Graphon Filters

1 Proof of the Convergence of filter response for Lipschitz continuous graph filters

We consider that, provided that the sequence of input graph signals $\{(\mathbf{G}_n, \mathbf{x}_n)\}$ converges to a graphon signal, the output signals obtained by appling the filters $\mathbf{H}_n(\mathbf{S}_n)$ converge in the same sense as $\{(\mathbf{G}_n, \mathbf{x}_n)\}$ in the vertex domain if the filter function h is *Lipschitz continuous*.

A function $h(\lambda) : [0, 1] \to \mathbb{R}$ is L-Lipschitz continuous in L_2 if, for all $\lambda, \lambda' \in [0, 1]$,

$$\|h(\lambda) - h(\lambda')\|_{L_2} \leq L\|\lambda - \lambda'\|_{L_2} . \tag{1}$$

This is equivalent to bounding $dh/d\lambda$ by L in absolute value. For filter functions satisfying (1), we can show that the filter response converges for any graphon signal.

We first recall the definition of convergent graph signal sequence:

Definition 1 (Convergent sequences of graph signals) *A sequence of graph signals $(\mathbf{G}_n, \mathbf{x}_n)$ is said to converge to the graphon signal $(\mathbf{W}, X)$ if there exists a sequence of permutations $\{\pi_n\}$ such that, for any simple graph $\mathbf{F}$,*

$$\lim_{n\to\infty} t(\mathbf{F}, \mathbf{G}_n) \to t(\mathbf{F}, \mathbf{W}) \tag{2}$$

i.e., $\mathbf{G}_n$ converges in the homomorphism density sense, and

$$\lim_{n\to\infty} \|X_{\pi_n(\mathbf{G}_n)} - X\|_{L_2} = 0 \tag{3}$$

where $(\mathbf{W}_{X_n(\mathbf{G}_n)}, X_{\pi_n(\mathbf{G}_n)})$ is the graphon signal induced by the permuted graph signal $(\pi_n(\mathbf{G}_n), \pi_n(\mathbf{x}_n))$.

We also need to formalize the relationship between the spectral properties of the graphon $\mathbf{W}_{\mathbf{G}_n}$ induced by $\mathbf{G}_n$ and graphon signal $X_{\mathbf{G}_n}$ induced by $\mathbf{x}_n$.

Lemma 1 *Let $(\mathbf{W_G}, X_\mathbf{G})$ be the graphon signal induced by the graph signal $(\mathbf{G}, \mathbf{x})$ on n nodes. Then, for $j \in \mathcal{L}$ we have*

$$\lambda_j(T_{\mathbf{W_G}}) = \frac{\lambda_j(\mathbf{S})}{n}$$
$$\varphi_j(T_{\mathbf{W_G}})(u) = [\mathbf{v}_j]_k \times \sqrt{n}\mathbb{I}\,(u \in I_k)$$
$$[\hat{X}_\mathbf{G}]_j = \frac{[\hat{\mathbf{x}}]_j}{\sqrt{n}}$$

where $\lambda_j(\mathbf{S})$ are the eigenvalues of the graph. For $j \notin \mathcal{L}$, $\lambda_j(T_{\mathbf{W_G}}) = [\hat{X}_\mathbf{G}]_j = 0$ and $\varphi_j(T_{\mathbf{W_G}}) = X_j$, such that $\{\varphi_j(T_{\mathbf{W_G}})\} \cup \{X_j\}$ forms an orthonormal basis of $L_2[0,1]$.

Proof can be found in [**?**].

In frequency domain, if the filter function h is the same, the frequency responses of graph filters and their graphon filters have the same expression. These frequency responses actually converge to one another as n goes to infinity which is stated in the following theorem.

Theorem 1 (Convergence of graph filter frequency response) *On the graph sequence $\{\mathbf{G}_n\}$, let $\mathbf{H}_n(\mathbf{S}_n)$ be a sequence of filters of the form $\mathbf{H}_n(\mathbf{S}_n) = \mathbf{V}_n h(\mathbf{\Lambda}_n(\mathbf{S}_n)/n)\mathbf{V}_n^H$; and, on the graphon $\mathbf{W}$, define the filter $T_\mathbf{H} : L_2[0,1] \to L_2[0,1]$ where $(T_\mathbf{H}\phi)(v) = \sum_{j\in\mathbb{Z}\setminus\{0\}} h(\lambda_j(T_\mathbf{W}))\hat{X}(\lambda_j)\varphi_j(v)$. If $\{\mathbf{G}_n\}$ converges to $\mathbf{W}$ and $h : [0,1] \to \mathbb{R}$ is continuous, then*

$$\lim_{n\to\infty} \hat{\mathbf{H}}_n(\lambda_j(\mathbf{S}_n)/n) = \hat{T}_\mathbf{H}(\lambda_j(T_\mathbf{W}))\ . \tag{4}$$

Proof can also be found in [**?**]. Combine the above theorem with the convergence of Fourier transform for graphon bandlimited signals, the following corollary gives account of the limit behavior of the graph filter response in the vertex domain.

Corollary 1 (Convergence of graph filter response for graphon bandlimited signals) *Let $\{(\mathbf{G}_n, \mathbf{y}_n)\}$ be the sequence of graph signals obtained by applying filters $\mathbf{H}_n(\mathbf{S}_n) = \mathbf{V}_n h(\mathbf{\Lambda}_n(\mathbf{S}_n)/n)\mathbf{V}_n^H$ to the sequence $\{(\mathbf{G}_n, \mathbf{x}_n)\}$, and let $(\mathbf{W}, Y)$ be the graphon signal obtained by applying the graphon filter $(T_\mathbf{H}X)(v) = \sum_{j\in\mathbb{Z}\setminus\{0\}} h(\lambda_j)\hat{X}(\lambda_j)\varphi_j(v)$ to the c-bandlimited signal $(\mathbf{W}, X)$, where $\mathbf{W}$ is non-derogatory. If $\{(\mathbf{G}_n, \mathbf{x}_n)\}$ converges to $(\mathbf{W}, X)$ and the function h is continuous, then $\{(\mathbf{G}_n, \mathbf{y}_n)\}$ converges to $(\mathbf{W}, Y)$ in the sense of Def. 1.*

The general convergence conclusion is stated in the following theorem for Lipschitz continuous graph filters and arbitrary graphons. We first consider the setting with a non-derogatory graphon $\mathbf{W}$.

Theorem 2 (Convergence of filter response for Lipschitz continuous graph filters) *Let $\{(\mathbf{G}_n, \mathbf{y}_n)\}$ be the sequence of graph signals obtained by applying filters $\mathbf{H}_n(\mathbf{S}_n) = \mathbf{V}_n h(\mathbf{\Lambda}_n(\mathbf{S}_n)/n)\mathbf{V}_n^H$ to* *the sequence $\{(\mathbf{G}_n, \mathbf{x}_n)\}$, and let $(\mathbf{W}, Y)$ be the graphon signal obtained by applying the graphon filter $(T_{\mathbf{H}}X)(v) = \sum_{i=1}^{\infty} h(\lambda_i)\hat{X}(\lambda_i)\varphi_i(v)$ to the signal $(\mathbf{W}, X)$, where $\mathbf{W}$ is non-derogatory. If $\{(\mathbf{G}_n, \mathbf{x}_n)\}$ converges to $(\mathbf{W}, X)$ and the function h is L-Lipschitz, then $\{(\mathbf{G}_n, \mathbf{y}_n)\}$ converges to $(\mathbf{W}, Y)$ in the sense of Def. 1.*

Proof: Our derivations use the graphon signals $(\mathbf{W}_{\mathbf{G}_n}, x_{\mathbf{G}_n})$ and $(\mathbf{W}_{\mathbf{G}_n}, Y_{\mathbf{G}_n})$ induced by the graph signals $(\mathbf{G}_n, \mathbf{x}_n)$ and $(\mathbf{G}_n, \mathbf{y}_n)$ to facilitate comparison with $(\mathbf{W}, X)$ and $(\mathbf{W}, Y)$. We simplify notation by writing $\mathbf{W}_n = \mathbf{W}_{\mathbf{G}_n}$, $X_n = X_{\mathbf{G}_n}$, $Y_n = Y_{\mathbf{G}_n}$ and $\lambda_i^n = \lambda_i(T_{\mathbf{W}_n})$. We will also drop the subscript L_2 when writing the L_2 norm as $\|\cdot\|_{L_2} \equiv \|\cdot\|$.

Without loss of generality, consider the normalized filter function $\bar{h}(\lambda) = h(\lambda)/\max_{\lambda\in[0,1]} h(\lambda)$. The signal $(\mathbf{W}, Y)$ obtained by applying $T_{\bar{\mathbf{H}}}$ to $(\mathbf{W}, X)$ can be written as

$$Y(v) = \sum_{i=1}^{\infty} \bar{h}(\lambda_i)\hat{X}(\lambda_i)\varphi_i(v) \tag{5}$$

and the graphon signal $(\mathbf{W}_n, Y_n)$ induced by $\mathbf{y}_n = \bar{\mathbf{H}}(\mathbf{S}_n)\mathbf{x}_n$ as

$$Y_n(v) = \sum_{i=1}^{n} \bar{h}(\lambda_i^n/n)\hat{X}_n(\lambda_i^n)\varphi_i(T_{\mathbf{W}_n})(v) \; . \tag{6}$$

where we make the dependence of the eigenfunctions $\varphi_i(T_{\mathbf{W}_n})$ on $T_{\mathbf{W}_n}$ explicit to differentiate between them and φ_i.

To show that $(\mathbf{W}_n, Y_n)$ converges to $(\mathbf{W}, Y)$, we start by writing their norm difference using (5) and (6),

$$\|Y - Y_n\| = \left\| \sum_{i=1}^{\infty} \bar{h}(\lambda_i)\hat{X}(\lambda_i)\varphi_i - \sum_{i=1}^{n} \bar{h}(\mu_i^n)\hat{X}_n(\mu_i^n)\varphi_i(T_{\mathbf{W}_n}) \right\| \tag{7}$$

where $\mu_i^n = \lambda_i^n/n$. Defining the set $\mathcal{C} = \{i \mid |\lambda_i| \geq c\}$, where $c = (1 - h_{\min})\epsilon/L\|X\|$ with $\epsilon > 0$ and $h_{\min} = \min_{[0,c]} \bar{h}(\lambda)$, these sums can be split up between $i \in \mathcal{C}$ and $i \notin \mathcal{C}$ such that

$$\begin{aligned}
&\left\| \sum_{i=1}^{\infty} \bar{h}(\lambda_i)\hat{X}(\lambda_i)\varphi_i - \sum_{i=1}^{n} \bar{h}(\mu_i^n)\hat{X}_n(\mu_i^n)\varphi_i(T_{\mathbf{W}_n}) \right\| \\
&\leq \left\| \sum_{i\in\mathcal{C}} \bar{h}(\lambda_i)\hat{X}(\lambda_i)\varphi_i - \sum_{i\in\mathcal{C}} \bar{h}(\mu_i^n)\hat{X}_n(\mu_i^n)\varphi_i(T_{\mathbf{W}_n}) \right\| \; \textbf{(i)} \\
&+ \left\| \sum_{i\notin\mathcal{C}} \bar{h}(\lambda_i)\hat{X}(\lambda_i)\varphi_i - \sum_{i\notin\mathcal{C}} \bar{h}(\mu_i^n)\hat{X}_n(\mu_i^n)\varphi_i(T_{\mathbf{W}_n}) \right\| \; \textbf{(ii)} \; .
\end{aligned} \tag{8}$$

We first derive a bound for **(i)** by noticing that this expression corresponds to the difference between two bandlimited graphon signals. By Corollary 1, there exists n_0 such that, for all $n > n_0$,

$$\left\| \sum_{i\in\mathcal{C}} \bar{h}(\lambda_i)\hat{X}(\lambda_i)\varphi_i - \sum_{i\in\mathcal{C}} \bar{h}(\mu_i^n)\hat{X}_n(\mu_i^n)\varphi_i(T_{\mathbf{W}_n}) \right\| < \epsilon \; . \tag{9}$$

For **(ii)**, we use the filter's Lipschitz property to derive

$$
\begin{aligned}
&\left\|\sum_{i\notin\mathcal{C}} \bar{h}(\lambda_i)\hat{X}(\lambda_i)\varphi_i - \sum_{i\notin\mathcal{C}} \bar{h}(\mu_i^n)\hat{X}_n(\mu_i^n)\varphi_i(T_{\mathbf{W}_n})\right\| \\
&\leq \left\|\sum_{i\notin\mathcal{C}} (h_{\min} + Lc)\hat{X}(\lambda_i)\varphi_i - \sum_{i\notin\mathcal{C}} h_{\min}\hat{X}_n(\mu_i^n)\varphi_i(T_{\mathbf{W}_n})\right\| \\
&= \left\|h_{\min}\sum_{i\notin\mathcal{C}}\left[\hat{X}(\lambda_i)\varphi_i - \hat{X}_n(\mu_i^n)\varphi_i(T_{\mathbf{W}_n})\right] + Lc\sum_{i\notin\mathcal{C}}\hat{X}(\lambda_i)\varphi_i\right\| .
\end{aligned} \tag{10}
$$

Note that because $\{\varphi_i\}$ and $\{\varphi_i(T_{\mathbf{W}_n})\}$ form complete bases of L_2, $\sum_{i\notin\mathcal{C}}\hat{X}(\lambda_i)\varphi_i$ and $\sum_{i\notin\mathcal{C}}\hat{X}_n(\mu_i^n)\varphi_i(T_{\mathbf{W}_n})$ can be written as

$$
\sum_{i\notin\mathcal{C}}\hat{X}(\lambda_i)\varphi_i = X - \sum_{i\in\mathcal{C}}\hat{X}(\lambda_i)\varphi_i \quad \text{and} \tag{11}
$$

$$
\sum_{i\notin\mathcal{C}}\hat{X}_n(\mu_i^n)\varphi_i(T_{\mathbf{W}_n}) = X_n - \sum_{i\in\mathcal{C}}\hat{X}(\mu_i^n)\varphi_i(T_{\mathbf{W}_n}) . \tag{12}
$$

Using these identities, we once again leverage the fact that X_n converges to X in L_2 and Corollary 1 to show that there exists n_1 such that, for all $n > n_1$,

$$
\left\|\sum_{i\notin\mathcal{C}}\hat{X}(\lambda_i)\varphi_i - \hat{X}_n(\mu_i^n)\varphi_i(T_{\mathbf{W}_n})\right\| = \left\|\sum_{i\in\mathcal{C}}\hat{X}_n(\mu_i^n)\varphi_i(T_{\mathbf{W}_n}) - \hat{X}(\lambda_i)\varphi_i\right\| < \epsilon . \tag{13}
$$

Applying the Cauchy-Schwarz and triangle inequalities and substituting (13) in (10), we arrive at a bound for **(ii)**,

$$
\begin{aligned}
\left\|\sum_{i\notin\mathcal{C}}\bar{h}(\lambda_i)\hat{X}(\lambda_i)\varphi_i - \sum_{i\notin\mathcal{C}}\bar{h}(\mu_i^n)\hat{X}_n(\mu_i^n)\varphi_i(T_{\mathbf{W}_n})\right\| &\leq h_{\min}\epsilon + Lc\left\|\sum_{i\notin\mathcal{C}}\hat{X}(\lambda_i)\varphi_i\right\| \\
&\leq h_{\min}\epsilon + Lc\|X\| = \epsilon .
\end{aligned} \tag{14}
$$

Putting (9) and (14) together, we have thus proved that for all $n > \max\{n_0, n_1\}$,

$$
\|Y - Y_n\| < 2\epsilon \tag{15}
$$

i.e., the output of $\bar{\mathbf{H}}(\mathbf{S}_n)$ converges to the output of $T_{\bar{\mathbf{H}}}$. ■

The Lipschitz condition on the filter h allows bounding the variability of the filter response for signal components associated with eigenvalues smaller than some $c \in [0,1]$. However, here we still need the graphon $\mathbf{W}$ to be derogatory. The main difference lies in that the WFT cannot be defined. Therefore we cannot utilize the theorem for convergence of WFT. However, this is extenuated by Proposition 1. As long as eigengaps between adjacent eigenspaces can be defined, this proposition ensures convergence not only of the eigenvectors of a graph sequence, but also of the finite-dimensional eigenspaces associated with the repeated eigenvalues of an arbitrary graphon.

Proposition 1 (Graphon subspace convergence) *Consider a sequence of graphs $\{\mathbf{G}_n\}$ with normalized eigenvalues $\mu_i^n = \lambda_i(\mathbf{S}_n)/n$, and let this sequence converge to the graphon $\mathbf{W}$ with eigenvalues λ_i. If λ_i has multiplicity m_i and $\{\mu_{i_k}^n\}_{k=1}^{m_i}$ are the eigenvalues of $\mathbf{G}_n$ converging to λ_i, then*

$$E_{T_{\mathbf{W}_n}}(\{\mu_{i_k}^n\}) \to E_{T_{\mathbf{W}}}(\lambda_i) \tag{16}$$

where $\mathbf{W}_n$ is the graphon induced by $\mathbf{G}_n$ and $E_T(\{\lambda\})$ denotes the spectral projection onto the subspace associated with the eigenvalues $\{\lambda\}$.

We conclude by presenting our most general result: filter response convergence for Lipschitz continuous graph filters and arbitrary graphons. This result is stated in Theorem 3.

Theorem **3 (Convergence of filter response for Lipschitz continuous graph filters)** *Let $\{(\mathbf{G}_n, \mathbf{y}_n)\}$ be the sequence of graph signals obtained by applying filters $\mathbf{H}_n(\mathbf{S}_n) = \mathbf{V}_n h(\mathbf{\Lambda}_n(\mathbf{S}_n)/n)\mathbf{V}_n^H$ to the sequence $\{(\mathbf{G}_n, \mathbf{x}_n)\}$, and let $(\mathbf{W}, Y)$ be the graphon signal obtained by applying the graphon filter $(T_{\mathbf{H}}X)(v) = \sum_{i=1}^{\infty} h(\lambda_i)\hat{X}(\lambda_i)\varphi_i(v)$ to the signal $(\mathbf{W}, X)$. If $\{(\mathbf{G}_n, \mathbf{x}_n)\}$ converges to $(\mathbf{W}, X)$ and the function h is L-Lipschitz, then $\{(\mathbf{G}_n, \mathbf{y}_n)\}$ converges to $(\mathbf{W}, Y)$ in the sense of Def. 1.*

Proof: In the following, we consider the normalized filter function $\bar{h}(\lambda) = h(\lambda)/\max_{\lambda \in [0,1]} h(\lambda)$ to simplify our derivations. We also consider the graphon signals $(\mathbf{W}_n, X_n)$ induced by the graph signals $(\mathbf{G}_n, \mathbf{x}_n)$ interchangeably, recalling that their spectral properties are preserved per Lemma 1. The L_2 norm is written as $\|\cdot\|_{L_2} \equiv \|\cdot\|$ to make notation less cumbersome.

In order to prove filter output convergence for sequences of graphs converging to arbitrary (possibly derogatory) graphons, we must separate the convergence analysis between spectral components associated with eigenvalues with multiplicity $m_i = 1$ and eigenvalues with multiplicity $m_i > 1$. We thus write the output graphon signal $(\mathbf{W}, Y)$ as $Y = Y^{(1)} + Y^{(2)}$, with

$$Y^{(1)} = \sum_{i \in \mathcal{M}_{=1}} \bar{h}(\lambda_i)\hat{X}(\lambda_i)\varphi_i \quad \text{and} \tag{17}$$

$$Y^{(2)} = \sum_{i \notin \mathcal{M}_{=1}} \bar{h}(\lambda_i)\Pi(X, \lambda_i) \tag{18}$$

and where $\mathcal{M}_{=1} = \{i \mid m_i = 1\}$ and $\Pi(X, \lambda_i)$ denotes the projection of $(\mathbf{W}, X)$ onto the eigenspace associated with λ_i.

As for the graphon signals $(\mathbf{W}_n, Y_n)$, their spectral decomposition is split between eigenvalues converging individually to different eigenvalues of $\mathbf{W}$, and eigenvalues that are part of a set converging to a common eigenvalue of $\mathbf{W}$, i.e., $Y_n = Y_n^{(1)} + Y_n^{(2)}$ where

$$Y_n^{(1)} = \sum_{i \in \mathcal{M}_{=1}} \bar{h}(\mu_i^n)\hat{X}_n(\mu_i^n)\varphi_i(T_{\mathbf{W}_n}) \quad \text{and} \tag{19}$$

$$Y_n^{(2)} = \sum_{i \notin \mathcal{M}_{=1}} \bar{h}(\mu_i^n)\Pi(X_n, \mu_i^n) \tag{20}$$

and where $\mathcal{M}_{=1} = \{i \mid \mu_i^n \to \lambda_j, m_j = 1\}$ and $\mu_i^n = \lambda_i(T_{\mathbf{W}_n})$.

From Theorem 2, we conclude that $Y_n^{(1)} \to Y^{(1)}$ as $n \to \infty$. It remains to show that $Y_n^{(2)} \to Y^{(2)}$. Using (17) and (19), we write

$$\|Y^{(2)} - Y_n^{(2)}\| = \left\| \sum_{i \notin \mathcal{M}_{=1}} \bar{h}(\lambda_i)\Pi(X, \lambda_i) - \sum_{i \notin \mathcal{M}_{=1}} \bar{h}(\mu_i^n)\Pi(X_n, \mu_i^n) \right\| . \tag{21}$$

These sums can be further split up by defining the set $\mathcal{C} = \{i \mid i \notin \mathcal{M}_{=1},\ |\lambda_i| \geq c\}$, where $c = (1 - h_{\text{min}})\epsilon / L\|X\|$, $h_{\text{min}} = \min_{[0,c]} \bar{h}(\lambda)$ and $\epsilon > 0$. This allows us to rewrite (21) as

$$\begin{aligned}
&\left\| \sum_{i \notin \mathcal{M}_{=1}} \bar{h}(\lambda_i)\Pi(X, \lambda_i) - \sum_{i \notin \mathcal{M}_{=1}} \bar{h}(\mu_i^n)\Pi(X_n, \mu_i^n) \right\| \\
&\leq \left\| \sum_{i \in \mathcal{C}} \bar{h}(\lambda_i)\Pi(X, \lambda_i) - \sum_{i \in \mathcal{C}} \bar{h}(\mu_i^n)\Pi(X_n, \mu_i^n) \right\| \ \textbf{(i)} \\
&+ \left\| \sum_{i \notin \mathcal{C}} \bar{h}(\lambda_i)\Pi(X, \lambda_i) - \sum_{i \notin \mathcal{C}} \bar{h}(\mu_i^n)\Pi(X_n, \mu_i^n) \right\| \ \textbf{(ii)} .
\end{aligned} \tag{22}$$

From Proposition 1 and Theorem 1, there exists n_0 such that

$$\left\| \sum_{i \in \mathcal{C}} \bar{h}(\lambda_i)\Pi(X, \lambda_i) - \sum_{i \in \mathcal{C}} \bar{h}(\mu_i^n)\Pi(X_n, \mu_i^n) \right\| < \epsilon \tag{23}$$

for all $n > n_0$, which gives a bound for **(i)**. Using the filter's Lipschitz continuity, we can also derive a bound for **(ii)**,

$$\begin{aligned}
&\left\| \sum_{i \notin \mathcal{C}} \bar{h}(\lambda_i)\Pi(X, \lambda_i) - \sum_{i \notin \mathcal{C}} \bar{h}(\mu_i^n)\Pi(X_n, \mu_i^n) \right\| \\
&\leq \left\| \sum_{i \notin \mathcal{C}} (h_{\text{min}} + Lc)\Pi(X, \lambda_i) - \sum_{i \notin \mathcal{C}} h_{\text{min}}\Pi(X_n, \mu_i^n) \right\| \\
&= \left\| h_{\text{min}} \sum_{i \notin \mathcal{C}} [\Pi(X, \lambda_i) - \Pi(X_n, \mu_i^n)] + Lc \sum_{i \notin \mathcal{C}} \Pi(X, \lambda_i) \right\| .
\end{aligned} \tag{24}$$

Noting that $\sum_{i \notin \mathcal{C}} \Pi(X, \lambda_i)$ can be written as

$$\sum_{i \notin \mathcal{C}} \Pi(X, \lambda_i) = X - \sum_{i \in \mathcal{C}} \Pi(X, \lambda_i) \tag{25}$$

and $\sum_{i \notin \mathcal{C}} \Pi(X_n, \mu_i^n)$ as

$$\sum_{i \notin \mathcal{C}} \Pi(X_n, \mu_i^n) = X - \sum_{i \in \mathcal{C}} \Pi(X_n, \mu_i^n) \tag{26}$$

and using Proposition 1 and the fact that X_n converges to X in the sense of Def. 1, we conclude that there exists n_1 such that, for all $n > n_1$,

$$\left\| \sum_{i \notin \mathcal{C}} \Pi(X, \lambda_i) - \Pi(X_n, \mu_i^n) \right\| = \left\| \sum_{i \in \mathcal{C}} \Pi(X_n, \mu_i^n) - \Pi(X, \lambda_i) \right\| < \epsilon . \tag{27}$$

Using the Cauchy-Schwarz and triangle inequalities and substituting (13) in (10), we get

$$\begin{aligned}&\left\|\sum_{i\notin\mathcal{C}}\bar{h}(\lambda_i)\Pi(X,\lambda_i)-\sum_{i\notin\mathcal{C}}\bar{h}(\mu_i^n)\Pi(X_n,\mu_i^n)\right\|\\&\leq h_{\text{min}}\epsilon+Lc\left\|\sum_{i\notin\mathcal{C}}\Pi(X,\lambda_i)\right\|\leq h_{\text{min}}\epsilon+Lc\|X\|=\epsilon\end{aligned}\tag{28}$$

by which we have thus proved

$$\|Y-Y_n\|<2\epsilon \text{ for all } n>\max\{n_0,n_1\}\tag{29}$$

i.e., the output of $\bar{\mathbf{H}}(\mathbf{S}_n)$ converges to the output of $T_{\bar{\mathbf{H}}}$. ■