Appendix A
Proof of Theorem 4

The Proof of Theorem 4. We first note that

$$\|\mathbf{\Phi}(\mathbf{H},\mathbf{L_n},\mathbf{P}_nf)-\mathbf{P_n}\mathbf{\Phi}(\mathbf{H},\mathcal{L},f)\|\leq\sum_{q=1}^{F_L}\|\mathbf{x}^q_{n,L}-\mathbf{P}_nf^q_L\|.$$

Next, we observe that for any $1\leq\ell\leq L$, $1\leq p\leq F_\ell$, we have

$$\begin{aligned}&\|\mathbf{x}^p_{n,\ell}-\mathbf{P}_nf^p_\ell\|\\ &\leq\sum_{q=1}^{F_{\ell-1}}\|h^{pq}_\ell(\mathbf{L}_n)\mathbf{x}^q_{n,\ell-1}-\mathbf{P}_nh^{pq}_\ell(\mathcal{L})f^q_{\ell-1}\|\\ &\leq\sum_{q=1}^{F_{\ell-1}}\|h^{pq}_\ell(\mathbf{L}_n)\mathbf{x}^q_{n,\ell-1}-h^{pq}_\ell(\mathbf{L}_n)\mathbf{P}_nf^q_{\ell-1}\|\\ &\quad+\sum_{q=1}^{F_{\ell-1}}\|h^{pq}_\ell(\mathbf{L}_n)\mathbf{P}_nf^q_{\ell-1}-\mathbf{P}_nh^{pq}_\ell(\mathcal{L})f^q_{\ell-1}\|.\end{aligned}\tag{7}$$

Since h^{pq}_ℓ is non-amplifying, the matrix $h^{pq}_\ell(\mathbf{L}_n)$ has operator norm at most one on $\mathbf{L}^2(\mathbf{G}_n)$. Therefore, we may bound the first term from (7) by

$$\begin{aligned}&\sum_{q=1}^{F_{\ell-1}}\|h^{pq}_\ell(\mathbf{L}_n)\mathbf{x}^q_{n,\ell-1}-h^{pq}_\ell(\mathbf{L}_n)\mathbf{P}_nf^q_{\ell-1}\|\\ &\qquad\leq F_{\ell-1}\max_{1\leq q\leq F_{L-1}}\|\mathbf{x}^q_{n,\ell-1}-\mathbf{P}_nf^q_{\ell-1}\|.\end{aligned}\tag{8}$$

For notational clarity, we will now temporarily drop the feature index q and the layer index l. To bound the second term, we note that:

$$\begin{aligned}&\|h(\mathbf{L}_n)\mathbf{P}_nf-\mathbf{P}_nh(\mathcal{L})f\|\\ &=\left\|\sum_{i=1}^{\kappa}\widehat{h}(\lambda^n_i)\langle\mathbf{P}_nf,\phi^n_i\rangle_{\mathbf{G}_n}\phi^n_i-\sum_{i=1}^{\kappa}\widehat{h}(\lambda_i)\langle f,\phi_i\rangle_{\mathcal{M}}\mathbf{P}_n\phi_i\right\|\\ &\leq\left\|\sum_{i=1}^{\kappa}(\widehat{h}(\lambda^n_i)-\widehat{h}(\lambda_i))\langle\mathbf{P}_nf,\phi^n_i\rangle_{\mathbf{G}_n}\phi^n_i\right\| &(9)\\ &\quad+\left\|\sum_{i=1}^{\kappa}\widehat{h}(\lambda_i)(\langle\mathbf{P}_nf,\phi^n_i\rangle_{\mathbf{G}_n}\phi^n_i-\langle f,\phi_i\rangle_{\mathcal{M}}\mathbf{P}_n\phi_i)\right\|. &(10)\end{aligned}$$

To bound the term from (9), we note that by the triangle inequality and the Cauchy-Schwarz inequality we have

$$\begin{aligned}&\left\|\sum_{i=1}^{\kappa}(\widehat{h}(\lambda^n_i)-\widehat{h}(\lambda_i))\langle\mathbf{P}_nf,\phi^n_i\rangle_{\mathbf{G}_n}\phi^n_i\right\|\\ &\qquad\leq\max_{1\leq i\leq\kappa}|\widehat{h}(\lambda^n_i)-\widehat{h}(\lambda_i)|\sum_{i=1}^{\kappa}\|\mathbf{P}_nf\|\|\phi^n_i\|^2.\end{aligned}$$

By Remark 1 and the assumption that all of the h^{pq}_ℓ are C-Lipschitz, we have

$$\max_{1\leq i\leq\kappa}|\widehat{h}(\lambda^n_i)-\widehat{h}(\lambda_i)|\leq C\mathcal{O}\left(n^{-\frac{2}{d+6}}\sqrt{\log n}\right),$$

and by Lemma 1, we have

$$\|\mathbf{P}_nf\|^2\leq\|f\|_2^2+\sqrt{\frac{18\log n}{n}}\|f\|^2_\infty.$$

Therefore, using fact that $\sqrt{a^2+b^2}\leq|a|+|b|$ for all $a,b\in\mathbb{R}$, we obtain

$$\begin{aligned}&\left\|\sum_{i=1}^{\kappa}(\widehat{h}(\lambda^n_i)-\widehat{h}(\lambda_i))\langle\mathbf{P}_nf,\phi^n_i\rangle_{\mathbf{G}_n}\phi^n_i\right\|\\ &\leq C\kappa\mathcal{O}\left(\frac{\sqrt{\log n}}{n^{2/(d+6)}}\right)\sqrt{\|f\|_2^2+\sqrt{\frac{18\log n}{n}}\|f\|^2_\infty}\\ &\leq C\kappa\mathcal{O}\left(\frac{\sqrt{\log n}}{n^{2/(d+6)}}\right)\left(\|f\|_2+\mathcal{O}\left(\left(\frac{\log n}{n}\right)^{1/4}\right)\|f\|_\infty\right).\end{aligned}\tag{11}$$

Now, turning our attention to the terms from (10), we have

$$\begin{aligned}&\left\|\sum_{i=1}^{\kappa}\widehat{h}(\lambda_i)(\langle\mathbf{P}_nf,\phi^n_i\rangle_{\mathbf{G}_n}\phi^n_i-\langle f,\phi_i\rangle_{\mathcal{M}}\mathbf{P}_n\phi_i)\right\|\\ &\leq\left\|\sum_{i=1}^{\kappa}\widehat{h}(\lambda_i)\langle\mathbf{P}_nf,\phi^n_i\rangle_{\mathbf{G}_n}(\phi^n_i-\mathbf{P}_n\phi_i)\right\| &(12)\\ &\quad+\left\|\sum_{i=1}^{\kappa}\widehat{h}(\lambda_i)(\langle\mathbf{P}_nf,\phi^n_i\rangle_{\mathbf{G}_n}-\langle f,\phi_i\rangle_{\mathcal{M}})\mathbf{P}_n\phi_i\right\|. &(13)\end{aligned}$$

By Remark 1, we have $\|\phi^n_i-\mathbf{P}_n\phi_i\|=\mathcal{O}\left(n^{-\frac{2}{d+6}}\sqrt{\log n}\right)$. Therefore, since the frequency responses of $\mathbf{H}$ are non-amplifying, the term from (12) can be bounded by

$$\begin{aligned}&\left\|\sum_{i=1}^{\kappa}\widehat{h}(\lambda_i)\langle\mathbf{P}_nf,\phi^n_i\rangle_{\mathbf{G}_n}(\phi^n_i-\mathbf{P}_n\phi_i)\right\|\\ &\leq\kappa\max_{1\leq i\leq\kappa}|\langle\mathbf{P}_nf,\phi^n_i\rangle_{\mathbf{G}_n}|\|\phi^n_i-\mathbf{P}_n\phi_i\|\\ &\leq\kappa\mathcal{O}\left(\frac{\sqrt{\log n}}{n^{2/(d+6)}}\right)\|\mathbf{P}_nf\|\\ &\leq\kappa\mathcal{O}\left(\frac{\sqrt{\log n}}{n^{2/(d+6)}}\right)\left(\|f\|_2+\mathcal{O}\left(\left(\frac{\log n}{n}\right)^{1/4}\right)\|f\|_\infty\right),\end{aligned}\tag{14}$$

where the final inequality again follows by using Lemma 1. Meanwhile, the term from (13) can be bounded by

$$\begin{aligned}&\left\|\sum_{i=1}^{\kappa}\widehat{h}(\lambda_i)(\langle\mathbf{P}_nf,\phi^n_i\rangle_{\mathbf{G}_n}-\langle f,\phi_i\rangle_{\mathcal{M}})\mathbf{P}_n\phi_i\right\|\\ &\leq\sum_{i=1}^{\kappa}|\widehat{h}(\lambda_i)||\langle\langle\mathbf{P}_nf,\phi^n_i\rangle_{\mathbf{G}_n}-\langle f,\phi_i\rangle_{\mathcal{M}}||\mathbf{P}_n\phi_i\|\\ &\leq\sum_{i=1}^{\kappa}|\langle\mathbf{P}_nf,\phi^n_i\rangle_{\mathbf{G}_n}-\langle f,\phi_i\rangle_{\mathcal{M}}||\mathbf{P}_n\phi_i\|\\ &\leq\sum_{i=1}^{\kappa}|\langle\mathbf{P}_nf,\phi^n_i\rangle_{\mathbf{G}_n}-\langle\mathbf{P}_nf,\mathbf{P}_n\phi_i\rangle_{\mathbf{G}_n}||\mathbf{P}_n\phi_i\|\\ &\qquad+\sum_{i=1}^{\kappa}|\langle\mathbf{P}_nf,\mathbf{P}_n\phi_i\rangle_{\mathbf{G}_n}-\langle f,\phi_i\rangle_{\mathcal{M}}||\mathbf{P}_n\phi_i\|.\end{aligned}$$

By the Cauchy-Schwarz inequality, Remark 1, and Lemma 1, we have

$$
\begin{aligned}
&|\langle \mathbf{P}_n f, \phi_i^n\rangle - \langle \mathbf{P}_n f, \mathbf{P}_n \phi_i\rangle| \\
&\leq \|\mathbf{P}_n f\| \|\phi_i^n - \mathbf{P}_n \phi_i\| \\
&\leq \mathcal{O}\left(\frac{\sqrt{\log n}}{n^{-2/d+6}}\right)\left(\|f\|_2 + \mathcal{O}\left(\left(\frac{\log n}{n}\right)^{1/4}\right)\|f\|_\infty\right).
\end{aligned}
$$

Again using Lemma 1, we have

$$
|\langle \mathbf{P}_n f, \mathbf{P}_n \phi_i\rangle - \langle f, \phi_i\rangle| \leq \mathcal{O}\left(\left(\frac{\log n}{n}\right)^{1/2}\right)\|f\|_\infty \|\phi_i\|_\infty
$$

and also that

$$
\|\mathbf{P}_n \phi_i\| \leq 1 + \mathcal{O}\left(\left(\frac{\log n}{n}\right)^{1/4}\right)\|\phi_i\|_\infty.
$$

It is known (see Appendix L of [17] and the references there) that $\|\phi_i\|_\infty \leq C_{\mathcal{M}} i^{(d-1)/2d} \leq C_{\mathcal{M}} i^{1/2}$. Therefore, for all $i \leq \kappa$ we have

$$
\begin{aligned}
|\langle \mathbf{P}_n f, \mathbf{P}_n \phi_i\rangle - \langle f, \phi_i\rangle| &\leq \mathcal{O}\left(\left(\frac{\log n}{n}\right)^{1/2}\right)\kappa^{1/2}\|f\|_\infty \\
&= \mathcal{O}\left(\left(\frac{\log n}{n}\right)^{1/2}\right)\|f\|_\infty,
\end{aligned}
$$

(where in the second inequality we used the fact that the constants implied by big-$\mathcal{O}$ notation depend on κ). Similarly, we also have

$$
\begin{aligned}
\|\mathbf{P}_n \phi_i\| &\leq 1 + \mathcal{O}\left(\left(\frac{\log n}{n}\right)^{1/4}\right)\kappa^{1/2} \\
&= 1 + \mathcal{O}\left(\left(\frac{\log n}{n}\right)^{1/4}\right) \\
&= \mathcal{O}(1).
\end{aligned}
$$

Therefore, when $d \geq 2$, we have

$$
\begin{aligned}
&\left\|\sum_{i=1}^{\kappa} \widehat{h}(\lambda_i)(\langle \mathbf{P}_n f, \phi_i^n\rangle_{\mathbf{G}_n} - \langle f_{\ell-1}^q, \phi_i\rangle_{\mathcal{M}})\mathbf{P}_n \phi_i\right\| \\
\leq &\sum_{i=1}^{\kappa} |\langle \mathbf{P}_n f, \phi_i^n\rangle - \langle \mathbf{P}_n f, \mathbf{P}_n \phi_i\rangle| \|\mathbf{P}_n \phi_i\| \\
&\quad + \sum_{i=1}^{\kappa} |\langle \mathbf{P}_n f, \mathbf{P}_n \phi_i\rangle - \langle f, \phi_i\rangle| \|\mathbf{P}_n \phi_i\| \\
\leq &\sum_{i=1}^{\kappa} \mathcal{O}\left(\frac{\sqrt{\log n}}{n^{2/(d+6)}}\right)\left(\|f\|_2 + \mathcal{O}\left(\left(\frac{\log n}{n}\right)^{1/4}\right)\|f\|_\infty\right) \\
&\quad + \sum_{i=1}^{\kappa} \mathcal{O}\left(\left(\frac{\log n}{n}\right)^{1/2}\right)\|f\|_\infty \\
\leq &\kappa\mathcal{O}\left(\frac{\sqrt{\log n}}{n^{2/(d+6)}}\right)\left(\|f\|_2 + \mathcal{O}\left(\left(\frac{\log n}{n}\right)^{1/4}\right)\|f\|_\infty\right) \\
&\quad + \kappa\mathcal{O}\left(\left(\frac{\log n}{n}\right)^{1/2}\right)\|f\|_\infty \qquad (15) \\
\leq &\kappa\left(\mathcal{O}\left(\frac{\sqrt{\log n}}{n^{2/(d+6)}}\right)\left(\|f\|_2 + \mathcal{O}\left(\left(\frac{\log n}{n}\right)^{1/4}\right)\|f\|_\infty\right)\right) \qquad (16)
\end{aligned}
$$

where in the last line we used the fact that $d \geq 2$.

Therefore, combining Equations (11) through (16) and reinstating the feature and layer indices, we have

$$
\begin{aligned}
&\|h_\ell^{pq}(\mathbf{L}_n)\mathbf{P}_n f_{\ell-1}^q - \mathbf{P}_n h_\ell^{pq}(\mathcal{L}) f_{\ell-1}^q\| \\
\leq &\left\|\sum_{i=1}^{\kappa} (\widehat{h}_\ell^{pq}(\lambda_i^n) - \widehat{h}_\ell^{pq}(\lambda_i))\langle \mathbf{P}_n f_{\ell-1}^q, \phi_i^n\rangle_{\mathbf{G}_n} \phi_i^n\right\| \\
&+ \left\|\sum_{i=1}^{\kappa} \widehat{h}_\ell^{pq}(\lambda_i)(\langle \mathbf{P}_n f_{\ell-1}^q, \phi_i^n\rangle_{\mathbf{G}_n} \phi_i^n - \langle f_{\ell-1}^q, \phi_i\rangle_{\mathcal{M}} \mathbf{P}_n \phi_i)\right\| \\
\leq &C\kappa\mathcal{O}\left(\frac{\sqrt{\log n}}{n^{2/(d+6)}}\right)\left(\|f_{\ell-1}^q\|_2 + \mathcal{O}\left(\left(\frac{\log n}{n}\right)^{1/4}\right)\|f_{\ell-1}^q\|_\infty\right) \\
&+ \kappa\mathcal{O}\left(\frac{\sqrt{\log n}}{n^{2/(d+6)}}\right)\left(\|f_{\ell-1}^q\|_2 + \mathcal{O}\left(\left(\frac{\log n}{n}\right)^{1/4}\right)\|f_{\ell-1}^q\|_\infty\right) \\
&+ \kappa\mathcal{O}\left(\frac{\sqrt{\log n}}{n^{2/(d+6)}}\right)\left(\|f_{\ell-1}^q\|_2 + \mathcal{O}\left(\left(\frac{\log n}{n}\right)^{1/4}\right)\|f_{\ell-1}^q\|_\infty\right) \\
\leq &\tilde{C}\mathcal{O}\left(\frac{\sqrt{\log n}}{n^{2/(d+6)}}\right)\left(\|f_{\ell-1}^q\|_2 + \mathcal{O}\left(\left(\frac{\log n}{n}\right)^{1/4}\right)\|f_{\ell-1}^q\|_\infty\right),
\end{aligned}
$$

where $\tilde{C} = \max\{C, 1\}$ and in the final line we have absorbed κ into the implied constant.

When $d = 1$, we repeat the same string of inequalities up to eq. (15) and obtain

$$\left\|\sum_{i=1}^{\kappa}\widehat{h}(\lambda_i)(\langle\mathbf{P}_nf,\phi_i^n\rangle_{\mathbf{G}_n}\mathbf{P}_n\phi_i-\langle f,\phi_i\rangle_{\mathcal{M}})\mathbf{P}_n\phi_i\right\|$$
$$\leq\kappa\mathcal{O}\left(\frac{\sqrt{\log n}}{n^{2/7}}\right)\left(\|f_{\ell-1}^q\|_2+\mathcal{O}\left(\left(\frac{\log n}{n}\right)^{1/4}\right)\|f_{\ell-1}^q\|_\infty\right)$$
$$+\kappa\mathcal{O}\left(\left(\frac{\log n}{n}\right)^{1/2}\right)\|f_{\ell-1}^q\|_\infty$$
$$\leq\kappa\left(\mathcal{O}\left(\frac{\sqrt{\log n}}{n^{2/7}}\right)\|f_{\ell-1}^q\|_2+\mathcal{O}\left(\frac{\sqrt{\log n}}{n^{1/2}}\right)\|f_{\ell-1}^q\|_\infty\right) \tag{17}$$

Then, combining eqs. (11), (14) and (17), and again absorbing κ into the implied constant we obtain

$$\|h_\ell^{pq}(\mathbf{L}_n)\mathbf{P}_nf_{\ell-1}^q-\mathbf{P}_nh_\ell^{pq}(\mathcal{L})f_{\ell-1}^q\|$$
$$\leq\left\|\sum_{i=1}^{\kappa}(\widehat{h}_\ell^{pq}(\lambda_i^n)-\widehat{h}_\ell^{pq}(\lambda_i))\langle\mathbf{P}_nf_{\ell-1}^q,\phi_i^n\rangle_{\mathbf{G}_n}\phi_i^n\right\|$$
$$+\left\|\sum_{i=1}^{\kappa}\widehat{h}_\ell^{pq}(\lambda_i)(\langle\mathbf{P}_nf_{\ell-1}^q,\phi_i^n\rangle_{\mathbf{G}_n}\phi_i^n-\langle f_{\ell-1}^q,\phi_i\rangle_{\mathcal{M}})\mathbf{P}_n\phi_i\right\|$$
$$\leq C\kappa\mathcal{O}\left(\frac{\sqrt{\log n}}{n^{2/7}}\right)\left(\|f_{\ell-1}^q\|_2+\mathcal{O}\left(\left(\frac{\log n}{n}\right)^{1/4}\right)\|f_{\ell-1}^q\|_\infty\right)$$
$$+\kappa\mathcal{O}\left(\frac{\sqrt{\log n}}{n^{2/7}}\right)\left(\|f_{\ell-1}^q\|_2+\mathcal{O}\left(\left(\frac{\log n}{n}\right)^{1/4}\right)\|f_{\ell-1}^q\|_\infty\right)$$
$$+\kappa\left(\mathcal{O}\left(\frac{\sqrt{\log n}}{n^{2/7}}\right)\|f_{\ell-1}^q\|_2+\mathcal{O}\left(\frac{\sqrt{\log n}}{n^{1/2}}\right)\|f_{\ell-1}^q\|_\infty\right)$$
$$\leq\tilde{C}\left(\mathcal{O}\left(\frac{\sqrt{\log n}}{n^{2/7}}\right)\|f_{\ell-1}^q\|_2+\mathcal{O}\left(\frac{\sqrt{\log n}}{n^{1/2}}\right)\|f_{\ell-1}^q\|_\infty\right).$$

Thus by (7) and (8), we have derived the relationship that, when $d\geq 2$,

$$\|\mathbf{x}_{n,\ell}^p-\mathbf{P}_nf_\ell^p\|\leq F_{\ell-1}\max_{1\leq q\leq F_{L-1}}\|\mathbf{x}_{n,\ell-1}^q-\mathbf{P}_nf_{\ell-1}^q\|$$
$$+F_{\ell-1}\tilde{C}\mathcal{O}\left(\frac{\sqrt{\log n}}{n^{2/(d+6)}}\right)\|f_{\ell-1}^q\|_2$$
$$+F_{\ell-1}\tilde{C}\mathcal{O}\left(\frac{(\log n)^{3/4}}{n^{1/4+2/(d+6)}}\right)\|f_{\ell-1}^q\|_\infty$$

and when $d=1$,

$$\|\mathbf{x}_{n,\ell}^p-\mathbf{P}_nf_\ell^p\|$$
$$\leq F_{\ell-1}\max_{1\leq q\leq F_{L-1}}\|\mathbf{x}_{n,\ell-1}^q-\mathbf{P}_nf_{\ell-1}^q\|$$
$$+F_{\ell-1}\tilde{C}\left(\mathcal{O}\left(\frac{\sqrt{\log n}}{n^{2/7}}\right)\|f_{\ell-1}^q\|_2+\mathcal{O}\left(\frac{\sqrt{\log n}}{n^{1/2}}\right)\|f_{\ell-1}^q\|_\infty\right).$$

Let $\epsilon_{n,\ell-1}=\max_{1\leq q\leq F_{L-1}}\|\mathbf{x}_{n,\ell-1}^q-\mathbf{P}_nf_{\ell-1}^q\|$ and define

$$\delta_n=\begin{cases}\tilde{C}\mathcal{O}\left(\frac{\sqrt{\log n}}{n^{2/(d+6)}}\right)\max\limits_{\ell,q}\|f_\ell^q\| & \\ \quad+\tilde{C}\mathcal{O}\left(\frac{(\log n)^{3/4}}{n^{1/4+2/(d+6)}}\right)\max\limits_{\ell,q}\|f_\ell^q\|_\infty & \text{if } d\geq 2\\ \tilde{C}\mathcal{O}\left(\frac{\sqrt{\log n}}{n^{2/7}}\right)\max\limits_{\ell,q}\|f_\ell^q\| & \\ \quad+\tilde{C}\mathcal{O}\left(\frac{\sqrt{\log n}}{n^{1/2}}\right)\max\limits_{\ell,q}\|f_\ell^q\|_\infty & \text{if } d=1\end{cases}.$$

Then, we have the recurrence relation

$$\epsilon_{n,\ell}\leq F_{\ell-1}(\epsilon_{n,\ell-1}+\delta_n.)$$

Therefore, one may verify by induction that for all $\ell\geq 1$ we have

$$\epsilon_{n,\ell}\leq\delta_n\sum_{k=1}^{\ell}\prod_{j=\ell-k}^{\ell-1}F_j.$$

The proof now follows from the definitions of $\epsilon_{n,\ell}$ and δ_n. □