



Sparse recovery in bounded Riesz systems with applications to numerical methods for PDEs



Simone Brugiapaglia^a, Sjoerd Dirksen^b, Hans Christian Jung^c, Holger Rauhut^{c,*}

^a Department of Mathematics and Statistics, Concordia University, Montreal, QC, Canada

^b Mathematical Institute, Utrecht University, Utrecht, the Netherlands

^c Chair for Mathematics of Information Processing, RWTH Aachen University, Aachen, Germany

ARTICLE INFO

Article history:

Received 19 May 2020

Received in revised form 28 January 2021

Accepted 28 January 2021

Available online 6 February 2021

Communicated by Charles K. Chui

Keywords:

Compressive sensing

Restricted isometry constants

Bounded Riesz systems

Numerical PDEs

CORSING method

Generic chaining

ABSTRACT

We study sparse recovery with structured random measurement matrices having independent, identically distributed, and uniformly bounded rows and with a nontrivial covariance structure. This class of matrices arises from random sampling of bounded Riesz systems and generalizes random partial Fourier matrices. Our main result improves the currently available results for the null space and restricted isometry properties of such random matrices. The main novelty of our analysis is a new upper bound for the expectation of the supremum of a Bernoulli process associated with a restricted isometry constant. We apply our result to prove new performance guarantees for the CORSING method, a recently introduced numerical approximation technique for partial differential equations (PDEs) based on compressive sensing.

© 2021 Elsevier Inc. All rights reserved.

1. Introduction

Compressive sensing [18,19,27,29,45] provides efficient methods that allow to recover (approximately) sparse vectors from a surprisingly small amount of random measurements. Although compressive sensing was initially motivated by signal processing applications, it has recently inspired a new generation of hybrid methodologies in computational mathematics. This includes compressive sensing techniques for polynomial interpolation [1,41,43], high-dimensional function approximation [2,3,22], the numerical solution of PDEs [15, 31,33], uncertainty quantification of PDEs with random inputs [28,38,42,50,51], dynamical systems [48], and inverse problems in PDEs [5]. In this paper, our main application of interest is the numerical approximation of PDEs based on compressive sensing via the CORSING (COmpressed SolvING) method [15,16].

* Corresponding author.

E-mail addresses: simone.brugiapaglia@concordia.ca (S. Brugiapaglia), s.dirksen@uu.nl (S. Dirksen), jung@mathc.rwth-aachen.de (H.C. Jung), rauhut@mathc.rwth-aachen.de (H. Rauhut).

A key theoretical task in compressive sensing is to provide estimates for the so-called restricted isometry constants or, strongly related to this, to derive the null space property, of random measurement matrices. Such estimates lead to bounds on the sufficient number of measurements for recovery in terms of the sparsity and the dimension of the vector to be recovered. We also note that, besides compressive sensing, the estimation of the restricted isometry constant of a matrix is a central problem in, e.g., the analysis of list-decodable linear codes [21], and Johnson-Lindenstrauss embeddings [34,39].

A class of structured measurement matrices, which plays an important role in compressive sensing as well as in the CORSING method for the numerical solution of PDEs, arises from random sampling function systems (such as the Fourier system). While corresponding bounds on the restricted isometry constants exist for orthonormal systems that are bounded in the L^∞ -norm [18,29,45,10], similar estimates for the more general class of bounded Riesz systems (where orthogonality does not necessary hold) are not yet available in the literature up to the best of the authors’ knowledge. The CORSING method, however, typically requires to work with Riesz systems rather than orthonormal systems, which raises the need for such a generalization.

The contribution of this paper is twofold. On the one hand, we provide a new analysis of the restricted isometry constants and null space property of matrices arising from random sampling in bounded Riesz systems, which also improves the available bounds for the orthonormal case. On the other hand, we take advantage of this result to obtain substantially improved theoretical guarantees for the CORSING method.

1.1. Main results

Recall that the restricted isometry constant of sparsity level s of a matrix A with N columns is defined by

$$\varepsilon_s := \sup_{f \in D_{s,N}} \left| \|Af\|_2^2 - 1 \right|,$$

where $D_{s,N} := \{f \in \mathbb{C}^N : \|f\|_0 \leq s, \|f\|_2 = 1\}$ is the set of unit-norm vectors with support size at most s . Our main theorem establishes a concentration inequality, which implies bounds for the restricted isometry constants (and establishes the null space property) of a random matrix whose rows are independent, identically distributed, and uniformly bounded random vectors.

Theorem 1.1. *There exist absolute constants $\kappa, c_0, c_1 > 0$ such that the following holds. Let X_1, \dots, X_m be independent copies of a random vector $X \in \mathbb{C}^N$ with bounded coordinates, i.e., for all $i = 1, \dots, N$ we have $|\langle X, e_i \rangle| \leq K$ for some $K > 0$ where e_1, \dots, e_N is the standard basis of \mathbb{C}^N . Let $T \subseteq \{f \in \mathbb{C}^N : \|f\|_1 \leq \sqrt{s}\}$, $\delta \in (0, \kappa)$ and assume that*

$$m \geq c_0 K^2 \delta^{-2} s \log(eN) \log^2(sK^2/\delta). \tag{1.1}$$

Then, with probability exceeding $1 - 2 \exp(-\delta^2 m/(sK^2))$,

$$\sup_{f \in T} \left| \frac{1}{m} \sum_{i=1}^m |\langle f, X_i \rangle|^2 - \mathbb{E}|\langle f, X \rangle|^2 \right| \leq c_1 \left(\delta + \delta \sup_{f \in T} \mathbb{E}|\langle f, X \rangle|^2 \right). \tag{1.2}$$

Remark 1.2. The values of the constants in the above theorem can be chosen as $\kappa = \frac{1}{28}(10 - 7\sqrt{2}) \approx 0.3066$, $c_0 = 1600(99 + 70\sqrt{2}) \approx 316792$ and $c_1 = 492$. We have not optimized these constants in our proof and believe that they can be further reduced.

Theorem 1.1 fits into a line of work [9,20,22,30,45] that studies the restricted isometry constants of random matrices associated with random sampling from a bounded orthonormal system. In the breakthrough work [20] it was established that the restricted isometry constant ε_s of such a matrix satisfies

$\varepsilon_s \leq \delta_0$ with probability at least $1 - N^{-c_0\tau}$ provided that the number of measurements m satisfies $m \geq c_1\tau s \log^6(N)$ (where c_0 and c_1 only depend on δ_0). In [45] generic chaining techniques were used to show that $m \geq c_0\varepsilon^{-2}s \log(s \log(N)) \log^2(s) \log(N)$ suffice to guarantee that $\varepsilon_s \leq \varepsilon$ with high probability. Using a different technique, [9] showed that $m \geq c_0\varepsilon^{-6}s \log(s) \log^2(N)$ measurements are sufficient for a discrete bounded orthonormal system. The latter result was generalized in [22] to continuous sampling scenarios. Inspired by the methods in [9] it was shown in [30] that for a discrete bounded orthonormal system $m \geq c_0\varepsilon^{-2} \log^2(s/\varepsilon) \log(eN) \log(1/\varepsilon)$ measurements are sufficient. Recently, the work [6] has shown that the condition $m \geq c_0s \log(s) \log(N/s)$ is necessary for a randomly subsampled Hadamard matrix to satisfy $\varepsilon_s \leq c_1$. Hence, for constant δ , K and for $s \leq N^\alpha$ for some $\alpha < 1$, say, the bound (1.1) is optimal up to a factor of $\log(s)$.

Theorem 1.1 improves on these state-of-the-art results in several ways. In the setting of continuous bounded orthonormal systems, it shows that $m \geq c_0\varepsilon^{-2}s \log(eN) \log^2(s/\varepsilon)$ measurements suffice to guarantee that $\varepsilon_s \leq \varepsilon$ with high probability. In particular, compared to [30], we remove a $\log(1/\varepsilon)$ factor and achieve a result in the setting of continuous sampling. Compared to [22], we improve the dependence on ε . Finally, Theorem 1.1 more generally yields estimates on the restricted isometry constants of matrices associated with random sampling from a bounded Riesz system. We refer to Theorem 2.3 and Remark 2.4 for details.

Our proof of Theorem 1.1 is inspired by the methods in [9] and [30], but in contrast to these references we develop our proof in the terminology of generic chaining [47]. We believe that this makes the proof more transparent, at least, for those familiar with generic chaining techniques. Moreover, our proof leads to the improved dependence in δ in (1.1) (resp. in ε in estimates for the RIP).

Although Theorem 1.1 is concerned with subsets of the ℓ^1 -ball, it is possible to extend it to subsets of a weighted ℓ^1 -ball, i.e., the unit ball of the norm $\|f\|_{\omega,1} := \sum_{j=1}^n \omega_j |f_j|$, where $\omega_j \geq 1$ is a sequence of weights. We refer the reader to Section 2.3 for more details on this extension and its applications to weighted compressive sensing and uncertainty quantification.

We apply Theorem 1.1 to obtain improved robust recovery guarantees for the CORSING method. This is a recently introduced numerical method for computing a sparse approximation of the solution to a PDE based on compressive sensing [12,15,16]. Given a PDE admitting a weak formulation, the idea of the method is to assemble a reduced Petrov-Galerkin discretization via random sampling and to solve this reduced system using a sparse recovery algorithm. Compared to other nonlinear approximation methods for PDEs, such as adaptive finite elements or adaptive wavelets methods (see, e.g., [49] and references therein), CORSING has the advantages that no *a posteriori* error indicators are needed and that the assembly of the discretization matrix as well as the sparse recovery step (here performed via Orthogonal Matching Pursuit (OMP)) can be easily parallelized. We refer to [16] for a more detailed discussion and to [12,14–16] for numerical experiments for multi-dimensional advection-diffusion-reaction equations and the Stokes problem.

The best available theoretical guarantees for the CORSING method [16] state that in order to recover the best s -term approximation of the PDE solution with respect to a Riesz basis of trial functions, it is sufficient to assemble a number of rows proportional to s^2 (up to logarithmic factors) under suitable assumptions involving the trial and test functions and the bilinear form defining the weak formulation of the PDE. This is highly suboptimal compared to standard compressive sensing results, where a number of measurements proportional to s (up to log factors) is usually sufficient. Indeed, numerical experiments show that the quadratic scaling s^2 is highly pessimistic (see [16, Figure 8]). Another limitation of the state-of-the-art results is the assumption that one solves an NP-hard problem *exactly* in the recovery phase.

In this paper, we bridge these gaps by showing robust recovery guarantees under optimal linear scaling between m and s (up to log factors) which cover sparse recovery via ℓ^1 -minimization and via orthogonal matching pursuit (OMP). The latter is preferred in practice. This result is stated in Theorem 3.9. The main technical challenge is to establish an improved bound on the sufficient size of the reduced Petrov-Galerkin

discretization by analyzing the restricted isometry constant of the corresponding random matrix, which does not have a trivial covariance. This challenge is overcome thanks to Theorem 1.1.

1.2. Organization of the paper

In Section 2 we present several consequences of Theorem 1.1 for sparse recovery via ℓ^1 -minimization in the case of subsampled bounded Riesz systems (Sections 2.1 and 2.2). Taking advantage of the theory presented in Section 2, we prove new recovery guarantees for the CORSING method in Section 3, focusing on the case where recovery is performed via orthogonal matching pursuit. Section 4 develops the proof of Theorem 1.1 and is the technical core of the paper.

1.3. Notation

Throughout the paper we will use the following notation. Given $N \in \mathbb{N}$, we define $[N] := \{1, \dots, N\}$. Moreover, given a vector $x \in \mathbb{C}^N$, we denote its ℓ^p -norm as $\|x\|_p = (\sum_{j=1}^N |x_j|^p)^{1/p}$ for $p > 0$ and its ℓ^0 -norm as $\|x\|_0 = |\{j \in [N] : x_j \neq 0\}|$. Moreover, we let $B_{\ell^p}^N = \{x \in \mathbb{C}^N : \|x\|_p \leq 1\}$ for $p > 0$. We denote the set of s -sparse vectors of \mathbb{C}^N as $\Sigma_{s,N} := \{x \in \mathbb{C}^N : \|x\|_0 \leq s\}$ and the set of s -sparse unit vectors of \mathbb{C}^N as $D_{s,N} := \{x \in \mathbb{C}^N : \|x\|_0 \leq s, \|x\|_2 = 1\}$. Finally, the notation $X \lesssim Y$ hides the presence of a constant $c > 0$ independent of X and Y such that $X \leq cY$. Given a set S , $|S|$ denotes its cardinality. Moreover, if $S \subseteq [N]$, then S^c is the complement set of S with respect to $[N]$. We use letters c, c' or c_k with $k \in \mathbb{N}$ to denote universal (or absolute) constants and it is understood that such numbers are not necessarily the same on every occurrence.

2. Sparse recovery in bounded Riesz systems

This section outlines several applications of our main result, Theorem 1.1, in compressive sensing. We focus on new sparse recovery guarantees for subsampled bounded Riesz systems, which extend previously known results for these settings (Sections 2.1 and 2.2). Most of the recovery guarantees rely on the notion of restricted isometry property.

Definition 2.1 (*Restricted isometry property*). The restricted isometry constant ε_s of a matrix $A \in \mathbb{C}^{m \times N}$ is defined by

$$\varepsilon_s := \sup_{f \in D_{s,N}} \left| \|Af\|_2^2 - 1 \right|. \tag{2.1}$$

If a matrix $A \in \mathbb{C}^{m \times N}$ satisfies $\varepsilon_s \leq \varepsilon$ for some $\varepsilon > 0$, then we say that A satisfies RIP(ε, s).

2.1. Subsampled bounded Riesz systems

The first application of Theorem 1.1 is concerned with the recovery of functions having a sparse expansion with respect to a Riesz system. We start by recalling the definition of a Riesz system.

Definition 2.2 (*Riesz system*). Let $(H, \langle \cdot, \cdot \rangle_H)$ denote a complex Hilbert space. A sequence $(\psi_j)_{j \in \mathbb{N}}$ with $\psi_j \in H$ is called a Riesz system if there exist constants $0 < c_\psi \leq C_\psi < \infty$, such that for every $f = (f_n)_{n \in \mathbb{N}} \in \ell^2(\mathbb{C})$,

$$c_\psi \|f\|_{\ell^2(\mathbb{C})}^2 \leq \left\| \sum_{j \in \mathbb{N}} f_j \psi_j \right\|_H^2 \leq C_\psi \|f\|_{\ell^2(\mathbb{C})}^2. \tag{2.2}$$

Let us point out that an orthogonal system is a Riesz system with constants $c_\psi = C_\psi = 1$.

In the following, we focus on Riesz systems for Hilbert spaces of the form $H = L^2(S, \mu)$, where $S \subseteq \mathbb{R}^r$ and μ is a probability measure on S . For the Hilbert space $L^2(S, \mu)$ the inner product $\langle \cdot, \cdot \rangle_{L^2(S, \mu)}$ and the corresponding norm $\| \cdot \|_{L^2(S, \mu)}$ are given by

$$\langle f, g \rangle_{L^2(S, \mu)} = \int_S f(\omega) \overline{g(\omega)} d\mu(\omega) \quad \text{and} \quad \|f\|_{L^2(S, \mu)} = \sqrt{\langle f, f \rangle_{L^2(S, \mu)}}.$$

We consider a sequence $(\psi_j : S \rightarrow \mathbb{C})_{j \in \mathbb{N}}$ of bounded, measurable functions satisfying (2.2). Let $F \in H$ be a function with a finite and sparse Riesz expansion, i.e.,

$$F = \sum_{j \in [N]} f_j \psi_j \quad \text{and} \quad f \in \Sigma_{s, N}. \tag{2.3}$$

In this setting, we are interested in the problem of recovering the coefficients f of the function $F \in L^2(S, \mu)$ with respect to $(\psi_j)_{j \in [N]}$ from a finite number of samples $F(\omega_1), \dots, F(\omega_m)$, where $\omega_1, \dots, \omega_m \in S$ are drawn i.i.d. at random with respect to the measure μ . We will outline two approaches to address this question. The first one is based on the restricted isometry property (Theorem 2.3) and the second one is based on the ℓ^2 -robust null space property (Theorem 2.6).

The first approach is to show that the matrix

$$A = \frac{1}{\sqrt{mC_\psi}} (\psi_j(\omega_i))_{i \in [m], j \in [N]}, \tag{2.4}$$

satisfies $\text{RIP}(\varepsilon, s)$ under suitable conditions on m, s , and ε . If A satisfies $\text{RIP}(\varepsilon, 2s)$ for $\varepsilon \leq 1/\sqrt{2}$, then the coefficients of F with respect to $(\psi_j)_{j \in [N]}$ can be recovered from noisy observations $y_1 = F(\omega_1) + e_1, \dots, y_m = F(\omega_m) + e_m$ with $\|e\|_2 \leq \zeta$ via the quadratically-constrained basis pursuit program (see, e.g., [17, Theorem 2.1])

$$\min_{z \in \mathbb{C}^N} \|z\|_1 \quad \text{subject to} \quad \|\sqrt{C_\psi m} Az - y\|_2 \leq \zeta. \tag{BP_\zeta}$$

This approach will lead to some limitations on the restricted isometry constant ε_s that are particularly restrictive when the ratio c_ψ/C_ψ is small.

Theorem 2.3. *There are absolute constants $c_0, c_1 > 0$, such that the following holds. Let $H = L^2(S, \mu)$ and let $(\psi_j)_{j \in [N]}$ be a Riesz system. Let $\varepsilon \in (1 - \frac{c_\psi}{C_\psi}, 1)$, $K_\psi = \max_{j \in [N]} \|\psi_j\|_{L^\infty(S)}$ and assume that*

$$m \geq c_0 \max\{C_\psi^{-2}, 1\} K_\psi^2 \eta^{-2} s \log^2(s K_\psi^2 \max\{C_\psi^{-2}, 1\} \eta^{-2}) \log(eN), \tag{2.5}$$

where $\eta = (\varepsilon - 1 + \frac{c_\psi}{C_\psi}) > 0$. Then, with probability at least

$$1 - 2 \exp(-c_1 \min\{C_\psi^2, 1\} \eta^2 m / (K_\psi^2 s))$$

the matrix A defined in (2.4) satisfies $\text{RIP}(\varepsilon, s)$.

Remark 2.4. As pointed out above, an orthonormal system is a special case of a Riesz system with $c_\psi = C_\psi = 1$. Therefore, Theorem 2.3 also recovers and extends known results for bounded orthonormal systems [9,30,22].

Proof of Theorem 2.3. Let $X(\omega) := (\psi_j(\omega))_{j \in [N]}$, where $\omega \in S$ is chosen at random with respect to μ , and let X_1, \dots, X_m denote independent copies of X . Then, recalling (2.2), (2.4) and that $\|f\|_2 = 1$ for $f \in D_{s,N}$ we see that

$$\mathbb{E}|\langle X, f \rangle|^2 = \|F\|_{L^2(S, \mu)}^2 \in [c_\psi \|f\|_2^2, C_\psi \|f\|_2^2] = [c_\psi, C_\psi]$$

and

$$\frac{1}{m} \sum_{j=1}^m |\langle X_j, f \rangle|^2 = C_\psi \|Af\|_2^2.$$

Further, let $\kappa, c_1 > 0$ denote the constants in Theorem 1.1 and observe that $D_{s,N} \subseteq \sqrt{s}B_{\ell^1}^N$. Therefore, setting $\eta = \varepsilon - 1 + \frac{c_\psi}{C_\psi}$ and applying Theorem 1.1 with $\delta = \frac{1}{2}c_1 \min\{C_\psi, 1\} \cdot \eta$, $T = \sqrt{s}B_{\ell^1}^N \cap S^{N-1}$ and noticing that with these parameters we have $\delta < \kappa$, shows that the following event occurs with probability exceeding $1 - 2 \exp(-c \min\{C_\psi^2, 1\} \eta^2 m / (sK_\psi^2))$:

$$\begin{aligned} \sup_{f \in D_{s,N}} \left| \|Af\|_2^2 - \frac{1}{C_\psi} \mathbb{E}|\langle f, X \rangle|^2 \right| &\leq \frac{\min\{C_\psi, 1\} \eta}{C_\psi} \frac{\eta}{2} + \frac{\min\{C_\psi, 1\} \eta}{C_\psi} \frac{\eta}{2} \sup_{f \in D_{s,N}} \mathbb{E}|\langle f, X \rangle|^2 \\ &\leq \min\{C_\psi, 1\} \left(1 + \frac{1}{C_\psi}\right) \frac{\eta}{2} \leq \eta, \end{aligned}$$

provided that (2.5) is satisfied for a suitable constant $c_0 > 0$. In the above event, we find that for all $f \in D_{s,N}$ the inequality

$$\frac{c_\psi}{C_\psi} - \eta \leq \|Af\|_2^2 \leq 1 + \eta, \tag{2.6}$$

holds. Recalling the definition of η , the inequality (2.6) reads

$$(1 - \varepsilon) \leq \|Af\|_2^2 \leq \left(1 + \varepsilon - 1 + \frac{c_\psi}{C_\psi}\right) \leq (1 + \varepsilon),$$

which implies RIP(ε, s) for $\varepsilon \in (1 - \frac{c_\psi}{C_\psi}, 1)$ and concludes the proof. \square

Remark 2.5. An inspection of the proof of Theorem 2.3 reveals that a sufficient assumption on $(\psi_j)_{j \in [N]}$ is that the relation (2.2) holds only for every $f \in \Sigma_{s,N}$. This relaxed assumption will be used for the analysis in Section 2.2 and for the application to the CORSING method in Section 3.

As pointed out above, the fact that the matrix A defined in (2.4) only satisfies RIP(ε, s) for $\varepsilon \in (1 - \frac{c_\psi}{C_\psi}, 1)$ might cause difficulties for recovering the coefficients of the function F , since, e.g., [29, Theorem 6.12] requires that A satisfies RIP($\varepsilon, 2s$) for $\varepsilon \leq 1/\sqrt{2}$ in order to recover all s -sparse signals via (BP $_{\zeta}$). Thus, if the ratio c_ψ/C_ψ satisfies $c_\psi/C_\psi \leq (\sqrt{2} - 1)/\sqrt{2}$, then the assumption on the restricted isometry constant cannot be satisfied. We note that it is possible to relax the assumption on ε and require that A satisfies RIP(ε, ts) with $\varepsilon < \sqrt{(t-1)/t}$ with $t \geq 4/3$ (see [17, Theorem 2.1]). Following this approach requires that $t \geq C_\psi/c_\psi$ and therefore requires the number of measurements m at rate C_ψ/c_ψ .

Below we provide an alternative analysis of the problem of recovering the coefficients of F based on the ℓ^2 -robust null space property. This will circumvent the limitations due to the restricted isometry property analysis and leads to a better measurement complexity.

Theorem 2.6. *There exist absolute constants $c_0, c_1 > 0$ such that the following holds. Let $s, m, N \in \mathbb{N}$ and $A \in \mathbb{C}^{m \times N}$ be defined as in (2.4). Let $\|e\|_2 \leq \zeta$, $f \in \mathbb{C}^N$ and set $y = Af + e \in \mathbb{C}^m$. Assume that*

$$m \geq c_0 \left(\frac{\max\{1, C_\psi\}}{c_\psi} \right)^2 K_\psi^2 s \log^2(s K_\psi^2 \max\{1, C_\psi\}/c_\psi) \log(eN). \tag{2.7}$$

Then, with probability at least $1 - 2 \exp(-c_1(c_\psi/\max\{1, C_\psi\})^2 m/(s K_\psi^2))$ any minimizer $f^\#$ of the program (BP $_\zeta$) satisfies

$$\begin{aligned} \|f - f^\#\|_1 &\leq \frac{9}{2} \min_{z \in \Sigma_{s,N}} \|f - z\|_1 + 14 \frac{C_\psi}{c_\psi} \sqrt{s} \zeta, \\ \|f - f^\#\|_2 &\leq \frac{9}{2\sqrt{s}} \min_{z \in \Sigma_{s,N}} \|f - z\|_1 + 14 \frac{C_\psi}{c_\psi} \zeta. \end{aligned}$$

Remark 2.7. If $f \in \Sigma_{s,N}$, then $\min_{z \in \Sigma_{s,N}} \|f - z\|_1 = 0$ and we conclude that for universal constants $c, c' > 0$ $\|f - f^\#\|_1 \leq c(C_\psi/c_\psi)\sqrt{s}\zeta$ and $\|f - f^\#\|_2 \leq c'(C_\psi/c_\psi)\zeta$.

As mentioned above, the proof of Theorem 2.6 relies on the notion of the ℓ^2 -robust null space property and its relation to sparse recovery via ℓ^1 -minimization.

Definition 2.8. A matrix $A \in \mathbb{C}^{m \times N}$ satisfies the ℓ^2 -robust null space property of order s with constants $\alpha \in (0, 1)$ and $\tau > 0$ if

$$\|v_S\|_2 \leq \frac{\alpha}{\sqrt{s}} \|v_{S^c}\|_1 + \tau \|Av\|_2,$$

holds for every $v \in \mathbb{C}^N$ and every set $S \subseteq [N]$ with $|S| \leq s$.

The relation between stable recovery of $f \in \mathbb{C}^N$ and the ℓ^2 -robust null space property is stated by the following theorem (see, e.g., [29, Theorem 4.22]).

Theorem 2.9. *Let $A \in \mathbb{C}^{m \times N}$ satisfy the ℓ^2 -robust null space property of order s with constants $\alpha \in (0, 1)$ and $\tau > 0$. Let $\|e\|_2 \leq \zeta$, $f \in \mathbb{C}^N$ and set $y = Af + e$. Then any minimizer $f^\#$ of the program (BP $_\zeta$) satisfies*

$$\begin{aligned} \|f - f^\#\|_1 &\leq c_0 \min_{z \in \Sigma_{s,N}} \|f - z\|_1 + c_1 \sqrt{s} \zeta, \\ \|f - f^\#\|_2 &\leq \frac{c_0}{\sqrt{s}} \min_{z \in \Sigma_{s,N}} \|f - z\|_1 + c_1 \zeta, \end{aligned}$$

with $c_0 = \frac{(1+\alpha)^2}{1-\alpha}$ and $c_1 = \frac{(3+\alpha)}{1-\alpha} \tau$.

Proof of Theorem 2.6. By Theorem 2.9 it suffices to show that A as defined in (2.4) satisfies the ℓ^2 -robust null space property of order s with $\alpha = \frac{1}{2}$ and $\tau = \frac{2C_\psi}{c_\psi}$. Define the cone

$$\mathcal{T}_{\nu,s} := \left\{ f \in \mathbb{C}^N : \forall S \subseteq [N] \text{ with } |S| \leq s : \|f_S\|_2 \geq \frac{\alpha}{\sqrt{s}} \|f_{S^c}\|_1 \right\}. \tag{2.8}$$

It is known (see e.g. [26,32]) that if

$$\inf_{z \in \mathcal{T}_{\nu,s} \cap S^{N-1}} \|Af\|_2 \geq \frac{1}{\tau},$$

with $S^{N-1} := \{f \in \mathbb{C}^N : \|f\|_2 = 1\}$, then A satisfies the ℓ^2 -robust null space property with constants τ and α . Moreover, by [26, Lemma 3],

$$\mathcal{T}_{\alpha,s} \subseteq (2 + \alpha^{-1}) \operatorname{conv}(D_{s,N}) \subseteq (2 + \alpha^{-1})\sqrt{s}B_{\ell^1}^N.$$

Hence, it is enough to show that

$$\inf_{z \in 4\sqrt{s}B_{\ell^1}^N \cap S^{N-1}} \|Af\|_2 \geq \frac{c_\psi}{2C_\psi}.$$

This lower bound is found by applying Theorem 1.1 with a minor modification, which corresponds to working with $s' = 16s$ and multiplying the right-hand side of (1.2) by 16. Since $\mathbb{E}\|Af\|_2^2 \geq \frac{c_\psi}{C_\psi}\|f\|_2^2$,

$$\begin{aligned} \inf_{z \in \sqrt{s'}B_{\ell^1}^N \cap S^{N-1}} \|Az\|_2^2 &\geq \inf_{z \in \sqrt{s'}B_{\ell^1}^N \cap S^{N-1}} \mathbb{E}\|Az\|_2^2 - \sup_{f \in \sqrt{s'}B_{\ell^1}^N \cap S^{N-1}} \left| \|Af\|_2^2 - \mathbb{E}\|Af\|_2^2 \right| \\ &\geq \frac{c_\psi}{C_\psi} - \sup_{f \in \sqrt{s'}B_{\ell^1}^N \cap S^{N-1}} \left| \|Af\|_2^2 - \mathbb{E}\|Af\|_2^2 \right|. \end{aligned}$$

By Theorem 1.1 applied for $\sqrt{s'}B_{\ell^1}^N \cap S^{N-1} \subseteq \sqrt{s'}B_{\ell^1}^N$ and $\delta = c_\psi/(64 \cdot c_0 \max\{1, C_\psi\})$, where $c_0 > 0$ is the constant from Theorem 1.1, the supremum is bounded with probability at least $1 - 2 \exp(-c_1(c_\psi/\max\{1, C_\psi\})^2 m/(sK_\psi^2))$ by

$$\begin{aligned} \sup_{f \in \sqrt{s'}B_{\ell^1}^N \cap S^{N-1}} \left| \|Af\|_2^2 - \mathbb{E}\|Af\|_2^2 \right| &\leq \frac{16 \cdot c_0 c_\psi}{64 \cdot c_0 C_\psi \max\{1, C_\psi\}} \left(1 + \sup_{f \in \sqrt{s'}B_{\ell^1}^N \cap S^{N-1}} \mathbb{E}|\langle f, X \rangle|^2 \right) \\ &\leq \frac{c_\psi}{4C_\psi \max\{1, C_\psi\}} (1 + C_\psi) \end{aligned} \tag{2.9}$$

provided that (recall that $s' = 16s$) for a constant $c > 0$ depending on c_0

$$m \geq c \cdot \left(\frac{\max\{1, C_\psi\}}{c_\psi} \right)^2 K_\psi^2 s \log(sK_\psi^2 \max\{1, C_\psi\}/c_\psi)^2 \log(eN).$$

We distinguish two cases on the event (2.9). For $C_\psi \leq 1$, we find

$$\frac{c_\psi}{4C_\psi \max\{1, C_\psi\}} (1 + C_\psi) \leq \frac{c_\psi}{4C_\psi} (1 + C_\psi) \leq \frac{c_\psi}{2C_\psi}. \tag{2.10}$$

On the other hand, for $C_\psi > 1$,

$$\frac{c_\psi}{4C_\psi \max\{1, C_\psi\}} (1 + C_\psi) \leq \frac{c_\psi}{4C_\psi^2} (1 + C_\psi) \leq \frac{c_\psi}{2C_\psi}. \tag{2.11}$$

Combining the estimates (2.9), (2.10) and (2.11) concludes the proof. \square

2.2. Coherence-based sampling

We conclude this section with an application of Theorem 1.1 that will be essential for our analysis of the CORSING method in Section 3. We consider vectors $\{b_1, \dots, b_M\} \subseteq \mathbb{C}^N$ with $M \geq N$, such that the matrix

$$B = (b_1 \mid b_2 \mid \cdots \mid b_M)^T \in \mathbb{C}^{M \times N}, \tag{2.12}$$

satisfies

$$c_B \|f\|_2^2 \leq \|Bf\|_2^2 \leq C_B \|f\|_2^2, \quad \text{for all } f \in \Sigma_{s,N}, \tag{2.13}$$

and suitable constants $0 < c_B \leq C_B$. The constants c_B and C_B are also known as the minimum and the maximum s -sparse eigenvalues of B^*B (see [46]). Our goal is to construct a matrix that satisfies the restricted isometry property by sampling a small number of rows of B . Following [35], the idea is to sample according to the local coherence. We say that $\nu \in \mathbb{R}^M$ is a local coherence for B if

$$\max_{n \in [N]} |B_{jn}|^2 \leq \nu_j \quad \text{for all } j \in [M]. \tag{2.14}$$

Define a random vector $X \in \mathbb{C}^N$ by

$$\mathbb{P}\left(X = \sqrt{\frac{\|\nu\|_1}{\nu_j}} b_j\right) = \frac{\nu_j}{\|\nu\|_1}. \tag{2.15}$$

Let X_1, \dots, X_m denote independent copies of X and consider

$$A = \frac{1}{\sqrt{mC_B}} (X_1 \mid \cdots \mid X_m)^T. \tag{2.16}$$

Then for any $f \in \mathbb{C}^N$,

$$\mathbb{E}|\langle X, f \rangle|^2 = \sum_{j=1}^M \frac{\nu_j}{\|\nu\|_1} \frac{\|\nu\|_1}{\nu_j} |\langle b_j, f \rangle|^2 = \|Bf\|_2^2 \quad \text{and} \quad \frac{1}{m} \sum_{i=1}^m |\langle X_i, f \rangle|^2 = C_B \|Af\|_2^2.$$

Moreover, for each $i \in [N]$ and e_i a standard basis vector of \mathbb{C}^m we have

$$|\langle X, e_i \rangle| \leq \sqrt{\|\nu\|_1} \max_{j \in [M]} \frac{1}{\sqrt{\nu_j}} \max_{n \in [N]} |B_{jn}| \leq \sqrt{\|\nu\|_1}. \tag{2.17}$$

Now, applying Theorem 1.1 and arguing similarly to Theorem 2.3, we obtain the following RIP result for the matrix A .

Theorem 2.10. *There exist universal constants $c_0, c_1 > 0$, such that the following holds. Let $B \in \mathbb{C}^{M \times N}$ be such that (2.13) holds. Let $s, m \in \mathbb{N}$ and $\varepsilon \in (1 - \frac{c_B}{C_B}, 1)$. Let $\nu \in \mathbb{R}^M$ be such that (2.14) holds and assume*

$$m \geq c_0 \max\{C_B^{-2}, 1\} \eta^{-2} s \|\nu\|_1 \log^2(s \|\nu\|_1 \max\{C_B^{-2}, 1\} \eta^{-2}) \log(eN),$$

with $\eta = \varepsilon - 1 + \frac{c_B}{C_B}$. Then, with probability at least $1 - 2 \exp(-c_1 \min\{C_B^2, 1\} \eta^2 m / (\|\nu\|_1 s))$ the matrix A defined in (2.16) satisfies $\text{RIP}(\varepsilon, s)$.

2.3. Extension to weighted ℓ^1 -minimization

The previous sections studied the impact of Theorem 1.1 on sparse recovery via ℓ^1 -minimization. The weighted ℓ^1 -minimization program (see $(\text{BP}_{w,\zeta})$ below) was suggested in [44] as a means of incorporating additional information on the smoothness of the function $F : S \rightarrow \mathbb{C}$, when trying to solve the interpolation

problem discussed in Section 2.1. Recall that we aim at finding the coefficients of a function $F : S \rightarrow \mathbb{C}$ with respect to a bounded Riesz system $(\psi_j)_{j \in \mathbb{N}}$ from a finite number noisy observations $F(\omega_1) + e_1, \dots, F(\omega_m) + e_m$. Our strategy is based on the assumption that the sequence of coefficients of $(f_j)_{i \in [N]}$, which represent the function $F = \sum_{j \in [N]} f_j \psi_j$ with respect to $(\psi_j)_{j \in [N]}$, is s -sparse. In real world scenarios smoothness of F often plays a crucial role. In this section we discuss the problem of amalgamating smoothness assumptions with sparsity assumptions on F . Weighted sparsity and weighted ℓ_1 -minimization is also crucial for compressive sensing approaches to function recovery in high-dimensions and, in particular, for solving parametric PDEs [42,22,8] appearing for instance in the context of uncertainty quantification. We refer the reader to [44,42,22,3] and the references therein for an introduction to the subject and examples. In this section our path to recovery guarantees for weighted ℓ^1 -minimization passes through a suitable weighted notion of the restricted isometry property as used in [44].

2.3.1. *Weighted ℓ^p -spaces and the weighted RIP*

In order to quantify smoothness information on the function that we are trying to interpolate, we follow the reference [44] and use the following weighted versions of ℓ^p -spaces for $0 < p \leq 2$. Let $w = (w_j)_{j \in [N]}$ denote a sequence of weights with $w_j \geq 1$ and for $0 < p \leq 2$ set

$$\ell_w^p := \left\{ f \in \mathbb{C}^N : \|f\|_{w,p} := \left(\sum_{j \in [N]} w_j^{2-p} |f_j|^p \right)^{\frac{1}{p}} \right\}. \tag{2.18}$$

Later the weight sequence $(w_j)_{j \in [N]}$ is chosen in a way, such that $w_j \geq \|\psi_j\|_{L^\infty(S)}$. The work [44] recognized that there is a notion of sparsity which is consistent with the weighted version of ℓ^p -spaces in (2.18). For each $f \in \mathbb{C}^N$ we define a suitable version of the ℓ^0 -norm given by setting

$$\|f\|_{w,0} = \sum_{j \in \{j : f_j \neq 0\}} w_j^2, \tag{2.19}$$

and calling a signal $f \in \mathbb{C}^N$ s -sparse respect to the sequence w , if $\|f\|_{w,0} \leq s$.

The notion of sparsity with respect to a weight sequence $w \in [1, \infty)^N$ gives rise to the following version of the restricted isometry property.

Definition 2.11 (*Weighted restricted isometry property*). Let

$$D_{s,N}(w) := \{f \in \mathbb{C}^N : \|f\|_{w,0} \leq s, \|f\|_2 = 1\}. \tag{2.20}$$

The weighted restricted isometry constant $\varepsilon_{s,w}$ of a matrix $A \in \mathbb{C}^{m \times N}$ is defined by

$$\varepsilon_{s,w} := \sup_{f \in D_{s,N}(w)} \left| \|Af\|_2^2 - 1 \right|. \tag{2.21}$$

If a matrix $A \in \mathbb{C}^{m \times N}$ satisfies $\varepsilon_{s,w} \leq \varepsilon$ for some $\varepsilon > 0$, then we say that A satisfies w -RIP(ε, s).

A crucial observation is that the Cauchy-Schwarz inequality implies that every $f \in D_{s,N}(w)$ satisfies

$$\|f\|_{w,1} \leq \left(\sum_{j=1}^N w_j^2 \right)^{1/2} \|f\|_2 \leq \sqrt{s}.$$

Hence, we find the familiar inclusion

$$D_{s,N}(w) \subset \sqrt{s} B_{\ell_w^1}^N \cap S^{N-1}. \tag{2.22}$$

2.3.2. Recovery by weighted ℓ^1 -minimization

In Section 2.1 we leveraged a connection between the restricted isometry property and recovery guarantees of ℓ^1 -minimization in order to use Theorem 1.1. A similar connection was observed for the weighted ℓ^1 -minimization program,

$$\min_{z \in \mathbb{C}^N} \|z\|_{w,1} \quad \text{subject to} \quad \|\sqrt{C_\psi} m A z - y\|_2 \leq \zeta. \tag{BP}_{w,\zeta}$$

Let us state the following simplified version of this connection from [44].

Theorem 2.12 ([44, Theorem 4.5 & Corollary 4.3]). *Let $A \in \mathbb{C}^{m \times N}$ such that A satisfies w -RIP(ε, s) for $\varepsilon \leq 1/3$ and for $s \geq 2\|w\|_\infty^2$. Let $f \in D_{s,N}(w)$ and set $y = Af + e$ with $\|e\|_2 \leq \zeta$. Further, let $f^\#$ denote the minimizer of $(BP_{w,\zeta})$, then there exist constants $c_0, c_1 > 0$ depending only on $\varepsilon > 0$, such that*

$$\begin{aligned} \|f - f^\#\|_{w,1} &\leq c_0 \sqrt{s} \zeta, \\ \|f - f^\#\|_2 &\leq c_1 \zeta. \end{aligned}$$

Using this result we can employ the same strategy as in Section 2.1 in order to show recovery guarantees for the weighted ℓ^1 -minimization program, provided that we have a suitable version of Theorem 1.1.

2.3.3. Weighted restricted isometry property for Riesz frames

Recalling the inclusion (2.22), we will use the following result in order to obtain recovery guarantees for the program $(BP_{w,\zeta})$.

Theorem 2.13. *There exist absolute constants $\kappa, c_0, c_1 > 0$ such that the following holds. Let X_1, \dots, X_m be independent copies of a random vector $X \in \mathbb{C}^N$ such that for all $j = 1, \dots, N$ we have $|\langle X, e_j \rangle| \leq K_j$ for some $K_j > 0$ where e_1, \dots, e_N is the standard basis of \mathbb{C}^N . Let $w_j \geq K_j$ for all $j = 1, \dots, N$ and let $T \subseteq \{f \in \mathbb{C}^N : \|f\|_{w,1} \leq \sqrt{s}\}$, $\delta \in (0, \kappa)$. Assume further that*

$$m \geq c_0 \delta^{-2} s \log(eN) \log^2(s/\delta). \tag{2.23}$$

Then, with probability exceeding $1 - 2 \exp(-\delta^2 m/s)$,

$$\sup_{f \in T} \left| \frac{1}{m} \sum_{i=1}^m |\langle f, X_i \rangle|^2 - \mathbb{E} |\langle f, X \rangle|^2 \right| \leq c_1 \left(\delta + \delta \sup_{f \in T} \mathbb{E} |\langle f, X \rangle|^2 \right). \tag{2.24}$$

This enables us to show the following result on the w -RIP(ε, s) for truncated Riesz sequences (and in particular for bounded orthonormal systems).

Theorem 2.14. *There are absolute constants $c_0, c_1 > 0$, such that the following holds. Let $H = L^2(S, \mu)$ and let $(\psi_j)_{j \in [N]}$ be a truncated Riesz system. Let $\varepsilon \in (1 - \frac{c_\psi}{C_\psi}, 1)$, $w_j \geq \|\psi_j\|_{L^\infty(S)}$ and assume that*

$$m \geq c_0 \max\{C_\psi^{-2}, 1\} \eta^{-2} s \log(s \max\{C_\psi^{-2}, 1\} \eta^{-2})^2 \log(eN), \tag{2.25}$$

where $\eta = (\varepsilon - 1 + C_\psi) > 0$. Then, with probability at least $1 - 2 \exp(-c_1 \min\{C_\psi, 1\} \eta^2 m/s)$ the matrix A defined in (2.4) satisfies w -RIP(ε, s).

Theorem 2.14 follows from Theorem 2.13 by chasing through to argument of Theorem 2.3, and replacing the application of Theorem 1.1 to $D_{s,N} \subseteq \sqrt{s} B_{\ell^1}^N$ by applying Theorem 2.13 to the weighted version $D_{s,N}(w) \subseteq \sqrt{s} B_{\ell^1_w}^N$.

We will outline how to deduce Theorem 2.13 after we have established the proof of Theorem 1.1. The difference of the proofs of Theorem 1.1 and of Theorem 2.13 is limited to technical details.

3. Application to numerical approximation of PDEs

As an application of the theory presented in Section 2, we consider the CORSING (COmpressed SolvING) method for the numerical approximation of solutions of PDEs based on compressive sensing [12,15,16]. The CORSING method is a general paradigm to compute a sparse approximation to the solution of a PDE that admits a weak formulation. It assembles a reduced Petrov-Galerkin discretization via compressive sensing. As discussed in [16], the advantages of this method compared to other nonlinear approximation methods for PDEs, such as adaptive finite elements or adaptive wavelets methods (see, e.g., [49] and references therein), are that (i) no *a posteriori* error indicators are needed and (ii) the assembly of the discretization matrix and the sparse recovery, here performed via Orthogonal Matching Pursuit (OMP), can be easily parallelized.

Here, we will focus on the restricted isometry analysis and sparse recovery guarantees for CORSING. For numerical experiments for multi-dimensional advection-diffusion-reaction equations and for the Stokes problem, we refer the reader to [15,12,16,14]. We also note in passing that the CORSING paradigm can be adapted to the framework of collocation techniques for PDEs (see [13]). For an overview of numerical methods for PDEs based on compressive sensing, we refer the reader to [13, Section 1.2].

Concerning the sparse recovery method, we resort to OMP as opposed to ℓ^1 -minimization for two main reasons, mainly related to computational efficiency considerations: (i) given a target sparsity level, using OMP we can easily control the number of iterations and, consequently, the computational cost of the recovery phase; (ii) OMP is easily parallelizable. For a numerical comparison between OMP and ℓ^1 -minimization in this context, we refer to [15, Section 5]. Finally, we note that the restricted isometry analysis of CORSING presented here can be applied when different sparse recovery procedures are considered.

The section is structured as follows. After describing the setting of weak problems in Hilbert spaces in Section 3.1, we recall the main elements of the CORSING method in Section 3.2. Then, we present an RIP analysis for the CORSING discretization matrix in Section 3.3 and discuss recovery guarantees for the method when the solution is approximated via OMP in Section 3.4.

3.1. Weak problems in Hilbert spaces

Let U and V be separable Hilbert spaces equipped with inner products $(\cdot, \cdot)_U$ and $(\cdot, \cdot)_V$, and norms $\|\cdot\|_U = (\cdot, \cdot)_U^{1/2}$ and $\|\cdot\|_V = (\cdot, \cdot)_V^{1/2}$. We consider a weak problem of the form

$$\text{find } u \in U : \quad a(u, v) = \mathcal{F}(v), \quad \forall v \in V, \quad (3.1)$$

where $a : U \times V \rightarrow \mathbb{R}$ is a bilinear form and $\mathcal{F} \in V^*$, the dual space of V . An important example of (3.1) is the weak formulation of the advection-diffusion-reaction equation with homogeneous boundary conditions, where $U = V = H_0^1(\mathcal{D})$, $\mathcal{D} \subseteq \mathbb{R}^d$ is the physical domain, and the bilinear form is

$$a(u, v) = \int_{\mathcal{D}} \mu \nabla u \cdot \nabla v + (\beta \cdot \nabla u)v + \rho uv \, dx, \quad \forall u, v \in H_0^1(\mathcal{D}), \quad (3.2)$$

where $\mu : \mathcal{D} \rightarrow \mathbb{R}$, $\beta : \mathcal{D} \rightarrow \mathbb{R}^d$, and $\rho : \mathcal{D} \rightarrow \mathbb{R}$ are the diffusion, advection, and reaction coefficients. The operator \mathcal{F} is defined by

$$\mathcal{F}(u) := \int_{\mathcal{D}} uF \, dx, \quad \forall u \in H_0^1(\mathcal{D}),$$

where $F : \mathcal{D} \rightarrow \mathbb{R}$ is the forcing term (see, e.g., [40] for more details). The analysis in this section will focus on abstract weak problems of the form (3.1), but it can be specified to particular PDEs, such as (3.2) (see [16]).

In order to guarantee the existence and uniqueness of the solution to (3.1), we assume that the bilinear form $a(\cdot, \cdot)$ satisfies the hypotheses of the classical Babuška-Nečas theory (see, e.g., [40, Theorem 5.1.2]), namely

$$\exists \alpha > 0 : \quad \inf_{u \in U \setminus \{0\}} \sup_{v \in V \setminus \{0\}} \frac{a(u, v)}{\|u\|_U \|v\|_V} \geq \alpha, \tag{3.3}$$

$$\exists \beta > 0 : \quad \sup_{u \in U \setminus \{0\}} \sup_{v \in V \setminus \{0\}} \frac{a(u, v)}{\|u\|_U \|v\|_V} \leq \beta, \tag{3.4}$$

$$\sup_{u \in U} a(u, v) > 0, \quad \forall v \in V \setminus \{0\}. \tag{3.5}$$

Consider two Riesz bases $(\varphi_j)_{j \in \mathbb{N}}$ and $(\xi_q)_{q \in \mathbb{N}}$ for U and V , respectively, i.e., satisfying relation (2.2). We denote the lower and the upper Riesz constants of $(\varphi_j)_{j \in \mathbb{N}}$ as c_φ and C_φ , respectively. Analogously, we denote the Riesz constants of the system $(\xi_q)_{q \in \mathbb{N}}$ as c_ξ and C_ξ . Finally, define the reconstruction and decomposition operators $\Phi : \ell^2(\mathbb{N}) \rightarrow U$ and $\Phi^* : U \rightarrow \ell^2(\mathbb{N})$, respectively, by

$$\Phi x = \sum_{j \in \mathbb{N}} x_j \varphi_j, \quad \forall x \in \ell^2(\mathbb{N}), \quad (\Phi^* u)_j = (u, \varphi_j^*)_U, \quad \forall j \in \mathbb{N},$$

where $(\varphi_j^*)_{j \in \mathbb{N}}$ is the biorthogonal basis of $(\varphi_j)_{j \in \mathbb{N}}$ (see [23]). The operators Ξ and Ξ^* are defined analogously. By the Riesz property,

$$c_\varphi \|x\|_2^2 \leq \|\Phi x\|_2^2 \leq C_\varphi \|x\|_2^2, \quad c_\xi \|x\|_2^2 \leq \|\Xi x\|_2^2 \leq C_\xi \|x\|_2^2. \tag{3.6}$$

We discretize problem (3.1) via a Petrov-Galerkin approach. Let $N, M \in \mathbb{N}$ and consider the finite dimensional truncated spaces

$$U^N := \text{span}(\varphi_j)_{j \in [N]}, \quad V^M := \text{span}(\xi_q)_{q \in [M]}, \tag{3.7}$$

called the *trial* and the *test space*, respectively. Accordingly we call $(\varphi_j)_{j \in [N]}$ and $(\xi_j)_{j \in [N]}$ trial and test basis functions, respectively. We associate with these spaces a finite dimensional formulation of (3.1), namely

$$\text{find } u \in U^N : \quad a(u, v) = \mathcal{F}(v), \quad \forall v \in V^M. \tag{3.8}$$

By the bilinearity of $a(\cdot, \cdot)$ and the linearity of \mathcal{F} , (3.8) can be discretized as a linear system

$$Bz = c, \tag{3.9}$$

where $B \in \mathbb{C}^{M \times N}$ and $c \in \mathbb{C}^M$ are defined by

$$B_{qj} = a(\varphi_j, \xi_q), \quad c_q = \mathcal{F}(\xi_q), \quad \forall j \in [N], \quad \forall q \in [M]. \tag{3.10}$$

The linear system (3.9) is usually referred to as a Petrov-Galerkin discretization of (3.1), or, in particular, as a Galerkin discretization when $N = M$ and $U^N = V^M$. This general class of discretizations contains many popular numerical approximation methods for PDEs, whose most prominent example is the finite element method (see, e.g., [40]).

Remark 3.1 (On the Riesz basis assumption). When the physical domain \mathcal{D} has dimension $d = 1$, it is easy to construct orthonormal bases for the trial and the test functions. For example, in the case of homogeneous boundary conditions, i.e., $U = V = H_0^1(\mathcal{D})$ with $\mathcal{D} = (0, 1)$, we can consider hierarchical hat functions (see, e.g., [25]) as trial basis functions and Fourier-like functions (e.g., sine functions) as test basis functions. This setting, considered in [15,16], leads to two orthonormal bases of $H_0^1(\mathcal{D})$. When $\mathcal{D} = (0, 1)^d$ with $d > 1$, generalizing the Fourier-like basis while preserving orthogonality is easily done via tensorization. However, tensorizing the hierarchical basis of hat functions does not preserve orthogonality. In order to obtain a stable discretization, one can instead resort to biorthogonal spline wavelets to obtain a Riesz basis (see [14] for more details).

3.2. The CORSING method

In order to take advantage of the compressive sensing paradigm, we consider a Petrov-Galerkin discretization of (3.1), where the bilinear form is evaluated at trial and test basis functions that satisfy suitable incoherence properties. Solving the discretized linear system in (3.9) can be expensive from a computational perspective. The idea of CORSING is to solve the discretized system (3.9) via compressive sensing, with the aim of computing an s -sparse approximation to the solution to (3.1), with $s \ll N$, i.e., the computed solution belongs to the space

$$U_s^N := \left\{ \sum_{j \in [N]} w_j \varphi_j : \|w\|_0 \leq s \right\}.$$

The quality of the approximation space U_s^N depends on the sparsity or compressibility of the solution with respect to the trial basis. When the trial basis is a Fourier-like basis, sparsity or compressibility is observed when the most important frequencies of the solution are clustered over different regions of the spectrum due to multiscale phenomena (see, e.g., [37]). On the other hand, using a hierarchical (or wavelet) basis as a trial basis leads to sparsity or compressibility of solutions with localized features or boundary layers (see [12,16]).

In order to reduce the dimensionality of the discretization (3.9), we pick $m \ll M$ test indices $\tau_1, \dots, \tau_m \in [M]$ i.i.d. at random according to a discrete probability measure $p \in \mathbb{R}^M$ on the index set $[M]$, i.e.,

$$\mathbb{P}(\tau_i = q) = p_q, \quad \text{for } i \in [m], q \in [M]. \tag{3.11}$$

Then, the resulting CORSING discretization of (3.8) is given by the underdetermined linear system

$$Az = y, \tag{3.12}$$

where $A \in \mathbb{C}^{m \times N}$ and $y \in \mathbb{C}^m$ are defined by

$$A_{ij} = a(\varphi_j, \xi_{\tau_i}), \quad y_i = \mathcal{F}(\xi_{\tau_i}), \quad \forall i \in [m], \forall j \in [N]. \tag{3.13}$$

Taking advantage of the compressive sensing paradigm, in the recovery phase we seek an s -sparse solution to (3.12), by considering an approximate solution $\hat{x} \in \mathbb{C}^N$ to the optimization program

$$\hat{x} \approx \arg \min_{z \in \Sigma_{s,N}} \|D(Az - y)\|_2, \tag{3.14}$$

where $D \in \mathbb{R}^{m \times m}$ is a diagonal preconditioning matrix with diagonal elements

$$D_{ii} = \frac{1}{\sqrt{mp_{\tau_i}}}, \quad \text{for } i \in [m]. \tag{3.15}$$

The CORSING solution is then defined by

$$\hat{u} := \Phi \hat{x} = \sum_{j \in [N]} \hat{x}_j \varphi_j. \tag{3.16}$$

A detailed discussion of the case where \hat{x} is computed via OMP will be carried out in Section 3.4.

3.3. Restricted isometry property

A first theoretical analysis of the CORSING method was proposed in [16]. The main tool employed is the local a -coherence, a generalization of the local coherence [4,35] to bilinear forms in Hilbert spaces that implements a preconditioning as suggested in [43].

Definition 3.2 (*local a -coherence*). Given $N \in \mathbb{N} \cup \{\infty\}$, the sequence $\mu^N := (\mu_q^N)_{q \in \mathbb{N}}$ defined by

$$\mu_q^N := \sup_{j \in [N]} |a(\varphi_j, \xi_q)|^2, \quad q \in \mathbb{N}, \tag{3.17}$$

is called the *local a -coherence* of $(\varphi_j)_{j \in [N]}$ with respect to $(\xi_q)_{q \in \mathbb{N}}$.

In the following, we will assume that $\mu^N \in \ell^1(\mathbb{N})$, for every $N \in \mathbb{N}$. The next proposition gives a sufficient condition on μ^N and M in order for U_s^N and V^M to satisfy the inf-sup condition. The result immediately follows from [16, Lemma 3.6] and [16, Remark 3.11].

Proposition 3.3. *Let $s, N, M \in \mathbb{N}$, with $s \leq N$, and $\gamma \in (0, 1)$ be such that*

$$\sum_{q > M} \mu_q^N \leq \frac{\alpha^2 \gamma c_\varphi c_\xi}{s}. \tag{3.18}$$

Then,

$$\inf_{u \in U_s^N} \sup_{v \in V^M} \frac{a(u, v)}{\|u\|_U \|v\|_V} \geq (1 - \gamma)^{\frac{1}{2}} \alpha. \tag{3.19}$$

In the following we will assume to have access to a computable (but not necessarily sharp) upper bound $\nu^N = (\nu_q^N)_{q \in \mathbb{N}}$ to the local a -coherence, i.e., $\mu_q^N \leq \nu_q^N$ for all $q \in \mathbb{N}$ and define

$$\nu^{N, M} := (\nu_q^N)_{q \in [M]} \in \mathbb{R}^M. \tag{3.20}$$

Throughout this section, we choose the probability density p in (3.11) over the test indices as

$$p := \frac{\nu^{N, M}}{\|\nu^{N, M}\|_1}. \tag{3.21}$$

The analysis carried out in [16] provides sufficient conditions on M and m that guarantee an optimal error estimate in expectation for CORSING. Yet, the recovery analysis in [16] has two main limitations: (i) m depends quadratically on s (up to logarithmic factors), whereas from the compressive sensing theory one expects this dependence to be linear (up to logarithmic factors); (ii) the vector \hat{x} is assumed to solve (3.14) *exactly*, whereas in practice OMP (or another approach for sparse recovery) has to be employed to approximate the solution to (3.14).

These two issues are fixed in the following. In particular, thanks to the analysis based on the restricted isometry property, we show that a linear dependence of m on s (up to logarithmic factors) is sufficient and prove a recovery error estimate for the case that the CORSING solution \hat{x} is approximated via OMP (Section 3.4).

We start by defining a condition number of the weak infinite-dimensional problem (3.1) by

$$\kappa := \frac{C_\varphi C_\xi \beta^2}{c_\varphi c_\xi \alpha^2}, \tag{3.22}$$

where α and β are the inf-sup and continuity constants of $a(\cdot, \cdot)$ defined in (3.3) and (3.4), respectively. Of course, in general we have $\kappa \geq 1$.

We also introduce the preconditioned and rescaled versions of A and y defined in (3.13) by

$$\tilde{A} := \frac{1}{\beta \sqrt{C_\varphi C_\xi}} DA, \quad \tilde{y} := \frac{1}{\beta \sqrt{C_\varphi C_\xi}} Dy, \tag{3.23}$$

where D is defined as in (3.15). This normalization ensures that $\mathbb{E} \|\tilde{A}t\|_2^2 \leq \|t\|_2^2$ for every $t \in \Sigma_{s,N}$. In this setting, the following RIP result holds.

Theorem 3.4 (RIP for the CORSING matrix). *There exist universal constants $c_0, c_1 > 0$ such that the following holds. Let $s, N \in \mathbb{N}$ with $s \leq N$, $\gamma \in (0, 1)$ and choose $M = M(s, N, \mu^N)$ such that*

$$\sum_{q>M} \mu_q^N \leq \frac{\alpha^2 \gamma c_\varphi c_\xi}{s}.$$

Further, let κ be defined as in (3.22) and let $\nu^{N,M}$ be defined as in (3.20). Then, for every

$$\varepsilon \in \left(1 - \frac{1-\gamma}{\kappa}, 1\right), \quad \text{and accordingly} \quad \eta = \varepsilon - 1 + \frac{1-\gamma}{\kappa} \tag{3.24}$$

with probability at least

$$1 - 2 \exp\left(-c_1 \frac{\eta^2 m \min\{1, C_\varphi^2 C_\xi^2 \beta^4\}}{s \|\nu^{N,M}\|_1^2}\right)$$

the rescaled CORSING matrix \tilde{A} defined in (3.23) satisfies the RIP(ε, s) provided that

$$m \geq c_0 \frac{\|\nu^{N,M}\|_1^2}{\min\{1, C_\varphi^2 C_\xi^2 \beta^4\} \eta^2} s \log(eN) \log^2\left(\frac{s \|\nu^{N,M}\|_1^2}{\min\{1, C_\varphi^2 C_\xi^2 \beta^4\} \eta^2}\right). \tag{3.25}$$

Proof. This theorem is a consequence of Theorem 2.10. In order to apply this result, we need to find constants c_B and C_B such that (2.13) holds.

Let us start by estimating c_B . Thanks to Proposition 3.3 and to the Riesz property (3.6) of $(\varphi_j)_{j \in \mathbb{N}}$ and $(\xi_q)_{q \in \mathbb{N}}$, we have

$$\begin{aligned} \inf_{x \in \Sigma_{s,N} \setminus \{0\}} \frac{\|Bx\|_2^2}{\|x\|_2^2} &\geq c_\varphi \inf_{x \in \Sigma_{s,N} \setminus \{0\}} \frac{\|Bx\|_2^2}{\|\Phi x\|_U^2} = c_\varphi \left(\inf_{x \in \Sigma_{s,N} \setminus \{0\}} \sup_{z \in \mathbb{R}^M \setminus \{0\}} \frac{z^T Bx}{\|\Phi x\|_U \|z\|_2} \right)^2 \\ &\geq c_\varphi c_\xi \left(\inf_{x \in \Sigma_{s,N} \setminus \{0\}} \sup_{z \in \mathbb{R}^M \setminus \{0\}} \frac{z^T Bx}{\|\Phi x\|_U \|\Xi z\|_2} \right)^2 \end{aligned}$$

$$= c_\varphi c_\xi \left(\inf_{u \in U_s^N \setminus \{0\}} \sup_{v \in V^M \setminus \{0\}} \frac{a(u, v)}{\|u\|_U \|v\|_V} \right)^2 \geq (1 - \gamma) c_\varphi c_\xi \alpha^2.$$

Let us now estimate C_B . Using the continuity (3.4) of $a(\cdot, \cdot)$ and the Riesz property, we see that

$$\begin{aligned} \sup_{x \in \Sigma_{s, N} \setminus \{0\}} \frac{\|Bx\|_2^2}{\|x\|_2^2} &= \sup_{x \in \Sigma_{s, N} \setminus \{0\}} \frac{1}{\|x\|_2^2} \sum_{j \in [N]} \sum_{k \in [N]} x_j x_k \sum_{q \in [M]} a(\varphi_j, \xi_q) a(\varphi_k, \xi_q) \\ &\leq \sup_{u \in U_s^N \setminus \{0\}} \frac{C_\varphi}{\|u\|_U^2} \sum_{q \in [M]} a(u, \xi_q)^2. \end{aligned}$$

Notice that for every $u \in U$,

$$\sum_{q \in [M]} a(u, \xi_q)^2 = \|a(u, \cdot) \circ (\Xi|_{\mathbb{R}^M})\|_2^2 \leq \|a(u, \cdot)\|_{V^*}^2 \|\Xi|_{\mathbb{R}^M}\|_{\mathbb{R}^M \rightarrow V}^2 \leq \|u\|_U^2 \beta^2 C_\xi, \tag{3.26}$$

where Ξ is the reconstruction operator associated with the test basis $(\xi_q)_{q \in \mathbb{N}}$ and where we have used that $\|\Xi\|_{\ell^2 \rightarrow V}^2 \leq C_\xi$ and that $\|a(u, \cdot)\|_{V^*} \leq \|u\|_U \beta$. Combining the above inequalities yields

$$\sup_{x \in \Sigma_{s, N} \setminus \{0\}} \frac{\|Bx\|_2^2}{\|x\|_2^2} \leq \beta^2 C_\varphi C_\xi.$$

The proof is concluded by applying Theorem 2.10 with $c_B = (1 - \gamma) c_\varphi c_\xi \alpha^2$ and $C_B = C_\varphi C_\xi \beta^2$. \square

Remark 3.5 (*Limitations on ε*). Relation (3.24) implies a lower bound for the RIP constant. In particular, by letting $\gamma \rightarrow 0^+$ in (3.24), we obtain the necessary condition

$$\varepsilon > 1 - \frac{1}{\kappa}. \tag{3.27}$$

This will imply restrictions on κ to guarantee the recovery via OMP, studied in the next section. Note that this restriction on the RIP constant is analogous to the one discussed in Section 2.1.

3.4. Recovery via orthogonal matching pursuit

We study the performance of the CORSING recovery scheme when the approximate solution \hat{x} to (3.14) is computed via OMP. We choose the OMP algorithm thanks to its parallelizability and its ability to control the number of iterations, and hence the resulting computational cost, when an estimate of the target sparsity is known. In order to prove a precise recovery estimate, we define the output of OMP (Algorithm 3.6) taking into account the ℓ^2 -normalization of the columns (lines 1 and 8 in Algorithm 3.6).

Algorithm 3.6. (Orthogonal Matching Pursuit)

Inputs: $A = (a_1 | \dots | a_N) \in \mathbb{C}^{m \times N}$, $b \in \mathbb{C}^m$, $s \in \mathbb{N}$

Output: $\hat{x} \in \Sigma_{s, N}$

Procedure: $\hat{x} = \text{OMP}(A, y, s)$

- 1: $B \leftarrow AR$, with $R_{jk} = \delta_{jk} / \|a_j\|_2$, for $j, k \in [N]$;
- 2: $S \leftarrow \emptyset$; $\hat{x} \leftarrow 0$;
- 3: **for** $i = 1, \dots, s$ **do**
- 4: $k \leftarrow \arg \max_{j \in [N]} |(B^*(y - B\hat{x}))_j|$;
- 5: $S \leftarrow S \cup \{k\}$;

- 6: $\hat{x} \leftarrow \operatorname{argmin}_{z \in \mathbb{C}^N} \|Bz - y\|_2$ s.t. $\operatorname{supp}(z) \subseteq S$;
- 7: **end for**
- 8: $\hat{x} \leftarrow R\hat{x}$;

The main tool employed here is [29, Theorem 6.25], a recovery theorem for OMP based on the RIP. This theorem is the generalization of a result first published in [52]; see also [24]. We give its version in a discrete setting and at the same time add a slight generalization regarding the ℓ^2 -normalization of the columns.

Theorem 3.7. *There exist constants $\overline{K} \in \mathbb{N}$, $C > 0$, and $\varepsilon^* \in (0, 1)$ such that for every $s \in \mathbb{N}$, the following holds. If $A \in \mathbb{C}^{m \times N}$ satisfies the $\operatorname{RIP}(\varepsilon, (\overline{K} + 1)s)$, with $\varepsilon < \varepsilon^*$ then, for any $y \in \mathbb{C}^m$, the output $\hat{x} = \operatorname{OMP}(A, y, \overline{K}s)$ of Algorithm 3.6 is such that*

$$\|A\hat{x} - y\|_2 \leq C \inf_{z \in \Sigma_{s,N}} \|Az - y\|_2.$$

Possible values for the constants are $\overline{K} = 12$, $C = 49$. Moreover,

$$\varepsilon^* = \begin{cases} 1/6 & \text{if the columns of } A \text{ are } \ell^2\text{-normalized,} \\ 1/13 & \text{otherwise.} \end{cases}$$

Proof. If A has ℓ^2 -normalized columns, then the theorem is a direct consequence of [29, Theorem 6.25] (with a minor modification of the proof in order to have $C \inf_{z \in \Sigma_{s,N}} \|Az - y\|_2$ instead of $C \|Ax_{S^c} - y\|_2$ in the right-hand side of the recovery error bound). The values of the constants $\overline{K} = 12$, $C = 49$ and $\varepsilon^* = 1/6$ are deduced by a direct inspection of the proof (see, in particular, [29, Proposition 6.24]).

Let us now assume that there exists some $j \in [N]$, with $\|a_j\|_2 \neq 1$. First, observe that if A satisfies the $\operatorname{RIP}(\varepsilon, (\overline{K} + 1)s)$ then $1 - \varepsilon \leq \|a_j\|_2^2 \leq 1 + \varepsilon$, for every $j \in [N]$. Then, considering the normalization matrix R defined in step 1 of Algorithm 3.6, a direct computation shows that, for every $(\overline{K} + 1)s$ -sparse vector z , it holds

$$\|ARz\|_2^2 \leq (1 + \varepsilon)\|Rz\|_2^2 \leq \frac{1 + \varepsilon}{1 - \varepsilon} \|z\|_2^2 = \left(1 + \frac{2\varepsilon}{1 - \varepsilon}\right) \|z\|_2^2,$$

and, similarly $\|ARz\|_2^2 \geq (1 - 2\varepsilon/(1 + \varepsilon))\|z\|_2^2$. Hence, the matrix AR satisfies the $\operatorname{RIP}(2\varepsilon/(1 - \varepsilon), (\overline{K} + 1)s)$. Now, defining $\hat{z} := \operatorname{OMP}(AR, y, \overline{K}s)$, we have $\hat{x} = R\hat{z}$. Applying [29, Theorem 6.25] to AR , we finally obtain

$$\|A\hat{x} - y\|_2 = \|AR\hat{z} - y\|_2 \leq C \inf_{z \in \Sigma_{s,N}} \|ARz - y\|_2 = C \inf_{z \in \Sigma_{s,N}} \|Az - y\|_2,$$

provided that the RIP constant of AR is less than or equal to $1/6$. This is equivalent to require $2\varepsilon/(1 - \varepsilon) < 1/6$, which is equivalent to $\varepsilon < 1/13 =: \varepsilon^*$. \square

Remark 3.8. An inspection of the argument employed in [29, Theorem 6.25] reveals that the constants \overline{K} , C , and ε^* of Theorem 3.7 are intertwined. For example, one could relax the condition on ε^* and consequently increase the values of C and \overline{K} .

Using the notation of Algorithm 3.6, we consider the CORSING solution defined by

$$\hat{u} := \Phi(\operatorname{OMP}(DA, Dy, k)), \tag{3.28}$$

where D is as defined in (3.15). Theorem 3.9 shows that, in order to achieve a CORSING error $\|\hat{u} - u\|_U$ comparable to the best approximation error of u in U_s^N in expectation, it is sufficient to choose $k = \mathcal{O}(s)$

iterations of OMP in (3.28). Notice that it is possible to show a version of this theorem in probability, analogously to [16].

Theorem 3.9 (Recovery guarantee for CORSING). *There exist constants $\bar{K} \in \mathbb{N}$, $C > 0$, and $\varepsilon^* \in (0, 1)$ such that the following holds. Let $\|u\|_U \leq L$ for some $L > 0$ and assume that the constant κ defined in (3.22) satisfies*

$$\kappa < \frac{1}{1 - \varepsilon^*}. \tag{3.29}$$

Let $s, N, M \in \mathbb{N}$, with $s \leq N$, and

$$\gamma \in (0, 1 - (1 - \varepsilon^*)\kappa), \quad \varepsilon \in \left(1 - \frac{1 - \gamma}{\kappa}, \varepsilon^*\right), \tag{3.30}$$

be such that the following truncation condition holds:

$$\sum_{q>M} \mu_q^N \leq \frac{\alpha^2 \gamma c_\varphi c_\xi}{(\bar{K} + 1)s}.$$

Then, provided that m satisfies (3.25), where $\eta = \varepsilon - 1 + \frac{1-\gamma}{\kappa}$, the CORSING solution \hat{u} computed via OMP as in (3.28) satisfies

$$\mathbb{E}[\|\mathcal{T}_L \hat{u} - u\|_U] \leq \left(1 + \frac{1 + C}{(1 - \varepsilon)^{\frac{1}{2}}}\right) \inf_{w \in U_s^N} \|u - w\|_U + 2L\zeta, \tag{3.31}$$

where $\mathcal{T}_L v := \min\{1, L/\|v\|_U\}v$ and where $\zeta = 2 \exp(-\min\{1, C_\varphi^2 C_\xi^2 \beta^4\} \eta^2 m / (5^{12} s \|\nu^{N,M}\|_1^2))$ bounds the failure probability of the RIP. Possible values for the constants are $\bar{K} = 12$, $C = 49$ and $\varepsilon^* = 1/13$.

Proof. The argument is analogous to that of [16, Theorem 3.13], where the role of the restricted inf-sup property is replaced by the RIP. Consider the constants \bar{K} , C , and ε^* as in Theorem 3.7 and define the event

$$\Omega_{\text{RIP}} := \{\tilde{A} \text{ satisfies RIP}(\varepsilon, \bar{K} + 1)s\}.$$

We split the expectation accordingly as

$$\mathbb{E}[\|\mathcal{T}_L(\hat{u}) - u\|_U] = \mathbb{E}[\mathbf{1}_{\Omega_{\text{RIP}}} \|\mathcal{T}_L(\hat{u}) - u\|_U] + \mathbb{E}[\mathbf{1}_{\Omega_{\text{RIP}}^c} \|\mathcal{T}_L(\hat{u}) - u\|_U].$$

The second term is bounded by $2L\zeta$, since the adopted choice of M and m guarantees $\mathbb{P}(\Omega_{\text{RIP}}^c) \leq \zeta$, thanks to Theorem 3.4, and due to the truncation via \mathcal{T}_L .

Now, consider a generic $w \in U_s^N$. Observing that \mathcal{T}_L is 1-Lipschitz with respect to $\|\cdot\|_U$ and using the triangle inequality, we find

$$\begin{aligned} \mathbb{E}[\mathbf{1}_{\Omega_{\text{RIP}}} \|\mathcal{T}_L(\hat{u}) - u\|_U] &= \mathbb{E}[\mathbf{1}_{\Omega_{\text{RIP}}} \|\mathcal{T}_L(\hat{u}) - \mathcal{T}_L(u)\|_U] \leq \mathbb{E}[\mathbf{1}_{\Omega_{\text{RIP}}} \|\hat{u} - u\|_U] \\ &\leq \mathbb{E}[\mathbf{1}_{\Omega_{\text{RIP}}} \|\hat{u} - w\|_U] + \mathbb{E}[\mathbf{1}_{\Omega_{\text{RIP}}} \|u - w\|_U]. \end{aligned} \tag{3.32}$$

Notice that the second expectation in (3.32) is trivially bounded by $\|u - w\|_U$. In order to bound the other expectation, we note that $\hat{x} = \text{OMP}(DA, Dy, n) = \text{OMP}(\tilde{A}, \tilde{y}, n)$ and that Theorem 3.7 holds on the event Ω_{RIP} . Therefore, denoting $z = \Phi^* w$, we have the chain of inequalities

$$\begin{aligned} \|\hat{u} - w\|_U^2 &\leq C_\varphi \|\hat{x} - z\|_2^2 \leq \frac{C_\varphi}{1 - \varepsilon} \|\tilde{A}(\hat{x} - z)\|_2^2 \leq \frac{C_\varphi}{1 - \varepsilon} (\|\tilde{A}\hat{x} - \tilde{y}\|_2 + \|\tilde{A}z - \tilde{y}\|_2)^2 \\ &\leq \frac{C_\varphi}{1 - \varepsilon} (1 + C)^2 \|\tilde{A}z - \tilde{y}\|_2^2. \end{aligned}$$

Hence, we estimate

$$\mathbb{E}[\mathbf{1}_{\Omega_{\text{RIP}}} \|\hat{u} - w\|_U] \leq \sqrt{\frac{C_\varphi}{1 - \varepsilon}} (1 + C) \mathbb{E}[\mathbf{1}_{\Omega_{\text{RIP}}} \|\tilde{A}z - \tilde{y}\|_2]. \tag{3.33}$$

Now, exploiting that the τ_i 's are i.i.d., the Riesz property of $(\xi_q)_{q \in \mathbb{N}}$ (see the definition in (3.23)), and the continuity (3.4) of $a(\cdot, \cdot)$, we have

$$\begin{aligned} \mathbb{E}[\|\tilde{A}z - \tilde{y}\|_2^2] &= \beta^{-2} C_\varphi^{-1} C_\xi^{-1} \mathbb{E}[\|D(Az - y)\|_2^2] = \beta^{-2} C_\varphi^{-1} C_\xi^{-1} \sum_{q \in [M]} [a(w, \xi_q) - \mathcal{F}(\xi_q)]^2 \\ &= \beta^{-2} C_\varphi^{-1} C_\xi^{-1} \sum_{q \in [M]} a(w - u, \xi_q)^2 \leq C_\varphi^{-1} \|w - u\|_U^2. \end{aligned}$$

Note that we have used inequality (3.26) in the last step. Now, applying Jensen's inequality to the previous relation we obtain $\mathbb{E}[\|\tilde{A}z - \tilde{y}\|_2] \leq C_\varphi^{-1/2} \|w - u\|_U$, which, combined with (3.33), yields

$$\int_{\Omega_{\text{RIP}}} \|\hat{u} - w\|_U \, d\mathbb{P} \leq \sqrt{\frac{C_\varphi}{1 - \varepsilon}} (1 + C) \mathbb{E}[\|\tilde{A}z - \tilde{y}\|_2] \leq \frac{1 + C}{(1 - \varepsilon)^{\frac{1}{2}}} \|u - w\|_U.$$

Combining the above inequalities completes the proof. \square

Remark 3.10. Plugging $\varepsilon^* = 1/13$ into relation (3.29), we obtain

$$\kappa < \frac{13}{12}, \tag{3.34}$$

which is a very restrictive condition. As already mentioned in Remark 3.8, the value of ε^* can be made larger, in price of larger values of \bar{K} and C . An interesting open question is whether ε^* can be made arbitrarily close to 1, or if there is a maximal admissible value strictly lower than 1 (see also the discussion in Section 2.1).

We believe that the sufficient condition (3.34) is too conservative. Indeed, numerical experiments show the success of the CORSING method in computing accurate sparse approximations via OMP in problems where (3.34) is not met (e.g., advection-dominated problems), see [15,12,16]. Bridging this gap between theory and practice and showing recovery guarantees for OMP without assuming (3.34) is left to future work.

4. Proof of Theorem 1.1

In this section we prove Theorem 1.1. Recall that $X \in \mathbb{C}^N$ is a random vector with bounded components, i.e., for all $i \leq N : |\langle X, e_i \rangle| \leq K$, where e_1, \dots, e_N denotes the standard basis of \mathbb{C}^N and that X_1, \dots, X_m are independent copies of X . We aim to bound

$$\sup_{f \in T} \left| \frac{1}{m} \sum_{i=1}^m |\langle f, X_i \rangle|^2 - \mathbb{E}|\langle f, X \rangle|^2 \right|, \tag{4.1}$$

for $T \subseteq \sqrt{s} B_{\ell^1}^N$. Let us recall the following deviation inequality for the empirical process [11, Theorem 2.3].

Theorem 4.1. Let \mathcal{F} be a class of functions $f : S \rightarrow \mathbb{R}$ on some set S and let X_1, \dots, X_m denote independent S -valued random variables, which are independent copies of random variable X . Set

$$Z_{\mathcal{F}} = \sup_{f \in \mathcal{F}} \left| \sum_{i=1}^m f(X_i) - \mathbb{E}f(X) \right|, \quad \sigma_{\mathcal{F}}^2 = \sup_{f \in \mathcal{F}} \mathbb{E}f(X)^2 \quad \text{and} \quad \beta_{\mathcal{F}} = \sup_{f \in \mathcal{F}} \|f\|_{L^\infty}.$$

Then, for any $u > 0$,

$$\mathbb{P}\left(Z_{\mathcal{F}} \geq \mathbb{E}Z_{\mathcal{F}} + \sqrt{2u(\sigma_{\mathcal{F}}^2 + 2\beta_{\mathcal{F}}\mathbb{E}Z_{\mathcal{F}})} + \frac{1}{3}\beta_{\mathcal{F}}u\right) \leq 2 \exp(-u). \tag{4.2}$$

Theorem 4.1 is related to Talagrand’s concentration inequality for the empirical process. We employ Theorem 4.1’s inequality by considering $\mathcal{F}_T = \{|\langle f, \cdot \rangle|^2 : f \in T\}$ for the set $T \subseteq \sqrt{s}B_{\ell^1}^N$. In this case, for any $u \geq 0$, the event of (4.2) reads

$$\begin{aligned} \sup_{f \in T} \left| \frac{1}{m} \sum_{i=1}^m |\langle f, X_i \rangle|^2 - \mathbb{E}|\langle f, X \rangle|^2 \right| &\leq \mathbb{E} \sup_{f \in T} \left| \frac{1}{m} \sum_{i=1}^m |\langle f, X_i \rangle|^2 - \mathbb{E}|\langle f, X \rangle|^2 \right| \\ &+ \sqrt{2u} \frac{1}{m} \left(\sigma_{\mathcal{F}_T}^2 m + 2\beta_{\mathcal{F}_T} \sup_{f \in T} \left| \sum_{i=1}^m |\langle f, X_i \rangle|^2 - \mathbb{E}|\langle f, X \rangle|^2 \right| \right)^{\frac{1}{2}} + \frac{1}{3}\beta_{\mathcal{F}_T} \frac{u}{m}. \end{aligned}$$

Now using the fact that X has bounded coordinates, we obtain that

$$\mathbb{E}|\langle f, X \rangle|^4 \leq sK^2 \sup_{f \in T} \mathbb{E}|\langle f, X \rangle|^2 \quad \text{and} \quad \sup_{f \in T} \|\langle f, X \rangle\|_{L^\infty} \leq K^2 s.$$

Hence, if we set $u = \delta^2 m / (K^2 s)$ for $\delta \in (0, 1)$, then Theorem 4.1 yields that with probability at least $1 - 2 \exp(-\delta^2 m / (K^2 s))$,

$$\begin{aligned} \sup_{f \in T} \left| \frac{1}{m} \sum_{i=1}^m |\langle f, X_i \rangle|^2 - \mathbb{E}|\langle f, X \rangle|^2 \right| &\leq \mathbb{E} \sup_{f \in T} \left| \frac{1}{m} \sum_{i=1}^m |\langle f, X_i \rangle|^2 - \mathbb{E}|\langle f, X \rangle|^2 \right| \\ &+ \left(2\delta^2 \sup_{f \in T} \mathbb{E}|\langle f, X_i \rangle|^2 + 4\delta^2 \sup_{f \in T} \left| \frac{1}{m} \sum_{i=1}^m |\langle f, X_i \rangle|^2 - \mathbb{E}|\langle f, X \rangle|^2 \right| \right)^{\frac{1}{2}} + \frac{\delta^2}{3}. \end{aligned} \tag{4.3}$$

It remains to estimate the expectation of the process in (4.1). By symmetrization [36, Lemma 6.3],

$$\mathbb{E} \sup_{f \in T} \left| \frac{1}{m} \sum_{i=1}^m |\langle f, X_i \rangle|^2 - \mathbb{E}|\langle f, X \rangle|^2 \right| \leq 2\mathbb{E} \sup_{f \in T} \left| \frac{1}{m} \sum_{i=1}^m |\langle f, X_i \rangle|^2 \varepsilon_i \right|, \tag{4.4}$$

where $(\varepsilon_i)_{i \in [m]}$ denotes a sequence of independent symmetric Bernoulli random variables, also independent of $(X_i)_{i \in [m]}$. The following theorem provides a bound for the right-hand-side of (4.4).

Theorem 4.2. Let $T \subseteq \sqrt{s}B_{\ell^1}^N$ and let $\delta \in (0, 1)$. Then, the following holds:

$$\begin{aligned} \mathbb{E} \sup_{f \in T} \left| \frac{1}{m} \sum_{i=1}^m |\langle f, X_i \rangle|^2 \varepsilon_i \right| &\leq (280 + 200\sqrt{2}) \sqrt{\frac{sK^2 \log^2(sK^2/\delta) \log(eN)}{m}} \left(\mathbb{E} \sup_{f \in T} \frac{1}{m} \sum_{i=1}^m |\langle f, X_i \rangle|^2 \right)^{1/2} \\ &+ (69 + 49\sqrt{2}) \delta \sum_{i=1}^m |\langle f, X_i \rangle|^2 + (643 + 468\sqrt{2}) \delta m. \end{aligned}$$

Let us observe that this bound implies a bound for the right-hand side of (4.4) provided that $\delta \in (0, 1)$ is small enough.

Corollary 4.3. *Let $T \subseteq \sqrt{s}B_{\ell^1}^N$ and let $\delta \in (0, \frac{1}{28}(10 - 7\sqrt{2}))$. Assume that*

$$m \geq 1600(99 + 70\sqrt{2}) \delta^{-2} s K^2 \log(sK^2/\delta)^2 \log(eN). \quad (4.5)$$

Then, the following holds:

$$\mathbb{E} \sup_{f \in T} \left| \frac{1}{m} \sum_{i=1}^m |\langle f, X_i \rangle|^2 - \mathbb{E} |\langle f, X \rangle|^2 \right| \leq 8(161 + 117\sqrt{2}) \delta + 14(10 + 7\sqrt{2}) \delta \sup_{f \in T} \mathbb{E} |\langle f, X \rangle|^2.$$

Proof. Let $T \subseteq \sqrt{s}B_{\ell^1}^N$ and $\delta \in (0, 1/218)$. Choosing $m \geq c\delta^{-2} s K^2 \log(sK^2/\delta)^2 \log(eN)$ for an appropriate absolute constant $c \geq 1600(99 + 70\sqrt{2})$, it follows from (4.4) and Theorem 4.2, that

$$\begin{aligned} \mathbb{E} \sup_{f \in T} \left| \frac{1}{m} \sum_{i=1}^m |\langle f, X_i \rangle|^2 - \mathbb{E} |\langle f, X_i \rangle|^2 \right| &\leq 2\delta \left(\mathbb{E} \sup_{f \in T} \left| \frac{1}{m} \sum_{i=1}^m |\langle f, X_i \rangle|^2 \right| \right)^{1/2} \\ &\quad + (69 + 49\sqrt{2}) \delta \sum_{i=1}^m |\langle f, X_i \rangle|^2 + (643 + 468\sqrt{2}) \delta m. \end{aligned}$$

By the arithmetic-mean-geometric-mean inequality we have

$$\begin{aligned} \delta \left(\mathbb{E} \sup_{f \in T} \left| \frac{1}{m} \sum_{i=1}^m |\langle f, X_i \rangle|^2 \right| \right)^{1/2} &= \sqrt{\delta} \left(\delta \mathbb{E} \sup_{f \in T} \left| \frac{1}{m} \sum_{i=1}^m |\langle f, X_i \rangle|^2 \right| \right)^{1/2} \\ &\leq \frac{\delta}{2} + \frac{\delta}{2} \mathbb{E} \sup_{f \in T} \left| \frac{1}{m} \sum_{i=1}^m |\langle f, X_i \rangle|^2 \right|. \end{aligned}$$

Combining this with the fact that

$$\mathbb{E} \sup_{f \in T} \left| \frac{1}{m} \sum_{i=1}^m |\langle f, X_i \rangle|^2 \right| \leq \mathbb{E} \sup_{f \in T} \left| \frac{1}{m} \sum_{i=1}^m |\langle f, X_i \rangle|^2 - \mathbb{E} |\langle f, X \rangle|^2 \right| + \sup_{f \in T} \mathbb{E} |\langle f, X \rangle|^2,$$

we obtain the inequality

$$\begin{aligned} \mathbb{E} \sup_{f \in T} \left| \frac{1}{m} \sum_{i=1}^m |\langle f, X_i \rangle|^2 - \mathbb{E} |\langle f, X_i \rangle|^2 \right| &\leq (70 + 49\sqrt{2}) \delta \mathbb{E} \sup_{f \in T} \left| \frac{1}{m} \sum_{i=1}^m |\langle f, X_i \rangle|^2 - \mathbb{E} |\langle f, X \rangle|^2 \right| \\ &\quad + (70 + 49\sqrt{2}) \delta \sup_{f \in T} \mathbb{E} |\langle f, X \rangle|^2 + (644 + 468\sqrt{2}) \delta. \end{aligned} \quad (4.6)$$

Assuming that $\delta \in (0, \frac{1}{28}(10 - 7\sqrt{2}))$, we conclude from (4.6), that

$$\sup_{f \in T} \left| \frac{1}{m} \sum_{i=1}^m |\langle f, X_i \rangle|^2 - \mathbb{E} |\langle f, X_i \rangle|^2 \right| \leq 8(161 + 117\sqrt{2}) \delta + 14(10 + 7\sqrt{2}) \delta \sup_{f \in T} \mathbb{E} |\langle f, X \rangle|^2.$$

This implies the desired estimate. \square

With this estimate all pieces are in place to deduce Theorem 1.1.

Proof of Theorem 1.1. Combining Corollary 4.3 with the event (4.3), it follows that with probability at least $1 - 2 \exp(-\delta^2 m / (K^2 s))$,

$$\begin{aligned} & \sup_{f \in T} \left| \frac{1}{m} \sum_{i=1}^m |\langle f, X_i \rangle|^2 - \mathbb{E} |\langle f, X \rangle|^2 \right| \\ & \leq 8(161 + 117\sqrt{2})\delta + 14(10 + 7\sqrt{2}) \sup_{f \in T} \mathbb{E} |\langle f, X \rangle|^2 \\ & \quad + \sqrt{2}\delta \left(\sup_{f \in T} \mathbb{E} |\langle f, X_i \rangle|^2 + 8(161 + 117\sqrt{2})\delta + 14(10 + 7\sqrt{2})\delta \sup_{f \in T} \mathbb{E} |\langle f, X \rangle|^2 \right)^{\frac{1}{2}} + \frac{\delta^2}{3}. \end{aligned}$$

Using the arithmetic-mean-geometric-mean inequality, we find

$$\begin{aligned} & \sqrt{2}\delta \left(\sup_{f \in T} \mathbb{E} |\langle f, X_i \rangle|^2 + 8(161 + 117\sqrt{2})\delta + 14(10 + 7\sqrt{2})\delta \sup_{f \in T} \mathbb{E} |\langle f, X \rangle|^2 \right)^{\frac{1}{2}} \\ & \leq \delta + \frac{\delta}{2} \sup_{f \in T} \mathbb{E} |\langle f, X_i \rangle|^2 + 8(161 + 117\sqrt{2})\frac{\delta^2}{2} + 14(10 + 7\sqrt{2})\frac{\delta^2}{2} \sup_{f \in T} \mathbb{E} |\langle f, X \rangle|^2 \end{aligned}$$

Hence, on the same event and using the fact that $\delta < 1$,

$$\sup_{f \in T} \left| \frac{1}{m} \sum_{i=1}^m |\langle f, X_i \rangle|^2 - \mathbb{E} |\langle f, X \rangle|^2 \right| \leq \frac{1}{6}(1457 + 1053\sqrt{2})\delta + \frac{1}{2}(421 + 294\sqrt{2})\delta \sup_{f \in T} \mathbb{E} |\langle f, X \rangle|^2.$$

This implies the theorem, since $\frac{1}{2}(421 + 294\sqrt{2}) < \frac{1}{6}(1457 + 1053\sqrt{2}) < 492$. \square

We are left with establishing Theorem 4.2. The proof of this theorem will occupy the rest of this section. In order to establish Theorem 4.2 we will start by introducing a general quantity, which can control the Bernoulli process. This quantity is a mixture of an ℓ^1 approximation term and a finite precision approximation of a γ_2 functional (the meaning of the term “finite precision” will be clarified in Section 4.2).

4.1. Generic chaining

Let (T, d) denote a (semi-)metric space. An increasing sequence $(A_n)_{n \geq 0}$ of subsets of T is called admissible if, for all $n \geq 0$, $|A_n| \leq 2^{2^n}$. For a set $A \subseteq T$ we set $d(A, x) = \inf_{a \in A} d(a, x)$. A central object of study in generic chaining are the functionals

$$\gamma_\alpha(T, d) := \inf_{\mathcal{A}} \sup_{t \in T} \sum_{n \geq 1} 2^{\frac{n}{\alpha}} d(A_n, t), \tag{4.7}$$

where the infimum is taken over all admissible sequences $\mathcal{A} = (A_n)_{n \geq 0}$ of subsets of T .

4.2. Bounding the Bernoulli process

In this section we consider subsets U of the space \mathbb{R}^m equipped with the Euclidean distance and study bounds for the Bernoulli process given by

$$\sup_{x \in U} \sum_{i=1}^m x_i \varepsilon_i, \tag{4.8}$$

where, as before, $(\varepsilon_i)_{i \leq m}$ is a sequence of independent symmetric Bernoulli random variables. The set U will later be replaced by the set of sequences $\{(|\langle f, X_i \rangle|^2)_{i \leq m} : f \in T\}$ for some $T \subseteq \sqrt{s}B_{\ell^1}^N$. With the intention of keeping the necessary notation as simple as possible, we formulate the results of this section in terms of subsets of \mathbb{R}^m .

Hoeffding’s inequality implies that increments of the process $\sup_{x \in U} |\sum_{i=1}^m x_i \varepsilon_i|$ are subgaussian with respect to the Euclidean metric on \mathbb{R}^m . Combining this observation with the trivial bound $\sup_{x \in U} |\sum_{i=1}^m x_i \varepsilon_i| \leq \sup_{x \in U} \|x\|_1$ and a standard generic chaining argument yields the bound (see [47], in particular the discussion at the beginning of chapter 5)

$$\mathbb{E} \sup_{x \in U} \sum_{i=1}^m x_i \varepsilon_i \leq b(U) := \inf_{U \subseteq U_1 + U_2} \left\{ \sup_{x \in U_1} \|x\|_1 + \gamma_2(U_2, \|\cdot\|_2) \right\}. \tag{4.9}$$

For our result, we are only interested in a finite precision approximation for the Bernoulli process $\sup_{x \in U} |\sum_{i=1}^m x_i \varepsilon_i|$. We therefore propose to substitute $b(U)$ by a finite precision version, which is able to control the left hand side of (4.9). The following result is implicitly stated in [47]. A proof is provided for the convenience of the reader.

Lemma 4.4. *Let $U \subseteq \mathbb{R}^m$. Let $n_0, \ell > 0$, let $(A_n)_{n \geq n_0}$ denote any admissible sequence for U and let $\pi_n : U \rightarrow A_n$ be generic maps for $n \geq n_0$. Moreover, let $\pi_n(x) = 0$ for every $x \in U$, for $n < n_0$. Then, there is a constant $c(n_0) \in (1, 2)$ such that*

$$\mathbb{E} \sup_{x \in U} \left| \sum_{i=1}^m x_i \varepsilon_i \right| \leq \sup_{x \in U} \|x - \pi_{n_0+\ell}(x)\|_1 + c(n_0) \sup_{x \in U} \sum_{n=n_0}^{n_0+\ell} 2^{\frac{n}{2}} \|\pi_n(x) - \pi_{n-1}(x)\|_2. \tag{4.10}$$

Proof. For $n < n_0$ we set $\pi_n(x) = 0$. Fix $\ell > 0$ and observe that

$$\begin{aligned} \left| \sum_{i=1}^m x_i \varepsilon_i \right| &= \left| \sum_{i=1}^m \varepsilon_i (x_i - \pi_{n_0+\ell}(x)_i) + \sum_{n=n_0}^{n_0+\ell} \sum_{i=1}^m \varepsilon_i (\pi_n(x) - \pi_{n-1}(x))_i \right| \\ &\leq \|x - \pi_{n_0+\ell}(x)\|_1 + \sum_{n=n_0}^{n_0+\ell} \left| \sum_{i=1}^m \varepsilon_i (\pi_n(x) - \pi_{n-1}(x))_i \right|. \end{aligned}$$

For $x \in \mathbb{R}^m$ let $Z_x = \sum_{i=1}^m \varepsilon_i x_i$ denote the associated random variable. Taking suprema on each side in the inequality above yields

$$\sup_{x \in U} \left| \sum_{i=1}^m \varepsilon_i x_i \right| \leq \sup_{x \in U} \|x - \pi_{n_0+\ell}(x)\|_1 + \sup_{x \in U} \sum_{n=n_0}^{n_0+\ell} |Z_{\pi_n(x)} - Z_{\pi_{n-1}(x)}|. \tag{4.11}$$

By Hoeffding’s inequality (see, e.g., [7]) we obtain the estimate

$$\mathbb{P} \left(|Z_{\pi_n(x)} - Z_{\pi_{n-1}(x)}| > 2^{\frac{n}{2}} u \|\pi_n(x) - \pi_{n-1}(x)\|_2 \right) \leq 2e^{-2^{n-1} u^2} \tag{4.12}$$

Since $(A_n)_{n \geq n_0}$ is admissible there are at most $|A_n| |A_{n-1}| \leq 2^{2^{n+1}}$ pairs of the form $(\pi_n(x), \pi_{n-1}(x))$ and it follows by a union bound that the event $\Omega_n(u)$

$$\forall (\pi_n(x), \pi_{n-1}(x)) \in A_n \times A_{n-1} : |Z_{\pi_n(x)} - Z_{\pi_{n-1}(x)}| \leq 2^{\frac{n}{2}} u \|\pi_n(x) - \pi_{n-1}(x)\|_2 \tag{4.13}$$

occurs with probability at least $1 - 2 \cdot 2^{2^{n+1}} e^{-2^{n-1}u^2}$. Therefore, a union bound over $n = n_0, \dots, n_0 + \ell$ yields, that the probability that the event $\Omega(u) = \bigcup_{n=n_0}^{n_0+\ell} \Omega_n(u)$ does not occur is bounded by

$$2 \sum_{n=n_0}^{n_0+\ell} 2^{2^{n+1}} e^{-2^{n-1}u^2} \leq 2e^{-u^2/2} \sum_{n=n_0}^{n_0+\ell} 2^{2^{n+1}} e^{-2^{n+1}} \leq 2\kappa(n_0)e^{-u^2/2}, \tag{4.14}$$

where $\kappa(n_0) = \sum_{n=n_0}^{n_0+\ell} (2/e)^n$ and where we used that $u^2 2^{n-1} \geq u^2/2 + u^2 2^{n-2} \geq u^2/2 + 2^{n+1}$ holds for $n \geq 1$ and $u \geq 1$. On the event $\Omega(u)$, the inequality (4.11) reads

$$\sup_{x \in U} |Z_x| \leq \sup_{x \in U} \|x - \pi_{n_0+\ell}(x)\|_1 + u \sup_{x \in U} \sum_{n=n_0}^{n_0+\ell} 2^{\frac{n}{2}} \|\pi_n(x) - \pi_{n-1}(x)\|_2.$$

This implies that the following tail bound is valid for $\sup_{x \in U} |Z_x|$, $n_0 \geq 1$ and $u \geq 1$,

$$\mathbb{P} \left(\sup_{x \in U} |Z_x| > \sup_{x \in U} \|x - \pi_{n_0+\ell}(x)\|_1 + u \sup_{x \in U} \sum_{n=n_0}^{n_0+\ell} 2^{\frac{n}{2}} \|\pi_n(x) - \pi_{n-1}(x)\|_2 \right) \leq 2\kappa(n_0) e^{-u^2/2}. \tag{4.15}$$

Set $d_1 = \sup_{x \in U} \|x - \pi_{n_0+\ell}(x)\|_1$ and $d_2 = \sup_{x \in U} \sum_{n=n_0}^{n_0+\ell} 2^{\frac{n}{2}} \|\pi_n(x) - \pi_{n-1}(x)\|_2$. Then, integrating the tail bound (4.15) yields

$$\begin{aligned} \mathbb{E} \left[\sup_{x \in T} |Z_x| \right] &= \int_0^\infty \mathbb{P} \left(\sup_{x \in T} |Z_x| \geq t \right) dt \leq d_1 + d_2 \int_0^\infty \mathbb{P} \left(\sup_{x \in U} |Z_x| \geq d_1 + ud_2 \right) du \\ &\leq d_1 + \left(2\kappa(n_0) \int_1^\infty e^{-u^2/2} du + 1 \right) d_2 \\ &\leq d_1 + \left(\frac{4}{5} \kappa(n_0) + 1 \right) d_2, \end{aligned}$$

where we used the change of variable $t = d_1 + ud_2$. This, shows that the desired estimate is true for $c(n_0) = \frac{4}{5} \kappa(n_0) + 1 > 0$. Now, observe that by the definition of $\kappa(n_0)$,

$$\kappa(n_0) \leq \left(\frac{2}{e} \right)^{2^{n_0+1}} \sum_{n=0}^\infty \left(\frac{2}{e} \right)^n = \left(\frac{2}{e} \right)^{2^{n_0+1}} \frac{e}{e-2} \leq \frac{111}{100}.$$

This finishes the proof. \square

This lemma leads us to introduce the following definition.

Definition 4.5. Let $U \subseteq \mathbb{R}^m$ and let \mathcal{A} denote the set of all admissible sequences $(A_n)_{n \geq n_0}$ in U . For each admissible sequence let $\pi_n : U \rightarrow A_n$ denote a generic map. For $n_0, \ell > 0$ we define

$$b_{n_0, \ell}(T) := \inf_{\mathcal{A}} \left\{ \sup_{x \in T} \|x - \pi_{n_0+\ell}(x)\|_1 + c(n_0) \sup_{x \in T} \sum_{n=n_0}^{n_0+\ell} 2^{\frac{n}{2}} \|\pi_n(x) - \pi_{n-1}(x)\|_2 \right\}.$$

In order to find a bound for $b_{n_0, \ell}(U)$ and hence for $\sup_{x \in U} |\sum_{i=1}^m x_i \varepsilon_i|$ given a fixed set $U \subseteq \mathbb{R}^m$ it suffices to identify a suitable admissible sequence $(A_n)_{n \geq n_0}$.

4.3. Estimating $b_{n_0, \ell}(T)$

Returning to our initial setting we are looking for estimates of

$$\sup_{f \in T} \sum_{i=1}^m |\langle f, X_i \rangle|^2 \varepsilon_i,$$

for a set $T \subseteq \sqrt{s}B_{\ell^1}^N$. By Lemma 4.4 it suffices to bound $b_{n_0, \ell}(\{(|\langle f, X_i \rangle|^2)_{i \leq m} : f \in T\})$. Such bounds can be achieved by constructing a suitable admissible sequence $(A_n)_{n \geq n_0}$ in the set $\{(|\langle f, X_i \rangle|^2)_{i \leq m} : f \in T\}$. In this section we provide such an admissible sequence for $b_{n_0, \ell}(\{(|\langle f, X_i \rangle|^2)_{i \leq m} : f \in T\})$ based on an initial covering of the set $\{(\langle f, X_i \rangle)_{i \in [m]} : f \in T\}$ with respect to the (semi-)norms $\|x\|_I$ defined below. The construction is based on the notion of a weak covering of a set $T \subseteq \mathbb{C}^N$, which is explained in the following subsection.

4.3.1. Weak coverings

We start by defining the empirical seminorms

$$\|x\|_{I, \mathbf{X}} := \max_{i \in [m] \setminus I} |\langle X_i, x \rangle|,$$

where $I \subseteq [m]$ and $\mathbf{X} = (X_1, \dots, X_m)$ denotes a realization of the random vectors in question. Let $B_{I, \mathbf{X}} := \{x \in \mathbb{C}^N : \|x\|_{I, \mathbf{X}} \leq 1\}$ denote the unit balls with respect to the seminorm $\|\cdot\|_{I, \mathbf{X}}$. In the following we will be interested in covering a set $T \subseteq \mathbb{C}^N$ by the following family of sets for a given width $\rho > 0$,

$$\mathcal{B}(\rho, M) := \{\rho B_{I, \mathbf{X}} : I \subseteq [m], |I| \leq M\}. \tag{4.16}$$

Definition 4.6. Let $\rho, M > 0$. We say that $T \subseteq \mathbb{C}^N$ is *weakly covered* by $\mathcal{B}(\rho, M)$, if for some $r \in \mathbb{N}$, there exists $x_1, \dots, x_r \in \mathbb{C}^N$ and sets $B_1, \dots, B_r \in \mathcal{B}(\rho, M)$, such that

$$T \subseteq \bigcup_{i=1}^r (x_i + B_i). \tag{4.17}$$

We call the smallest $r \in \mathbb{N}$ such that (4.17) is satisfied the *weak covering number* of T and denote this number by $\mathcal{N}^*(T, \rho, k)$. Moreover, we refer to the set $\{x_1, \dots, x_r\}$ as *weak covering* or *net*.

For sets $T \subseteq \sqrt{s}B_{\ell^1}^N \subset \mathbb{C}^N$, Maurey’s empirical method can be used to estimate the size of a weak covering of T .

Lemma 4.7 (Maurey’s empirical method). Let $T \subseteq \sqrt{s}B_{\ell^1}^N$, $\delta \in (0, 1)$, $\rho > 0$ and X_1, \dots, X_m denote random vectors with bounded coordinates. Then,

$$\log \mathcal{N}^*\left(T, \rho, \frac{4\delta m}{sK^2 \log_2(sK^2/\delta)}\right) \leq \frac{2 \log_2(sK^2/\delta) \log_2(2N) sK^2}{\rho^2}. \tag{4.18}$$

Proof. We observe that every $x \in T \subseteq \sqrt{s}B_{\ell^1}^N$ is a convex combination of $V = \{\pm\sqrt{s}e_j \pm \sqrt{s}ie_j : j \in [N]\}$, i.e. there is a sequence $(\lambda_v)_{v \in V}$ with $\lambda_v \geq 0$ and $\sum_{v \in V} \lambda_v = 1$, such that $\sum_{v \in V} \lambda_v v = x$. For a fixed $x \in T \subseteq \sqrt{s}B_{\ell^1}^N$ the associated sequence $\lambda = (\lambda_v)_{v \in V}$ defines a probability distribution on V . Let Z denote a random vector in V with distribution λ and let Z_l for $l = 1, \dots, L$ denote independent copies of Z . Set $L = \lfloor K^2 s \rho^{-2} \log_2(sK^2 \log_2(sK^2/\delta)/\delta) \rfloor$ and consider the random set

$$J = \left\{ i \in [m] : \left| \langle X_i, x \rangle - \frac{1}{L} \sum_{l \leq L} \langle X_i, Z_l \rangle \right| > \rho \right\}.$$

For this set we will bound the probability of the event, that $|J| \geq \delta m / (sK^2 \log_2(sK^2/\delta))$. Recall that for $l = 1, \dots, L$ we have $\mathbb{E}_{Z_l} \langle X_i, Z_l \rangle = \langle X_i, x \rangle$. By Hoeffding’s inequality it follows that for every $i \in [m]$

$$\mathbb{P}_{Z_1, \dots, Z_L} (i \in J) \leq 2 \exp(-L\rho^2 / (s\|X_i\|_\infty^2)) \leq \frac{2\delta}{sK^2 \log_2(sK^2/\delta)},$$

where we used that $|\langle X_i, Z_l \rangle| \leq \sqrt{s}\|X_i\|_\infty \leq \sqrt{s}K$ for every $i \in [m]$ and $l \in [L]$. Hence, we have $\mathbb{E}_{Z_1, \dots, Z_L} |J| \leq \frac{2\delta m}{sK^2 \log_2(sK^2/\delta)}$. Using Chebyshev’s inequality it follows that

$$\mathbb{P}_{Z_1, \dots, Z_L} \left(|J| \geq \frac{4\delta m}{s \log(sK^2/\delta)} \right) \leq \frac{\mathbb{E}_{Z_1, \dots, Z_L} |J| s \log_2(sK^2/\delta)}{4\delta m} \leq \frac{1}{2}.$$

Therefore, we can find a realization of Z_1, \dots, Z_L , such that $|J| \leq \frac{4\delta m}{s \log_2(s/\delta)}$. For this realization we have

$$\left\| x - L^{-1} \sum_{l \leq L} Z_l \right\|_{J, \mathbf{X}} = \max_{i \in [m] \setminus J} \left| \langle X_i, L^{-1} \sum_{l \leq L} Z_l - x \rangle \right| \leq \rho.$$

Since $x \in T$ is arbitrary, by considering all possible realizations of $L^{-1} \sum_{l \leq L} Z_l$ we obtain a weak covering of T . For the choice of $L = \lfloor \rho^{-2} sK^2 \log_2(sK^2 \log_2(sK^2/\delta)/\delta) \rfloor \leq 2 \lfloor \rho^{-2} sK^2 \log_2(sK^2/\delta) \rfloor$ we have at most $|V|^L = (2N)^L \leq (2N)^{2\rho^{-2} K^2 s \log_2(sK^2/\delta)}$ realizations of Z_1, \dots, Z_L . This implies the desired result. \square

Based on this estimate we derive the following result, which provides the promised admissible sequence.

Theorem 4.8. *Let $T \subseteq \sqrt{s}B_{\ell^1}^N$ and let $\delta \in (0, 1)$. For $n_0 = \lceil \log_2 \log_2(2N) + \log_2 \log_2(sK^2/\delta) \rceil$ and $\ell = \lceil \log_2(sK^2/\delta) \rceil$ there is an admissible sequence $(A_n)_{n \geq n_0}$ for T , such that*

- For all $f \in T$ we have

$$\begin{aligned} & \| (|\langle f, X_i \rangle|^2)_{i \in [m]} - \pi_{n_0 + \ell} (|\langle f, X_i \rangle|^2)_{i \in [m]} \|_1 \\ & \leq (69 + 49\sqrt{2}) \delta \sum_{i=1}^m |\langle f, X_i \rangle|^2 + (643 + 468\sqrt{2}) \delta m. \end{aligned} \tag{4.19}$$

- For all $f \in T$ we have

$$\begin{aligned} & \sum_{n=n_0}^{n_0 + \ell - 1} 2^{\frac{n}{2}} \| (|\langle \pi_{n+1}(f), X_i \rangle|^2)_{i \in [m]} - (|\langle \pi_n(f), X_i \rangle|^2)_{i \in [m]} \|_2 \\ & \leq (200 + 140\sqrt{2}) \sqrt{\ell s} K 2^{\frac{n_0}{2}} \left(\sum_{i=1}^m |\langle f, X_i \rangle|^2 \right)^{1/2}. \end{aligned} \tag{4.20}$$

Theorem 4.8 directly implies a bound for $b_{n_0, \ell}(\{(|\langle f, X_i \rangle|^2)_{i \in [m]} : f \in T\})$ with $T \subseteq \sqrt{s}B_{\ell^1}^N$ and henceforth it implies a bound for the expectation of (4.1). We remark that the size of n_0 is determined by the initial estimate given in Lemma 4.7, while the parameter ℓ depends on the desired approximation accuracy $\delta \in (0, 1)$ and the ℓ^∞ -diameter of the set $\{(|\langle f, X_i \rangle|^2)_{i \in [m]} : f \in T\}$.

4.3.2. Construction of an admissible sequence and proof of Theorem 4.8

The goal of this section is ultimately to bound $b_{n_0, \ell}(T)$ for a subset $T \subset \sqrt{s}B_{\ell_1}^N$, thereby identifying $n_0, \ell > 0$. The plan is to achieve a bound for $b_{n_0, \ell}(T)$ by means of constructing an admissible sequence $(A_n)_{n \geq n_0}$ (see (4.27) below). The main point here is to ensure that the admissible sequence $(A_n)_{n \geq n_0}$ balances the bounds for the terms

$$\sup_{x \in T} \|x - \pi_{n_0 + \ell}(x)\|_1 \quad \text{and} \quad \sup_{x \in T} \sum_{n=n_0}^{n_0 + \ell - 1} 2^{\frac{n}{2}} \|\pi_{n+1}(x) - \pi_n(x)\|_2.$$

The approach of this section is based on ideas in [9,30]. The argument we propose is based on the observation that we have to deal with sequences $(|\langle f, X_i \rangle|^2)_{i \in [m]}$ for $f \in T$ and aims at organizing the admissible sequence $(A_n)_{n \geq n_0}$ around this fact. A vital observation is that for a fixed $i \in [m]$ we have

$$\left| |\langle f, X_i \rangle|^2 - |\langle g, X_i \rangle|^2 \right| \leq 2 \max\{|\langle f, X_i \rangle|, |\langle g, X_i \rangle|\} |\langle f, X_i \rangle - \langle g, X_i \rangle|.$$

We leverage this by coupling the approximation of $|\langle f, X_i \rangle|^2$ in the ℓ^∞ -norm to the size of the coefficients of the sequence $(|\langle f, X_i \rangle|^2)_{i \in [m]}$. This gives us a quasi- ℓ^∞ -control along the chain. This ℓ^∞ -control allows us to simultaneously bound the differences $\|\pi_{n+1}(x) - \pi_n(x)\|_2$ and, if ℓ is sufficiently large, to bound $\|x - \pi_{n_0 + \ell}(x)\|_1$.

Henceforth, we will use the following notation. For an element $f \in T \subseteq \sqrt{s}B_{\ell_1}^N$ and a realization of X_1, \dots, X_m we write $f_{\mathbf{X}} = (\langle f, X_i \rangle)_{i \in [m]}$ and $|f_{\mathbf{X}}|^2 = (|\langle f, X_i \rangle|^2)_{i \in [m]}$. Further, for the rest of this section we fix an approximation accuracy $\delta \in (0, 1)$. Let k_0 denote a positive integer satisfying $k_0 = \lceil \log_2(1/\delta^2) \rceil$. Moreover, we let $\ell = \lceil \log_2(sK^2/\delta) \rceil$ and $n_0 = \lceil \log_2 \log_2(2N) + \log_2 \log_2(sK^2/\delta) \rceil$ as in the setting of Theorem 4.8.

Let us start by observing that Lemma 4.7 shows that for $T \subset \sqrt{s}B_{\ell_1}^N$ there is a sequence of weak coverings for the parametrized families of sets $\mathcal{B}(\rho_n, M)$ with

$$\rho_n = \sqrt{s}K2^{-n/2} \quad \text{for } n = 0, \dots, \ell + k_0, \quad \text{and} \quad M = \frac{4\delta m}{sK^2 \log_2(sK^2/\delta)}. \tag{4.21}$$

We denote the associated nets as \tilde{A}_{n_0+n} , for $n = 0, \dots, \ell + k_0$. By (4.18) and since $n_0 \geq \log_2 \log_2(2N) + \log_2 \log_2(sK^2/\delta)$, the weak nets $\tilde{A}_{n_0}, \dots, \tilde{A}_{n_0+\ell+k_0}$ corresponding to $\rho_0, \dots, \rho_{\ell+k_0}$ and M as in (4.21) satisfy

$$|\tilde{A}_{n_0+n}| \leq 2^{\log_2(2N) \log_2(sK^2/\delta) K^2 s / \rho_n^2} = 2^{\log_2(2N) \log_2(sK^2/\delta) \cdot 2^n} \leq 2^{2^{n_0+n}},$$

and are therefore admissible sequences in T . For each weak covering $A_{n_0+n} = \{x_1, \dots, x_r\}$ any point x_i for $i = 1, \dots, r$ comes with its own unitball. Hence, each point in x_i for $i = 1, \dots, r$ is only correctly described by the tuple $(x_i, \|\cdot\|_{\mathbf{X}, I_i})$ for $i = 1, \dots, r$. Let us therefore agree on the following notation. For each $f \in T$ we denote by $\tilde{\pi}_{n_0+n}(f)$ any element $x \in \tilde{A}_{n_0+n}$ of the weak net $\tilde{A}_{n_0+n} = \{x_1, \dots, x_r\}$ that satisfies

$$\|x - \tilde{\pi}_{n_0+n}(f)\|_{\tilde{I}_{n_0+n}(f), \mathbf{X}} = \min_{i=1, \dots, r} \|f - x_i\|_{I_i, \mathbf{X}},$$

for a suitable $\tilde{I}_{n_0+n}(f) \subseteq [m]$ with $|\tilde{I}_{n_0+n}(f)| \leq M$ such that the pair $(\tilde{\pi}_{n_0+n}(f), \|\cdot\|_{\tilde{I}_{n_0+n}(f)})$ is an element of $\{(x_i, \|\cdot\|_{\tilde{I}_{n_0+n}(f)})\}_{i \in [r]}$. By the same token we also introduce another piece of notation, which is necessary because of the nature of the seminorms $\|\cdot\|_{\mathbf{X}, I}$. As pointed out above, by the definition of $\|\cdot\|_{\mathbf{X}, I}$ each approximant $\tilde{\pi}_{n_0+n}(f)$ is associated with a set I satisfying $|I| \leq M$, such that $|\langle f - \tilde{\pi}_{n_0+n}(f), X_i \rangle| > \rho_n$ for all $i \in I$. Let us therefore define,

$$I_{n_0+n}(f) := \{i \in [m] : |\langle f - \tilde{\pi}_{n_0+n}(f), X_i \rangle| > \rho_n\}. \tag{4.22}$$

We denote the sequence $\tilde{\pi}_{n_0}(f), \dots, \tilde{\pi}_{n_0+n}(f)$ by $\mathcal{I}_n(f_{\mathbf{X}})$.

We refine the admissible sequence $(\tilde{A}_{n_0+n})_{n=0}^{\ell+k_0}$ based on our chosen approximation accuracy $\delta > 0$. For each $f \in T$ and some $n \in \mathbb{N}$ we consider the tuple $(\tilde{\pi}_{n_0}(f), \dots, \tilde{\pi}_{n_0+n}(f))$ and inductively define sets $(E_{n_0+n})_{n=0}^{\ell}$ as $E_{n_0-1} := \emptyset$ and, for $n = 0, \dots, \ell$, as

$$E_{n_0+n} = E_{n_0+n}(f_{\mathbf{X}}) := \{i \in [m] : |\langle \tilde{\pi}_{n_0+n}(f), X_i \rangle| \geq \sqrt{2} \cdot \rho_n\} \setminus \bigcup_{k < n} E_{n_0+k}(f_{\mathbf{X}}). \tag{4.23}$$

We define an admissible sequence of approximations for an element $f \in T$ based on the data $(\tilde{\pi}_{n_0+n}(f))_{n=0}^{\ell+k_0}$ and the associated sets $(E_{n_0+n}(f_{\mathbf{X}}))_{n=0}^{\ell}$. For each sequence $|f_{\mathbf{X}}|^2$, we set

$$\pi_{n,k}^* |f_{\mathbf{X}}|^2 := \begin{cases} (|\langle \tilde{\pi}_{n_0+k}(f), X_i \rangle|^2 \mathbf{1}_{E_{n_0+n}}(i))_{i \in [m]}, & \text{if } k \in [n, n + k_0], \\ 0, & \text{if } k \notin [n, n + k_0]. \end{cases} \tag{4.24}$$

We also introduce the notation $\pi_{n,k}^* |\langle f, X_i \rangle|^2$ for the i -th component of $\pi_{n,k}^* |f_{\mathbf{X}}|^2$. Given these maps and using this shorthand notation, we set

$$\pi_{k,l}^\# |\langle f, X_i \rangle|^2 := \begin{cases} \pi_{k,l}^* |\langle f, X_i \rangle|^2 - \pi_{k,l-1}^* |\langle f, X_i \rangle|^2 =: \theta, & \text{if } |\theta| \leq (11 + 9\sqrt{2}) \rho_k \rho_l, \\ 0, & \text{otherwise,} \end{cases} \tag{4.25}$$

for every $i \in [m]$ and

$$\pi_{n_0+n+1} |\langle f, X_i \rangle|^2 := \sum_{k \leq n} \sum_{l \in [k, k+k_0]} \pi_{k,l}^\# |\langle f, X_i \rangle|^2, \tag{4.26}$$

for all $i \in [m]$. The associated sequence of sets A_{n_0+n+1} is the range of the maps $\pi_{n_0+n+1} : T \rightarrow A_{n_0+n+1}$ for $n = 0, \dots, \ell - 1$, i.e.

$$A_{n_0+n+1} := \left\{ \left(\sum_{k \leq n} \sum_{l \in [k, k+k_0]} \pi_{k,l}^\# |\langle f, X_i \rangle|^2 \right)_{i \in [m]} : f \in T \right\}. \tag{4.27}$$

Moreover, for $n = -1$, note that $\pi_{n_0} |f_{\mathbf{X}}| = 0$ and $A_{n_0} = \{0\}$. To show that the sequence of sets $(A_{n_0+n})_{n=0}^{\ell-1}$ we need to provide a bound for the size of A_{n_0+n} for each $n = 0, \dots, \ell - 1$. We recall that by construction $|\tilde{A}_n| \leq 2^{2^n}$ provided that $n \geq n_0$. Moreover, for each $n \in [0, \ell]$ and every $f \in T$ we have that the elements in A_{n_0+n} are determined by the first n elements $\tilde{\pi}_{n_0}(f), \dots, \tilde{\pi}_{n_0+n-1}(f)$. Hence,

$$|A_{n_0+n}| \leq \prod_{k \leq n-1} |\tilde{A}_{n_0+k}| \leq 2^{\sum_{k \leq n-1} 2^k} \leq 2^{2^n - 1} \leq 2^{2^n}.$$

Therefore, the sequence $(A_n)_{n=n_0}^{n_0+\ell}$ is an admissible sequence for T . Before we turn to the proof of Theorem 4.8 let us observe the following fact concerning the ℓ^∞ -norm of the approximants, which will come handy in proving Theorem 4.8. For each $f \in T$ the definition (4.25) implies that,

$$\begin{aligned}
 \max_{i \in [m]} \pi_{n_0+n} |\langle f, X_i \rangle|^2 &\leq \sum_{k=0}^{n-1} \sum_{l \in [k, k+k_0]} \pi_{k,l}^\# |\langle f, X_i \rangle|^2 \\
 &\leq (11 + 9\sqrt{2})\sqrt{s}K^2 \sum_{k \leq n-1} 2^{-k} \sum_{l=0}^{k_0} 2^{-l/2} \\
 &\leq (11 + 9\sqrt{2})\sqrt{s}K^2 \sum_{k=0}^{\infty} 2^{-k} \sum_{l=0}^{\infty} 2^{-l/2} \\
 &\leq (80 + 58\sqrt{2})\sqrt{s}K^2.
 \end{aligned} \tag{4.28}$$

Now that these estimates are in place we are left with ensuring the properties of this sequence claimed by Theorem 4.8. We have split this task into several lemmas, which can be combined to three leading principles for the admissible sequence, ℓ^2 -stability, γ_2 -boundedness and ℓ^1 -approximation.

- ℓ^2 -stability. Lemma 4.9 contains the observation that for $i \in E_{n_0+n}$ the approximants $\langle \tilde{\pi}_{n_0+n}(f), X_i \rangle$ have size roughly equal to ρ_n . In line with this observation is Lemma 4.11, which states that the sequence $(|E_{n_0+n} \rho_n^2|)_{n=0}^\ell$ captures essentially the ℓ^2 -norm of $(\langle f, X_i \rangle)_{i \in [m]}$.
- γ_2 -boundedness. Lemma 4.10 provides the key bound for $\|\pi_{n_0+n+1}|f_{\mathbf{X}}|^2 - \pi_{n_0+n}|f_{\mathbf{X}}|^2\|_2$ and therefore a bound for the rightmost term in $b_{n_0,\ell}(T)$.
- ℓ^1 -approximation. Lemma 4.12 finally provides a bound for the ℓ^1 -approximation term and is the last step towards proving Theorem 4.8.

We start with the observation regarding the interplay between the approximants $\tilde{\pi}_{n_0+n}(f)$ and the sets E_{n_0+n} .

Lemma 4.9. *Let $\ell > 0$, $f \in T \subseteq \sqrt{s}B_{\ell^1}^N$ and let X_1, \dots, X_m denote realizations of X . Let $\pi_{n+1} : T \rightarrow A_{n+1}$ be defined as in (4.26), (4.25), and (4.24) and let $E_{n_0}, \dots, E_{n_0+\ell}$ denote the associated sets defined as in (4.23). Then, for all $n = 0, \dots, \ell$, $k \in [n, n + k_0]$ and for all $i \in [m] \setminus \bigcup_{n \leq \ell} I_{n_0+n}(f)$ we have $\mathbf{1}_{E_{n_0+n}}(i) |\langle \tilde{\pi}_{n_0+k}(f), X_i \rangle| \leq (3 + \sqrt{2})\rho_n$.*

Proof. We first observe that if $n = 0$ the definition of the nets \tilde{A}_{n_0+n} implies that for each $i \in [m]$, $|\langle \tilde{\pi}_{n_0}(f), X_i \rangle| \leq \|\tilde{\pi}_{n_0}(f)\|_1 K \leq \rho_0$ and, if $n \geq 1$, the definition (4.23) of the set E_{n_0+n} implies that $\mathbf{1}_{E_{n_0+n}}(i) |\langle \tilde{\pi}_{n_0+n-1}(f), X_i \rangle| \leq \sqrt{2} \cdot \rho_{n-1}$. Further, by definition (4.22) of the sets $I_{n_0+n}(f)$, for all $i \in [m]$ outside of the set $\bigcup_{n \leq \ell} I_{n_0+n}(f)$ the estimates $|\langle \tilde{\pi}_{n_0+k}(f), X_i \rangle| - |\langle f, X_i \rangle| \leq \rho_k$ and $|\langle \tilde{\pi}_{n_0+n-1}(f), X_i \rangle| - |\langle f, X_i \rangle| \leq \rho_{n-1}$ hold simultaneously. Recalling the definition (4.21) of ρ_n , it follows that

$$\begin{aligned}
 |\langle \tilde{\pi}_{n_0+k}(f), X_i \rangle| \mathbf{1}_{E_{n_0+n}}(i) &\leq \mathbf{1}_{E_{n_0+n}}(i) (|\langle f, X_i \rangle| + \rho_k) \\
 &\leq \mathbf{1}_{E_{n_0+n}}(i) (|\langle \tilde{\pi}_{n_0+n-1}(f), X_i \rangle| + \rho_{n-1} + \rho_k) \\
 &\leq (3 + \sqrt{2})\rho_n.
 \end{aligned} \tag{4.29}$$

This concludes the proof. \square

Let us further observe that by a similar argument as in the lemma we find that for each $i \in [m] \setminus \bigcup_{n \leq \ell} I_{n_0+n}(f)$, and for each $n = 0, \dots, \ell$ we have the bounds

$$\sqrt{2}\rho_n \leq \mathbf{1}_{E_{n_0+n}}(i) |\langle \tilde{\pi}_{n_0+n}(f), X_i \rangle| \leq \mathbf{1}_{E_{n_0+n}}(i) (|\langle f, X_i \rangle| + \rho_n)$$

and

$$\mathbf{1}_{E_{n_0+n}}(i)|\langle f, X_i \rangle| \leq \mathbf{1}_{E_{n_0+n}}(i)|\langle \tilde{\pi}_{n_0+n}(f), X_i \rangle| + \rho_n \leq (4 + \sqrt{2})\rho_n .$$

In conjunction with Lemma 4.9 these bounds imply that for $i \in [m] \setminus \bigcup_{n \leq \ell} I_{n_0+n}(f)$ and $n = 0, \dots, \ell$,

$$(4.30) \text{ if } i \in E_{n_0+n}, \text{ then for all } k \in [n, n + k_0], |\langle \tilde{\pi}_{n_0+k}(f), X_i \rangle|^2 \leq (3 + \sqrt{2})^2 \rho_n^2,$$

$$(4.31) \text{ if } i \in E_{n_0+n}, \text{ then } (\sqrt{2} - 1)^2 \rho_n^2 \leq |\langle f, X_i \rangle|^2 \leq (\sqrt{2} + 4)^2 \rho_n^2.$$

Lemma 4.10. *Let $(A_n)_{n \geq n_0}$ be as in (4.27), $f \in T \subseteq \sqrt{s}B_{\ell^1}^N$ and $E_{n_0}, \dots, E_{n_0+\ell}$ defined as in (4.23). Then,*

$$\|\pi_{n_0+n+1}|f_{\mathbf{X}}|^2 - \pi_{n_0+n}|f_{\mathbf{X}}|^2\|_2 \leq (40 + 29\sqrt{2}) \sqrt{|E_{n_0+n}|} \rho_n^2 .$$

Proof. Let $f \in T$. Expanding the short-hand notation $|f_{\mathbf{X}}|^2$ and the ℓ^2 norm, we have

$$\|\pi_{n_0+n+1}|f_{\mathbf{X}}|^2 - \pi_{n_0+n}|f_{\mathbf{X}}|^2\|_2 = \left(\sum_{i=1}^m (\pi_{n_0+n+1}|\langle f, X_i \rangle|^2 - \pi_{n_0+n}|\langle f, X_i \rangle|^2)^2 \right)^{1/2} . \tag{4.32}$$

Recalling the definition (4.26) of $\pi_n|\langle f, X_i \rangle|^2$ for each $i \in [m]$, we find that the right hand side of (4.32) is equal to

$$\begin{aligned} & \left(\sum_{i=1}^m \left(\sum_{k \leq n} \sum_{l \in [k, k+k_0]} \pi_{k,l}^\# |\langle f, X_i \rangle|^2 - \sum_{k \leq n-1} \sum_{l \in [k, k+k_0]} \pi_{k,l}^\# |\langle f, X_i \rangle|^2 \right)^2 \right)^{1/2} \\ & = \left(\sum_{i=1}^m \left(\sum_{l \in [n, n+k_0]} \pi_{n,l}^\# |\langle f, X_i \rangle|^2 \right)^2 \right)^{1/2} . \end{aligned} \tag{4.33}$$

By definition (4.25) and recalling (4.24), we see that the support of the approximant $(\pi_{n,l}^\# |\langle f, X_i \rangle|^2)_{i \in [m]}$ is contained in the intersection of E_{n_0+n} and the set

$$J_{n,l} := \{i \in [m] : |\langle \tilde{\pi}_{n_0+l}(f), X_i \rangle|^2 - |\langle \tilde{\pi}_{n_0+l-1}(f), X_i \rangle|^2| \leq (11 + 9\sqrt{2}) \rho_n \rho_l\} .$$

Hence, we can rewrite the above sum in (4.33) as

$$\begin{aligned} & \left(\sum_{i=1}^m (\mathbf{1}_{E_{n_0+n}}(i) \sum_{l \in [n, n+k_0]} \mathbf{1}_{J_{n,l}}(i) (|\langle \tilde{\pi}_{n_0+l}(f), X_i \rangle|^2 - |\langle \tilde{\pi}_{n_0+l-1}(f), X_i \rangle|^2)^2) \right)^{1/2} \\ & \leq (11 + 9\sqrt{2}) \left(\sum_{i=1}^m (\mathbf{1}_{E_{n_0+n}}(i) \rho_n \sum_{l \in [n, n+k_0]} \mathbf{1}_{J_{n,l}}(i) \rho_l)^2 \right)^{1/2} \end{aligned}$$

Recalling that $\rho_l = K\sqrt{s}2^{-l/2}$ this implies

$$\begin{aligned} \sum_{l \in [n, n+k_0]} \rho_l & = K\sqrt{s} \sum_{l \in [n, n+k_0]} 2^{-l/2} = K\sqrt{s}2^{-n/2} \sum_{l \in [0, k_0]} 2^{-l/2} \\ & \leq K\sqrt{s}2^{-n/2} \sum_{l=0}^{\infty} 2^{-l/2} \leq \frac{\sqrt{2}}{\sqrt{2} - 1} K\sqrt{s}2^{-n/2} . \end{aligned}$$

Hence, the right hand side of (4.32) is bounded by

$$(40 + 29\sqrt{2}) \rho_n \left(\sum_{i=1}^m \mathbf{1}_{E_{n_0+n}}(i) \underbrace{(2K\sqrt{s}2^{-n/2})^2}_{=\rho_n} \right)^{1/2} \leq (40 + 29\sqrt{2}) \sqrt{|E_{n_0+n}|} \rho_n^2 .$$

Combining the above inequalities, we obtain the desired estimate. \square

Lemma 4.11. *Let $E_{n_0}, \dots, E_{n_0+\ell}$ be the sets associated with $\pi_{n_0+\ell}|f_{\mathbf{X}}|^2$, defined as in (4.23). Then, for every realization of X_1, \dots, X_m and for every $f \in T \subseteq \sqrt{s}B_{\ell^1}^N$, we have*

$$(\sqrt{2} - 1)^2 \sum_{n \leq \ell} |E_{n_0+n}| \rho_n^2 \leq \sum_{i=1}^m |\langle f, X_i \rangle|^2. \tag{4.34}$$

Proof. The sets $(E_{n_0+n})_{n=0}^{\ell}$ together with the set $E_* := \{i \in [m] : |\langle \tilde{\pi}_{n_0+\ell}(f), X_i \rangle| < \sqrt{2}\rho_{\ell}\}$ form a partition of the set $[m]$. Let $\mathcal{I}_{\ell}(f_{\mathbf{X}}) = (I_{n_0}(f), \dots, I_{n_0+\ell}(f))$ denote the sets associated with the initial approximants defined in (4.22). Then, by definitions (4.23) and (4.22), and using the fact that the sets $(E_{n_0+n})_{n=0}^{\ell}$ are disjoint, for every $i \in [m] \setminus \bigcup_{n \leq \ell} I_{n_0+n}(f)$, we see that

$$\begin{aligned} |\langle f, X_i \rangle| &\geq \sum_{n \leq \ell} |\langle f, X_i \rangle| \mathbf{1}_{E_{n_0+n}}(i) \\ &\geq \sum_{n \leq \ell} |\langle \tilde{\pi}_{n_0+n}(f), X_i \rangle| \mathbf{1}_{E_{n_0+n}}(i) - \rho_n \mathbf{1}_{E_{n_0+n}}(i) \\ &\geq \sum_{n \leq \ell} (\sqrt{2}\rho_n - \rho_n) \mathbf{1}_{E_{n_0+n}}(i) \geq (\sqrt{2} - 1) \sum_{n \leq \ell} \rho_n \mathbf{1}_{E_{n_0+n}}(i). \end{aligned}$$

Squaring both sides of the inequality above, summing over all indices $i \in [m] \setminus \bigcup_{n \leq \ell} I_{n_0+n}(f)$ and, again, using the fact that the sets E_{n_0+n} are disjoint, it follows that

$$\sum_{i=1}^m |\langle f, X_i \rangle|^2 \geq \sum_{i \in [m] \setminus \bigcup_{n \leq \ell} I_{n_0+n}(f)} |\langle f, X_i \rangle|^2 \geq (\sqrt{2} - 1)^2 \sum_{n \leq \ell} \rho_n^2 |E_{n_0+n}|. \tag{4.35}$$

This is the desired estimate. \square

Next we study the problem of bounding the difference $\| |f_{\mathbf{X}}|^2 - \pi_{n_0+\ell}|f_{\mathbf{X}}|^2 \|_1$. The following lemma provides a key step towards Theorem 4.8.

Lemma 4.12. *Let $T \subseteq \sqrt{s}B_{\ell^1}^N$ and let $\pi_{n_0+\ell} : T \rightarrow A_{n_0+\ell}$ be as defined in (4.26). Further, let $n_0 \geq \log_2 \log_2(2N) + \log_2 \log_2(sK^2/\delta)$ with $\delta \in (0, 1)$, $\ell = \lceil \log_2(sK^2/\delta) \rceil$, $k_0 = \lceil \log_2(1/\delta^2) \rceil$. Then, for every realization of X_1, \dots, X_m and every $f \in T$ we have*

$$\sum_{i=1}^m \left| \pi_{n_0+\ell} |\langle f, X_i \rangle|^2 - |\langle f, X_i \rangle|^2 \right| \leq (69 + 49\sqrt{2}) \delta \sum_{i=1}^m |\langle f, X_i \rangle|^2 + (643 + 468\sqrt{2}) \delta m. \tag{4.36}$$

Proof. For the approximation $\pi_{n_0+\ell} |\langle f, X_i \rangle|^2$ we let $\tilde{\pi}_{n_0}(f), \dots, \tilde{\pi}_{n_0+\ell}(f)$ denote the initial approximants in T and $E_{n_0}, \dots, E_{n_0+\ell}$ the corresponding sets of indices defined in (4.23). Further, let $\mathcal{I}_{\ell}(f_{\mathbf{X}}) = (I_{n_0}(f), \dots, I_{n_0+\ell}(f))$ denote the sets associated with the approximants, defined by (4.22). Moreover, we define the sets

$$J := [m] \setminus \bigcup_{n \leq \ell} I_{n_0+n}(f) \quad \text{and} \quad J^c = \bigcup_{n \leq \ell} I_{n_0+n}(f).$$

We split the proof into two main parts. First, we show that the contribution from the set J^c to the approximation error, i.e., $\| (|f_{\mathbf{X}}|^2 - \pi_{n_0+\ell}|f_{\mathbf{X}}|^2)|_{J^c} \|_1$, is bounded by $16(40 + 29\sqrt{2}) \delta m$. Afterwards, we study the contributions of the set J , namely $\| (|f_{\mathbf{X}}|^2 - \pi_{n_0+\ell}|f_{\mathbf{X}}|^2)|_J \|_1$.

For the first part, recalling from (4.21) that $M = 4\delta m/(sK^2 \log_2(sK^2/\delta))$, we observe that by the definition of weak covering and by the definition (4.22) of the sets I_{n_0+n} we have $|I_{n_0+n}(f)| \leq M$ and, therefore,

$$|J^c| \leq (\ell + 1) \max_{n \leq \ell} |I_{n_0+n}(f)| \leq \frac{4\delta m(\ell + 1)}{sK^2 \log_2(sK^2/\delta)}. \tag{4.37}$$

Hence, using (4.28) we find that for each $f \in T \subseteq \sqrt{s}B_{\ell}^N$,

$$\begin{aligned} & \sum_{i \in J^c} \left| |\langle f, X_i \rangle|^2 - \pi_{n_0+\ell} |\langle f, X_i \rangle|^2 \right| \\ & \leq \left(\max_{i \in [m]} |\langle f, X_i \rangle|^2 + \max_{i \in [m]} \pi_{n_0+\ell} |\langle f, X_i \rangle|^2 \right) \frac{4m(\ell + 1)\delta}{sK^2 \log_2(sK^2/\delta)} \\ & \leq 16(40 + 29\sqrt{2}) \delta m. \end{aligned} \tag{4.38}$$

In (4.38) we have used that $\ell = \lceil \log_2(sK^2/\delta) \rceil$. Therefore, the contribution from the indices $i \in J^c$ is bounded by $16(40 + 29\sqrt{2}) \delta m$.

In the second part of the proof, we estimate the approximation error corresponding to the set J . We start by expressing the approximant $\pi_{n_0+\ell} |f_{\mathbf{X}}|^2$ in an equivalent form (see (4.40) below) obtained by simplifying a telescoping sum. Namely, for every $i \in J$, we observe that, assuming that

$$\mathbf{1}_{E_{n_0+n}}(i) (|\pi_{n,k}^* |\langle f, X_i \rangle|^2 - \pi_{n,k-1}^* |\langle f, X_i \rangle|^2) \leq (11 + 9\sqrt{2}) \cdot \rho_n \rho_k, \tag{4.39}$$

for every $k \in [n, n + k_0]$ and $n \in [0, \ell]$, then for each $i \in [m]$ the telescoping sum in the definition (4.26) satisfies

$$\begin{aligned} \pi_{n_0+\ell} |\langle f, X_i \rangle|^2 &= \sum_{n=0}^{\ell-1} \sum_{k \in [n, n+k_0]} (|\pi_{n,k}^*(f), X_i \rangle|^2 - |\langle \pi_{n,k-1}^*(f), X_i \rangle|^2) \\ &= \sum_{n=0}^{\ell-1} |\langle \pi_{n,n+k_0}^* f, X_i \rangle|^2 = \sum_{n=0}^{\ell-1} \mathbf{1}_{E_{n_0+n}}(i) |\langle \tilde{\pi}_{n_0+n+k_0}(f), X_i \rangle|^2. \end{aligned} \tag{4.40}$$

Let us now verify the validity of (4.39). This is trivial if $i \notin E_{n_0+n}$. For $i \in E_{n_0+n}$, it is implied by the following estimate: For each $i \in J$ and $k \in [n, n + k_0]$, we have

$$\left| |\langle f, X_i \rangle|^2 - \pi_{n,k}^* |\langle f, X_i \rangle|^2 \right| = \left| |\langle f, X_i \rangle|^2 - \tilde{\pi}_{n_0+k} |\langle f, X_i \rangle|^2 \right| \mathbf{1}_{E_{n_0+n}}(i) \leq (7 + 2\sqrt{2}) \rho_k \rho_n. \tag{4.41}$$

Let us show the validity of (4.41) for $i \in E_{n_0+n}$. Recall that for $i \in J$ and $k \in [n, n + k_0]$ the definition of (4.21) implies that $||\langle f, X_i \rangle| - |\langle \tilde{\pi}_{n_0+k}(f), X_i \rangle|| \leq \rho_k$. Therefore, by using (4.30) and (4.31) we find,

$$\begin{aligned} & \left| |\langle f, X_i \rangle|^2 - |\langle \tilde{\pi}_{n_0+k}(f), X_i \rangle|^2 \right| \mathbf{1}_{E_{n_0+n}}(i) \\ &= \left| |\langle f, X_i \rangle| - |\langle \tilde{\pi}_{n_0+k}(f), X_i \rangle| \right| \cdot \left| |\langle f, X_i \rangle| + |\langle \tilde{\pi}_{n_0+k}(f), X_i \rangle| \right| \mathbf{1}_{E_{n_0+n}}(i) \\ &\leq \rho_k \left| |\langle f, X_i \rangle| + |\langle \tilde{\pi}_{n_0+k}(f), X_i \rangle| \right| \mathbf{1}_{E_{n_0+n}}(i) \\ &\leq (7 + 2\sqrt{2}) \rho_k \rho_n \end{aligned}$$

This proves the claim (4.41). With this we observe that for each $i \in J \cap E_{n_0+n}$ we have

$$|\pi_{n,k}^*|\langle f, X_i \rangle|^2 - \pi_{n,k-1}^*|\langle f, X_i \rangle|^2| \leq (11 + 9\sqrt{2}) \rho_n \rho_k . \tag{4.42}$$

Indeed, for $k \geq n + 1$ we employ (4.41) and the triangle inequality to see that

$$\begin{aligned} & |\pi_{n,k}^*|\langle f, X_i \rangle|^2 - \pi_{n,k-1}^*|\langle f, X_i \rangle|^2| \\ &= \left| \pi_{n,k}^*|\langle f, X_i \rangle|^2 - |\langle f, X_i \rangle|^2 + |\langle f, X_i \rangle|^2 - \pi_{n,k-1}^*|\langle f, X_i \rangle|^2 \right| \\ &\leq \left| \pi_{n,k}^*|\langle f, X_i \rangle|^2 - |\langle f, X_i \rangle|^2 \right| + \left| \pi_{n,k-1}^*|\langle f, X_i \rangle|^2 - |\langle f, X_i \rangle|^2 \right| \\ &\leq ((7 + 2\sqrt{2})\rho_n \rho_k + (7 + 2\sqrt{2})\rho_n \rho_{k-1}) \\ &\leq (7 + 2\sqrt{2})(1 + \sqrt{2}) \rho_n \rho_k = (11 + 9\sqrt{2}) \rho_n \rho_k . \end{aligned}$$

Further, for $k = n$ the desired estimate follows from Lemma 4.9 and for $k < n$ we have $\pi_{n,k}^*|\langle f, X_i \rangle|^2 = 0$. This proves (4.42) and, consequently, (4.39) and, in turn, (4.40).

Taking advantage of the representation (4.40), the proof is now concluded by estimating $\|(|f_{\mathbf{X}}|^2 - \phi_{n_0,\ell}|f_{\mathbf{X}}|^2)|_J\|_1$, where we define

$$\phi_{n_0,\ell}|\langle f, X_i \rangle|^2 := \sum_{n=0}^{\ell-1} \mathbf{1}_{E_{n_0+n}}(i) |\langle \tilde{\pi}_{n_0+n+k_0}(f), X_i \rangle|^2 . \tag{4.43}$$

For any $i \in J$, there are two possibilities. Either i belongs to a set E_{n_0+n} for exactly one value of $n = 0, \dots, \ell$ (recall that the sets $(E_{n_0+n})_{n \leq \ell}$ are disjoint), or i does not belong to any of these sets. In the first case, i.e., $i \in E_{n_0+n} \cap J$ for some $n = 0, \dots, \ell$, the estimate (4.41) applied for $k = n + k_0$ together with (4.31), i.e., $(\sqrt{2} - 1)^2 \rho_n^2 \leq |\langle f, X_i \rangle|^2$ and the fact that $k_0 = \lceil \log_2(1/\delta^2) \rceil$ imply that

$$\begin{aligned} \left| |\langle \tilde{\pi}_{n_0+n+k_0}(f), X_i \rangle|^2 - |\langle f, X_i \rangle|^2 \right| &\leq (11 + 9\sqrt{2}) \rho_n \rho_{n+k_0} \\ &\leq (11 + 9\sqrt{2}) \rho_n^2 2^{-k_0/2} \leq (69 + 49\sqrt{2}) \delta |\langle f, X_i \rangle|^2 . \end{aligned}$$

Hence, since the sets $(E_{n_0+n})_{n \leq \ell}$ are disjoint, we obtain the estimate

$$\sum_{i \in J \cap \bigcup_{n \leq \ell} E_{n_0+n}} \left| |\langle f, X_i \rangle|^2 - \phi_{n_0,\ell}|\langle f, X_i \rangle|^2 \right| \leq (69 + 49\sqrt{2}) \delta \sum_{i=1}^m |\langle f, X_i \rangle|^2 . \tag{4.44}$$

Next, let us consider the case that for all $n \leq \ell$ we have $i \notin E_{n_0+n}$. Under this condition, for all $n \leq \ell$ we have $\phi_{n_0,\ell}|\langle f, X_i \rangle|^2 = 0$. From the definition of E_{n_0+n} and the fact that $\tilde{\pi}_{n_0+\ell}(f)$ is associated with a weak covering of parameter ρ_ℓ , it follows that, for any $i \in J$,

$$|\langle f, X_i \rangle| \leq \rho_\ell + |\langle \tilde{\pi}_{n_0+\ell}(f), X_i \rangle| \leq \rho_\ell + \sqrt{2}\rho_\ell \leq (1 + \sqrt{2})\rho_\ell \leq (1 + \sqrt{2})\sqrt{\delta} .$$

Therefore, if $i \notin E_n$ for all $n \leq \ell$, we obtain

$$\begin{aligned} \sum_{i \in J \cap \bigcap_{n \leq \ell} E_{n_0+n}^c} \left| \phi_{n_0,\ell}|\langle f, X_i \rangle|^2 - |\langle f, X_i \rangle|^2 \right| &= \sum_{i \in J \cap \bigcap_{n \leq \ell} E_{n_0+n}^c} |\langle f, X_i \rangle|^2 \\ &\leq (1 + \sqrt{2})^2 \delta |J \cap \bigcap_{n \leq \ell} E_{n_0+n}^c| \\ &\leq (1 + \sqrt{2})^2 m \delta . \end{aligned} \tag{4.45}$$

Finally, since $[m] = (J \cap \bigcup_{n \leq \ell} E_{n_0+n}) \cup (J \cap \bigcap_{n \leq \ell} E_{n_0+n}^c) \cup J^c$ is a partition of $[m]$, combining the estimates (4.44), (4.45) and (4.38) we obtain,

$$\begin{aligned} & \sum_{i=1}^m \left| \pi_{n_0+\ell} |\langle f, X_i \rangle|^2 - |\langle f, X_i \rangle|^2 \right| \\ &= \sum_{i \in J \cap \bigcup_{n \leq \ell} E_{n_0+n}} \left| \pi_{n_0+\ell} |\langle f, X_i \rangle|^2 - |\langle f, X_i \rangle|^2 \right| \\ & \quad + \sum_{i \in J \cap \bigcap_{n \leq \ell} E_{n_0+n}^c} \left| \pi_{n_0+\ell} |\langle f, X_i \rangle|^2 - |\langle f, X_i \rangle|^2 \right| \\ & \quad + \sum_{i \in J^c} \left| \pi_{n_0+\ell} |\langle f, X_i \rangle|^2 - |\langle f, X_i \rangle|^2 \right| \\ & \leq (69 + 49\sqrt{2}) \delta \sum_{i \in J \cap \bigcup_{n \leq \ell} E_{n_0+n}} |\langle f, X_i \rangle|^2 + (1 + \sqrt{2})^2 \delta m + 16(40 + 29\sqrt{2}) \delta m . \end{aligned}$$

This estimate finishes the proof. \square

Having established these results we are ready to give the proof of Theorem 4.8.

Proof of Theorem 4.8. The inequality (4.19) coincides with the claim of Lemma 4.12.

Let us verify the second claim (4.20) of Theorem 4.8. By Lemma 4.10 we have the bound

$$\sum_{n=n_0}^{n_0+\ell-1} 2^{\frac{n}{2}} \|\pi_{n_0+n+1} |f_{\mathbf{X}}|^2 - \pi_{n_0+n} |f_{\mathbf{X}}|^2\|_2 \leq (40 + 29\sqrt{2}) \sum_{n=n_0}^{n_0+\ell-1} 2^{\frac{n}{2}} \sqrt{|E_{n_0+n}|} \rho_n^2$$

Recalling that $\rho_n = \sqrt{s} K 2^{-\frac{n}{2}}$, we see that the left hand side of (4.20) is bounded by

$$(40 + 29\sqrt{2}) \sum_{n=n_0}^{n_0+\ell-1} 2^{\frac{n}{2}} \sqrt{|E_{n_0+n}|} \rho_n^2 \leq (40 + 29\sqrt{2}) \sqrt{s} K 2^{\frac{n_0}{2}} \sum_{n=0}^{\ell-1} \sqrt{|E_{n_0+n}|} \rho_n .$$

By using Cauchy-Schwarz to estimate the sum over $n = 0, \dots, \ell - 1$, it follows that this term is bounded by

$$(40 + 29\sqrt{2}) \cdot 2^{\frac{n_0}{2}} \sqrt{\ell s} K \left(\sum_{n=0}^{\ell-1} |E_{n_0+n}| \rho_n^2 \right)^{1/2} . \tag{4.46}$$

Applying Lemma 4.11 to this estimate we obtain that (4.46) is bounded by

$$(200 + 140\sqrt{2}) \sqrt{\ell s} K 2^{\frac{n_0}{2}} \left(\sum_{i=1}^m |\langle f, X_i \rangle|^2 \right)^{1/2} .$$

Summarizing these estimates we obtain (4.20). \square

The last step in this section is to establish Theorem 4.2 based on Theorem 4.8.

Proof of Theorem 4.2. Using Lemma 4.4 together with Theorem 4.8 we obtain that for $T \subseteq \sqrt{s} B_{\ell^1}^N$, $\delta \in (0, 1)$, $n_0 = \lceil \log_2 \log_2(2N) + \log_2 \log_2(sK^2/\delta) \rceil$ (thus, we obtain $2^{\frac{n_0}{2}} \leq \sqrt{2} \sqrt{\log_2(2N) \log_2(sK^2/\delta)}$) and $\ell = \lceil \log_2(sK^2/\delta) \rceil \leq 2 \log_2(sK^2/\delta)$ the following bound holds:

$$\begin{aligned} \mathbb{E} \sup_{f \in T} \left| \frac{1}{m} \sum_{i=1}^m |\langle f, X_i \rangle|^2 \varepsilon_i \right| &\leq (69 + 49\sqrt{2}) \delta \sum_{i=1}^m |\langle f, X_i \rangle|^2 + (643 + 468\sqrt{2}) \delta m \\ &+ (280 + 200\sqrt{2}) \sqrt{\frac{sK^2 \log^2(sK^2/\delta) \log(eN)}{m}} \left(\mathbb{E} \sup_{f \in T} \frac{1}{m} \sum_{i=1}^m |\langle f, X_i \rangle|^2 \right)^{1/2}. \end{aligned}$$

This is the desired estimate. \square

4.4. Extension to Theorem 2.13

In this section we discuss the changes of the proof of Theorem 1.1 that are necessary in order to obtain its weighted version, Theorem 2.13. The result in Section 2.3 can be obtained by the same argument as presented in this section with minor modifications. For the convenience of the reader this section discusses the necessary changes and their impact on the argument.

As a first step let us note that the arguments stated up to the subsection 4.3 are also valid in the context of Theorem 2.13 and can be applied without changes after recognizing that under the assumption that $w_j \geq \|\langle X_i, e_j \rangle\|_{L^\infty}$ for all $j \in [N]$, then for each $f \in \mathbb{C}^N$ with $\|f\|_{w,1} \leq \sqrt{s}$ we have

$$|\langle f, X_j \rangle| \leq \|f\|_{w,1} \max_{j \in [N]} w_j^{-1} |X_j| \leq \|f\|_{w,1} \leq \sqrt{s}. \tag{4.47}$$

Therefore, the main content of this subsection is to show that we can construct an admissible sequence $(A_n)_{n \geq n_0}$, which mimics the behavior of the sequence we constructed in Theorem 4.8. In order to do this, let us start by introducing some notation.

The argument that we would like to adopt is concerned with a weighted version of the ℓ^1 -ball. Recall from (2.18) that the definition of the weighted ℓ^p -spaces gives the following definition for the weighted ℓ^1 -norm. For a sequence of weights $w \in [1, \infty)^N$ we have

$$\|f\|_{w,1} = \sum_{j=1}^N w_j |f_j|. \tag{4.48}$$

The unit ball of this norm is given by

$$B_{\ell_w^1}^N := \{f \in \mathbb{C}^N : \|f\|_{w,1} \leq 1\}. \tag{4.49}$$

Let us now indicate how the argument of Subsection 4.3 has to be adjusted in order to cover the set $\{(|\langle f, X_i \rangle|^2)_{i \in [m]} : f \in T\}$, where $T \subseteq \sqrt{s} B_{\ell_w^1}^N$.

The first point is to find a weak covering of $\sqrt{s} B_{\ell_w^1}^N$ in order to get the argument started.

Lemma 4.13. *Let $T \subset \sqrt{s} B_{\ell_w^1}^N$, $\delta \in (0, 1)$ and $\rho > 0$. Assume that $w \in [1, \infty)^N$ satisfies $w_j \geq \|\langle X, e_j \rangle\|_{L^\infty}$. Then, for every realization of X_1, \dots, X_m we have*

$$\log \mathcal{N}^* \left(T, \rho, \frac{4\delta m}{s \log_2(s/\delta)} \right) \leq \frac{2 \log_2(s/\delta) \log_2(2N)s}{\rho^2}$$

Sketch of Proof. The lemma is a straightforward adaption of Lemma 4.7. First observe that every $x \in B_{\ell_w^1}^N$ is a convex combination of $V = \{\pm w_j \sqrt{s} e_j, \pm i w_j \sqrt{s} e_j : j \in [N]\}$, i.e., there are coefficients $\lambda_v \geq 0$ with $\sum_{v \in V} \lambda_v = 1$, such that

$$x = \sum_{v \in V} \lambda_v v.$$

By arguing as in the proof of Lemma 4.7 we can construct random variables Z_l , for $l = 1, \dots, L$, such that $\mathbb{P}(Z_l = v) = \lambda_v$. This sequence of random variables satisfies $\mathbb{E}\langle Z_l, X_i \rangle = \langle x, X_i \rangle$ and for each $i \in [m]$,

$$\mathbb{P}\left(\left|\sum_{l=1}^L \langle Z_l, X_i \rangle - \langle x, X_i \rangle\right| \geq \rho\right) \leq 2 \exp(-L\rho^2 / (\max_{i \in [L]} |\langle X_i, Z_l \rangle|)) \leq 2 \exp(-L\rho^2/s),$$

since $|\langle X_i, Z_l \rangle| \leq \|Z_l\|_{w,1} \max_{j \in [N]} w_j^{-1} |\langle X_i, e_j \rangle| \leq \sqrt{s}$. The rest of the argument is identical to the arguments in the proof of Lemma 4.7. \square

4.4.1. Adaption of Theorem 4.8

Let us state a version of Theorem 4.8, which can be used in the weighted context. The theorem and its proof only contain minor changes compared to the result in Section 4.3.

Theorem 4.14. *Let $T \subseteq \sqrt{s}B_{\ell_1}^N$ and let $\delta \in (0, 1)$. Assume that for all $j \in [N]$ we have $w_j \geq \|\langle X, e_j \rangle\|_{L^\infty}$. Then, for $n_0 = \lceil \log_2 \log_2(eN) + \log_2 \log_2(s/\delta) \rceil$ and $\ell = \lceil \log_2(s/\delta) \rceil$ there is an admissible sequence $(A_n)_{n \geq n_0}$ for T such that*

- For all $f \in T$ we have

$$\| |f_{\mathbf{X}}|^2 - \pi_{n_0+\ell} |f_{\mathbf{X}}|^2 \|_1 \leq (69 + 49\sqrt{2})\delta \sum_{i=1}^m |\langle f, X_i \rangle|^2 + (643 + 468\sqrt{2})\delta. \tag{4.50}$$

- For all $f \in T$ we have

$$\sum_{n=n_0}^{n_0+\ell-1} 2^{\frac{n}{2}} \|\pi_{n+1} |f_{\mathbf{X}}|^2 - \pi_n |f_{\mathbf{X}}|^2\|_2 \leq (200 + 140\sqrt{2}) \cdot \sqrt{\ell s} 2^{\frac{n_0}{2}} \|(\langle f, X_i \rangle)_{i \in [m]}\|_2. \tag{4.51}$$

Let us summarize the necessary changes in the argument presented in Section 4.3. As in the covering argument above, the following assumption is crucial for the deduction of Theorem 4.14:

$$\text{for all } j \in [N] \text{ we have } w_j \geq \|\langle X, e_j \rangle\|_{L^\infty}. \tag{4.52}$$

The remainder is organized around the principles used in Section 4.3 and only contains remarks on the minor changes that need to be applied in order to obtain Theorem 4.14.

As a general note in order to understand the adaptations we remark that the estimates in Section 4.3 mostly depend on the inner product $|\langle f, X_i \rangle|$ for elements $f \in T \subseteq \sqrt{s}B_{\ell_1}^N$. However, under the assumption (4.52) the inner product $|\langle f, X_i \rangle|$ satisfies the same bounds as before.

The initial admissible sequence. In order to set up the initial admissible sequence we consider the norm $\|f\|_{\mathbf{X}} := \max_{i \in [m]} |\langle f, X_i \rangle|$ and observe that for $T \subseteq \sqrt{s}B_{\ell_1}^N$ we have the estimate $\sup_{f \in T} \|f\|_{\mathbf{X}} \leq \sqrt{s}$ provided that (4.52) is satisfied. Therefore, the initial sequence is defined as in (4.21) by setting

$$\rho_n = \sqrt{s}2^{-n/2} \quad \text{and} \quad M = \frac{4\delta m}{s \log_2(s/\delta)}. \tag{4.53}$$

The above lemma then guarantees the existence of an admissible sequence for $n = 0, \dots, \ell + k_0$. The necessary definitions in (4.25) and (4.27) can be adapted without changes.

ℓ^2 -stability. Versions of Lemma 4.9 and Lemma 4.11 for the weighted setting can be deduced by the same arguments as in Subsection 4.3.

The γ_2 -boundedness. A version of Lemma 4.10 can be deduced by exactly the same arguments. Further, the bound for the γ_2 functional in the proof of Theorem 4.8 can be used up to replacing $\sqrt{s}K$ by $\sup_{f \in T} \|f\|_{\mathbf{X}}$.

The ℓ^1 -approximation. In order to deduce a version of Lemma 4.12 we can again use the argument presented in Subsection 4.3. This is possible, since all arguments in the proof only depend on $\max_{i \in [m]} |\langle f, X_i \rangle|^2$ and the differences $\left| |\langle f, X_i \rangle| - |\langle \tilde{\pi}_{n_0+k}(f), X_i \rangle| \right| \mathbf{1}_{E_{n_0+n}}(i)$. Further, by the choice of ρ_n , Assumption (4.52), it follows for $f \in \sqrt{s}B_{\ell_w^1}^N$ we have the bound $\|f\|_{\mathbf{X}} \leq \sqrt{s}$.

Acknowledgments

S.B. acknowledges the Postdoctoral Training Centre in Stochastics of the Pacific Institute for the Mathematical Sciences (PIMS), the Natural Sciences and Engineering Research Council of Canada (NSERC) through grant number 611675 and RGPIN-2020-06766, the Centre for Advanced Modelling Science (CAD-MOS), and the Faculty of Arts and Science of Concordia University for their financial support. S.D., H.C.J., and H.R. acknowledge funding by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under the priority program SPP 1798 (COSIP -Compressed Sensing in Information Processing) through the project Quantized Compressive Spectrum Sensing (project no. 273202924).

References

- [1] B. Adcock, Infinite-dimensional compressed sensing and function interpolation, *Found. Comput. Math.* 18 (3) (2018) 661–701.
- [2] B. Adcock, A. Bao, S. Brugiapaglia, Correcting for unknown errors in sparse high-dimensional function approximation, *Numer. Math.* 142 (3) (2019) 667–711.
- [3] B. Adcock, S. Brugiapaglia, C.G. Webster, Compressed sensing approaches for polynomial approximation of high-dimensional functions, in: *Compressed Sensing and Its Applications*, Springer, 2017, pp. 93–124.
- [4] B. Adcock, A.C. Hansen, C. Poon, B. Roman, *Breaking the Coherence Barrier: A New Theory for Compressed Sensing*, *Forum Math. Sigma*, vol. 5, Cambridge University Press, 2017.
- [5] G.S. Alberti, M. Santacesaria, Infinite dimensional compressed sensing from anisotropic measurements and applications to inverse problems in PDE, *Appl. Comput. Harmon. Anal.* (2019).
- [6] J. Błasiok, P. Lopatto, K. Luh, J. Marcinek, S. Rao, An improved lower bound for sparse reconstruction from subsampled Hadamard matrices, Preprint, arXiv:1903.12135, 2019.
- [7] S. Boucheron, G. Lugosi, P. Massart, *Concentration Inequalities*, Oxford University Press, Oxford, 2013.
- [8] J.L. Bouchot, H. Rauhut, C. Schwab, Multi-level compressive sensing Petrov-Galerkin discretization of high-dimensional parametric PDEs, Preprint arXiv:1701.01671, 2017.
- [9] J. Bourgain, An improved estimate in the restricted isometry problem, in: B. Klartag, E. Milman (Eds.), *Geometric Aspects of Functional Analysis*, in: *Lect. Notes Math.*, vol. 2116, Springer International Publishing, 2014, pp. 65–70.
- [10] J. Bourgain, S. Dilworth, K. Ford, S. Konyagin, D. Kutzarova, Explicit constructions of RIP matrices and related problems, *Duke Math. J.* 159 (1) (2011) 145–185.
- [11] Olivier Bousquet, A Bennett concentration inequality and its application to suprema of empirical processes, *C. R. Math.* 334 (6) (2002) 495–500.
- [12] S. Brugiapaglia, *COmpressed SolvING: sparse approximation of PDEs based on compressed sensing*, PhD thesis, MOX - Politecnico di Milano, 2016.
- [13] S. Brugiapaglia, A compressive spectral collocation method for the diffusion equation under the restricted isometry property, arXiv preprint arXiv:1807.06606, 2018.
- [14] S. Brugiapaglia, S. Micheletti, F. Nobile, S. Perotto, Wavelet-Fourier CORSING techniques for multi-dimensional advection-diffusion-reaction equations, *IMA J. Numer. Anal.* (2020), <https://doi.org/10.1093/imanum/draa036>, in press.
- [15] S. Brugiapaglia, S. Micheletti, S. Perotto, Compressed solving: a numerical approximation technique for elliptic PDEs based on compressed sensing, *Comput. Math. Appl.* 70 (6) (2015) 1306–1335.
- [16] S. Brugiapaglia, F. Nobile, S. Micheletti, S. Perotto, A theoretical study of COmpressed SolvING for advection-diffusion-reaction problems, *Math. Comput.* 87 (309) (2018) 1–38.
- [17] T.T. Cai, A. Zhang, Sparse representation of a polytope and recovery of sparse signals and low-rank matrices, *IEEE Trans. Inf. Theory* 60 (1) (January 2014) 122–132.
- [18] E.J. Candès, T. Tao, J.K. Romberg, Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information, *IEEE Trans. Inf. Theory* 52 (2) (2006) 489–509.
- [19] E.J. Candès, T. Tao, Decoding by linear programming, *IEEE Trans. Inf. Theory* 51 (12) (Dec 2005) 4203–4215.
- [20] E.J. Candès, T. Tao, Near optimal signal recovery from random projections: universal encoding strategies?, *IEEE Trans. Inf. Theory* 52 (12) (2006) 5406–5425.
- [21] M. Cheraghchi, V. Guruswami, A. Velingker, Restricted isometry of Fourier matrices and list decodability of random linear codes, *SIAM J. Comput.* 42 (5) (2013) 1888–1914.

- [22] A. Chkifa, N. Dexter, H. Tran, C.G. Webster, Polynomial approximation via compressed sensing of high-dimensional functions on lower sets, *Math. Comput.* 87 (311) (2018) 1415–1450.
- [23] O. Christensen, *An Introduction to Frames and Riesz Bases*, Appl. Numer. Harmon. Anal., Birkhäuser, Boston, 2002.
- [24] A. Cohen, W. Dahmen, R. DeVore, Orthogonal matching pursuit under the restricted isometry property, *Constr. Approx.* 45 (1) (2017) 113–127.
- [25] W. Dahmen, *Multiscale and wavelet methods for operator equations*, in: *Multiscale Problems and Methods in Numerical Simulations*, Springer, 2003, pp. 31–96.
- [26] S. Dirksen, G. Lecue, H. Rauhut, On the gap between restricted isometry properties and sparse recovery conditions, *IEEE Trans. Inf. Theory* 64 (8) (2018) 5478–5487.
- [27] D.L. Donoho, Compressed sensing, *IEEE Trans. Inf. Theory* 52 (4) (2006) 1289–1306.
- [28] A. Doostan, H. Owhadi, A non-adapted sparse approximation of PDEs with stochastic inputs, *J. Comput. Phys.* 230 (8) (2011) 3015–3034.
- [29] S. Foucart, H. Rauhut, A mathematical introduction to compressive sensing, in: *Appl. Numer. Harmon. Anal.*, Birkhäuser, 2013.
- [30] I. Haviv, O. Regev, The restricted isometry property for subsampled Fourier matrices, in: *Geometric Aspects of Functional Analysis*, Springer, 2017, pp. 163–179.
- [31] S. Jokar, V. Mehrmann, M.E. Pfetsch, H. Yserentant, Sparse approximate solution of partial differential equations, *Appl. Numer. Math.* 60 (4) (2010) 452–472.
- [32] M. Kabanava, H. Rauhut, Analysis ℓ^1 -recovery with frames and Gaussian measurements, *Acta Appl. Math.* 140 (1) (December 2015) 173–195.
- [33] H. Kang, M.-J. Lai, X. Li, An economical representation of PDE solution by using compressive sensing approach, *Comput. Aided Des.* 115 (2019) 78–86.
- [34] F. Kraher, R. Ward, New and improved Johnson-Lindenstrauss embeddings via the restricted isometry property, *SIAM J. Math. Anal.* 43 (3) (2011) 1269–1281.
- [35] F. Kraher, R. Ward, Stable and robust sampling strategies for compressive imaging, *IEEE Trans. Image Process.* 23 (2) (2014) 612–622.
- [36] M. Ledoux, M. Talagrand, *Probability in Banach Spaces*, Springer-Verlag, Berlin, 1991.
- [37] A. Mackey, H. Schaeffer, S. Osher, On the compressive spectral method, *Multiscale Model. Simul.* 12 (4) (2014) 1800–1827.
- [38] L. Mathelin, K.A. Gallivan, A compressed sensing approach for partial differential equations with random input data, *Commun. Comput. Phys.* 12 (4) (2012) 919–954.
- [39] S. Oymak, B. Recht, M. Soltanolkotabi, Isometric sketching of any set via the restricted isometry property, *Inf. Inference* 7 (4) (2018) 707–726.
- [40] A. Quarteroni, A. Valli, *Numerical Approximation of Partial Differential Equations*, Springer Ser. Comput. Math., vol. 23, Springer-Verlag, Berlin, 2008.
- [41] H. Rauhut, Random sampling of sparse trigonometric polynomials, *Appl. Comput. Harmon. Anal.* 22 (1) (2007) 16–42.
- [42] H. Rauhut, C. Schwab, Compressive sensing Petrov-Galerkin approximation of high-dimensional parametric operator equations, *Math. Comput.* 86 (304) (2017) 661–700.
- [43] H. Rauhut, R. Ward, Sparse Legendre expansions via ℓ_1 -minimization, *J. Approx. Theory* 164 (5) (2012) 517–533.
- [44] H. Rauhut, R. Ward, Interpolation via weighted ℓ^1 -minimization, *Appl. Comput. Harmon. Anal.* 40 (2) (2016) 321–351.
- [45] M. Rudelson, Roman Vershynin, On sparse reconstruction from Fourier and Gaussian measurements, *Commun. Pure Appl. Math.* 61 (2008) 1025–1045.
- [46] M. Rudelson, S. Zhou, Reconstruction from anisotropic random measurements, in: *Conference on Learning Theory*, 2012, pp. 10.1–10.24.
- [47] M. Talagrand, *Upper and Lower Bounds for Stochastic Processes: Modern Methods and Classical Problems*, Vol. 60, Springer Science & Business Media, 2014.
- [48] G. Tran, R. Ward, Exact recovery of chaotic systems from highly corrupted data, *Multiscale Model. Simul.* 15 (3) (2017) 1108–1129.
- [49] K. Urban, *Wavelet Methods for Elliptic Partial Differential Equations*, Oxford University Press, 2009.
- [50] L. Yan, L. Guo, D. Xiu, Stochastic collocation algorithms using ℓ_1 -minimization, *Int. J. Uncertain. Quantificat.* 2 (3) (2012).
- [51] X. Yang, G.E. Karniadakis, Reweighted ℓ_1 minimization method for stochastic elliptic differential equations, *J. Comput. Phys.* 248 (2013) 87–108.
- [52] T. Zhang, Sparse recovery with orthogonal matching pursuit under RIP, *IEEE Trans. Inf. Theory* 57 (9) (2011) 6215–6221.