Chapter 10

Final arguments

We are now ready to prove the main results, Theorems 3.4–3.12. In several of these proofs, we require the following definition. Let $s \in \mathbb{N}$ and set

$$k(s) = \max\{|S|_{\boldsymbol{u}} : S \subset \mathbb{N}_{\mathbf{0}}^{d},\ |S| \leq s,\ S \text{ lower}\}, \tag{10.1}$$

where $\boldsymbol{u}$ are the intrinsic weights (4.7) (recall the definition of a lower set from Definition 2.8). It can be shown that

$$k(s) = s^2, \quad \text{(Legendre)}, \quad k(s) \leq \min\left\{2^d s, s^{\log(3)/\log(2)}\right\}, \quad \text{(Chebyshev)}. \tag{10.2}$$

See, e.g., [8, equation (7.42), Propositions 5.13 and 5.17]. We will use this property several times in what follows.

10.1 Algebraic rates of convergence, finite dimensions

Proof of Theorem 3.4. The mapping was described in Table 4.1. As shown therein, we can write the corresponding approximation as $\hat{f} = \sum_{\boldsymbol{\nu} \in \Lambda} \hat{c}_{\boldsymbol{\nu}} \Psi_{\boldsymbol{\nu}}$, where $\hat{\boldsymbol{c}} = (\hat{c}_{\boldsymbol{\nu}})_{\boldsymbol{\nu} \in \Lambda}$ is a minimizer of (4.6). Next, due to the various assumptions made, we may apply Theorem 8.2. Setting $\tilde{f} = \hat{f}$ and $\tilde{\boldsymbol{c}} = \hat{\boldsymbol{c}}$, we deduce that

$$\|f - \hat{f}\|_{L^2_{\varrho}(\mathcal{U};\mathcal{V})} \leq c_1 \cdot \xi, \quad \|f - \hat{f}\|_{L^\infty(\mathcal{U};\mathcal{V})} \leq c_2 \cdot \sqrt{k} \cdot \xi, \tag{10.3}$$

where (after writing out the term $E_{h,\infty}(f)$ explicitly)

$$\begin{aligned} \xi = {} & \frac{\sigma_k(\boldsymbol{c}_\Lambda)_{1,\boldsymbol{u};\mathcal{V}}}{\sqrt{k}} + \frac{E_{\Lambda,\infty}(f)}{\sqrt{k}} + E_{\Lambda,2}(f) + \|f - \mathcal{P}_h(f)\|_{L^\infty(\mathcal{U};\mathcal{V})} \\ & + \mathcal{G}(\hat{\boldsymbol{c}}) - \mathcal{G}(\mathcal{P}_h(\boldsymbol{c}_\Lambda)) + \frac{\|\boldsymbol{n}\|_{2;\mathcal{V}}}{\sqrt{m}}, \end{aligned} \tag{10.4}$$

and $k = m/(c_0 L)$ with $c_0 \geq 1$ a universal constant. We now bound each term separately.

Step 1. The terms $\sigma_k(\boldsymbol{c}_\Lambda)_{1,\boldsymbol{u};\mathcal{V}}/\sqrt{k}$, $E_{\Lambda,\infty}(f)/\sqrt{k}$ and $E_{\Lambda,2}(f)$. The term

$$\sigma_k(\boldsymbol{c}_\Lambda)_{1,\boldsymbol{u};\mathcal{V}}/\sqrt{k}$$

is estimated via (ii) of Theorem A.1 with $q = 1$. This gives

$$\frac{\sigma_k(\boldsymbol{c}_\Lambda)_{1,\boldsymbol{u};\mathcal{V}}}{\sqrt{k}} \leq C(d,p,\boldsymbol{\rho}) \cdot k^{1/2-1/p} = C(d,p,\boldsymbol{\rho}) \cdot \left(\frac{m}{c_0 L}\right)^{1/2-1/p}. \tag{10.5}$$

We estimate the term $E_{\Lambda,2}(f)$ by first recalling that $\Lambda = \Lambda^{\mathsf{HC}}_{n,d}$ is the union of all lower sets (see Definition 2.8) of size at most $n = \lceil m/L \rceil$ (see Section 3.2). Hence, using (i) of Theorem A.1 with $s = n$ and $q = 2$, we get

$$\begin{aligned} E_{\Lambda,2}(f) = \|\boldsymbol{c} - \boldsymbol{c}_\Lambda\|_{2;\mathcal{V}} \leq \|\boldsymbol{c} - \boldsymbol{c}_S\|_{2;\mathcal{V}} &\leq C(d,p,\boldsymbol{\rho}) \cdot n^{1/2-1/p} \\ &\leq C(d,p,\boldsymbol{\rho}) \cdot \left(\frac{m}{c_0 L}\right)^{1/2-1/p}. \end{aligned} \tag{10.6}$$

Here, in the last step we recall that $n = \lceil m/L \rceil$ and $c_0 \geq 1$.

It remains to consider $E_{\Lambda,\infty}(f)/\sqrt{k}$. Due to the choice of weights, we have $E_{\Lambda,\infty}(f) \leq \|\boldsymbol{c} - \boldsymbol{c}_\Lambda\|_{1,\boldsymbol{u};\mathcal{V}}$. We now apply (i) of Theorem A.1 once more, with $s = n$ and $q = 1$, to get

$$E_{\Lambda,\infty}(f) \leq \|\boldsymbol{c} - \boldsymbol{c}_S\|_{1,\boldsymbol{u};\mathcal{V}} \leq C(d,p,\boldsymbol{\rho}) \cdot n^{1-1/p}.$$

Since $n = \lceil m/L \rceil \geq m/(c_0 L) = k$, we obtain

$$\frac{E_{\Lambda,\infty}(f)}{\sqrt{k}} \leq C(d,p,\boldsymbol{\rho}) \cdot \left(\frac{m}{c_0 L}\right)^{1/2-1/p}. \tag{10.7}$$

Step 2. The term $\mathcal{G}(\hat{\boldsymbol{c}}) - \mathcal{G}(\mathcal{P}_h(\boldsymbol{c}_\Lambda))$. Since $\hat{\boldsymbol{c}}$ is a minimizer of (4.6) and $\mathcal{P}_h(\boldsymbol{c}_\Lambda) \in \mathcal{V}_h^N$ is feasible for (4.6), this term satisfies

$$\mathcal{G}(\hat{\boldsymbol{c}}) - \mathcal{G}(\mathcal{P}_h(\boldsymbol{c}_\Lambda)) \leq 0. \tag{10.8}$$

Step 3. Conclusion. We now substitute the bounds (10.5)–(10.8) into (10.4). Since $k \leq m/L$, we deduce that $\xi \leq \zeta$, where ζ is given by (3.10). This completes the proof. ∎

Proof of Theorem 3.5. The argument is similar to that of the previous theorem. Recall from Section 4.5 that, in this case the approximation $\hat{f} = \sum_{\boldsymbol{\nu}\in\Lambda} \tilde{c}_{\boldsymbol{\nu}} \Psi_{\boldsymbol{\nu}}$, where $\hat{\boldsymbol{c}} = \bar{\boldsymbol{c}}^{(T)}$ is the ergodic sequence obtained after T steps of the primal-dual iteration applied to (4.6). Hence, the only difference is the estimation of $\mathcal{G}(\hat{\boldsymbol{c}}) - \mathcal{G}(\mathcal{P}_h(\boldsymbol{c}_\Lambda))$ in Step 2.

We now do this using Lemma 9.2. In order to apply this lemma we first need to estimate $\|\boldsymbol{A}\|_{\mathcal{B}(\mathcal{V}_h^N,\mathcal{V}_h^m)}$. Let $\boldsymbol{x} = (x_{\boldsymbol{\nu}})_{\boldsymbol{\nu}\in\Lambda} \in \mathcal{V}_h^N$ and define $p(\boldsymbol{y}) = \sum_{\boldsymbol{\nu}\in\Lambda} x_{\boldsymbol{\nu}} \Psi_{\boldsymbol{\nu}}$. Then

$$\|\boldsymbol{A}\boldsymbol{x}\|_{2;\mathcal{V}} = \sqrt{\frac{1}{m}\sum_{i=1}^m \|p(\boldsymbol{y}_i)\|_{\mathcal{V}}^2} \leq \sup_{\boldsymbol{y}\in\mathcal{U}} \|p(\boldsymbol{y})\|_{\mathcal{V}} \leq \sum_{\boldsymbol{\nu}\in\Lambda} \|x_{\boldsymbol{\nu}}\|_{\mathcal{V}} u_{\boldsymbol{\nu}} \leq \|\boldsymbol{x}\|_{2;\mathcal{V}} \sqrt{|\Lambda|_{\boldsymbol{u}}}.$$

Now the set Λ is lower and of cardinality $|\Lambda| = \Theta(n,d)$. Hence, by (10.2) with $s = N$, we have $|\Lambda|_{\boldsymbol{u}} \leq (\Theta(n,d))^{2\alpha}$, where α is as in (3.7). Since $\boldsymbol{x}$ was arbitrary, we get

$$\|\boldsymbol{A}\|_{2;\mathcal{V}} \leq (\Theta(n,d))^{\alpha}. \tag{10.9}$$

Since the primal-dual iteration in Section 4.5 is used with $\tau = \sigma = (\Theta(n,d))^{-\alpha}$, we have that

$$\|A\|_{2;\mathcal{V}}^2 \le (\tau\sigma)^{-1}.$$

Hence, we may apply Lemma 9.2. Since the iteration is also initialized with the zero vector and run for a total of $T = \lceil 2(\Theta(n,d))^\alpha t \rceil$ iterations (see Section 4.5 once more), this gives

$$\mathcal{G}(\hat{c}) - \mathcal{G}(\mathcal{P}_h(c_\Lambda)) \le (\Theta(n,d))^\alpha \frac{\|\mathcal{P}_h(c_\Lambda)\|_{2;\mathcal{V}}^2 + 1}{T}.$$

Observe that

$$\|\mathcal{P}_h(c_\Lambda)\|_{2;\mathcal{V}} \le \|c_\Lambda\|_{2;\mathcal{V}} \le \|c\|_{c;\mathcal{V}} = \|f\|_{L^2_\varrho(\mathcal{U};\mathcal{V})} \le 1.$$

Here, in the last step, we use the fact that $f \in \mathcal{B}(\rho)$, and therefore

$$\|f\|_{L^2_\varrho(\mathcal{U};\mathcal{V})} \le \|f\|_{L^\infty(\mathcal{U};\mathcal{V})} \le 1.$$

Using this and the value of T, we deduce that

$$\mathcal{G}(\hat{c}) - \mathcal{G}(\mathcal{P}_h(c_\Lambda)) \le \frac{1}{t}.$$

Substituting this into (10.4) and combining with the other estimates (10.5)–(10.7) derived in Step 2 of the proof of Theorem 3.4 now gives the desired error bound.

It remains to estimate the computational cost. We do this via Lemmas 4.3 and 4.4. First observe that the value k in Lemma 4.4 is equal to $k = d$ in this case, since the index set $\Lambda = \Lambda^{\mathsf{HC}}_{n,d}$ is a d-dimensional hyperbolic cross index set. Similarly, the value n in Lemma 4.4 is bounded by the order n of this hyperbolic cross. As Λ is a lower set, we also have $n \le N$. Hence, the computational cost for forming the matrix A is bounded by $c \cdot m \cdot N \cdot d$. We now use Lemma 4.3 to bound the computational cost of the algorithm. Finally, we recall that $N = \Theta(n,d)$ and $T = \lceil 2(\Theta(n,d))^\alpha t \rceil$ in this case. ∎

Proof of Theorem 3.6. As in the previous proof, we only need to estimate the term $\mathcal{G}(\hat{c}) - \mathcal{G}(\mathcal{P}_h(c_\Lambda))$. Recall from Table 4.3 that in this case $\hat{c} = \tilde{c}^{(R)}$ is the output of the restarted primal-dual iteration with R restarts. Our goal is to use Theorem 9.4 applied to the problem (4.6) with weights $w = u$ as in (4.7), $\lambda = (4\sqrt{m/L})^{-1}$ and $x = \mathcal{P}_h(c_\Lambda)$.

We first show that the conditions of this theorem hold. Recall from Step 2 of the proof of Theorem 8.2 that the matrix A has the weighted rNSP of order (k, u) over $\mathcal{V}_h$ with constants $\rho = 2\sqrt{2}/3$ and $\gamma = 2\sqrt{5}/3$. In particular,

$$\frac{(1+\rho)^2}{(3+\rho)\gamma} \ge 0.64.$$

We now use (8.5) to see that

$$\lambda = \frac{1}{4\sqrt{c_0}} \frac{1}{\sqrt{k}} \leq \frac{(1+\rho)^2}{(3+\rho)\gamma} \frac{1}{\sqrt{k}},$$

for a sufficiently large choice of c_0.

Next, with this choice of $\boldsymbol{x}$, we see that

$$\xi(\boldsymbol{x}, \boldsymbol{b}) = \frac{\sigma_k(\mathcal{P}_h(\boldsymbol{c}_\Lambda))_{1,\boldsymbol{u};\mathcal{V}}}{\sqrt{k}} + \|\boldsymbol{A}\mathcal{P}_h(\boldsymbol{c}_\Lambda) - \boldsymbol{b}\|_{2;\mathcal{V}}.$$

Using (8.6) and (8.8), we get

$$\xi(\boldsymbol{x}, \boldsymbol{b}) \leq \frac{\sigma_k(\boldsymbol{c}_\Lambda)_{1,\boldsymbol{w};\mathcal{V}}}{\sqrt{k}} + \sqrt{2}\left(\frac{E_{\Lambda,\infty}(f)}{\sqrt{k}} + E_{\Lambda,2}(f)\right) + E_{h,\infty}(f) + \frac{\|\boldsymbol{n}\|_{2;\mathcal{V}}}{\sqrt{m}},$$

with probability at least $1-\epsilon$. Using (10.5)–(10.7), we deduce that

$$\xi(\boldsymbol{x}, \boldsymbol{b}) \leq \zeta,$$

with probability at least $1-\epsilon$, where ζ is as in (3.10). Hence, $\xi(\boldsymbol{x}, \boldsymbol{b}) \leq \zeta'$.

Next, recall from Table 4.3 that $\tau = \sigma = (\Theta(n,d))^{-\alpha}$ in this case. Due to (10.9), we see that $\|\boldsymbol{A}\|_{2;\mathcal{V}} \leq (\tau\sigma)^{-1}$ as well.

Now consider the constant C defined in (9.4). The values for ρ and γ give that $C_1' \leq C_2' \leq 103$. Since $\lambda = c/\sqrt{k}$ with $c = 1/(4\sqrt{c_0})$, we see that

$$4C \leq 812/c = 3296\sqrt{c_0} := c^\star. \tag{10.10}$$

Therefore, recalling that $r = 1/2$ and $\tau = \sigma = (\Theta(n,d))^{-\alpha}$, we see that

$$\left\lceil \frac{2C}{r\sqrt{\sigma\tau}} \right\rceil = \lceil (\Theta(n,d))^\alpha c^\star \rceil = T,$$

where T is as specified in Table 4.3, and

$$\frac{1}{2} r\sigma(\varepsilon_k + \zeta')T = \frac{(\Theta(n,d))^\alpha T}{4}\varepsilon_{k+1} = s\varepsilon_{k+1} = a_k,$$

where s and a_k are as specified in Table 4.3 and Algorithm 4, respectively.

With this in hand, we are now finally in a position to apply Theorem 9.4. We deduce that

$$\mathcal{G}(\hat{\boldsymbol{c}}) - \mathcal{G}(\mathcal{P}_h(\boldsymbol{c}_\Lambda)) = \mathcal{E}(\tilde{\boldsymbol{c}}^{(R)}, \mathcal{P}_h(\boldsymbol{c}_\Lambda), \boldsymbol{b}) \leq \varepsilon_k = \mathrm{e}^{-R}\|\boldsymbol{b}\|_{2;\mathcal{V}} + \zeta'.$$

To complete the proof of the error bound (3.12), we simply note that $\|\boldsymbol{b}\|_{2;\mathcal{V}} \leq \|f\|_{L^\infty(\mathcal{U};\mathcal{V})} \leq 1$, since $f \in \mathcal{B}(\boldsymbol{\rho})$.

It remains to estimate the computational cost. As before, the computational cost for forming the matrix $\boldsymbol{A}$ is bounded by $c \cdot m \cdot N \cdot d$. Next, by construction, we observe that the algorithm consists of $R = t$ primal-dual iterations, each involving $T = \lceil (\Theta(n,d))^{\alpha} c^{\star} \rceil$ steps. Therefore, by Lemma 4.3 the computational cost for the algorithm is

$$c \cdot (m \cdot N \cdot K + (m + N) \cdot (F(\boldsymbol{G}) + K)) \cdot \lceil (\Theta(n,d))^{\alpha} c^{\star} \rceil \cdot t.$$

Since $N = \Theta(n,d)$ and $c^{\star}$ is a universal constant, the result follows. ■

10.2 Algebraic rates of convergence, infinite dimensions

Proof of Theorem 3.7. The proof is similar to that of Theorem 3.4, except that it uses Theorem 8.3 in place of Theorem 8.2. In particular, we see that (10.3) also holds in this case with ξ as in (10.4) and $k = m/(c_0 L)$.

Step 2 is identical. The only differences occur in Step 1. We now describe the changes needed in this case. First consider the term $\sigma_k(\boldsymbol{c}_{\Lambda})_{1,\boldsymbol{u};\mathcal{V}}/\sqrt{k}$. To bound this, we use (ii) of Theorem A.3 with $q = 1 > p$. This gives

$$\frac{\sigma_k(\boldsymbol{c}_{\Lambda})_{1,\boldsymbol{u};\mathcal{V}}}{\sqrt{k}} \leq C(\boldsymbol{b},\varepsilon,p) \cdot k^{1/2-1/p} = C(\boldsymbol{b},\varepsilon,p) \cdot \left(\frac{m}{c_0 L}\right)^{1/2-1/p}.$$

To estimate $E_{\Lambda,2}(f)$, recall that $\Lambda = \Lambda_n^{\mathsf{HCI}}$ contains all anchored sets (see Definition 2.8) of size at most $n = \lceil m/L \rceil$ (see Section 3.2). Hence, using (iii) of Theorem A.3 with $s = n$ and $q = 2 > p$, we get

$$\begin{aligned} E_{\Lambda,2}(f) = \|\boldsymbol{c} - \boldsymbol{c}_{\Lambda}\|_{2;\mathcal{V}} \leq \|\boldsymbol{c} - \boldsymbol{c}_S\|_{2;\mathcal{V}} &\leq C(\boldsymbol{b},\varepsilon,p) \cdot n^{1/2-1/p} \\ &\leq C(\boldsymbol{b},\varepsilon,p) \cdot \left(\frac{m}{c_0 L}\right)^{1/2-1/p}. \end{aligned}$$

Finally, for $E_{\Lambda,\infty}(f)$, we use (iii) of Theorem A.3 once more (with $q = 1 > p$) to get

$$\begin{aligned} \frac{E_{\Lambda,\infty}(f)}{\sqrt{k}} \leq \frac{\|\boldsymbol{c} - \boldsymbol{c}_S\|_{1,\boldsymbol{u};\mathcal{V}}}{\sqrt{k}} &\leq C(\boldsymbol{b},\varepsilon,p) \cdot k^{1/2-1/p} \\ &= C(\boldsymbol{b},\varepsilon,p) \cdot \left(\frac{m}{c_0 L}\right)^{1/2-1/p}. \end{aligned}$$

Having done this, we also observe that $\mathcal{G}(\hat{\boldsymbol{c}}) - \mathcal{G}(\mathcal{P}_h(\boldsymbol{c}_{\Lambda})) \leq 0$ in this case, since $\hat{\boldsymbol{c}}$ is once more an exact minimizer. Using this and the previously derived bounds, we conclude that $\xi \leq \zeta$, where ζ is as in (3.14). This gives the result. ■

Proof of Theorem 3.8. The argument is similar to that of Theorem 3.5. Here $\hat{\boldsymbol{c}} = \bar{\boldsymbol{c}}^{(T)}$ is the ergodic sequence obtained after T steps of the primal-dual iteration applied to (4.6) as well.

We recall that the set Λ is lower and of cardinality $|\Lambda| = \Theta(n,d)$ with $d = \infty$. Hence, by (10.2) with $s = N$, we have $|\Lambda|_{\boldsymbol{u}} \leq (\Theta(n,d))^{2\alpha}$, where α is as in (3.7). Using this, we get

$$\|\boldsymbol{A}\|_{2;\mathcal{V}} \leq (\Theta(n,d))^{\alpha},$$

as before. Since the primal-dual iteration in Table 4.3 is used with

$$\tau = \sigma = (\Theta(n,d))^{-\alpha},$$

we have that $\|\boldsymbol{A}\|_{2;\mathcal{V}}^2 \leq (\tau\sigma)^{-1}$. Hence, following the same steps we deduce that

$$\mathcal{G}(\hat{\boldsymbol{c}}) - \mathcal{G}(\mathcal{P}_h(\boldsymbol{c}_\Lambda)) \leq \frac{1}{t}.$$

Substituting this into (10.4) and combining with the other estimates (10.5)–(10.7) derived in Step 2 of the proof of Theorem 3.4 now gives the desired error bound.

The computational cost estimate is similar to the that in the proof of Theorem 3.5. In this case, observe that the value k in Lemma 4.4 is equal to n. Hence, the computational cost of forming $\boldsymbol{A}$ is bounded by $c \cdot m \cdot N \cdot n$ in this case. The computational cost for the algorithm is given by Lemma 4.3. To complete the estimate, we substitute the values $N = \Theta(n,d)$ and $T = \lceil 2(\Theta(n,d))^{\alpha} t \rceil$, as before. ∎

Proof of Theorem 3.9. The proof is similar to that of Theorem 3.6 and involves estimating the term $\mathcal{G}(\hat{\boldsymbol{c}}) - \mathcal{G}(\mathcal{P}_h(\boldsymbol{c}_\Lambda))$. Using the same steps, we deduce that

$$\xi(\boldsymbol{x},\boldsymbol{b}) \leq \zeta,$$

with probability at least $1 - \epsilon/2$, where ζ is as in (3.14). Hence, $\xi(\boldsymbol{x},\boldsymbol{b}) \leq \zeta'$.

Next, recall from Table 4.3 that $\tau = \sigma = (\Theta(n,d))^{-\alpha}$ with $d = \infty$ in this case. Due to (10.9), we see that $\|\boldsymbol{A}\|_{2;\mathcal{V}} \leq (\tau\sigma)^{-1}$ holds. We now apply Theorem 9.4 to obtain

$$\mathcal{G}(\hat{\boldsymbol{c}}) - \mathcal{G}(\mathcal{P}_h(\boldsymbol{c}_\Lambda)) = \mathcal{E}(\tilde{\boldsymbol{c}}^{(R)}, \mathcal{P}_h(\boldsymbol{c}_\Lambda), \boldsymbol{b}) \leq \varepsilon_R = \mathrm{e}^{-R}\|\boldsymbol{b}\|_{2;\mathcal{V}} + \zeta'.$$

To complete the proof of the error bound (3.12), we simply note that $\|\boldsymbol{b}\|_{2;\mathcal{V}} \leq \|f\|_{L^\infty(\mathcal{U};\mathcal{V})} \leq 1$, since $f \in \mathcal{B}(\boldsymbol{b},\varepsilon)$.

The computational cost estimate is as in the previous proof. ∎

10.3 Exponential rates of convergence, finite dimensions

Proof of Theorem 3.10. The proof has the same structure to that of Theorem 3.4, the only differences being the use of Theorem 8.4 instead of Theorem 8.2 and the estimation of the various terms in Step 1. Suppose first that $m \geq c_0 2^{d+2} L$ and define the following:

$$s = \begin{cases} \lceil \sqrt{m/(4c_0 L)} \rceil & \text{Legendre}, \\ \lceil m/(4c_0 2^d L) \rceil & \text{Chebyshev}. \end{cases}$$

Observe that

$$s \leq \begin{cases} \sqrt{m/(c_0 L)} & \text{Legendre}, \\ m/(c_0 2^d L) & \text{Chebyshev}, \end{cases}$$

and therefore the quantity $k(s)$ defined in (10.1) satisfies

$$k(s) \leq \frac{m}{c_0 L} = k.$$

Now consider the term $\sigma_k(\boldsymbol{c}_\Lambda)_{1,\boldsymbol{u};\mathcal{V}}/\sqrt{k}$. Using this and (iii) of Theorem A.1 with $p = 1$ we have

$$\frac{\sigma_k(\boldsymbol{c}_\Lambda)_{1,\boldsymbol{u};\mathcal{V}}}{\sqrt{k}} \leq \frac{\sigma_{k(s)}(\boldsymbol{c})_{1,\boldsymbol{u};\mathcal{V}}}{\sqrt{k}} \leq \frac{C(d,\gamma,\boldsymbol{\rho}) \cdot \exp(-\gamma s^{1/d})}{\sqrt{k}}.$$

Note that this is possible since any lower set S of size at most s satisfies $|S|_{\boldsymbol{u}} \leq k(s)$ by definition.

Now consider $E_{\Lambda,\infty}(f)$. Recall that $\Lambda = \Lambda^{\mathsf{HC}}_{n,d}$, where n is as in (3.15). Clearly $n \geq s$, since $c_0 \geq 1$. Hence Λ contains all lower sets of size at most s. We deduce that

$$E_{\Lambda,\infty}(f) \leq \|\boldsymbol{c} - \boldsymbol{c}_S\|_{1,\boldsymbol{u};\mathcal{V}},$$

for any lower set of size s. We now use (iii) of Theorem A.1 with $p = 1$ once more, to get

$$E_{\Lambda,\infty}(f) \leq C(d,\gamma,\boldsymbol{\rho}) \cdot \exp(-\gamma s^{1/d}).$$

We now combine this with the previous bound to deduce that the quantity ξ in Theorem 8.4 satisfies

$$\xi \leq C(d,\gamma,\boldsymbol{\rho}) \cdot \exp(-\gamma s^{1/d}) + E_{h,\infty}(f) + \frac{\|\boldsymbol{n}\|_{2;\mathcal{V}}}{\sqrt{m}},$$

(here, we also recall that the term $\mathcal{G}(\hat{\boldsymbol{c}}) - \mathcal{G}(\mathcal{P}_h(\boldsymbol{c}_\Lambda)) \leq 0$, as in the proof of Theorem 3.4). Using the value of s and recalling that $m \geq c_0 2^{d+2} L$, we deduce that

$$\xi \leq C(d,\gamma,\boldsymbol{\rho}) \cdot \begin{cases} \exp\left(-\frac{\gamma}{2}\left(\frac{m}{4c_0 L}\right)^{\frac{1}{d}}\right) & \text{Chebyshev} \\ \exp\left(-\gamma\left(\frac{m}{4c_0 L}\right)^{\frac{1}{2d}}\right) & \text{Legendre} \end{cases} + \frac{\|\boldsymbol{n}\|_{2;\mathcal{V}}}{\sqrt{m}}$$
$$+ \|f - \mathcal{P}_h(f)\|_{L^\infty(\mathcal{U};\mathcal{V})},$$

for $m \geq c_0 2^{d+2} L$. However, this bound also clearly holds for all $m \geq 1$, up to a change in the constant $C(d,\gamma,\boldsymbol{\rho})$. After relabeling the universal constant $4c_0$ as c_0, we deduce that $\xi \leq \zeta$, where ζ is as in (3.17). This concludes the proof. ∎

Proof of Theorem 3.11. The argument is the same as the proof of Theorem 3.5. The difference relies on the fact that now ζ has the following bound

$$\xi \leq C(d,\gamma,\rho) \cdot \begin{cases} \exp\Big(-\frac{\gamma}{2}\big(\frac{m}{4c_0 L}\big)^{\frac{1}{d}}\Big) & \text{Chebyshev} \\ \exp\Big(-\gamma\big(\frac{m}{4c_0 L}\big)^{\frac{1}{2d}}\Big) & \text{Legendre} \end{cases} + \frac{\|\boldsymbol{n}\|_{2;\mathcal{V}}}{\sqrt{m}}$$
$$+ \|f - \mathcal{P}_h(f)\|_{L^\infty(\mathcal{U};\mathcal{V})} + \mathcal{G}(\hat{\boldsymbol{c}}) - \mathcal{G}(\mathcal{P}_h(\boldsymbol{c}_\Lambda)).$$

To estimate the final term, we argue exactly as in the proof of Theorem 3.5. The computational cost estimate is likewise identical. ∎

Proof of Theorem 3.12. The proof is similar to that of Theorem 3.6, except we use Theorem 8.4 instead. Recall from Step 2 of the proof of Theorem 8.4 that the matrix $\boldsymbol{A}$ has the weighted rNSP of order $(k,\boldsymbol{u})$ over $\mathcal{V}_h$ with constants $\rho = 2\sqrt{2}/3$ and $\gamma = 2\sqrt{5}/3$ with probability $1-\epsilon$. In particular,

$$\frac{(1+\rho)^2}{(3+\rho)\gamma} \geq 0.64.$$

We now use (8.5) to see that

$$\lambda = \frac{1}{4\sqrt{c_0}}\frac{1}{\sqrt{k}} \leq \frac{(1+\rho)^2}{(3+\rho)\gamma}\frac{1}{\sqrt{k}},$$

for a sufficiently large choice of c_0, as before.

Next, with the choice $\boldsymbol{x} = \mathcal{P}_h(\boldsymbol{c}_\Lambda)$ as before, we see that

$$\xi(\boldsymbol{x},\boldsymbol{b}) = \frac{\sigma_k(\mathcal{P}_h(\boldsymbol{c}_\Lambda))_{1,\boldsymbol{u};\mathcal{V}}}{\sqrt{k}} + \|\boldsymbol{A}\mathcal{P}_h(\boldsymbol{c}_\Lambda) - \boldsymbol{b}\|_{2;\mathcal{V}}.$$

Using (8.11), we get

$$\xi(\boldsymbol{x},\boldsymbol{b}) \leq \frac{\sigma_k(\boldsymbol{c}_\Lambda)_{1,\boldsymbol{w};\mathcal{V}}}{\sqrt{k}} + E_{\Lambda,\infty}(f) + E_{h,\infty}(f) + \frac{\|\boldsymbol{n}\|_{2;\mathcal{V}}}{\sqrt{m}},$$

with probability $1-\epsilon$. It now follows from the proof of Theorem 3.10 that

$$\xi(\boldsymbol{x},\boldsymbol{b}) \leq \zeta,$$

with probability at least $1-\epsilon$, where ζ is as in (3.17). Hence, $\xi(\boldsymbol{x},\boldsymbol{b}) \leq \zeta'$.

The rest of the proof follows the same steps as the proof of Theorem 3.6. ∎