Graph-Graphon Filters Approximation Theorem

Alejandro Ribeiro, Luana Ruiz and Luiz Chamon

1 Assumptions and Preliminary Results

Before delving deeper into the proof of the approximation results for graph and graphon filters, we state clearly the assumptions under which the result holds. We also present some propositions we will need in order to prove our main result.

1.1 Assumptions

Assumption 1 *The graphon* $\mathbf{W}$ *is* L_1*-Lipschitz, i.e.* $|\mathbf{W}(u_2, v_2) - \mathbf{W}(u_1, v_1)| \leq L_1(|u_2 - u_1| + |v_2 - v_1|)$.

Assumption 2 *The convolutional filters* h *are* L_2*-Lipschitz and non-amplifying, i.e.* $|h(\lambda)| < 1$.

Assumption 3 *The graphon signal* X *is* L_3*-Lipschitz.*

1.2 Propositions

To prove Theorem 1, we need the following four propositions.

Proposition 1 *Let* $\mathbf{W} : [0,1]^2 \to [0,1]$ *be an* L_1*-Lipschitz graphon, and let* $\mathbf{W}_n$ *be the graphon induced by the deterministic graph* $\mathbf{G}_n$ *obtained from* $\mathbf{W}$ *as*

$$[\mathbf{S}_n]_{ij} = s_{ij} = \mathbf{W}(u_i, u_j). \tag{1}$$

The L_2 *norm of* $\mathbf{W} - \mathbf{W}_n$ *satisfies*

$$\|\mathbf{W} - \mathbf{W}_n\|_{L_2([0,1]^2)} \leq \sqrt{\|\mathbf{W} - \mathbf{W}_n\|_{L_1([0,1]^2)}} \leq \frac{\sqrt{L_1}}{\sqrt{n}}.$$

Proof: See section 4.1. ∎

Proposition 2 *Let* T *and* T' *be two self-adjoint operators on a separable Hilbert space* $\mathcal{H}$ *whose spectra are partitioned as* $\gamma \cup \Gamma$ *and* $\omega \cup \Omega$ *respectively, with* $\gamma \cap \Gamma = \emptyset$ *and* $\omega \cap \Omega = \emptyset$*. If there exists* $d > 0$ *such that* $\min_{x \in \gamma, y \in \Omega} |x - y| \geq d$ *and* $\min_{x \in \omega, y \in \Gamma} |x - y| \geq d$*, then*

$$\|E_T(\gamma) - E_{T'}(\omega)\| \leq \frac{\pi}{2} \frac{\|T - T'\|}{d}$$

Proof: See [1]. ∎

Proposition 3 *Let* $X \in L_2([0,1])$ *be an* L_3*-Lipschitz graphon signal, and let* X_n *be the graphon signal induced by the deterministic graph signal* $\mathbf{x}_n$ *obtained from* X *as in*

$$\begin{aligned} &[\mathbf{S}_n]_{ij} = \mathbf{W}(u_i, u_j) \quad \textit{and} \\ &[\mathbf{x}_n]_i = X(u_i) \end{aligned} \tag{2}$$

and

$$\begin{aligned} &[\mathbf{S}_n]_{ij} = \mathbf{W}(u_i, u_j) \\ &[\mathbf{x}_n]_i = X(u_i) \\ &\mathbf{H}_n(\mathbf{S}_n)\mathbf{x}_n = \mathbf{V}_n^{\mathsf{H}} h(\mathbf{\Lambda}_n) \mathbf{V}_n^{\mathsf{H}} \mathbf{x}_n. \end{aligned} \tag{3}$$

The L_2 *norm of* $X - X_n$ *satisfies*

$$\|X - X_n\|_{L_2([0,1])} \leq \frac{L_3}{\sqrt{3n}}.$$

Proof: See section 4.2. ■

Proposition 4 *Let* $\mathbf{W} : [0,1]^2 \to [0,1]$ *and* $\mathbf{W}' : [0,1]^2 \to [0,1]$ *be two graphons with eigenvalues given by* $\{\lambda_i(T_{\mathbf{W}})\}_{i\in\mathbb{Z}\setminus\{0\}}$ *and* $\{\lambda_i(T_{\mathbf{W}'})\}_{i\in\mathbb{Z}\setminus\{0\}}$*, ordered according to their sign and in decreasing order of absolute value. Then, for all* $i \in \mathbb{Z}\setminus\{0\}$*, the following inequalities hold*

$$|\lambda_i(T_{\mathbf{W}'}) - \lambda_i(T_{\mathbf{W}})| \le \|T_{\mathbf{W}'-\mathbf{W}}\| \le \|\mathbf{W}' - \mathbf{W}\|_{L_2} .$$

Proof: See section 4.3. ■

2 Approximation of Graphon Filters by Graph Filters

Theorem 1 *Consider the graphon convolution given by* $Y = T_{\mathbf{H}}X$ *as in*

$$\begin{aligned}(T_{\mathbf{H}}X)(v) &= \sum_{i\in\mathbb{Z}\setminus\{0\}} \sum_{k=0}^{K-1} h_k \lambda_i^k \varphi_i(v) \int_0^1 \varphi_i(u)X(u)du \\ &= \sum_{i\in\mathbb{Z}\setminus\{0\}} h(\lambda_i)\varphi_i(v) \int_0^1 \varphi_i(u)X(u)du, \end{aligned} \quad (4)$$

where for $|\lambda| < c$ *we have that* $h(\lambda_j) = K$*, and* $h(\lambda)$ *has total variation* $TV(h)$*. For the graph convolution instantiated from* $T_{\mathbf{H}}$ *as* $\mathbf{y}_n = \mathbf{H}_n(\mathbf{S}_n)\mathbf{x}_n$ *[cf. (3)], under Assumptions 1 through 3 it holds*

$$\|Y - Y_n\|_{L_2} \le \sqrt{L_1}\left(L_2 + \frac{\pi B_{nc}}{\delta_{nc}}\right) n^{-\frac{1}{2}} \|X\|_{L_2} + \frac{2L_3}{\sqrt{3}} n^{-\frac{1}{2}} + TV(h)\|X\|_{L_2} \quad (5)$$

where $Y_n = T_{\mathbf{H}_n}X_n$ *is the graph convolution induced by* $\mathbf{y}_n = \mathbf{H}_n(\mathbf{S}_n)\mathbf{x}_n$*:*

$$\begin{aligned} &\mathbf{W}_n(u,v) = [\mathbf{S}_n]_{ij} \times \mathbb{I}(u \in I_i)\mathbb{I}(v \in I_j), \\ &X_n(u) = [\mathbf{x}_n]_i \times \mathbb{I}(u \in I_i), \\ &(T_{\mathbf{H}_n}X_n)(v) = \sum_{i\in\mathbb{Z}\setminus\{0\}} h(\lambda_i^n)\varphi_i^n(v) \int_0^1 \varphi_i^n(u)X_n(u)du, \end{aligned} \quad (6)$$

B_{nc} is the cardinality of the set $\mathcal{C} = \{i \mid |\lambda_i^n| \geq c\}$, and $\delta_{nc} = \min_{i \in \mathcal{C}}(|\lambda_i - \lambda^n_{i+sgn(i)}|, |\lambda_{i+sgn(i)} - \lambda_i^n|, |\lambda_1 - \lambda^n_{-1}|, |\lambda_1^n - \lambda_{-1}|)$, with λ_i and λ_i^n denoting the eigenvalues of $\mathbf{W}$ and $\mathbf{W}_n$ respectively. In particular, if $X = X_n$ we have

$$\|Y - Y_n\|_{L_2} \leq \sqrt{L_1}\left(L_2 + \frac{\pi B_{nc}}{\delta_{nc}}\right) n^{-\frac{1}{2}} \|X\|_{L_2}. \tag{7}$$

3 Proof of Filter Approximation Theorem

Proof: [Proof of Theorem 1]

We first prove the result of Theorem 1 for filters $h(\lambda)$ satisfying $h(\lambda) = 0$ for $|\lambda| < c$. Using the triangle inequality, we can write the norm difference $\|Y - Y_n\|_{L_2}$ as

$$\begin{aligned} \|Y - Y_n\|_{L_2} &= \|T_{\mathbf{H}} X - T_{\mathbf{H}_n} X_n\|_{L_2} = \|T_{\mathbf{H}} X + T_{\mathbf{H}_n} X - T_{\mathbf{H}_n} X - T_{\mathbf{H}_n} X_n\|_{L_2} \\ &\leq \|T_{\mathbf{H}} X - T_{\mathbf{H}_n} X\|_{L_2} \ \textbf{(1)} \ + \|T_{\mathbf{H}_n}(X - X_n)\|_{L_2} \ \textbf{(2)} \end{aligned}$$

where the LHS is split between terms **(1)** and **(2)**.

Writing the inner products $\int_0^1 X(u)\varphi_i(u)du$ and $\int_0^1 X(u)\varphi_i^n(u)du$ as $\hat{X}(\lambda_i)$ and $\hat{X}(\lambda_i^n)$ for simplicity, we can then express **(1)** as

$$\begin{aligned} \|T_{\mathbf{H}} X - T_{\mathbf{H}_n} X\|_{L_2} &= \left\| \sum_i h(\lambda_i)\hat{X}(\lambda_i)\varphi_i - \sum_i h(\lambda_i^n)\hat{X}(\lambda_i^n)\varphi_i^n \right\|_{L_2} \\ &= \left\| \sum_i h(\lambda_i)\hat{X}(\lambda_i)\varphi_i - h(\lambda_i^n)\hat{X}(\lambda_i^n)\varphi_i^n \right\|_{L_2}. \end{aligned}$$

Using the triangle inequality, this becomes

$$\begin{aligned}
\|T_{\mathbf{H}}X - T_{\mathbf{H}_n}X\|_{L_2} &= \left\| \sum_i h(\lambda_i)\hat{X}(\lambda_i)\varphi_i - h(\lambda_i^n)\hat{X}(\lambda_i^n)\varphi_i^n \right\|_{L_2} \\
&= \left\| \sum_i h(\lambda_i)\hat{X}(\lambda_i)\varphi_i + h(\lambda_i^n)\hat{X}(\lambda_i)\varphi_i \right. \\
&\left. -h(\lambda_i^n)\hat{X}(\lambda_i)\varphi_i - h(\lambda_i^n)\hat{X}(\lambda_i^n)\varphi_i^n \right\|_{L_2} \\
&\leq \left\| \sum_i \left(h(\lambda_i) - h(\lambda_i^n)\right)\hat{X}(\lambda_i)\varphi_i \right\|_{L_2} \quad \textbf{(1.1)} \\
&+ \left\| \sum_i h(\lambda_i^n)\left(\hat{X}(\lambda_i)\varphi_i - \hat{X}(\lambda_i^n)\varphi_i^n\right) \right\|_{L_2} \quad \textbf{(1.2)}
\end{aligned}$$

where we have now split **(1)** between **(1.1)** and **(1.2)**.

Focusing on **(1.1)**, note that the filter's Lipschitz property allows writing $h(\lambda_i) - h(\lambda_i^n) \leq L_2|\lambda_i - \lambda_i^n|$. Hence, using Proposition 4 together with the Cauchy-Schwarz inequality, we get

$$\left\| \sum_i \left(h(\lambda_i) - h(\lambda_i^n)\right)\hat{X}(\lambda_i)\varphi_i \right\|_{L_2} \leq L_2\|\mathbf{W} - \mathbf{W}_n\|_{L_2} \left\| \sum_i \hat{X}(\lambda_i)\varphi_i \right\|_{L_2}$$

and, from Proposition 1,

$$\left\| \sum_i \left(h(\lambda_i) - h(\lambda_i^n)\right)\hat{X}(\lambda_i)\varphi_i \right\|_{L_2} \leq \frac{L_2\sqrt{L_1}}{\sqrt{n}}\|X\|_{L_2}. \tag{8}$$

For **(1.2)**, we use the triangle and Cauchy-Schwarz inequalities to write

$$\begin{aligned}
&\left\| \sum_i h(\lambda_i^n)\left(\hat{X}(\lambda_i)\varphi_i - \hat{X}(\lambda_i^n)\varphi_i^n\right) \right\|_{L_2} \\
&\quad = \left\| \sum_i h(\lambda_i^n)\left(\hat{X}(\lambda_i)\varphi_i + \hat{X}(\lambda_i)\varphi_i^n - \hat{X}(\lambda_i)\varphi_i^n - \hat{X}(\lambda_i^n)\varphi_i^n\right) \right\|_{L_2} \\
&\quad \leq \left\| \sum_i h(\lambda_i^n)\hat{X}(\lambda_i)(\varphi_i - \varphi_i^n) \right\|_{L_2} + \left\| \sum_i h(\lambda_i^n)\varphi_i^n\langle X, \varphi_i - \varphi_i^n\rangle \right\|_{L_2} \\
&\quad \leq 2\sum_i \|h(\lambda_i^n)\|_{L_2}\|X\|_{L_2}\|\varphi_i - \varphi_i^n\|_{L_2}.
\end{aligned}$$

Using Proposition 2 with $\gamma = \lambda_i$ and $\omega = \lambda_i^n$, we then get

$$\left\| \sum_i h(\lambda_i^n) \left(\hat{X}(\lambda_i)\varphi_i - \hat{X}(\lambda_i^n)\varphi_i^n \right) \right\|_{L_2} \leq \|X\|_{L_2} \sum_i \|h(\lambda_i^n)\|_{L_2} \frac{\pi \|T_{\mathbf{W}} - T_{\mathbf{W}_n}\|}{d_i}$$

where d_i is the minimum between $\min(|\lambda_i - \lambda_{i+1}^n|, |\lambda_i - \lambda_{i-1}^n|)$ and $\min(|\lambda_i^n - \lambda_{i+1}|, |\lambda_i^n - \lambda_{i-1}|)$ for each i. Since $\delta_{nc} \leq d_i$ for all i and $\|T_{\mathbf{W}} - T_{\mathbf{W}_n}\| \leq \|\mathbf{W} - \mathbf{W}_n\|_{L_2}$ (i.e., the Hilbert-Schmidt norm dominates the operator norm), this becomes

$$\left\| \sum_i h(\lambda_i^n) \left(\hat{X}(\lambda_i)\varphi_i - \hat{X}(\lambda_i^n)\varphi_i^n \right) \right\|_{L_2} \leq \frac{\pi \|\mathbf{W} - \mathbf{W}_n\|_{L_2}}{\delta_{nc}} \|X\|_{L_2} \sum_i \|h(\lambda_i^n)\|_{L_2}$$

and, using Proposition 1,

$$\left\| \sum_i h(\lambda_i^n) \left(\hat{X}(\lambda_i)\varphi_i - \hat{X}(\lambda_i^n)\varphi_i^n \right) \right\|_{L_2} \leq \frac{\pi \sqrt{L_1}}{\delta_{nc}\sqrt{n}} \|X\|_{L_2} \sum_i \|h(\lambda_i^n)\|_{L_2}.$$

The final bound for **(1.2)** is obtained by noting that $|h(\lambda)| < 1$ and $h(\lambda) = 0$ for $|\lambda| < c$. Since there are a total of B_{nc} eigenvalues λ_i^n for which $|\lambda_i^n| \geq c$, we get

$$\left\| \sum_i h(\lambda_i^n) \left(\hat{X}(\lambda_i)\varphi_i - \hat{X}(\lambda_i^n)\varphi_i^n \right) \right\|_{L_2} \leq \frac{\pi \sqrt{L_1}}{\delta_{nc}\sqrt{n}} \|X\|_{L_2} B_{nc}. \tag{9}$$

A bound for **(2)** follows immediately from Proposition 3. Since $|h(\lambda)| < 1$, the norm of the operator $T_{\mathbf{H}_n}$ is bounded by 1. Using the Cauchy-Schwarz inequality, we then have $\|T_{\mathbf{H}_n}(X - X_n)\|_{L_2} \leq \|X - X_n\|_{L_2}$ and therefore

$$\|T_{\mathbf{H}_n}(X - X_n)\|_{L_2} \leq \frac{L_3}{\sqrt{3n}} \tag{10}$$

which completes the bound on $\|Y - Y_n\|_{L_2}$ when $h(\lambda) = 0$ for $|\lambda| < c$. For filters in which $h(\lambda)$ satisfies $h(\lambda_j) = K$ for $|\lambda| < c$, we obtain a bound by observing that $h(\lambda)$ can be constructed as the sum of two filters: an L_2-Lipschitz filter $f(\lambda)$ with $f(\lambda) = 0$ for $|\lambda| < c$, and a slow variation filter $g(\lambda_j) = K$. Hence, by the triangle inequality

$$\|Y - Y_n\|_{L_2} = \|T_{\mathbf{H}} X - T_{\mathbf{H}_n}\|_{L_2} \leq \|T_{\mathbf{F}} X - T_{\mathbf{F}_n} X_n\|_{L_2} + \|T_{\mathbf{G}} X - T_{\mathbf{G}_n} X_n\|_{L_2}.$$

The bound on $\|T_{\mathbf{F}} X - T_{\mathbf{F}_n}\|_{L_2}$ is the one we have derived, and for $\|T_{\mathbf{G}} X - T_{\mathbf{G}_n} X_n\|_{L_2}$, we use $|g(\lambda)| \leq 1$ to obtain a first component of the bound given by

$$\|T_{\mathbf{G}_n}(X - X_n)\|_{L_2} \leq \|X - X_n\|_{L_2} \leq \frac{L_3}{\sqrt{3n}}. \tag{11}$$

To obtain the second component we start from

$$\begin{aligned}\|T_{\mathbf{G}} X - T_{\mathbf{G}_n} X\|_{L_2} &= \left\| \sum_i g(\lambda_i)\hat{X}(\lambda_i)\varphi_i - \sum_i g(\lambda_i^n)\hat{X}(\lambda_i^n)\varphi_i^n \right\|_{L_2} \\ &= \left\| \sum_i \left(g(\lambda_i)\hat{X}(\lambda_i)\varphi_i - g(\lambda_i^n)\hat{X}(\lambda_i^n)\varphi_i^n \right) \right\|_{L_2},\end{aligned}$$

Rearranging the terms we obtain

$$\begin{aligned}\left\| \sum_i \left(g(\lambda_i)\hat{X}(\lambda_i)\varphi_i - g(\lambda_i^n)\hat{X}(\lambda_i^n)\varphi_i^n \right) \right\|_{L_2} &\leq \left\| - \sum_i \left(g(\lambda_i^n) - g(\lambda_i) \right) \hat{X}(\lambda_i^n)\varphi_i^n \right\|_{L_2} \\ &+ \left\| \sum_i g(\lambda_i) \left(\hat{X}(\lambda_i)\varphi_i - \hat{X}(\lambda_i^n)\varphi_i^n \right) \right\|_{L_2}\end{aligned}$$

then taking into account that $g(\lambda_j) = K$, that $g(\lambda)$ has total variation $TV(h)$ and applying Cauchy-Schwartz we have

$$\begin{aligned}\left\| \sum_i \left(h(\lambda_i)\hat{X}(\lambda_i)\varphi_i - h(\lambda_i^n)\hat{X}(\lambda_i^n)\varphi_i^n \right) \right\|_{L_2} &\leq TV(h)\|X\|_{L_2} \\ &+ K \left\| \sum_i \left(\hat{X}(\lambda_i)\varphi_i - \hat{X}(\lambda_i^n)\varphi_i^n \right) \right\|_{L_2}\end{aligned} \tag{12}$$

where the second term in eqn. (12) equals zero. Notice that if $h(\lambda)$ is constant $TV(h) = 0$.

Putting together (8), (9), (10) and (11), we arrive at the first result of the theorem as stated in (5). The second result [cf. (7)] is obtained by observing that, for $X = X_n$, bound **(2)** in (10) simplifies to $\|T_{\mathbf{H}_n}(X - X_n)\|_{L_2} = 0$; and, similarly in (11), $\|T_{\mathbf{G}} X - T_{\mathbf{G}_n} X_n\|_{L_2} \leq \|X - X_n\|_{L_2} = 0$. ■

4 Proof of Propositions

4.1 Proof of Proposition 1

Proof: Partitioning the unit interval as $I_i = [(i-1)/n, i/n]$ for $1 \leq i \leq n$ (the same partition used to obtain $\mathbf{S}_n$, and thus $\mathbf{W}_n$, from $\mathbf{W}$), we can use the graphon's Lipschitz property to derive

$$\begin{aligned}
&\|\mathbf{W} - \mathbf{W}_n\|_{L_1(I_i \times I_j)} \\
&\quad \leq L_1 \int_0^{1/n} \int_0^{1/n} |u| du dv + L_1 \int_0^{1/n} \int_0^{1/n} |v| dv du = \frac{L_1}{2n^3} + \frac{L_1}{2n^3} \\
&\quad = \frac{L_1}{n^3}. \qquad (13)
\end{aligned}$$

We can then write

$$\|\mathbf{W} - \mathbf{W}_n\|_{L_1([0,1]^2)} = \sum_{i,j} \|\mathbf{W} - \mathbf{W}_n\|_{L_1(I_i \times I_j)} \leq n^2 \frac{L_1}{n^3} = \frac{L_1}{n}$$

which, since $\mathbf{W} - \mathbf{W}_n : [0,1]^2 \to [-1,1]$, implies

$$\|\mathbf{W} - \mathbf{W}_n\|_{L_2([0,1]^2)} \leq \sqrt{\|\mathbf{W} - \mathbf{W}_n\|_{L_1([0,1]^2)}} \leq \frac{\sqrt{L_1}}{\sqrt{n}}.$$

■

4.2 Proof of Proposition 3

Proof: Partitioning the unit interval as $I_i = [(i-1)/n, i/n]$ for $1 \leq i \leq n$ (the same partition used to obtain $\mathbf{x}_n$, and thus X_n, from X), we can use the Lipschitz property of X to derive

$$\|X - X_n\|_{L_2(I_i)} \leq \sqrt{L_3^2 \int_0^{1/n} u^2 du} = \sqrt{\frac{L_3^2}{3n^3}} + \frac{L_3}{n\sqrt{3n}}.$$

We can then write

$$\|X - X_n\|_{L_2([0,1])} = \sum_i \|X - X_n\|_{L_2(I_i)} \leq n \frac{L_3}{n\sqrt{3n}} = \frac{L_3}{\sqrt{3n}}.$$

■

4.3 Proof of Proposition 4

Proof: Let $\mathbf{A} := \mathbf{W}' - \mathbf{W}$ and let S_k denote a k-dimensional subspace of $L_2([0,1])$. Using the minimax principle [2, Chapter 1.6.10], we can write

$$\lambda_k(T_{\mathbf{W}}) = \min_{S_{k-1}} \max_{X \in S_{k-1}^{\perp}, \|X\|_{L_2}=1} \langle T_{\mathbf{W}} X, X \rangle \ .$$

Therefore, it holds that

$$\begin{aligned}
\lambda_k(T_{\mathbf{W}'}) &= \lambda_k(T_{\mathbf{W}+\mathbf{A}}) \\
&= \min_{S_{k-1}} \max_{X \in S_{k-1}^{\perp}, \|X\|_{L_2}=1} \langle T_{\mathbf{W}+\mathbf{A}} X, X \rangle \\
&= \min_{S_{k-1}} \max_{X \in S_{k-1}^{\perp}, \|X\|_{L_2}=1} \langle T_{\mathbf{W}} + T_{\mathbf{A}} X, X \rangle \\
&= \min_{S_{k-1}} \max_{X \in S_{k-1}^{\perp}, \|X\|_{L_2}=1} \left(\langle T_{\mathbf{W}} X, X \rangle + \langle T_{\mathbf{A}} X, X \rangle \right) \\
&\leq \min_{S_{k-1}} \left(\max_{X \in S_{k-1}^{\perp}, \|X\|_{L_2}=1} \langle T_{\mathbf{W}} X, X \rangle + \max_{X \in S_{k-1}^{\perp}, \|X\|_{L_2}=1} \langle T_{\mathbf{A}} X, X \rangle \right) \\
&\leq \min_{S_{k-1}} \left(\max_{X \in S_{k-1}^{\perp}, \|X\|_{L_2}=1} \langle T_{\mathbf{W}} X, X \rangle + \max_{\ell} \lambda_\ell(T_{\mathbf{A}}) \right) \\
&= \min_{S_{k-1}} \max_{X \in S_{k-1}^{\perp}, \|X\|_{L_2}=1} \langle T_{\mathbf{W}} X, X \rangle + \max_{\ell} \lambda_\ell(T_{\mathbf{A}}) = \lambda_k(T_{\mathbf{W}}) + \max_{\ell} \lambda_\ell(T_{\mathbf{A}}) \ .
\end{aligned}$$

where the first inequality follows from $\max(a+b) \leq \max(a) + \max(b)$ and the second from the fact that $\langle T_{\mathbf{A}} X, X \rangle \leq \max_\ell \lambda_\ell(T_{\mathbf{A}})$ for all unitary X. Rearranging terms and using the definition of the operator norm, we get

$$\lambda_k(T_{\mathbf{W}'}) - \lambda_k(T_{\mathbf{W}}) \leq \max_{\ell} \lambda_\ell(T_{\mathbf{A}}) \leq \max_{\ell} |\lambda_\ell(T_{\mathbf{A}})| = \|T_{\mathbf{A}}\| \leq \|A\|_{L_2}. \tag{14}$$

where we have also used the fact that the Hilbert-Schmidt norm dominates the operator norm.

To prove that this inequality holds in absolute value, let $\mathbf{A}' = -\mathbf{A}$. Following the same reasoning as before, we get

$$\lambda_k(T_{\mathbf{W}}) = \lambda_k(T_{\mathbf{W}'+A'}) \leq \lambda_k(T_{\mathbf{W}'}) + \|T_{\mathbf{A}'}\| \leq \lambda_k(T_{\mathbf{W}}) + \|A'\|_{L_2}$$

and since $\|T_{\mathbf{A}'}\| = \|T_{\mathbf{A}}\|$ and $\|A'\|_{L_2} = \|A\|_{L_2}$,

$$\lambda_k(T_{\mathbf{W}}) - \lambda_k(T_{\mathbf{W}'}) \leq \|T_{\mathbf{A}}\| \leq \|A\|_{L_2} . \tag{15}$$

Putting (14) and (15) together completes the proof. ■

References

[1] A. Seelmann, "Notes on the sin 2Θ theorem," *Integral Equations and Operator Theory*, vol. 79, no. 4, pp. 579–597, 2014.

[2] T. Kato, *Perturbation theory for linear operators*. Springer Science & Business Media, 2013, vol. 132.