilde{H}_n^s) - \mathrm{e}^{-\mu_r s})\}_{n=1}^{\infty}$$

converges weakly to $\oplus \mathrm{e}^{-s\mu_\ell}\tilde{\Gamma}_\ell$. Since

$$\begin{aligned} \mu_{r,n}^s &= -\frac{1}{s}\log(\lambda_r^{\downarrow}(\tilde{H}_n^s)) \\ &= \mu_r - \frac{1}{s}\log\left(1 + \frac{\lambda_r^{\downarrow}(\tilde{H}_n^s) - \mathrm{e}^{-\mu_r s}}{\mathrm{e}^{-\mu_r s}}\right) \\ &= \mu_r - s^{-1}\frac{\lambda_r^{\downarrow}(\tilde{H}_n^s) - \mathrm{e}^{-s\mu_r}}{\mathrm{e}^{-\mu_r s}} + o\left(\frac{\lambda_r^{\downarrow}(\tilde{H}_n^s) - \mathrm{e}^{-s\mu_r}}{\mathrm{e}^{-s\mu_r}}\right), \end{aligned}$$

the result follows. □

It can also be proved as a consequence of Theorem 7.2 that, if in addition to the hypotheses of the previous corollary, $h_{2s}(x, x)$ is uniformly bounded for some $s > 0$ and condition (7.17) holds for some $a > 0$, then, for all $\epsilon > 0$,

$$\mathbb{E}\sup_{r \geqslant 1} \mathrm{e}^{-2\mu_r s}(\mu_{r,n}^s - \mu_r)^2 = O(n^{-1+\epsilon}). \tag{7.36}$$

The eigenvalues of L can also be estimated by the eigenvalues of random matrices associated with Green functions as kernels. This is more difficult to justify from our results because Green functions in general are much less well behaved than heat kernels. Consider the operator $G_{\alpha,s} := (L + sI)^{-\alpha}$ for some $s > 0$ and $\alpha > 0$. It can be represented as $G_{\alpha,s} = \varphi(L)$ with

$$\varphi(\mu) := (\mu + s)^{-\alpha} = \Gamma(\alpha)^{-1}\int_0^{+\infty} u^{\alpha-1}\mathrm{e}^{-u(\mu+s)}\,\mathrm{d}u = \int_0^{+\infty}\mathrm{e}^{-\mu u}\,\mathrm{d}\nu(u),$$

where

$$\mathrm{d}\nu(u) := \Gamma(\alpha)^{-1}u^{\alpha-1}\mathrm{e}^{-su}\,\mathrm{d}u.$$

Assuming that

$$\sum_{r=1}^{\infty}(s + \mu_r)^{-2\alpha} = \int_0^{+\infty}\int_0^{+\infty}\tau(u+v)u^{\alpha-1}v^{\alpha-1}\mathrm{e}^{-s(u+v)}\,\mathrm{d}u\,\mathrm{d}v < +\infty,$$

the operator $G_{\alpha,s}$ is Hilbert–Schmidt with kernel

$$g_{\alpha,s}(x, y) := \sum_{r=1}^{\infty} \lambda_r \phi_r(x) \phi_r(y),$$

where $\lambda_r := (s + \mu_r)^{-\alpha}$; $g_{\alpha,s}$ is a Green function of L. Setting

$$\tilde{G}_n^{\alpha,s} = \left(\frac{g_{\alpha,s}(X_i, X_j)}{n} \right)_{i,j=1}^{n}$$

and letting $G_n^{\alpha,s}$ denote the matrix obtained from $\tilde{G}_n^{\alpha,s}$ by deleting the diagonal, we have the following consequence of Theorem 7.2:

Corollary 7.7. *If, for some $\rho > 0$ and $a < 2\alpha$, conditions (7.16) and (7.17) hold, then*

$$\mathbb{E}\delta_2^2(\lambda(G_n^{\alpha,s}), \lambda(G_{\alpha,s})) = O(n^{-(2\alpha-a)/(2\alpha-a+\rho)}).$$

It is not difficult to also extract a CLT for $\lambda(G_n^{\alpha,s})$ from the above results.

We now apply these corollaries to some specific examples of differential operators.

Strictly elliptic operators on bounded regions

We take S to be a bounded open connected subset of $\mathbb{R}^d$ and let P be the uniform distribution on S. Consider the operator

$$L := -\sum_{i,j=1}^{d} \frac{\partial}{\partial x_i} \left\{ a_{ij}(x) \frac{\partial}{\partial x_j} \right\}$$

with Dirichlet boundary condition, where the functions $x \mapsto a(x) := (a_{ij}(x))_{i,j=1}^{\mathrm{d}}$ are locally integrable in S with values in the set of non-negative real symmetric matrices and satisfy $\alpha \leqslant a(x) \leqslant A$ for some constants $\alpha > 0$ and $A < \infty$ and for all $x \in S$ (see, for example, Davies, 1989, p. 10). Then (see Davies 1989, Corollary 3.2.8) the semigroup e^{-Lt}, $t > 0$, has a kernel h_t, which satisfies the bound

$$0 \leqslant h_t(x, y) \leqslant Ct^{-d/2} \exp\{-(x-y)^2/4(1+\delta)At\}$$

for all $t > 0$, $x, y \in S$, $\delta \in (0, 1)$, and a constant $C = C(\delta, \alpha) < \infty$. It follows from this bound that $\psi(t) = \int_G h_t^2(x, x)\,\mathrm{d}P(x) < +\infty$ for all $t > 0$. Therefore, Corollary 7.6 applies, thus concluding that the finite-dimensional distributions of the sequence

$$\{(n^{1/2}(\mu_{r,n}^s - \mu_r) : r \geqslant 1)\}_{n=1}^{\infty}$$

converge weakly to the corresponding finite-dimensional distributions of the random vector $s^{-1} \oplus_{\ell=1}^{\infty} \lambda^{\downarrow}(\Gamma_\ell)$ for all $s > 0$.

Moreover (Davies 1989, Theorem 3.3.4),

$$\int_G h_t(x; x)\,\mathrm{d}x \asymp \sup_{x \in G} h_t(x; x) \asymp t^{-d/2} \tag{7.37}$$

as $t \to 0$, so that condition (7.16) holds with $\rho = d$. If ν is a finite measure on $[0, \infty)$, (7.37) yields

$$\gamma(t) \leqslant Ct^{-d}\int_0^\infty\int_0^\infty \left(1+\frac{u+v}{t}\right)^{-d/2} \mathrm{d}\nu(u)\,\mathrm{d}\nu(v) \leqslant Ct^{-d},$$

and, as a consequence, condition (7.17) holds with $a = d/2$. Thus Corollary 7.7, for $\alpha > d/4$ and any $s > 0$, gives

$$\mathbb{E}\delta_2^2(\lambda(\mathrm{G}_n^{\alpha,s}),\, \lambda(G_{\alpha,s})) = O(n^{-(4\alpha-d)/(4\alpha+d)}). \tag{7.38}$$

The case of Neumann boundary condition can be treated in a similar way.

The Laplace–Beltrami operator on a compact Riemannian manifold

We let $S = M$ be a compact d-dimensional Rimmannian manifold and let P be the uniform distribution on M. Let Δ be the Laplace–Beltrami operator (defined on the space $C^\infty(M)$ of all infinitely differentiable functions on M) and let L be the closure – more precisely, form closure, using the terminology of Davies (1989) – of $-\Delta$. Then e^{-Lt} is a positivity-preserving contraction semigroup of $L_2(M)$ which has a strictly positive C^∞ kernel (heat kernel) $(t, x, y) \mapsto h_t(x, y)$ on $(0, \infty) \times M \times M$, so that

$$(\mathrm{e}^{-Lt}f)(x) = \int_S h_t(x, y) f(y)\,\mathrm{d}P(y), \qquad f \in L_2(M)$$

(Davies 1989, Theorem 5.2.1). Since, M being compact, h_t is bounded for each t, the operator e^{-Lt} is Hilbert–Schmidt and the heat kernel can be represented as

$$h_t(x, y) = \sum_{r=1}^\infty \mathrm{e}^{-\mu_r t}\phi_r(x)\phi_r(y),$$

where $0 < \mu_1 \leqslant \mu_2 \leqslant \ldots$ and $\{\phi_r,\, r \geqslant 1\}$ is an orthonormal system in $L_2(M)$. Again by boundedness, $\psi(t) = \int h_t^2(x, x) P(\mathrm{d}x) < +\infty$ for all $t > 0$, and Corollary 7.6 gives the weak convergence of the finite-dimensional distributions of the sequence

$$\{(n^{1/2}(\mu_{r,n}^s - \mu_r) : r \geqslant 1)\}_{n=1}^\infty$$

for all $s > 0$. Estimates similar to (7.38) can also be obtained, using the bounds on the heat kernel of Laplace–Beltrami operators, in Davies (1989, Chapter 5).

Schrödinger operators

We now assume that $L := -\Delta + V$ on $L_2(\mathbb{R}^d)$. Suppose the potential V is continuous and $V(x) \to +\infty$ as $|x| \to \infty$. Let ϕ be the eigenfunction of L corresponding to the minimal eigenvalue E (the so-called 'ground state' which, under the above conditions on V, exists and is a strictly positive C^2 function). Suppose, in addition, that, for some constants $a_1 > 2$, $2 < a_2 < 2a_1 - 2$ and $c_1 > 0$, $c_2 > 0$, $c_3 > 0$, $c_4 > 0$, V satisfies the inequalities

$$c_1|x|^{a_1} - c_2 \leqslant V(x) - E \leqslant c_3|x|^{a_2} + c_4$$

for all $x \in \mathbb{R}^d$.

Then (Davies 1989, Theorem 4.5.4) the kernel $\bar{h}_t$ of the semigroup e^{-Lt}, $t > 0$, satisfies the bound

$$0 \leqslant \bar{h}_t(x, y) \leqslant b_1 \exp\{-b_2 t^{-b}\}\phi(x)\phi(y),$$

for all x, y. Moreover, under the above condition on the potential V, there exists $\delta > 0$ such that $|\phi(x)| = O(e^{-\delta|x|})$ (see, for example, Simon 1982, Theorem C.3.4). It follows that $\bar{h}_t(x, x) = O(e^{-\delta|x|})$. Let p be a continuous density on $\mathbb{R}^d$ such that $p(x) > 0$ for all x and $p(x) \geqslant c|x|^{-\gamma}$ for some $\gamma > 0$ and $c > 0$, and all x of large enough norm, and let P be the Borel measure with density p. Define

$$h_t(x, y) = \frac{\bar{h}_t(x, y)}{p^{1/2}(x)p^{1/2}(y)}.$$

Then we have $h_t(x, x) = O(e^{-\delta/2}|x|)$, which implies

$$\int_{\mathbb{R}^d} h_t^2(x, x)\, dP(x) = \int_{\mathbb{R}^d} \frac{\bar{h}_t^2(x, x)}{p(x)}\, dx < +\infty$$

for all $t > 0$. Therefore, Corollary 7.6 gives weak convergence of the finite-dimensional distributions of the sequence

$$\{(n^{1/2}(\mu_{r,n}^s - \mu_r) : r \geqslant 1)\}_{n=1}^{\infty}$$

for all $s > 0$.

Acknowledgements

The research of the first author was partially supported by Sandia Laboratories grants nos AR-2011 and AU-4005. The research of the second author was partially supported by National Science Foundation grant no. DMS-96-25457.

We thank Joel Zinn for several conversations on the law of large numbers, particularly for improving and simplifying our previous treatment of an example (see Proposition 3.4) and for posing the problem of necessity in the law of large numbers, which is not yet completely solved. We acknowledge interesting conversations on different aspects of this work with I. Koltracht and with S. Efromovich. We are also grateful to J. Epperson for pointing out the book of Davies (1989) and to J. Wellner for pointing out the paper of Eaton and Tyler (1991).

References

Anderson, T.W. (1948) The asymptotic distributions of the roots of certain determinantal equations. *J. Roy. Statist. Soc. Ser. B*, **10**, 132–139.

Bai, Z.D. (1993a) Convergence rate of expected spectral distribution of large random matrices. Part I. Wigner matrices. *Ann. Probab.*, **21**, 625–648.

Bai, Z.D. (1993b) Convergence rate of expected spectral distribution of large random matrices. Part II. Sample covariance matrices. *Ann. Probab.* **21**, 649–673.

Cycon, H.L., Froese, R.G., Kirsch, W. and Simon, B. (1987) *Schrödinger Operators with Applications to Quantum Mechanics and Global Geometry.* New York: Springer-Verlag.

Davies, E.B. (1989) *Heat Kernels and Spectral Theory.* Cambridge: Cambridge University Press.

Dauxois, J., Pousse, A. and Romain, Y. (1982) Asymptotic theory for the principal component analysis of a vector random function: some applications to statistical inference. *J. Multivariate Anal.*, **12**, 136–154.

Dehling, H. and Mikosch, T. (1994) Random quadratic forms and the bootstrap for U-statistics. *J. Multivariate Anal.*, **51**, 392–413.

Dunford, N. and Schwartz, J.T. (1963) *Linear Operators*. New York: Wiley Interscience.

Dyson, F.J. (1962a) Statistical theory of energy levels of complex systems, I. *J. Math. Phys.*, **3**, 140–156.

Dyson, F.J. (1962b) Statistical theory of energy levels of complex systems, II *J. Math. Phys.* **3**, 157–160.

Dyson, F.J. (1962c) Statistical theory of energy levels of complex systems, III. *J. Math. Phys.* **3**, 166–175.

Eaton, M.L. and Tyler, D.E. (1991) On Wielandt's inequality and its application to the asymptotic distribution of the eigenvalues of a random symmetric matrix. *Ann. Statist.*, **19**, 260–271.

Geman, S. (1980) A limit theorem for the norm of random matrices. *Ann. Probab.*, **8**, 252–261.

Giné, E. and Zinn, J. (1992) On Hoffmann-Jørgensen's inequality for U-processes. In R.M. Dudley, M.G. Hahn and J. Kuelbs (eds), *Probability in Banach Spaces 8*, pp. 80–91. Boston: Birkhäuser.

Giné, E. and Zinn, J. (1994) A remark on convergence in distribution of U-statistics. *Ann. Probab.*, **22**, 117–125.

Giné, E. and Zhang, C.-H. (1996) On integrability in the LIL for degenerate U-statistics. *J. Theoret. Probab.*, **9**, 385–412.

Girko, V.L. (1990) *Theory of Random Determinants*. Dordrecht: Kluwer Academic.

Gohberg, I. and Krein, M. (1968) *Introduction to the Theory of Linear Nonselfadjoint Operators*. Providence, RI: American Mathematical Society.

Grenander, U. (1963) *Probabilities on Algebraic Structures*. Stockholm: Almqvist & Wiksell.

Hoffman, A.J. and Wielandt, H.W. (1953) The variation of the spectrum of a normal matrix. *Duke Math. J.*, **20**, 37–39.

Hsu, P.L. (1939) On the distribution of the roots of certain determinantal equations. *Ann. Eugenics*, **9**, 250–258.

James, A.T. (1954) Normal multivariate analysis and the orthogonal group. *Ann. Math. Statist.*, **25**, 40–75.

Kato, T. (1982) *A Short Introduction to Perturbation Theory for Linear Operators*. New York: Springer-Verlag.

Latala, R. (1998) On the almost sure boundedness of norms of some empirical operators. *Statist Probab. Lett.*, **38**, 177–182.

Lidskii, V.B. (1950) The proper values of the sum and product of symmetric matrices. *Dokl. Akad. Nauk SSSR*, **75**, 769–772.

Mehta, M.L. (1991) *Random Matrices*. New York: Academic Press.

Ozawa, S. (1987) Fluctuation of spectra in random media. In K. Ito and N. Ikeda (eds), *Probabilistic Methods in Mathematical Physics*, pp. 335–361. Boston: Academic Press.

Ozawa, S. (1993) Spectra of random media with many randomly distributed obstacles. *Osaka J. Math.*, **30**, 1–27.

Pastur, L. (1973) Spectra of random selfadjoint operators. *Uspekhi Mat. Nauk*, **28**, 3–64. [in Russian]. English translation: *Russian Math. Surveys*, **28**(1), 1–67.

Silverstein, J.W. (1985) The largest eigenvalue of a large dimensional Wishart matrix. *Ann. Probab.*, **13**, 1364–1368.

Simon, B. (1979) *Functional Integration and Quantum Physics*. New York: Academic Press.

Simon, B. (1982) Schrödinger semigroups. *Bull. Amer. Math. Soc.*, **7**, 447–526.

Serfling, R.J. (1980) *Approximation Theorems in Mathematical Statistics*. New York: Wiley.

Vardi, Y., Shepp, L.A. and Kaufman, L. (1985) A statistical model for positron emission tomography. *J. Amer. Statist. Assoc.*, **80**, 8–22.

Voiculescu, D. (1991) Limit laws for random matrices and free products. *Invent. Math.*, **104**, 201–220.

Wielandt, H. (1967) *Topics in the Analytic Theory of Matrices* (lecture notes prepared by R.R. Meyer). Madison: University of Wisconsin Press.

Wigner, E.P. (1955) Characteristic vectors of bordered matrices with infinite dimensions. *Ann. Math.*, **62**, 548–564.

Received March 1997 and revised July 1998