Chapter 8

Error bounds for polynomial approximation via the Hilbert-valued, weighted SR-LASSO

Having developed the necessary tools for Hilbert-valued compressed sensing, we now specialize to the case introduced in Section 4.1 of polynomial approximation via the Hilbert-valued, weighted SR-LASSO problem (4.6). Our main results in this chapter, Theorems 8.2–8.4, yield error bounds for (inexact) minimizers of this problem in terms of the best polynomial approximation error, the Hilbert space discretization error and the noise.

8.1 The weighted RIP for the polynomial approximation problem

In order to obtain these results, we first need to assert conditions on m under which the relevant measurement matrix satisfies the weighted RIP. As in Section 4.1, we let $\{\Psi_{\boldsymbol{\nu}}\}_{\boldsymbol{\nu}\in\mathcal{F}} \subset L^2_{\varrho}(\mathcal{U})$ be either the tensor Chebyshev or Legendre polynomial basis,

$$\Lambda = \begin{cases} \Lambda^{\mathsf{HC}}_{n,d} & d < \infty, \\ \Lambda^{\mathsf{HCI}}_{n} & d = \infty, \end{cases} \tag{8.1}$$

be the hyperbolic cross index set and draw $\boldsymbol{y}_1, \ldots, \boldsymbol{y}_m$ independently and identically from the measure ϱ. Then we define the measurement matrix $\boldsymbol{A}$ exactly as in (4.3).

Lemma 8.1 (Weighted RIP for orthogonal polynomials). *Let $\{\Psi_{\boldsymbol{\nu}}\}_{\boldsymbol{\nu}\in\mathbb{N}_0^d}$ be the orthonormal tensor Legendre or Chebyshev polynomial basis of $L^2_{\varrho}(\mathcal{U})$, Λ be as in (8.1) for some $n \geq 1$ and $\boldsymbol{y}_1, \ldots, \boldsymbol{y}_m$ be drawn independently and identically from the measure ϱ. Let $0 < \epsilon < 1$, $k > 0$, $\boldsymbol{u}$ be the intrinsic weights (4.7), $L' = L'(k, n, d, \epsilon)$ be given by*

$$L' = \begin{cases} \log(2k)\cdot(\log(2k)\cdot\min\{\log(n)+d, \log(\mathrm{e}d)\cdot\log(2n)\}+\log(\epsilon^{-1})) & d < \infty, \\ \log(2k)\cdot(\log(2k)\cdot\log^2(2n)+\log(\epsilon^{-1})) & d = \infty, \end{cases}$$

and suppose that

$$m \geq c \cdot k \cdot L'(k, n, d, \epsilon), \tag{8.2}$$

where $c > 0$ is a universal constant. Then, with probability at least $1-\epsilon$, the matrix $\boldsymbol{A}$ defined in (4.3) satisfies the weighted RIP of order $(k, \boldsymbol{u})$ with constant $\delta_{k,\boldsymbol{u}} \leq 1/4$.

Proof. The proof uses ideas that are now standard. The matrix $\boldsymbol{A}$ is a specific type of measurement matrix associated to the *bounded orthonormal system* $\{\Psi_{\boldsymbol{\nu}}\}_{\boldsymbol{\nu}\in\Lambda}$ (see, e.g., [8, Section 6.4.3] or [61, Chapter 12]). Such a matrix satisfies the weighted RIP of order $k > 0$ with constant $\delta_{k,\boldsymbol{u}} \leq \delta$ whenever

$$m \geq c \cdot k \cdot \delta^{-2} \cdot \log\left(\frac{2k}{\delta^2}\right) \cdot \left[\frac{1}{\delta^4} \log\left(\frac{2k}{\delta^2}\right) \cdot \log(2N) + \frac{1}{\delta} \log(\epsilon^{-1})\right], \tag{8.3}$$

where $c > 0$ is a universal constant. See, e.g., [8, Theorem 6.27 and equation (6.36)] (this result is based on [37]). To obtain the result, we set $\delta = 1/4$. Hence, (8.3) is implied by

$$m \geq c \cdot k \cdot \log(2k) \cdot [\log(2k) \cdot \log(2N) + \log(\epsilon^{-1})],$$

for a potentially different universal constant c. Next, we use (3.6) (and recall that $|\Lambda_n^{\mathsf{HCI}}| = |\Lambda_{n,n}^{\mathsf{HC}}|$) to estimate

$$\log(2N) \leq c \begin{cases} \min\{d + \log(n), \log(2d) \cdot \log(2n)\} & d < \infty, \\ \log^2(2n) & d = \infty, \end{cases}$$

for a potentially different universal constant. The result now follows after substituting this into the previous expression. ∎

Note that the choice of $1/4$ in this lemma is arbitrary. Any value less than

$$(2\sqrt{2} - 1)/7 \approx 0.261$$

(see Lemma 7.6) will suffice.

8.2 Bounds for polynomial approximations obtained as inexact minimizers

We now present the main results of this chapter. These three results provide error bounds for polynomial approximations that are obtained as (inexact) minimizers to the weighted SR-LASSO problem (4.6). Each theorem corresponds to one of the three scenarios in our main results in Section 3.3. Hence, we label them accordingly as algebraic and finite-dimensional, algebraic and infinite-dimensional, and exponential. In order to state these results, we now define some additional notation. Given $f \in L^2_\varrho(\mathcal{U}; \mathcal{V})$ and $\Lambda \subseteq \mathcal{F}$, where $\mathcal{F}$ is as in (2.1)–(2.2), we let

$$E_{\Lambda,2}(f) = \|f - f_\Lambda\|_{L^2_\varrho(\mathcal{U};\mathcal{V})}, \quad E_{\Lambda,\infty}(f) = \|f - f_\Lambda\|_{L^\infty(\mathcal{U};\mathcal{V})},$$

where f_Λ is as in (4.1), and, given a subspace $\mathcal{V}_h \subseteq L^2_\varrho(\mathcal{U}; \mathcal{V})$, we let

$$E_{h,\infty}(f) = \|f - \mathcal{P}_h(f)\|_{L^\infty(\mathcal{U};\mathcal{V})},$$

where $\mathcal{P}_h(f)$ is as in (2.5).

Theorem 8.2 (Error bounds for inexact minimizers, algebraic and finite-dimensional case). *Let $d \in \mathbb{N}$, $m \geq 3$, $0 < \epsilon < 1$, $\{\Psi_{\boldsymbol{\nu}}\}_{\boldsymbol{\nu} \in \mathbb{N}_0^d} \subset L^2_{\varrho}(\mathcal{U})$ be either the orthonormal Chebyshev or Legendre basis, $\mathcal{V}_h \subseteq L^2_{\varrho}(\mathcal{U})$ be a subspace of $L^2_{\varrho}(\mathcal{U})$ and $\Lambda = \Lambda^{\mathsf{HC}}_{n,d}$ be the hyperbolic cross index set with $n = \lceil m/L \rceil$ where $L = L(m, d, \epsilon)$ is as in (3.8). Let $f \in L^2_{\varrho}(\mathcal{U}; \mathcal{V})$, draw $\boldsymbol{y}_1, \ldots, \boldsymbol{y}_m$ randomly and independently according to ϱ and suppose that $\boldsymbol{A}$, $\boldsymbol{b}$ and $\boldsymbol{e}$ are as in (4.3) and (4.4). Consider the Hilbert-valued, weighted SR-LASSO problem (4.6) with weights $\boldsymbol{w} = \boldsymbol{u}$ as in (4.7) and $\lambda = (4\sqrt{m/L})^{-1}$. Then there exists universal constants $c_0, c_1, c_2 \geq 1$ such that the following holds with probability at least $1 - \epsilon$. Any $\tilde{\boldsymbol{c}} = (\tilde{c}_{\boldsymbol{\nu}})_{\boldsymbol{\nu} \in \Lambda} \in \mathbb{C}^N$ satisfies*

$$\|f - \tilde{f}\|_{L^2_{\varrho}(\mathcal{U};\mathcal{V})} \leq c_1 \cdot \xi, \quad \|f - \tilde{f}\|_{L^\infty(\mathcal{U};\mathcal{V})} \leq c_2 \cdot \sqrt{k} \cdot \xi, \quad \tilde{f} := \sum_{\boldsymbol{\nu} \in \Lambda} \tilde{c}_{\boldsymbol{\nu}} \Psi_{\boldsymbol{\nu}},$$

where

$$\xi = \frac{\sigma_k(\boldsymbol{c}_\Lambda)_{1,\boldsymbol{u};\mathcal{V}}}{\sqrt{k}} + \frac{E_{\Lambda,\infty}(f)}{\sqrt{k}} + E_{\Lambda,2}(f) + E_{h,\infty}(f) + \mathcal{G}(\tilde{\boldsymbol{c}}) - \mathcal{G}(\mathcal{P}_h(\boldsymbol{c}_\Lambda)) + \frac{\|\boldsymbol{n}\|_{2;\mathcal{V}}}{\sqrt{m}},$$

$\boldsymbol{c}_\Lambda$ is as in (4.2), $\mathcal{P}_h(\boldsymbol{c}_\Lambda) = (\mathcal{P}_h(c_{\boldsymbol{\nu}}))_{\boldsymbol{\nu} \in \Lambda}$, $k = m/(c_0 L)$ for $L = L(m, d, \epsilon)$ as in (3.8), and $\boldsymbol{n}$ is as in (4.4).

Proof. We divide the proof into several steps.

Step 1: Splitting the error into separate terms. Consider the $L^2_{\varrho}(\mathcal{U}; \mathcal{V})$-norm error first. By the triangle inequality and the fact that $\mathcal{P}_h$ is a projection, we have

$$\begin{aligned}
&\|f - \tilde{f}\|_{L^2_{\varrho}(\mathcal{U};\mathcal{V})} \\
&\quad \leq \|f - \mathcal{P}_h(f)\|_{L^2_{\varrho}(\mathcal{U};\mathcal{V})} + \|\mathcal{P}_h(f) - \mathcal{P}_h(f_\Lambda)\|_{L^2_{\varrho}(\mathcal{U};\mathcal{V})} + \|\mathcal{P}_h(f_\Lambda) - \tilde{f}\|_{L^2_{\varrho}(\mathcal{U};\mathcal{V})} \\
&\quad \leq \|f - \mathcal{P}_h(f)\|_{L^\infty(\mathcal{U};\mathcal{V})} + \|f - f_\Lambda\|_{L^2_{\varrho}(\mathcal{U};\mathcal{V})} + \|\mathcal{P}_h(f_\Lambda) - \tilde{f}\|_{L^2_{\varrho}(\mathcal{U};\mathcal{V})} \\
&\quad = E_{h,\infty}(f) + E_{\Lambda,2}(f) + \|\mathcal{P}_h(f_\Lambda) - \tilde{f}\|_{L^2_{\varrho}(\mathcal{U};\mathcal{V})}.
\end{aligned}$$

Then, by orthonormality, we have

$$\|f - \tilde{f}\|_{L^2_{\varrho}(\mathcal{U};\mathcal{V})} \leq E_{h,\infty}(f) + E_{\Lambda,2}(f) + \|\mathcal{P}_h(\boldsymbol{c}_\Lambda) - \tilde{\boldsymbol{c}}\|_{2;\mathcal{V}}.$$

Similarly, for the $L^\infty(\mathcal{U}; \mathcal{V})$-norm error, we have

$$\begin{aligned}
&\|f - \tilde{f}\|_{L^\infty(\mathcal{U};\mathcal{V})} \\
&\quad \leq \|f - \mathcal{P}_h(f)\|_{L^\infty(\mathcal{U};\mathcal{V})} + \|\mathcal{P}_h(f) - \mathcal{P}_h(f_\Lambda)\|_{L^\infty(\mathcal{U};\mathcal{V})} + \|\mathcal{P}_h(f_\Lambda) - \tilde{f}\|_{L^\infty(\mathcal{U};\mathcal{V})} \\
&\quad \leq \|f - \mathcal{P}_h(f)\|_{L^\infty(\mathcal{U};\mathcal{V})} + \|f - f_\Lambda\|_{L^\infty(\mathcal{U};\mathcal{V})} + \|\mathcal{P}_h(f_\Lambda) - \tilde{f}\|_{L^\infty(\mathcal{U};\mathcal{V})} \\
&\quad = E_{h,\infty}(f) + E_{\Lambda,\infty}(f) + \|\mathcal{P}_h(f_\Lambda) - \tilde{f}\|_{L^\infty(\mathcal{U};\mathcal{V})}.
\end{aligned}$$

Using the definition (4.7) of the weights $\boldsymbol{u}$, we deduce that

$$\|f - \tilde{f}\|_{L^\infty(\mathcal{U};\mathcal{V})} \leq E_{h,\infty}(f) + E_{\Lambda,\infty}(f) + \|\mathcal{P}_h(\boldsymbol{c}_\Lambda) - \tilde{\boldsymbol{c}}\|_{1,\boldsymbol{u};\mathcal{V}}.$$

Therefore, the rest of the proof is devoted to showing the following bounds:

$$\|\mathcal{P}_h(\boldsymbol{c}_\Lambda) - \tilde{\boldsymbol{c}}\|_{2;\mathcal{V}} \leq c_1 \cdot \xi, \quad \|\mathcal{P}_h(\boldsymbol{c}_\Lambda) - \tilde{\boldsymbol{c}}\|_{1,\boldsymbol{u};\mathcal{V}} \leq c_2 \cdot \sqrt{k} \cdot \xi. \tag{8.4}$$

We do this in the next two steps by first asserting that $\boldsymbol{A}$ has the weighted rNSP (Step 2) and then by applying the error bounds of Lemma 7.4 (Steps 3 and 4).

Step 2: Asserting the weighted rNSP. We now show that $\boldsymbol{A}$ has the weighted rNSP over $\mathcal{V}_h$ of order $(k, \boldsymbol{u})$ with probability at least $1 - \epsilon/2$. This is based on Lemma 8.1. First observe that

$$L = L(m, d, \epsilon) \geq \log^2(3) \cdot \min\{\log(3) + 1, \log(3) \cdot \log(\mathrm{e})\} \geq 1,$$

since $m \geq 3$. This implies that $m \geq m/L \geq m/(c_0 L) = k$ since $c_0 \geq 1$ as well. Since $n = \lceil m/L \rceil \leq m/L + 1 \leq 2m$, we get

$$\begin{aligned} &\log(4k) \cdot (\log(4k) \cdot \min\{\log(n) + d, \log(\mathrm{e}d) \cdot \log(2n)\} + \log(2/\epsilon)) \\ &\quad \leq \log(4m) \cdot (\log(4m) \cdot \min\{\log(2m) + d, \log(\mathrm{e}d) \cdot \log(4m)\} + \log(2/\epsilon)) \\ &\quad \leq c_0 L(m, d, \epsilon)/2 \end{aligned}$$

for a suitably large choice of c_0. Hence

$$m = c_0 k L(m, d, \epsilon) \geq 2c_0 k L'(2k, d, \epsilon/2),$$

where L' is defined as in Lemma 8.1, and therefore (again assuming a suitably large choice of c_0) (8.2) holds with k replaced by $2k$. We deduce that $\boldsymbol{A}$ satisfies the weighted RIP over $\mathbb{C}$ of order $(2k, \boldsymbol{u})$ with constant $\delta_{2k,\boldsymbol{u}} \leq 1/4$, with probability at least $1 - \epsilon/2$. Then, we deduce from Lemmas 7.5 and 7.6 that $\boldsymbol{A}$ has (with the same probability) the weighted rNSP over $\mathcal{V}_h$ of order $(k, \boldsymbol{u})$ with constants $\rho = 2\sqrt{2}/3$ and $\gamma = 2\sqrt{5}/3$.

Step 3: Bounding $\mathcal{P}_h(\boldsymbol{c}_\Lambda) - \tilde{\boldsymbol{c}}$ using the weighted rNSP. We use Lemma 7.4. First, consider the value of λ. Since $c_0 \geq 1$ we have $m/L \geq m/(c_0 L) = k$. Hence, recalling the values for ρ and γ obtained in the previous step, we have

$$\frac{1}{4\sqrt{c_0}} \frac{1}{\sqrt{k}} = \frac{1}{4\sqrt{m/L}} = \lambda \leq \frac{1}{4\sqrt{k}} < \frac{(1+\rho)^2}{(3+\rho)\gamma} \frac{1}{\sqrt{k}}. \tag{8.5}$$

Therefore, (7.3) holds. We now apply this lemma with $\mathcal{V} = \mathcal{V}_h$, $\boldsymbol{x} = \mathcal{P}_h(\boldsymbol{c}_\Lambda)$, $\tilde{\boldsymbol{x}} = \tilde{\boldsymbol{c}}$ and $\boldsymbol{e} = \boldsymbol{A}\mathcal{P}_h(\boldsymbol{c}_\Lambda) - \boldsymbol{b}$. Notice first that the best $(k, \boldsymbol{u})$-approximation error (7.1) satisfies

$$\sigma_k(\mathcal{P}_h(\boldsymbol{c}_\Lambda))_{1,\boldsymbol{u};\mathcal{V}} = \inf\Big\{ \sum_{\boldsymbol{\nu}\in\Lambda\setminus S} u_{\boldsymbol{\nu}} \big\|\mathcal{P}_h(c_{\boldsymbol{\nu}})\big\|_{\mathcal{V}} : S \subseteq \Lambda,\ |S|_{\boldsymbol{u}} \leq k \Big\} \\ \leq \sigma_k(\boldsymbol{c}_\Lambda)_{1,\boldsymbol{u};\mathcal{V}}, \tag{8.6}$$

since $\mathcal{P}_h$ is a projection. Hence, applying Lemma 7.4 and using the lower bound in (8.5), we get

$$\|\tilde{\boldsymbol{c}}-\mathcal{P}_h(\boldsymbol{c}_\Lambda)\|_{2;\mathcal{V}} \leq c_1\left[\frac{\sigma_k(\boldsymbol{c}_\Lambda)_{1,\boldsymbol{w};\mathcal{V}}}{\sqrt{k}} + \mathcal{G}(\tilde{\boldsymbol{c}})-\mathcal{G}(\mathcal{P}_h(\boldsymbol{c}_\Lambda)) + \big\|\boldsymbol{A}\mathcal{P}_h(\boldsymbol{c}_\Lambda)-\boldsymbol{b}\big\|_{2;\mathcal{V}}\right],$$
$$\|\tilde{\boldsymbol{c}} - \mathcal{P}_h(\boldsymbol{c}_\Lambda)\|_{1,\boldsymbol{u};\mathcal{V}} \leq c_2\big[\sigma_k(\boldsymbol{c}_\Lambda)_{1,\boldsymbol{w};\mathcal{V}} + \sqrt{k}(\mathcal{G}(\tilde{\boldsymbol{c}}) - \mathcal{G}(\mathcal{P}_h(\boldsymbol{c}_\Lambda))) \\ + \sqrt{k}\big\|\boldsymbol{A}\mathcal{P}_h(\boldsymbol{c}_\Lambda) - \boldsymbol{b}\big\|_{2;\mathcal{V}}\big], \tag{8.7}$$

with probability at least $1-\epsilon/2$. Therefore, to show (8.4) and therefore complete the proof, it suffices to show that the following holds with probability at least $1-\epsilon/2$:

$$\|\boldsymbol{A}\mathcal{P}_h(\boldsymbol{c}_\Lambda) - \boldsymbol{b}\|_{2;\mathcal{V}} \leq \sqrt{2}\left(\frac{E_{\Lambda,\infty}(f)}{\sqrt{k}} + E_{\Lambda,2}(f)\right) + E_{h,\infty}(f) + \frac{\|\boldsymbol{n}\|_{2;\mathcal{V}}}{\sqrt{m}}. \tag{8.8}$$

The overall result then follows by the union bound.

Step 4: Showing that (8.8) *holds.* Observe that

$$\sqrt{m}\|(\boldsymbol{A}\mathcal{P}_h(\boldsymbol{c}_\Lambda) - \boldsymbol{b})_i\|_{\mathcal{V}} \\ \leq \|\mathcal{P}_h(f_\Lambda)(\boldsymbol{y}_i) - f(\boldsymbol{y}_i)\|_{\mathcal{V}} + \big\|n_i\big\|_{\mathcal{V}} \\ \leq \|\mathcal{P}_h(f_\Lambda)(\boldsymbol{y}_i) - \mathcal{P}_h(f)(\boldsymbol{y}_i)\|_{\mathcal{V}} + \|f(\boldsymbol{y}_i) - \mathcal{P}_h(f)(\boldsymbol{y}_i)\|_{\mathcal{V}} + \big\|n_i\big\|_{\mathcal{V}} \\ \leq \|f(\boldsymbol{y}_i) - f_\Lambda(\boldsymbol{y}_i)\|_{\mathcal{V}} + E_{h,\infty}(f) + \|n_i\|_{\mathcal{V}}.$$

Therefore,

$$\|\boldsymbol{A}\mathcal{P}_h(\boldsymbol{c}_\Lambda) - \boldsymbol{b}\|_{\mathcal{V};2} \leq E_{\Lambda,\mathrm{disc}}(f) + E_{h,\infty}(f) + \frac{\|\boldsymbol{n}\|_{2;\mathcal{V}}}{\sqrt{m}}, \tag{8.9}$$

where

$$E_{\Lambda,\mathrm{disc}}(f) = \sqrt{\frac{1}{m}\sum_{i=1}^{m}\|f(\boldsymbol{y}_i) - f_\Lambda(\boldsymbol{y}_i)\|_{\mathcal{V}}^2}. \tag{8.10}$$

For this final step, we follow near-identical arguments to those found in [8, Lemma 7.11]. This shows that

$$E_{\Lambda,\mathrm{disc}}(f) \leq \sqrt{2}\left(\frac{E_{\Lambda,\infty}(f)}{\sqrt{k}} + E_{\Lambda,2}(f)\right),$$

with probability at least $1-\epsilon/2$, provided $m \geq 2k\log(2/\epsilon)$. However, this follows due to the assumptions on m and the arguments given in Step 2. Thus we obtain (8.8) and the proof is complete. ∎

Theorem 8.3 (Error bounds for inexact minimizers, algebraic and infinite-dimensional case). *Let $d = \infty$, $m \geq 3$, $0 < \epsilon < 1$, $\{\Psi_{\boldsymbol{\nu}}\}_{\boldsymbol{\nu} \in \mathcal{F}} \subset L^2_\varrho(\mathcal{U})$ be either the orthonormal Chebyshev or Legendre basis, $\mathcal{V}_h \subseteq L^2_\varrho(\mathcal{U})$ be a subspace of $L^2_\varrho(\mathcal{U})$ and $\Lambda = \Lambda^{\mathsf{HCI}}_n$ be the hyperbolic cross index set with $n = \lceil m/L \rceil$ where $L = L(m, d, \epsilon)$ is as in (3.8). Let $f \in L^2_\varrho(\mathcal{U}; \mathcal{V})$, draw $\boldsymbol{y}_1, \ldots, \boldsymbol{y}_m$ randomly and independently according to ϱ and suppose that $\boldsymbol{A}$, $\boldsymbol{b}$ and $\boldsymbol{e}$ are as in (4.3) and (4.4). Consider the Hilbert-valued, weighted SR-LASSO problem (4.6) with weights $\boldsymbol{w} = \boldsymbol{u}$ as in (4.7) and $\lambda = (4\sqrt{m/L})^{-1}$. Then there exists universal constants $c_0, c_1, c_2 \geq 1$ such that the following holds with probability at least $1 - \epsilon$. Any $\tilde{\boldsymbol{c}} = (\tilde{c}_{\boldsymbol{\nu}})_{\boldsymbol{\nu} \in \Lambda} \in \mathbb{C}^N$ satisfies*

$$\|f - \tilde{f}\|_{L^2_\varrho(\mathcal{U};\mathcal{V})} \leq c_1 \cdot \xi, \quad \|f - \tilde{f}\|_{L^\infty(\mathcal{U};\mathcal{V})} \leq c_2 \cdot \sqrt{k} \cdot \xi, \quad \tilde{f} := \sum_{\boldsymbol{\nu} \in \Lambda} \tilde{c}_{\boldsymbol{\nu}} \Psi_{\boldsymbol{\nu}},$$

where

$$\xi = \frac{\sigma_k(\boldsymbol{c}_\Lambda)_{1,\boldsymbol{u};\mathcal{V}}}{\sqrt{k}} + \frac{E_{\Lambda,\infty}(f)}{\sqrt{k}} + E_{\Lambda,2}(f) + E_{h,\infty}(f) + \mathcal{G}(\tilde{\boldsymbol{c}}) - \mathcal{G}(\mathcal{P}_h(\boldsymbol{c}_\Lambda)) + \frac{\|\boldsymbol{n}\|_{2;\mathcal{V}}}{\sqrt{m}},$$

$\boldsymbol{c}_\Lambda$ is as in (4.2), $\mathcal{P}_h(\boldsymbol{c}_\Lambda) = (\mathcal{P}_h(c_{\boldsymbol{\nu}}))_{\boldsymbol{\nu} \in \Lambda}$, $k = m/(c_0 L)$ for $L = L(m, d, \epsilon)$ as in (3.8), and $\boldsymbol{n}$ is as in (4.4).

Proof. The proof has the same structure as that of the previous theorem. Steps 1, 3 and 4 are identical. The only differences occur in Step 2. We now describe these changes. Once more we observe that $L = L(m, \infty, \epsilon) \geq 1$ since $m \geq 3$. Hence, $m \geq m/L \geq m/(c_0 L) = k$ since $c_0 \geq 1$. We also have $n = \lceil m/L \rceil \leq 2m$. Therefore,

$$\begin{aligned} \log(4k) \cdot (\log(4k) \cdot \log^2(2n) + \log(2/\epsilon)) &\leq \log(4m) \cdot (\log^3(4m) + \log(2/\epsilon)) \\ &\leq c_0 L(m, \infty, \epsilon)/2 \end{aligned}$$

for a suitably large choice of c_0. We deduce that

$$m = c_0 k L(m, \infty, \epsilon) \geq 2c_0 k L'(2k, \infty, \epsilon/2),$$

where L' is as in Lemma 8.1. An application of this lemma now shows that $\boldsymbol{A}$ has the weighted RIP of order $(2k, \boldsymbol{u})$ with constant $\delta_{2k,\boldsymbol{u}} \leq 1/4$, as required. ∎

Theorem 8.4 (Error bounds for inexact minimizers, exponential case). *Let $d \in \mathbb{N}$, $m \geq 3$, $0 < \epsilon < 1$, $\{\Psi_{\boldsymbol{\nu}}\}_{\boldsymbol{\nu} \in \mathbb{N}_0^d} \subset L^2_\varrho(\mathcal{U})$ be either the orthonormal Chebyshev or Legendre basis, $\mathcal{V}_h \subseteq L^2_\varrho(\mathcal{U})$ be a subspace of $L^2_\varrho(\mathcal{U})$ and $\Lambda = \Lambda^{\mathsf{HC}}_{n,d}$ be the hyperbolic cross index set with n as in (3.15). Draw $\boldsymbol{y}_1, \ldots, \boldsymbol{y}_m$ randomly and independently according to ϱ. Then, with probability at least $1 - \epsilon$, the following holds. Let $f \in L^2_\varrho(\mathcal{U}; \mathcal{V})$ and suppose that $\boldsymbol{A}$, $\boldsymbol{b}$ and $\boldsymbol{e}$ are as in (4.3) and (4.4). Consider the Hilbert-valued, weighted SR-LASSO problem (4.6) with weights $\boldsymbol{w} = \boldsymbol{u}$ as in (4.7) and $\lambda = (4\sqrt{m/L})^{-1}$. Then there exists universal constants $c_0, c_1, c_2 \geq 1$ such that* *any* $\tilde{\boldsymbol{c}} = (\tilde{c}_{\boldsymbol{\nu}})_{\boldsymbol{\nu}\in\Lambda} \in \mathbb{C}^N$ *satisfies*

$$\|f - \tilde{f}\|_{L^2_{\varrho}(\mathcal{U};\mathcal{V})} \leq c_1 \cdot \xi, \quad \|f - \tilde{f}\|_{L^\infty(\mathcal{U};\mathcal{V})} \leq c_2 \cdot \sqrt{k} \cdot \xi, \quad \tilde{f} := \sum_{\boldsymbol{\nu}\in\Lambda} \tilde{c}_{\boldsymbol{\nu}} \Psi_{\boldsymbol{\nu}},$$

where

$$\xi = \frac{\sigma_k(\boldsymbol{c}_\Lambda)_{1,\boldsymbol{u};\mathcal{V}}}{\sqrt{k}} + E_{\Lambda,\infty}(f) + E_{h,\infty}(f) + \mathcal{G}(\tilde{\boldsymbol{c}}) - \mathcal{G}(\mathcal{P}_h(\boldsymbol{c}_\Lambda)) + \frac{\|\boldsymbol{n}\|_{2;\mathcal{V}}}{\sqrt{m}},$$

$\boldsymbol{c}_\Lambda$ *is as in* (4.2)*,* $\mathcal{P}_h(\boldsymbol{c}_\Lambda) = (\mathcal{P}_h(c_{\boldsymbol{\nu}}))_{\boldsymbol{\nu}\in\Lambda}$*,* $k = m/(c_0 L)$ *for* $L = L(m, d, \epsilon)$ *as in* (3.8)*, and* $\boldsymbol{n}$ *is as in* (4.4)*.*

Proof. The proof has the same structure as that of Theorem 8.2. Step 1 is identical, and reduces the proof to showing that (8.4) holds. We now describe the modifications needed in Steps 2–4.

Step 2: Asserting the weighted rNSP. We now show that $\boldsymbol{A}$ has the weighted rNSP over $\mathcal{V}_h$ of order $(k, \boldsymbol{u})$ with probability at least $1 - \epsilon$. This step is essentially the same, except for the choice of n and the probability $1 - \epsilon$ instead of $1 - \epsilon/2$.

Step 3: Bounding $\mathcal{P}_h(\boldsymbol{c}_\Lambda) - \tilde{\boldsymbol{c}}$ *using the weighted rNSP.* Since λ and k are the same as in Theorem 8.2, the bound (8.5) also holds in this case. We then follow the same arguments, leading to (8.7) holding with probability at least $1 - \epsilon$. Finally, rather than (8.8), we ask for the slightly modified bound

$$\|\boldsymbol{A}\mathcal{P}_h(\boldsymbol{c}_\Lambda) - \boldsymbol{b}\|_{2;\mathcal{V}} \leq E_{\Lambda,\infty}(f) + E_{h,\infty}(f) + \frac{\|\boldsymbol{n}\|_{2;\mathcal{V}}}{\sqrt{m}}, \tag{8.11}$$

to hold with probability one.

Step 4: Showing (8.11) *holds.* By the same argument, we see that (8.9) holds. Instead of the probabilistic bound for $E_{\Lambda,\text{disc}}(f)$, we now simply bound it as

$$E_{\Lambda,\text{disc}}(f) \leq \|f - f_\Lambda\|_{L^\infty(\mathcal{U};\mathcal{V})} = E_{\Lambda,\infty}(f).$$

This immediately implies (8.11).

Finally, we observe that we can simplify the previous estimates in this case using the bound $E_{\Lambda,2}(f) \leq E_{\Lambda,\infty}(f)$. ∎