ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

MATHICSE

Mathematics Institute of Computational Science and Engineering

School of Basic Sciences - Section of Mathematics

MATHICSE Technical Report

Nr. 23.2015

September 2015

A theoretical study of COmpRessed SolvING for advection-diffusion-reaction problems

Simone Brugiapaglia, Fabio Nobile, Stefano Micheletti, Simona Perotto

http://mathicse.epfl.ch

Phone: +41 21 69 37648

Address:
EPFL - SB - MATHICSE (Bâtiment MA)
Station 8 - CH-1015 - Lausanne - Switzerland

Fax: +41 21 69 32545

A theoretical study of COmpRessed SolvING for advection-diffusion-reaction problems

S. Brugiapaglia[♯,*], F. Nobile[♭], S. Micheletti[♯] and S. Perotto[♯]

September 7, 2015

[♯] MOX – Dipartimento di Matematica
Politecnico di Milano
P.zza Leonardo da Vinci 32, 20133 Milano, Italy

[♭] MATHICSE – CSQI
École Polytechnique Fédérale de Lausanne
Lausanne, CH-1015, Switzerland

Keywords: Compressed sensing, Petrov-Galerkin formulation, advection-diffusion-reaction equation, inf-sup property, local coherence

AMS Subject Classification (2010): Primary 65N30, 65Y20, 94A20; Secondary 65T40, 65K10, 42A61

Abstract

We present a theoretical analysis of the CORSING (COmpRessed SolvING) method for the numerical approximation of partial differential equations based on compressed sensing. In particular, we show that the best s-term approximation of the weak solution of a PDE with respect to an orthonormal system of N trial functions, can be recovered via a Petrov-Galerkin approach using $m \ll N$ orthonormal test functions. This recovery is guaranteed if the local a-coherence associated with the bilinear form and the selected trial and test bases fulfills suitable decay properties. The fundamental tool of this analysis is the restricted inf-sup property, i.e., a combination of the classical inf-sup condition and the well-known restricted isometry property of compressed sensing.

*Corresponding author: simone.brugiapaglia@polimi.it

1 Introduction

Compressed Sensing (CS) is an extremely powerful tool of signal processing employed to recover a sparse signal using far fewer measurements than those required by the Nyquist-Shannon sampling theorem. In particular, expanding the signal with respect to a basis of N vectors, it is possible to recover the best s-term approximation to the signal, with $s \ll N$, by means of m random measurements, with $s < m \ll N$ [16, 7, 18].

In [6], we introduced an application of CS to the numerical approximation of Partial Differential Equations (PDEs). For this purpose, we rely on an analogy between the sampling process of a signal and the evaluation of the bilinear form associated with a Petrov-Galerkin discretization ([4, 17, 26]) of the PDE against randomly chosen test functions. We named the resulting numerical method CORSING, acronym for COmpRessed SolvING. In particular, we showed through an extensive numerical assessment that CORSING can successfully reduce the computational cost of a full Petrov-Galerkin discretization of an elliptic problem.

Comparison with other techniques. The CORSING method aims at computing the best s-term approximation to the solution to a PDE. Therefore, it can be classified among nonlinear approximation methods ([15, 30]) for PDEs. Although the framework for CORSING is very general and can accommodate many different choices of trial and test spaces, when considering hierarchical piecewise polynomials over an initial coarse triangulation as trial basis functions, a possible competitor approach is the Adaptive Finite Element Method (AFEM) (see, e.g., [24] and the references therein). AFEM and CORSING are, however, thoroughly different: in AFEM, the solution is iteratively computed according to the loop

$$\text{SOLVE} \rightarrow \text{ESTIMATE} \rightarrow \text{MARK} \rightarrow \text{REFINE},$$

and exploiting suitable *a posteriori* error estimators. On the contrary, with CORSING we employ a reduced Petrov-Galerkin discretization, using a fixed trial space of dimension N (which corresponds ideally to a very fine uniform refinement, expressed in a hierarchical basis) and performing a fixed number of random measurements in the test space. In particular:

(1) the trial space is not iteratively enlarged, but fixed initially;

(2) the measurements in the test space are performed non-adaptively;

(3) no *a posteriori* error estimators/indicators are needed.

The CORSING procedure then recovers an s-sparse solution (with $s \ll N$), which can be compared with the AFEM solution on the same grounds. We consider (1) as a possible drawback of CORSING, whereas (2) and (3) are upsides. In principle (1) requires a higher computational cost in the recovery phase, whereas (2) allows for full parallelization and (3) significantly reduces the implementation complexity.

From a different perspective, CORSING can be considered as a variant of the infinite-dimensional CS, where CS is applied to infinite-dimensional Hilbert spaces [1, 2]. This is achieved by subsampling a given isometry of the Hilbert space, usually associated with an inner product and a change of basis (e.g., from a wavelet basis to the Fourier basis). The main idea behind CORSING is different, since it deals with the bilinear form arising from the weak formulation, that can be even nonsymmetric. Nevertheless, we think that the theory developed in [1, 2] could play a significant role for a deeper understanding of the CORSING technique and this will be a subject of future investigation.

Main contributions of the paper. The goal of this paper is to set up a theoretical analysis of CORSING, providing sufficient conditions for convergence, and formalizing the empirical recipes given in [6]. With this aim, we introduce a novel variant of the classical inf-sup condition [5], where the infimum is considered among the sparse elements of the trial space and the supremum over a small test space. We refer to this condition as *Restricted Inf-Sup Property* (RISP), since it combines the inf-sup condition and the Restricted Isometry Property (RIP), a well-known tool in the CS literature. Another important tool of the analysis is the concept of *local a-coherence*, a generalization of the *local coherence* to bilinear forms on Hilbert spaces. In particular, we have been inspired by [19], where an optimal recovery result for compressed sensing, with non-uniform random subsampling based on the local coherence, is proved for the Haar and Fourier discrete bases in dimension one and two.

The main results of the paper can be thus summarized. First, we prove sufficient conditions for the RISP, depending on suitable hypotheses on the local a-coherence. Then, recovery error estimates for the CORSING algorithm are provided. In particular, in Theorem 3.8 we show that a sufficient condition for the RISP to hold with high probability in a given s-sparse set is that m and s be linear dependent, up to logarithmic factors. On the contrary, at the moment we are only able to prove (Theorem 3.9) a *uniform* RISP (i.e., a RISP holding in *all* possible s-sparse sets) assuming a quadratic dependence between m and s, although we conjecture that, as in CS, the dependence on s should be linear. Exploiting these theorems, we prove a recovery result in expectation (Theorem 3.13) and one in probability (Theorem 3.14). In particular, we check the hypotheses on the local a-coherence in the case of a one-dimensional advection-diffusion-reaction equation employing the hierarchical multiscale basis in [33, 13] and the Fourier sine basis.

Outline of the paper. In Section 2, we formally introduce the CORSING, defining all the input/output variables involved in the algorithm. The theoretical analysis based on the RISP is presented in Section 3, and an application of the theory to a one-dimensional advection-diffusion-reaction equation is discussed in Section 4. In Section 5, we provide some numerical results, and we draw some conclusions in Section 6.

2 CORSING

In this section, after setting up the notation, we describe the COmpRessed SolvING procedure, in short, CORSING, first introduced in [6].

2.1 Notation

Let $\mathbb{N} := \{1, 2, 3, \ldots\}$ be the set of positive natural numbers and $\mathbb{N}_0 := \mathbb{N} \cup \{0\}$. Consider two separable Hilbert spaces, $U = \text{span}\{\psi_j\}_{j\in\mathbb{N}}$ and $V = \text{span}\{\varphi_q\}_{q\in\mathbb{N}}$, generated by the bases $\{\psi_j\}_{j\in\mathbb{N}}$ and $\{\varphi_q\}_{q\in\mathbb{N}}$, respectively, and equipped with the inner products $(\cdot,\cdot)_U$ and $(\cdot,\cdot)_V$. Given two positive integers N and M, we define the finite dimensional truncations of U and V, which represent the *trial* and *test* space, respectively, as

$$U^N := \text{span}\{\psi_j\}_{j\in[N]} \quad \text{and} \quad V^M := \text{span}\{\varphi_q\}_{q\in[M]},$$

where $[k] := \{1, \ldots, k\}$ for every $k \in \mathbb{N}$. In particular, $[\infty] = \mathbb{N}$. We denote the span of the basis functions relative to a given subset of indices $\mathcal{S} \subseteq [N]$ as

$$U^N_{\mathcal{S}} := \text{span}\{\psi_j\}_{j\in\mathcal{S}}.$$

Given a positive integer $s \leq N$, we also define the set U^N_s of s-sparse functions of U^N with respect to the basis $\{\psi_j\}_{j\in[N]}$ as the set of all functions that are linear combinations of at most s basis functions, namely

$$U^N_s := \bigcup_{\mathcal{S}\subseteq[N];\,|\mathcal{S}|=s} U^N_{\mathcal{S}}.$$

We stress that U^N_s is not a vector space. Indeed, the sum of two s-sparse elements is in general $2s$-sparse. The sets $V^M_{\mathcal{T}}$ and V^M_m are defined analogously, for every $\mathcal{T} \subseteq [M]$ and $m \leq M$.

We denote by U^* and V^* the dual spaces of U and V, respectively.

We also introduce the *reconstruction* and *decomposition* operators associated with a basis, that allow us to switch between functions and the corresponding coefficients in the basis expansion.

Definition 2.1. The *reconstruction operator* $\Psi : \ell^2 \to U$ related to the basis $\{\psi_j\}_{j\in\mathbb{N}}$ of U associates with a sequence $\mathbf{u} = (u_j)_{j\in\mathbb{N}} \in \ell^2$ the linear combination

$$u = \Psi \mathbf{u} := \sum_{j=1}^{\infty} u_j \psi_j.$$

The *decomposition operator* $\Psi^* : U \to \ell^2$ applied to a given function $u \in U$ is defined component-wise as

$$(\Psi^* u)_k := (u, \psi_k^*)_U, \quad \forall k \in \mathbb{N},$$

where $\{\psi_k^*\}_{k\in\mathbb{N}}$ is the basis biorthogonal to $\{\psi_j\}_{j\in\mathbb{N}}$, namely, $(\psi_j, \psi_k^*)_U = \delta_{j,k}$, $\forall j, k \in \mathbb{N}$.

The reconstruction operator Φ and the decomposition operator Φ^* associated with the basis $\{\varphi_q\}_{q\in\mathbb{N}}$ of V are defined analogously.

Remark 2.2. We observe that $\Psi\Psi^* = Id_U$ and $\Psi^*\Psi = Id_{\ell^2}$.

2.2 The general reference problem

Consider the following problem

$$\text{find } u \in U : a(u, v) = \mathcal{F}(v), \quad \forall v \in V, \tag{1}$$

where $a : U \times V \to \mathbb{R}$ is a bilinear form and $\mathcal{F} \in V^*$. We assume $a(\cdot,\cdot)$ to fulfill the following three conditions

$$\exists \alpha > 0 : \quad \inf_{u\in U} \sup_{v\in V} \frac{a(u, v)}{\|u\|_U \|v\|_V} \geq \alpha, \tag{2}$$

$$\exists \beta > 0 : \quad \sup_{u\in U} \sup_{v\in V} \frac{|a(u, v)|}{\|u\|_U \|v\|_V} \leq \beta, \tag{3}$$

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

These assumptions imply the existence and uniqueness of the solution to (1), thanks to a generalization of the Lax-Milgram lemma due to Nečas [23], [26, Theorem 5.1.2].

To simplify the notation, when an infimum or a supremum of a fraction $f(x)/g(x)$ over a given set X is considered, the zeros of $g(x)$ are understood to be removed from X.

Our goal is to approximate the solution to (1), by merging the classical Petrov-Galerkin formulation (sometimes also called non-standard Galerkin method) [4, 26, 17] with Compressed Sensing techniques [16, 7]. The adopted procedure corresponds to the R-CORSING method, recently introduced in [6], simply denoted by CORSING in the following developments.

2.3 Main hypotheses

We will use three assumptions throughout the article.

Hypothesis 1 (Orthonormal tests)**.** *The test basis $\{\varphi_q\}_{q\in\mathbb{N}}$ is an orthonormal system of V.*

We generalize the notion of local coherence (see, e.g., [19]) to bilinear forms defined over Hilbert spaces.

Definition 2.3 (Local a-coherence $\boldsymbol{\mu}^N$)**.** Given $N \in \mathbb{N}\cup\{\infty\}$, the real-valued sequence $\boldsymbol{\mu}^N$ defined as

$$\mu_q^N := \sup_{j\in[N]} |a(\psi_j, \varphi_q)|^2, \quad \forall q \in \mathbb{N},$$

is called *local a-coherence of $\{\psi_j\}_{j\in[N]}$ with respect to $\{\varphi_q\}_{q\in\mathbb{N}}$*.

The second hypothesis concerns the local a-coherence.

Hypothesis 2 (Summability of $\boldsymbol{\mu}^N$)**.** *The local a-coherence of $\{\psi_j\}_{j\in[N]}$ with respect to $\{\varphi_q\}_{q\in\mathbb{N}}$ fulfills the summability condition*

$$\|\boldsymbol{\mu}^N\|_1 < +\infty,$$

or, equivalently, $\boldsymbol{\mu}^N \in \ell^1$.

Notice that Hypothesis 2 does not hinge on the ordering considered for the elements of the truncated trial basis $\{\psi_j\}_{j\in[N]}$.

The last hypothesis concerns an explicit upper bound to the local a-coherence.

Hypothesis 3 (Upper bound $\boldsymbol{\nu}^N$)**.** *For every $N \in \mathbb{N}$, we assume to have a computable componentwise upper bound $\boldsymbol{\nu}^N$ to the local a-coherence $\boldsymbol{\mu}^N$, i.e., a real-valued sequence such that*

$$\mu_q^N \leq \nu_q^N, \quad \forall q \in \mathbb{N}.$$

For every $M \in \mathbb{N}$, we define the vector $\boldsymbol{\nu}^{N,M} \in \mathbb{R}^M$ as the restriction of $\boldsymbol{\nu}^N$ to the first M components. Moreover, we require that

- *the vector $\boldsymbol{\nu}^{N,M}/\|\boldsymbol{\nu}^{N,M}\|_1$ is efficiently computable for every $N, M \in \mathbb{N}$;*
- *there exists a real bivariate polynomial P such that*

$$\|\boldsymbol{\nu}^{N,M}\|_1 \lesssim P(\log N, \log M).$$

The upper bound $\boldsymbol{\nu}^N$ need not be sharp.

As usual, with notation $x \sim y$, $x \lesssim y$ or $x \gtrsim y$, it is understood that there exists a constant $C > 0$ not depending on x and y, such that $x = Cy$, $x \leq Cy$ or $x \geq Cy$, respectively.

Algorithm 2.1

PROCEDURE $\widehat{u} = \mathsf{CORSING}$ (N, s, $\boldsymbol{\nu}^N$, γ_M, C_M, γ_m, C_m)

1. **Definition of** M **and** m

$>$ $M \leftarrow C_M s^{\gamma_M} N$;

$>$ $m \leftarrow C_m s^{\gamma_m} \|\boldsymbol{\nu}^{N,M}\|_1 \log(N/s)$;

2. **Test selection**

$>$ $\mathbf{p} \leftarrow \boldsymbol{\nu}^{N,M} / \|\boldsymbol{\nu}^{N,M}\|_1$;

$>$ Draw $\tau_1, \ldots, \tau_m$ independently at random from $[M]$ according to the probability $\mathbf{p}$;

3. **Assembly**

$>$ Build $\mathbf{A}$, $\mathbf{f}$ and $\mathbf{D}$, defined in (5) and (6), respectively;

4. **Recovery**

$>$ Find a solution $\widehat{\mathbf{u}}$ to $\min_{\mathbf{u}\in\mathbb{R}^N} \|\mathbf{D}(\mathbf{A}\mathbf{u} - \mathbf{f})\|_2^2$, s.t. $\|\mathbf{u}\|_0 \leq s$;

$>$ $\widehat{u} \leftarrow \Psi\widehat{\mathbf{u}}$.

2.4 The CORSING procedure

The CORSING procedure is summarized in Algorithm 2.1. Let us now describe in more detail the input/output variables and the main steps of the method.

INPUT

- N: dimension of the trial space;
- $s \ll N$: number of trial coefficients to recover;
- upper bound $\boldsymbol{\nu}^N$ in Hypothesis 3 and four positive constants γ_M, C_M, γ_m, and C_m used to select the dimension M of the test space and the m tests to perform.

OUTPUT

- $\widehat{u} \in U_s^N$: approximate s-sparse solution to (1).

1. Definition of M and m. The test space dimension M and the number m of tests to perform are chosen as functions of N and s as

$$M = C_M s^{\gamma_M} N, \quad m = C_m s^{\gamma_m} \|\boldsymbol{\nu}^{N,M}\|_1 \log(N/s).$$

In Section 3, we prove the existence of suitable values for the constants γ_M, C_M, γ_m that ensure the CORSING algorithm to recover the best s-term approximation to u in expectation and in probability. In Section 4, we perform a sensitivity analysis on the constants C_M and C_m for some specific differential problems and with $\gamma_m = 1, 2$. Numerical evidence shows that $\gamma_m = 1$ is a valid choice, but proving this from a theoretical viewpoint still remains an open problem. On the contrary, the value of γ_M seems to depend on the trial and test bases considered (see Section 4).

2. Test selection. In order to formalize the test selection procedure, we introduce a probability space $(\Omega, \mathcal{E}, \mathbb{P})$ and consider $\tau_1, \ldots, \tau_m$ as i.i.d. discrete random variables taking values in $[M]$, namely

$$\tau_i : \Omega \to [M], \quad \forall i \in [m].$$

Moreover, given a vector $\mathbf{p} = (p_q)_{q \in [M]} \in [0,1]^M$ such that $\|\mathbf{p}\|_1 = 1$, the probability law is defined as

$$\mathbb{P}\{\tau_i = q\} = p_q, \quad \forall q \in [M].$$

Throughout the paper, the vector $\mathbf{p}$ will be assumed to be of the form

$$\mathbf{p} := \frac{\boldsymbol{\nu}^{N,M}}{\|\boldsymbol{\nu}^{N,M}\|_1}, \tag{4}$$

where the values for $\boldsymbol{\nu}^{N,M}$ are known from Hypothesis 3.

3. Assembly. In this phase, we build the *stiffness matrix* $\mathbf{A} \in \mathbb{R}^{m \times N}$ and the *load vector* $\mathbf{f} \in \mathbb{R}^m$ associated with the Petrov-Galerkin discretization of (1), defined as

$$A_{ij} := a(\psi_j, \varphi_{\tau_i}), \quad f_i := \mathcal{F}(\varphi_{\tau_i}), \quad \forall j \in [N], \forall i \in [m]. \tag{5}$$

Moreover, the matrix $\mathbf{D} \in \mathbb{R}^{m \times m}$ is a diagonal preconditioner, depending on the vector $\mathbf{p}$ as

$$D_{ik} := \frac{\delta_{ik}}{\sqrt{m p_{\tau_i}}}, \quad \forall i \in [m]. \tag{6}$$

4. Recovery. The vector of trial coefficients $\widehat{\mathbf{u}}$ of the approximate solution is recovered as

$$\widehat{\mathbf{u}} := \arg\min_{\mathbf{v}\in\mathbb{R}^N} \|\mathbf{D}(\mathbf{A}\mathbf{v}-\mathbf{f})\|_2^2, \quad \text{s.t.} \quad \|\mathbf{v}\|_0 \leq s, \tag{7}$$

where $\|\mathbf{u}\|_0 = |\{j : u_j \neq 0\}|$ is the so called ℓ^0-*norm*. Consequently, the approximate solution is defined as $\widehat{u} := \Psi\widehat{\mathbf{u}}$. An equivalent functional formulation of (7) is

$$\widehat{u} \equiv \arg\min_{v\in U_s^N} \sum_{i=1}^{m} \frac{1}{mp_{\tau_i}}(a(v,\varphi_{\tau_i}) - \mathcal{F}(\varphi_{\tau_i}))^2. \tag{8}$$

In practice, problem (7) is approximately solved through the greedy algorithm Orthogonal Matching Pursuit (OMP), [20, 25].

The procedure defined by (7) (or, equivalently, (8)) has been proved to be generally NP-hard, [21], but fortunately, there are several ways to efficiently and accurately approximate its solutions under particular circumstances, e.g., when the RIP holds. These strategies can be divided in two main families: convex relaxation techniques, such as the well known ℓ^1-minimization, also known as Basis Pursuit (BP) [8], and greedy algorithms [31, 22]. In this paper, we focus on greedy techniques and, in particular, we employ the OMP algorithm. For recent results concerning its accuracy, we refer to [34, 11].

The reason for this choice is twofold. First, using OMP we can easily control the parameter s, i.e., the sparsity of the compressed solution $\widehat{u}$. Second, the time complexity of the OMP algorithm is easily estimated, namely $\mathcal{O}(smN)$ for basic implementations, while the complexity of BP depends on the particular algorithm used to solve the corresponding Linear Programming and it is not easily quantifiable. All the numerical experiments made in this work are performed using the OMP-BOX MATLAB® package, version 10 - see [29, 28].[1]

3 Theoretical analysis

3.1 Preliminary results

The main statistical tools employed in this paper are Chernoff's bounds for matrices. They were introduced by H. Chernoff during the early 50's in the scalar form [9], and generalized to the matrix setting by R. Ahlswede and A. Winter in 2003 [3]. These bounds have been recently refined in 2012 by J. Tropp in [32].

First, we present the main result employed in our analysis. The proof of the following theorem can be found in [32, Corollary 5.2].

[1]The reader interested in the algorithmic issues can find many comparisons between the OMP and BP approaches in [6, Section 5].

Theorem 3.1 (Matrix Chernoff's bounds)**.** *Consider a finite sequence of i.i.d. random, symmetric $s \times s$ real matrices $\mathbf{X}^1, \dots, \mathbf{X}^m$ such that*

$$0 \leq \lambda_{\min}(\mathbf{X}^i) \text{ and } \lambda_{\max}(\mathbf{X}^i) \leq R \quad \text{almost surely}, \quad \forall i \in [m].$$

Define $\overline{\mathbf{X}} := \frac{1}{m}\sum_{i=1}^{m} \mathbf{X}^i$, $E_{\min} := \lambda_{\min}(\mathbb{E}[\mathbf{X}^i])$ *and* $E_{\max} := \lambda_{\max}(\mathbb{E}[\mathbf{X}^i])$. *Then,*

$$\mathbb{P}\{\lambda_{\min}(\overline{\mathbf{X}}) \leq (1-\delta)E_{\min}\} \leq s \exp\left(-\frac{m\rho_\delta E_{\min}}{R}\right), \quad \forall \delta \in [0,1], \tag{9}$$

$$\mathbb{P}\{\lambda_{\max}(\overline{\mathbf{X}}) \geq (1+\delta)E_{\max}\} \leq s \exp\left(-\frac{m\widetilde{\rho}_\delta E_{\max}}{R}\right), \quad \forall \delta \geq 0,$$

with

$$\rho_\delta := (1-\delta)\log(1-\delta) + \delta, \quad \widetilde{\rho}_\delta := (1+\delta)\log(1+\delta) - \delta. \tag{10}$$

□

Notice that both constants $\rho_\delta, \widetilde{\rho}_\delta \sim \delta^2$ when $\delta \to 0$.

We conclude this section by recalling few results that will be repeatedly used in the next proofs.

Lemma 3.2. *If $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{d\times d}$ are symmetric and $\mathbf{B}$ is also positive definite, it holds*

$$\lambda_{\min}(\mathbf{B}^{-\frac{1}{2}}\mathbf{A}\mathbf{B}^{-\frac{1}{2}}) = \inf_{\mathbf{u}\in\mathbb{R}^d} \frac{\mathbf{u}^\intercal\mathbf{A}\mathbf{u}}{\mathbf{u}^\intercal\mathbf{B}\mathbf{u}}, \tag{11}$$

$$\lambda_{\max}(\mathbf{B}^{-\frac{1}{2}}\mathbf{A}\mathbf{B}^{-\frac{1}{2}}) = \sup_{\mathbf{u}\in\mathbb{R}^d} \frac{\mathbf{u}^\intercal\mathbf{A}\mathbf{u}}{\mathbf{u}^\intercal\mathbf{B}\mathbf{u}}. \tag{12}$$

Lemma 3.3. *Consider a generic set X. The infimum and the supremum on X fulfill the following properties*

$$\sup_{x\in X} 1/f(x) = 1/\inf_{x\in X} f(x), \qquad \forall f : X \to (0,+\infty), \tag{13}$$

$$\sup_{x\in X} f(x)g(x) \leq \sup_{x\in X} f(x) \sup_{x\in X} g(x), \qquad \forall f,g : X \to [0,+\infty), \tag{14}$$

$$\inf_{x\in X}(f(x)-g(x)) \geq \inf_{x\in X} f(x) - \sup_{x\in X} g(x), \qquad \forall f,g : X \to \mathbb{R}. \tag{15}$$

3.2 Non-uniform restricted inf-sup property

In this section, we deal with the core of our paper, namely an analysis of the CORSING algorithm.

We denote the space of vectors of $\mathbb{R}^N$ supported in $\mathcal{S} \subseteq [N]$ as $\mathbb{R}^N_{\mathcal{S}}$, namely

$$\mathbb{R}^N_{\mathcal{S}} := \{\mathbf{u} \in \mathbb{R}^N : u_j = 0,\ \forall j \notin \mathcal{S}\}.$$

Moreover, we introduce some further notation.

Definition 3.4 (Matrices $\mathbf{K}$, $\mathbf{K}_{\mathcal{S}}$ and $\mathbf{A}_{\mathcal{S}}$)**.** We define the matrix $\mathbf{K} \in \mathbb{R}^{N \times N}$ as

$$K_{jk} := (\psi_j, \psi_k)_U.$$

and its restriction $\mathbf{K}_{\mathcal{S}} \in \mathbb{R}^{s \times s}$ to $\mathcal{S} := \{\sigma_1, \ldots, \sigma_s\} \subseteq [N]$ as

$$(K_{\mathcal{S}})_{jk} := (\psi_{\sigma_j}, \psi_{\sigma_k})_U.$$

Moreover, we denote by $\mathbf{A}_{\mathcal{S}} \in \mathbb{R}^{m \times s}$ the submatrix of $\mathbf{A}$ consisting only of the columns with indices in $\mathcal{S}$.

We observe that $\mathbf{K}$ is symmetric and positive definite (s.p.d.) and fulfills

$$\mathbf{u}^\mathsf{T}\mathbf{K}\mathbf{u} = \|\Psi\mathbf{u}\|_U^2, \quad \forall \mathbf{u} \in \mathbb{R}^N, \tag{16}$$

where the reconstruction operator in (16) is implicitly restricted from ℓ^2 to $\mathbb{R}^N$ (equivalently, the vector $\mathbf{u}$ is extended to ℓ^2 by adding zeros for $j > N$). The matrix $\mathbf{K}_{\mathcal{S}}$ is also s.p.d. and it satisfies the relation

$$\mathbf{u}_{\mathcal{S}}^\mathsf{T}\mathbf{K}_{\mathcal{S}}\mathbf{u}_{\mathcal{S}} = \mathbf{u}^\mathsf{T}\mathbf{K}\mathbf{u}, \quad \forall \mathbf{u} \in \mathbb{R}^N_{\mathcal{S}},$$

where $\mathbf{u}_{\mathcal{S}} \in \mathbb{R}^s$ is the restriction of $\mathbf{u}$ to $\mathcal{S}$, namely $(u_{\mathcal{S}})_j = u_{\sigma_j}$, for every $j \in [s]$. In this section, we fix a subset $\mathcal{S} := \{\sigma_1, \ldots, \sigma_s\} \subseteq [N]$ of cardinality s.

We introduce the Gram matrix $\mathbf{G}^\infty$ relative to the restriction of $a(\cdot, \cdot)$ to $U^N_{\mathcal{S}} \times V^\infty$.

Definition 3.5 (Matrix $\mathbf{G}^\infty$)**.** Define the matrix $\mathbf{G}^\infty \in \mathbb{R}^{s \times s}$ such that

$$G^\infty_{jk} := \sum_{q=1}^\infty a(\psi_{\sigma_j}, \varphi_q) a(\psi_{\sigma_k}, \varphi_q), \quad \forall j, k \in [s],$$

where the series are well defined thanks to Hypothesis 2 and $G^\infty_{jk} \leq \|\boldsymbol{\mu}^N\|_1$, for every $j, k \in [s]$.

The first lemma provides a relation between the inf-sup constant α associated with the bilinear form $a(\cdot, \cdot)$ and the Gram matrix $\mathbf{G}^\infty$.

Lemma 3.6. *Suppose that the bilinear form $a(\cdot, \cdot)$ fulfills the inf-sup property* (2)*. Then, it holds*

$$\lambda_{\min}(\mathbf{K}_{\mathcal{S}}^{-\frac{1}{2}}\mathbf{G}^\infty\mathbf{K}_{\mathcal{S}}^{-\frac{1}{2}}) \geq \alpha^2.$$

Proof. The following chain of inequalities holds

$$
\begin{aligned}
\alpha &\leq \inf_{u\in U}\sup_{v\in V}\frac{a(u,v)}{\|u\|_U\|v\|_V} \leq \inf_{u\in U^N_{\mathcal{S}}}\sup_{v\in V}\frac{a(u,v)}{\|u\|_U\|v\|_V}\\
&= \inf_{\mathbf{u}\in\mathbb{R}^N_{\mathcal{S}}}\sup_{\mathbf{v}\in\ell^2}\frac{1}{\|\mathbf{K}^{\frac{1}{2}}\mathbf{u}\|_2\|\mathbf{v}\|_2}\sum_{q=1}^{\infty}a(\Psi\mathbf{u},\varphi_q)v_q = \inf_{\mathbf{u}\in\mathbb{R}^N_{\mathcal{S}}}\frac{1}{\|\mathbf{K}^{\frac{1}{2}}\mathbf{u}\|_2}\left[\sum_{q=1}^{\infty}a(\Psi\mathbf{u},\varphi_q)^2\right]^{\frac{1}{2}}.
\end{aligned}
$$

The first inequality is property (2), while the second inequality follows from taking the infimum over a subset of U. The first equality is obtained by expanding u and v with respect to the bases $\{\psi_j\}_{j\in\mathcal{S}}$ and $\{\varphi_q\}_{q\in\mathbb{N}}$, respectively; moreover, we use relations (16) and $\|\mathbf{v}\|_2 = \|v\|_V$ implied by Hypothesis 1. The last equality can be deduced by applying the definition of operator norm

$$
\sup_{\mathbf{v}\in\ell^2}\frac{1}{\|\mathbf{v}\|_2}\sum_{q=1}^{\infty}a(\Psi\mathbf{u},\varphi_q)v_q = \|(a(\Psi\mathbf{u},\varphi_q))_{q\in\mathbb{N}}\|_{(\ell^2)^*}
$$

and by identifying $(\ell^2)^*$ with ℓ^2. Now, since all the quantities involved in the chain of inequalities are positive, we can square the terms

$$
\begin{aligned}
\alpha^2 &\leq \inf_{\mathbf{u}\in\mathbb{R}^N_{\mathcal{S}}}\frac{1}{\mathbf{u}^\intercal\mathbf{K}\mathbf{u}}\sum_{q=1}^{\infty}a(\Psi\mathbf{u},\varphi_q)^2 = \inf_{\mathbf{u}\in\mathbb{R}^s}\frac{1}{\mathbf{u}^\intercal\mathbf{K}_{\mathcal{S}}\mathbf{u}}\sum_{q=1}^{\infty}\left[\sum_{j=1}^{s}u_j a(\psi_{\sigma_j},\varphi_q)\right]^2\\
&= \inf_{\mathbf{u}\in\mathbb{R}^s}\frac{1}{\mathbf{u}^\intercal\mathbf{K}_{\mathcal{S}}\mathbf{u}}\sum_{q=1}^{\infty}\sum_{j=1}^{s}\sum_{k=1}^{s}u_j u_k a(\psi_{\sigma_j},\varphi_q)a(\psi_{\sigma_k},\varphi_q)\\
&= \inf_{\mathbf{u}\in\mathbb{R}^s}\frac{1}{\mathbf{u}^\intercal\mathbf{K}_{\mathcal{S}}\mathbf{u}}\sum_{j=1}^{s}\sum_{k=1}^{s}u_j u_k\sum_{q=1}^{\infty}a(\psi_{\sigma_j},\varphi_q)a(\psi_{\sigma_k},\varphi_q)\\
&= \inf_{\mathbf{u}\in\mathbb{R}^s}\frac{\mathbf{u}^\intercal\mathbf{G}^{\infty}\mathbf{u}}{\mathbf{u}^\intercal\mathbf{K}_{\mathcal{S}}\mathbf{u}} = \lambda_{\min}(\mathbf{K}_{\mathcal{S}}^{-\frac{1}{2}}\mathbf{G}^{\infty}\mathbf{K}_{\mathcal{S}}^{-\frac{1}{2}}).
\end{aligned}
$$

We have expanded $\Psi\mathbf{u}$ and identified $\mathbf{u}$ with its restriction to $\mathcal{S}$. Then, we have exchanged the summations thanks to Hypothesis 2 and Fubini-Tonelli's theorem. Successively, we have used the definition of $\mathbf{G}^{\infty}$ together with relation (11). □

The second lemma provides a recipe on how to choose the truncation level M on the tests, after selecting N and s.

Lemma 3.7. *Under the same hypotheses as in Lemma 3.6, we fix a real number $\delta_M \in [0,1)$. Then, if $M \in \mathbb{N}$ satisfies the truncation condition*

$$
s\sum_{q>M}\mu_q^N \leq \alpha^2\lambda_{\min}(\mathbf{K}_{\mathcal{S}})\delta_M, \tag{17}
$$

the following inequality holds

$$\lambda_{\min}(\mathbf{K}_{\mathcal{S}}^{-\frac{1}{2}}\mathbf{G}^M\mathbf{K}_{\mathcal{S}}^{-\frac{1}{2}}) \geq (1-\delta_M)\alpha^2,$$

where $\mathbf{G}^M \in \mathbb{R}^{s\times s}$ *is the truncated version of* $\mathbf{G}^\infty$*, namely*

$$G_{jk}^M := \sum_{q=1}^{M} a(\psi_{\sigma_j},\varphi_q)a(\psi_{\sigma_k},\varphi_q).$$

Proof. First, consider the splitting $\mathbf{G}^\infty = \mathbf{G}^M + \mathbf{T}^M$, where $\mathbf{T}^M$ corresponds to the tail of the series identifying $\mathbf{G}^\infty$,

$$T_{jk}^M = \sum_{q>M} a(\psi_{\sigma_j},\varphi_q)a(\psi_{\sigma_k},\varphi_q).$$

Now, notice that

$$\begin{aligned}\lambda_{\min}(\mathbf{K}_{\mathcal{S}}^{-\frac{1}{2}}\mathbf{G}^M\mathbf{K}_{\mathcal{S}}^{-\frac{1}{2}}) &= \lambda_{\min}(\mathbf{K}_{\mathcal{S}}^{-\frac{1}{2}}(\mathbf{G}^\infty - \mathbf{T}^M)\mathbf{K}_{\mathcal{S}}^{-\frac{1}{2}}) \\ &\geq \lambda_{\min}(\mathbf{K}_{\mathcal{S}}^{-\frac{1}{2}}\mathbf{G}^\infty\mathbf{K}_{\mathcal{S}}^{-\frac{1}{2}}) - \lambda_{\max}(\mathbf{K}_{\mathcal{S}}^{-\frac{1}{2}}\mathbf{T}^M\mathbf{K}_{\mathcal{S}}^{-\frac{1}{2}})\end{aligned}$$

The inequality can be proved using Lemma 3.2 and exploiting property (15). Applying Lemma 3.6, we obtain

$$\lambda_{\min}(\mathbf{K}_{\mathcal{S}}^{-\frac{1}{2}}\mathbf{G}^M\mathbf{K}_{\mathcal{S}}^{-\frac{1}{2}}) \geq \alpha^2(1-\lambda_{\max}(\mathbf{K}_{\mathcal{S}}^{-\frac{1}{2}}\mathbf{T}^M\mathbf{K}_{\mathcal{S}}^{-\frac{1}{2}})/\alpha^2).$$

Thus, the thesis is proved if we bound the maximum eigenvalue of the tail as follows

$$\lambda_{\max}(\mathbf{K}_{\mathcal{S}}^{-\frac{1}{2}}\mathbf{T}^M\mathbf{K}_{\mathcal{S}}^{-\frac{1}{2}}) \leq \delta_M\alpha^2.$$

For this purpose, we compute

$$\begin{aligned}\lambda_{\max}(\mathbf{K}_{\mathcal{S}}^{-\frac{1}{2}}\mathbf{T}^M\mathbf{K}_{\mathcal{S}}^{-\frac{1}{2}}) &= \sup_{\mathbf{u}\in\mathbb{R}^s}\frac{\mathbf{u}^\intercal\mathbf{T}^M\mathbf{u}}{\mathbf{u}^\intercal\mathbf{K}_{\mathcal{S}}\mathbf{u}} \\ &= \sup_{\mathbf{u}\in\mathbb{R}^s}\frac{1}{\mathbf{u}^\intercal\mathbf{K}_{\mathcal{S}}\mathbf{u}}\sum_{j=1}^{s}\sum_{k=1}^{s}u_ju_k\sum_{q>M}a(\psi_{\sigma_j},\varphi_q)a(\psi_{\sigma_k},\varphi_q) \\ &= \sup_{\mathbf{u}\in\mathbb{R}^s}\frac{1}{\mathbf{u}^\intercal\mathbf{K}_{\mathcal{S}}\mathbf{u}}\sum_{q>M}\left[\sum_{j=1}^{s}u_ja(\psi_{\sigma_j},\varphi_q)\right]^2 \\ &\leq \sup_{\mathbf{u}\in\mathbb{R}^s}\frac{\mathbf{u}^\intercal\mathbf{u}}{\mathbf{u}^\intercal\mathbf{K}_{\mathcal{S}}\mathbf{u}}s\sum_{q>M}\mu_q^N = \frac{1}{\lambda_{\min}(\mathbf{K}_{\mathcal{S}})}s\sum_{q>M}\mu_q^N.\end{aligned}$$

We start from definition (12). Then, by exploiting Hypothesis 2 and Fubini-Tonelli's theorem, combined with Cauchy-Schwarz inequality, the definition of $\boldsymbol{\mu}^N$, of (11) and of (13), we obtain the desired result under hypothesis (17). □

This lemma provides a sufficient condition on the truncation parameter M that ensures an arbitrarily small decrease of the inf-sup constant α by a factor $(1-\delta_M)^{\frac{1}{2}}$. Moreover, a value M that fulfills (17) always exists thanks to Hypothesis 2. Relation (17) can be also interpreted as a sufficient condition for the space V^M to be δ-proximal for $U_{\mathcal{S}}^N$, with constant $\delta = \sqrt{\delta_M}$ (see [14]).

Now, we prove the main result of this section.

Theorem 3.8 (Non-uniform RISP)**.** *Let the truncation condition in Lemma 3.7 hold. Then, for every $0 < \varepsilon < 1$ and $\delta_m \in [0,1)$, provided that*

$$m \geq \widetilde{C}_{\mathcal{S}}\, s \|\boldsymbol{\nu}^{N,M}\|_1 \log(s/\varepsilon),$$

where $\widetilde{C}_{\mathcal{S}} := [\rho_{\delta_m}(1-\delta_M)\alpha^2 \lambda_{\min}(\mathbf{K}_{\mathcal{S}})]^{-1}$ and ρ_{δ_m} is defined according to (10)*, the following non-uniform RISP holds with probability greater than or equal to $1-\varepsilon$*

$$\inf_{\mathbf{u}\in\mathbb{R}^s} \sup_{\mathbf{v}\in\mathbb{R}^m} \frac{\mathbf{v}^\intercal \mathbf{D}\mathbf{A}_{\mathcal{S}}\mathbf{u}}{\|\mathbf{K}_{\mathcal{S}}^{\frac{1}{2}}\mathbf{u}\|_2 \|\mathbf{v}\|_2} > \widetilde{\alpha} > 0, \tag{18}$$

where $\widetilde{\alpha} := (1-\delta_M)^{\frac{1}{2}}(1-\delta_m)^{\frac{1}{2}}\alpha$ and $\mathbf{D}$ is defined in (6)*.*

Proof. The proof is organized as follows. First, we show that the inf-sup in (18) can be interpreted as the square root of the minimum eigenvalue of the sample mean of a sequence of certain i.i.d. random matrices $\mathbf{X}^{\tau_1},\ldots,\mathbf{X}^{\tau_m}$. Then, we compute the expectation of $\mathbf{X}^{\tau_i}$ and show that the maximum eigenvalue of $\mathbf{X}^{\tau_i}$ is uniformly bounded. Finally, we apply the matrix Chernoff bound (9).

Let us discuss each step of the proof in detail. First, we compute

$$\begin{aligned}
\inf_{\mathbf{u}\in\mathbb{R}^s} \sup_{\mathbf{v}\in\mathbb{R}^m} \frac{\mathbf{v}^\intercal \mathbf{D}\mathbf{A}_{\mathcal{S}}\mathbf{u}}{\|\mathbf{K}_{\mathcal{S}}^{\frac{1}{2}}\mathbf{u}\|_2 \|\mathbf{v}\|_2} &= \inf_{\mathbf{u}\in\mathbb{R}^s} \frac{1}{\|\mathbf{K}_{\mathcal{S}}^{\frac{1}{2}}\mathbf{u}\|_2} \sup_{\mathbf{v}\in\mathbb{R}^m} \frac{\mathbf{v}^\intercal \mathbf{D}\mathbf{A}_{\mathcal{S}}\mathbf{u}}{\|\mathbf{v}\|_2} \\
&= \inf_{\mathbf{u}\in\mathbb{R}^s} \frac{\|\mathbf{D}\mathbf{A}_{\mathcal{S}}\mathbf{u}\|_2}{\|\mathbf{K}_{\mathcal{S}}^{\frac{1}{2}}\mathbf{u}\|_2} = [\lambda_{\min}(\mathbf{K}_{\mathcal{S}}^{-\frac{1}{2}}\mathbf{A}_{\mathcal{S}}^\intercal \mathbf{D}^2 \mathbf{A}_{\mathcal{S}} \mathbf{K}_{\mathcal{S}}^{-\frac{1}{2}})]^{\frac{1}{2}}.
\end{aligned}$$

The second equality hinges on the definition of operator norm combined with the identification of $(\mathbb{R}^m)^*$ with $\mathbb{R}^m$ while the third one exploits (11).

Relying on the following relation,

$$(\mathbf{A}_{\mathcal{S}}^\intercal \mathbf{D}^2 \mathbf{A}_{\mathcal{S}})_{jk} = \frac{1}{m}\sum_{i=1}^m \frac{1}{p_{\tau_i}} a(\psi_{\sigma_j},\varphi_{\tau_i}) a(\psi_{\sigma_k},\varphi_{\tau_i})$$

we define the matrices $\mathbf{H}^{\tau_i} \in \mathbb{R}^{s\times s}$ with $H_{jk}^{\tau_i} := \frac{1}{p_{\tau_i}} a(\psi_{\sigma_j},\varphi_{\tau_i}) a(\psi_{\sigma_k},\varphi_{\tau_i})$ and

$$\mathbf{X}^{\tau_i} := \mathbf{K}_{\mathcal{S}}^{-\frac{1}{2}} \mathbf{H}^{\tau_i} \mathbf{K}_{\mathcal{S}}^{-\frac{1}{2}},$$

so that

$$\overline{\mathbf{X}} := \frac{1}{m}\sum_{i=1}^{m}\mathbf{X}^{\tau_i} = \mathbf{K}_{\mathcal{S}}^{-\frac{1}{2}}\mathbf{A}_{\mathcal{S}}^{\intercal}\mathbf{D}^2\mathbf{A}_{\mathcal{S}}\mathbf{K}_{\mathcal{S}}^{-\frac{1}{2}}.$$

Thus, it holds

$$\inf_{\mathbf{u}\in\mathbb{R}^s}\sup_{\mathbf{v}\in\mathbb{R}^m}\frac{\mathbf{v}^{\intercal}\mathbf{D}\mathbf{A}_{\mathcal{S}}\mathbf{u}}{\|\mathbf{K}_{\mathcal{S}}^{\frac{1}{2}}\mathbf{u}\|_2\|\mathbf{v}\|_2} = [\lambda_{\min}(\overline{\mathbf{X}})]^{\frac{1}{2}}. \tag{19}$$

With a view to the Chernoff bounds, we estimate $\mathbb{E}[\mathbf{X}^{\tau_i}]$ and the corresponding minimum eigenvalue. A direct computation yields

$$\mathbb{E}[H_{jk}^{\tau_i}] = \sum_{q=1}^{M}\mathbb{P}\{\tau_i = q\}H_{jk}^{q} = \sum_{q=1}^{M}p_q\frac{1}{p_q}a(\psi_{\sigma_j},\varphi_q)a(\psi_{\sigma_k},\varphi_q) = G_{jk}^{M}.$$

As a consequence, we have

$$\mathbb{E}[\mathbf{X}^{\tau_i}] = \mathbb{E}[\mathbf{K}_{\mathcal{S}}^{-\frac{1}{2}}\mathbf{H}^{\tau_i}\mathbf{K}_{\mathcal{S}}^{-\frac{1}{2}}] = \mathbf{K}_{\mathcal{S}}^{-\frac{1}{2}}\mathbb{E}[\mathbf{H}^{\tau_i}]\mathbf{K}_{\mathcal{S}}^{-\frac{1}{2}} = \mathbf{K}_{\mathcal{S}}^{-\frac{1}{2}}\mathbf{G}^{M}\mathbf{K}_{\mathcal{S}}^{-\frac{1}{2}},$$

i.e., from Lemma 3.7

$$\lambda_{\min}(\mathbb{E}[\mathbf{X}^{\tau_i}]) \geq (1-\delta_M)\alpha^2. \tag{20}$$

Our aim is now to bound $\lambda_{\max}(\mathbf{X}^{\tau_i})$ from above. We have

$$\begin{aligned}
\lambda_{\max}(\mathbf{X}^{\tau_i}) &= \sup_{\mathbf{u}\in\mathbb{R}^s}\frac{\mathbf{u}^{\intercal}\mathbf{H}^{\tau_i}\mathbf{u}}{\mathbf{u}^{\intercal}\mathbf{K}_{\mathcal{S}}\mathbf{u}} \leq \sup_{\mathbf{u}\in\mathbb{R}^s}\frac{\mathbf{u}^{\intercal}\mathbf{u}}{\mathbf{u}^{\intercal}\mathbf{K}_{\mathcal{S}}\mathbf{u}}\sup_{\mathbf{u}\in\mathbb{R}^s}\frac{\mathbf{u}^{\intercal}\mathbf{H}^{\tau_i}\mathbf{u}}{\mathbf{u}^{\intercal}\mathbf{u}} \\
&= [\lambda_{\min}(\mathbf{K}_{\mathcal{S}})]^{-1}\sup_{\mathbf{u}\in\mathbb{R}^s}\frac{1}{\mathbf{u}^{\intercal}\mathbf{u}}\sum_{j=1}^{s}\sum_{k=1}^{s}u_j u_k\frac{1}{p_{\tau_i}}a(\psi_{\sigma_j},\varphi_{\tau_i})a(\psi_{\sigma_k},\varphi_{\tau_i}) \\
&= [\lambda_{\min}(\mathbf{K}_{\mathcal{S}})]^{-1}\frac{1}{p_{\tau_i}}\sup_{\mathbf{u}\in\mathbb{R}^s}\frac{1}{\mathbf{u}^{\intercal}\mathbf{u}}\left[\sum_{j=1}^{s}u_j a(\psi_{\sigma_j},\varphi_{\tau_i})\right]^2 \\
&\leq [\lambda_{\min}(\mathbf{K}_{\mathcal{S}})]^{-1}\frac{\|\boldsymbol{\nu}^{N,M}\|_1}{\nu_{\tau_i}^{N}}\sum_{j=1}^{s}a(\psi_{\sigma_j},\varphi_{\tau_i})^2 \leq [\lambda_{\min}(\mathbf{K}_{\mathcal{S}})]^{-1}\, s\,\|\boldsymbol{\nu}^{N,M}\|_1.
\end{aligned} \tag{21}$$

The first line follows from (12) and property (14). The last line exploits Cauchy-Schwarz inequality combined with definition (4) of $\mathbf{p}$, and Hypothesis 3.

Now, we compute the probability of failure of satisfying (18), i.e.,

$$\mathbb{P}\left\{\inf_{\mathbf{u}\in\mathbb{R}^s}\sup_{\mathbf{v}\in\mathbb{R}^m}\frac{\mathbf{v}^\intercal\mathbf{D}\mathbf{A}_{\mathcal{S}}\mathbf{u}}{\|\mathbf{K}_{\mathcal{S}}^{\frac{1}{2}}\mathbf{u}\|_2\|\mathbf{v}\|_2}\leq\widetilde{\alpha}\right\}=\mathbb{P}\left\{\lambda_{\min}(\overline{\mathbf{X}})\leq(1-\delta_m)(1-\delta_M)\alpha^2\right\}$$
$$\leq\mathbb{P}\{\lambda_{\min}(\overline{\mathbf{X}})\leq(1-\delta_m)\lambda_{\min}(\mathbb{E}[\mathbf{X}^{\tau_i}])\}\leq s\exp\left(-\frac{m\rho_{\delta_m}\lambda_{\min}(\mathbb{E}[\mathbf{X}^{\tau_i}])}{s\|\boldsymbol{\nu}^{N,M}\|_1[\lambda_{\min}(\mathbf{K}_{\mathcal{S}})]^{-1}}\right)$$
$$\leq s\exp\left(-\frac{m\rho_{\delta_m}(1-\delta_M)\alpha^2}{s\|\boldsymbol{\nu}^{N,M}\|_1[\lambda_{\min}(\mathbf{K}_{\mathcal{S}})]^{-1}}\right).\tag{22}$$

The first equality relies on (19) and on the definition of $\widetilde{\alpha}$. The first inequality in the second line hinges on (20), while the second inequality is the first matrix Chernoff bound (9), where the uniform estimate (21) has been employed. The final inequality follows from (20).

The thesis is finally proved on estimating that

$$s\exp\left(-\frac{m\rho_{\delta_m}(1-\delta_M)\alpha^2}{s\|\boldsymbol{\nu}^{N,M}\|_1[\lambda_{\min}(\mathbf{K}_{\mathcal{S}})]^{-1}}\right)\leq\varepsilon\iff m\geq\widetilde{C}_{\mathcal{S}}\; s\|\boldsymbol{\nu}^{N,M}\|_1\log(s/\varepsilon),$$

with $\widetilde{C}_{\mathcal{S}}:=[\rho_{\delta_m}(1-\delta_M)\alpha^2\lambda_{\min}(\mathbf{K}_{\mathcal{S}})]^{-1}$.

□

3.3 Uniform restricted inf-sup property

We extend the results in the previous Section to the uniform case, i.e., we aim at proving the RISP over U_s^N, instead of $U_{\mathcal{S}}^N$, for a fixed subset $\mathcal{S}\subseteq[N]$ with $|\mathcal{S}|=s$. For this purpose, we use the non-uniform Theorem 3.8 and a union bound.

First, we introduce the set Σ_s^N of s-sparse vectors of $\mathbb{R}^N$, namely

$$\Sigma_s^N:=\{\mathbf{x}\in\mathbb{R}^N:\|\mathbf{x}\|_0\leq s\}\equiv\bigcup_{\mathcal{S}\subseteq[N];\,|\mathcal{S}|=s}\mathbb{R}^N_{\mathcal{S}}.$$

The following theorem provides a sufficient condition for the uniform RISP to hold.

Theorem 3.9 (Uniform RISP)**.** *Given* $\delta_M\in[0,1)$*, choose* $M\in\mathbb{N}$ *such that the following truncation condition is fulfilled*

$$s\sum_{q>M}\mu_q^N\leq\alpha^2\kappa_s\delta_M,\tag{23}$$

where

$$\kappa_s:=\min_{\mathcal{S}\subseteq[N];\,|\mathcal{S}|=s}\lambda_{\min}(\mathbf{K}_{\mathcal{S}}).\tag{24}$$

Then, for every $0 < \varepsilon < 1$ *and* $\delta_m \in [0,1)$, *provided*

$$m \geq \widetilde{C}_s \, s\|\boldsymbol{\nu}^{N,M}\|_1[s\log(eN/s) + \log(s/\varepsilon)], \tag{25}$$

with

$$\widetilde{C}_s := [\rho_{\delta_m}(1-\delta_M)\alpha^2\kappa_s]^{-1} \tag{26}$$

and ρ_{δ_m} *as in* (10), *the following uniform* s*-sparse RISP holds with probability greater than or equal to* $1-\varepsilon$

$$\inf_{\mathbf{u}\in\Sigma_s^N} \sup_{\mathbf{v}\in\mathbb{R}^m} \frac{\mathbf{v}^\intercal \mathbf{D}\mathbf{A}\mathbf{u}}{\|\mathbf{K}^{\frac{1}{2}}\mathbf{u}\|_2\|\mathbf{v}\|_2} > \widetilde{\alpha} > 0,$$

where $\widetilde{\alpha} := (1-\delta_M)^{\frac{1}{2}}(1-\delta_m)^{\frac{1}{2}}\alpha$.

Proof. First, we define the event where the RISP holds non-uniformly over a single subset $\mathcal{S} \subseteq [N]$ with $|\mathcal{S}| = s$:

$$\Omega_{\mathcal{S}} := \left\{ \omega \in \Omega : \inf_{\mathbf{u}\in\mathbb{R}^s} \sup_{\mathbf{v}\in\mathbb{R}^m} \frac{\mathbf{v}^\intercal \mathbf{D}(\omega)\mathbf{A}_{\mathcal{S}}(\omega)\mathbf{u}}{\|\mathbf{K}_{\mathcal{S}}^{\frac{1}{2}}\mathbf{u}\|_2\|\mathbf{v}\|_2} > \widetilde{\alpha} \right\},$$

where the dependence of $\mathbf{A}_{\mathcal{S}}$ and $\mathbf{D}$ on ω has been highlighted. Analogously, we define the event where the RISP holds uniformly

$$\Omega_s := \left\{ \omega \in \Omega : \inf_{\mathbf{u}\in\Sigma_s^N} \sup_{\mathbf{v}\in\mathbb{R}^m} \frac{\mathbf{v}^\intercal \mathbf{D}(\omega)\mathbf{A}(\omega)\mathbf{u}}{\|\mathbf{K}^{\frac{1}{2}}\mathbf{u}\|_2\|\mathbf{v}\|_2} > \widetilde{\alpha} \right\}. \tag{27}$$

In particular, the following relation holds

$$\Omega_s = \bigcap_{\mathcal{S}\subseteq[N];\,|\mathcal{S}|=s} \Omega_{\mathcal{S}},$$

and, thanks to the subadditivity of $\mathbb{P}$ and De Morgan's laws, we have

$$\mathbb{P}(\Omega_s^c) = \mathbb{P}\left(\left(\bigcap \Omega_{\mathcal{S}}\right)^c\right) = \mathbb{P}\left(\bigcup \Omega_{\mathcal{S}}^c\right) \leq \sum_{\mathcal{S}\subseteq[N];\,|\mathcal{S}|=s} \mathbb{P}(\Omega_{\mathcal{S}}^c), \tag{28}$$

where the superindex c denotes the complement of a set. Now, the non-uniform inequality (22) and the definition (24) of κ_s, yield the following uniform upper bound

$$\mathbb{P}(\Omega_{\mathcal{S}}^c) \leq s\exp\left(-\frac{m\rho_{\delta_m}(1-\delta_M)\alpha^2}{s\|\boldsymbol{\nu}^{N,M}\|_1[\lambda_{\min}(\mathbf{K}_{\mathcal{S}})]^{-1}}\right) \leq s\exp\left(-\frac{m\rho_{\delta_m}(1-\delta_M)\alpha^2}{s\|\boldsymbol{\nu}^{N,M}\|_1\kappa_s^{-1}}\right). \tag{29}$$

Moreover, Stirling's formula furnishes the following upper bound

$$|\{\mathcal{S}\subseteq[N] : |\mathcal{S}| = s\}| = \binom{N}{s} = \frac{N!}{s!(N-s)!} \leq \frac{N^s}{s!} \leq \left(\frac{eN}{s}\right)^s. \tag{30}$$

Combining (28), (29) and (30), we finally obtain the uniform estimate

$$\mathbb{P}(\Omega_s^c) \leq \left(\frac{eN}{s}\right)^s s \exp\left(-\frac{m\rho_{\delta_m}(1-\delta_M)\alpha^2}{s\|\boldsymbol{\nu}^{N,M}\|_1 \kappa_s^{-1}}\right). \tag{31}$$

Simple algebraic manipulations show that the right hand-side of (31) is less than or equal to ε if and only if relation (25) holds. □

We note that the sufficient condition (25) is, in general, too pessimistic. Indeed, in the classical literature on compressed sensing, e.g., [16, 7], the optimal asymptotically dependence of m on s is linear. Likely, this lack of optimality is due to the union bound, that is a very rough estimate. We expect that it is possible to achieve the optimal behavior by using more advanced techniques, such as those described in [18, Chapter 12] and [27] in the case of Bounded Orthonormal Systems. This will be investigated in the future.

3.4 Recovery error analysis

In this section, we deal with the analysis of the recovery error associated with the CORSING procedure, computed with respect to the trial norm $\|\cdot\|_U$, i.e., the quantity $\|\widehat{u}-u\|_U$. Notice that this error is a random variable, depending on the extracted indices $\tau_1, \ldots, \tau_m$. Our aim is to compare the recovery error with the best s-term approximation error of the exact solution u in U^N, i.e., the quantity $\|u^s - u\|_U$, where

$$u^s := \arg\min_{w \in U_s^N} \|w - u\|_U. \tag{32}$$

Due to the s-sparsity constraint in the recovery procedure (7), u^s is the best result that CORSING can ideally provide.[2]

For this purpose, we show that the uniform $2s$-sparse RISP implies a recovery result, depending on a random preconditioned residual (Lemma 3.10), whose second moment is controlled by the square of the best s-term approximation error (Lemma 3.11). Afterwards, in Theorem 3.13, we prove that the best s-term approximation error dominates the first moment of the error associated with a truncated version of the CORSING solution and, finally, we provide a recovery error estimate that holds with high probability in Theorem 3.14.

In the following, a key quantity is the preconditioned random residual

$$\mathcal{R}(v) := \left[\frac{1}{m}\sum_{i=1}^{m} \frac{1}{p_{\tau_i}}[a(v,\varphi_{\tau_i}) - \mathcal{F}(\varphi_{\tau_i})]^2\right]^{\frac{1}{2}}, \quad \forall v \in U. \tag{33}$$

Now, we prove the two lemmas.

[2]The quantity in (32) is actually a minimum and not an infimum, since the function $w \mapsto \|w - u\|_U$ is convex and U_s^N is a finite union of linear subspaces.

Lemma 3.10. *If the uniform $2s$-sparse RISP*

$$\inf_{\mathbf{u}\in\Sigma_{2s}^N} \sup_{\mathbf{v}\in\mathbb{R}^m} \frac{\mathbf{v}^\intercal \mathbf{D}\mathbf{A}\mathbf{u}}{\|\mathbf{K}^{\frac{1}{2}}\mathbf{u}\|_2\|\mathbf{v}\|_2} > \widetilde{\alpha} > 0, \tag{34}$$

holds, then the CORSING *procedure computes a solution $\widehat{u}$ such that*

$$\|\widehat{u} - u^s\|_U < \frac{2}{\widetilde{\alpha}}\mathcal{R}(u^s).$$

Proof. Define $\widehat{\mathbf{u}} := \Psi^*\widehat{u}$ and $\mathbf{u}^s := \Psi^* u^s$. Then, casting (27) in Ω_{2s}, since $\widehat{\mathbf{u}} - \mathbf{u}^s$ is at most $2s$-sparse and thanks to the RISP property (34), and the definition of operator norm, we have

$$\|\widehat{u} - u^s\|_U = \|\mathbf{K}^{\frac{1}{2}}(\widehat{\mathbf{u}} - \mathbf{u}^s)\|_2 < \frac{1}{\widetilde{\alpha}} \sup_{\mathbf{v}\in\mathbb{R}^m} \frac{\mathbf{v}^\intercal\mathbf{D}\mathbf{A}(\widehat{\mathbf{u}} - \mathbf{u}^s)}{\|\mathbf{v}\|_2} = \frac{1}{\widetilde{\alpha}}\|\mathbf{D}\mathbf{A}(\widehat{\mathbf{u}} - \mathbf{u}^s)\|_2.$$

Moreover, the last norm can be bounded as

$$\begin{aligned}
\|\mathbf{D}\mathbf{A}(\widehat{\mathbf{u}} - \mathbf{u}^s)\|_2^2 &= \frac{1}{m}\sum_{i=1}^m \frac{1}{p_{\tau_i}} a(\widehat{u} - u^s, \varphi_{\tau_i})^2 \\
&= \frac{1}{m}\sum_{i=1}^m \frac{1}{p_{\tau_i}} [a(\widehat{u}, \varphi_{\tau_i}) - \mathcal{F}(\varphi_{\tau_i}) - a(u^s, \varphi_{\tau_i}) + \mathcal{F}(\varphi_{\tau_i})]^2 \\
&\leq \frac{2}{m}\sum_{i=1}^m \frac{1}{p_{\tau_i}} \{[a(\widehat{u}, \varphi_{\tau_i}) - \mathcal{F}(\varphi_{\tau_i})]^2 + [a(u^s, \varphi_{\tau_i}) - \mathcal{F}(\varphi_{\tau_i})]^2\} \\
&\leq \frac{4}{m}\sum_{i=1}^m \frac{1}{p_{\tau_i}} [a(u^s, \varphi_{\tau_i}) - \mathcal{F}(\varphi_{\tau_i})]^2 = 4\mathcal{R}(u^s)^2,
\end{aligned}$$

where the last inequality exploits the optimality of $\widehat{u}$. $\square$

Lemma 3.11. *The following upper bound holds*

$$\mathbb{E}[\mathcal{R}(u^s)^2] \leq \beta^2\|u^s - u\|_U^2, \tag{35}$$

where β is the continuity constant of $a(\cdot,\cdot)$ defined in (3).

Proof. Thanks to (1), the residual (33) becomes

$$\mathcal{R}(u^s)^2 = \frac{1}{m}\sum_{i=1}^m p_{\tau_i}^{-1} a(u^s - u, \varphi_{\tau_i})^2,$$

Thus, in expectation, we obtain

$$\mathbb{E}[\mathcal{R}(u^s)^2] = \frac{1}{m}\sum_{i=1}^m \mathbb{E}[p_{\tau_i}^{-1} a(u^s - u, \varphi_{\tau_i})^2]. \tag{36}$$

Each term in the last summation can be bounded as

$$\mathbb{E}[p_{\tau_i}^{-1} a(u^s - u, \varphi_{\tau_i})^2] = \sum_{q=1}^{M} p_q^{-1} a(u^s - u, \varphi_q)^2 p_q \leq \sum_{q=1}^{\infty} a(u^s - u, \varphi_q)^2. \quad (37)$$

Now, exploiting Hypothesis 1, we have

$$\begin{aligned} \|a(u^s - u, \cdot)\|_{V^*} &= \sup_{v \in V} \frac{|a(u^s - u, v)|}{\|v\|_V} \\ &= \sup_{\mathbf{v} \in \ell^2} \frac{|\sum_{q=1}^{\infty} v_q a(u^s - u, \varphi_q)|}{\|\mathbf{v}\|_2} = \left[\sum_{q=1}^{\infty} a(u^s - u, \varphi_q)^2 \right]^{\frac{1}{2}}. \end{aligned}$$

Plugging this equality and (37) in (36), and thanks to (3), we have

$$\mathbb{E}[\mathcal{R}(u^s)^2] \leq \|a(u^s - u, \cdot)\|_{V^*}^2 \leq \beta^2 \|u^s - u\|_U^2.$$

□

If an upper bound of the form $\|u\|_U \leq \mathcal{K}$ is known, a near-optimal recovery result holds in expectation for a truncation of the CORSING solution. This truncation is obtained through the operator $\mathcal{T}_{\mathcal{K}} : U \to U$ defined as

$$\mathcal{T}_{\mathcal{K}} w := \begin{cases} w & \text{if } \|w\|_U \leq \mathcal{K}, \\ \mathcal{K} w / \|w\|_U & \text{if } \|w\|_U > \mathcal{K}, \end{cases} \qquad \forall w \in U. \quad (38)$$

Using (1) and (2), a possible choice of $\mathcal{K}$ is $\|\mathcal{F}\|_{V^*}/\alpha$.

Then, we have the following lemma whose proof is straightforward.

Lemma 3.12. *$\mathcal{T}_{\mathcal{K}}$ is 1-Lipschitz, with respect to $\|\cdot\|_U$, for every $\mathcal{K} > 0$.*

Employing an argument similar to that used in [12, 10], we show an upper bound to the error associated with the truncated CORSING solution.

Theorem 3.13 (Error estimate in expectation)**.** *Let $\mathcal{K} > 0$ be such that $\|u\|_U \leq \mathcal{K}$. Given $\delta_M \in [0, 1)$, choose $M \in \mathbb{N}$ such that the truncation condition (23) is fulfilled and fix $\delta_m \in [0, 1)$.*

Then, for every $0 < \varepsilon < 1$, provided

$$m \geq 2\, \widetilde{C}_{2s}\, s \|\boldsymbol{\nu}^{N,M}\|_1 [2s \log(eN/(2s)) + \log(2s/\varepsilon)], \quad (39)$$

with $\widetilde{C}_{2s}$ defined analogously to (26) and $\widetilde{\alpha} = (1 - \delta_M)^{\frac{1}{2}} (1 - \delta_m)^{\frac{1}{2}} \alpha$, the truncated CORSING solution $\mathcal{T}_{\mathcal{K}} \widehat{u}$ fulfills

$$\mathbb{E}[\|\mathcal{T}_{\mathcal{K}} \widehat{u} - u\|_U] < \left(1 + \frac{2\beta}{\widetilde{\alpha}}\right) \|u^s - u\|_U + 2\mathcal{K}\varepsilon,$$

where β is the continuity constant of $a(\cdot, \cdot)$ defined in (3).

Proof. First, recalling the definition (27) of the event Ω_s, and considering the partitioning $\Omega = \Omega_{2s} \cup \Omega_{2s}^c$, we have the splitting

$$\mathbb{E}[\|\mathcal{T}_\mathcal{K}\widehat{u} - u\|_U] = \int_{\Omega_{2s}} \|\mathcal{T}_\mathcal{K}(\widehat{u} - u)\|_U \,\mathrm{d}\mathbb{P} + \int_{\Omega_{2s}^c} \|\mathcal{T}_\mathcal{K}\widehat{u} - u\|_U \,\mathrm{d}\mathbb{P}.$$

Then, the second term is easily bounded as

$$\int_{\Omega_{2s}^c} \|\mathcal{T}_\mathcal{K}\widehat{u} - u\|_U \,\mathrm{d}\mathbb{P} \leq 2\mathcal{K}\varepsilon.$$

Indeed, thanks to the adopted choice of m, Theorem 3.9 guarantees $\mathbb{P}(\Omega_{2s}^c) \leq \varepsilon$. Moreover, $\|\mathcal{T}_\mathcal{K}\widehat{u} - u\|_U \leq 2\mathcal{K}$, since both $\|\mathcal{T}_\mathcal{K}\widehat{u}\|_U$ and $\|u\|_U$ are less than or equal to $\mathcal{K}$.

Now, employing Lemma 3.12 and the triangle inequality, we have

$$\int_{\Omega_{2s}} \|\mathcal{T}_\mathcal{K}(\widehat{u}-u)\|_U \,\mathrm{d}\mathbb{P} \leq \int_{\Omega_{2s}} \|\widehat{u}-u\|_U \,\mathrm{d}\mathbb{P} \leq \int_{\Omega_{2s}} \|\widehat{u}-u^s\|_U \,\mathrm{d}\mathbb{P} + \int_{\Omega_{2s}} \|u^s-u\|_U \,\mathrm{d}\mathbb{P}.$$

The second integral on the right hand side is less than or equal to the best s-term approximation error $\|u^s - u\|_U$. In order to bound the first integral, we apply Lemmas 3.10 and 3.11, obtaining

$$\int_{\Omega_{2s}} \|\widehat{u} - u^s\|_U \,\mathrm{d}\mathbb{P} < \frac{2}{\widetilde{\alpha}} \int_{\Omega_{2s}} \mathcal{R}(u^s) \,\mathrm{d}\mathbb{P} \leq \frac{2}{\widetilde{\alpha}} \mathbb{E}[\mathcal{R}(u^s)] \leq \frac{2\beta}{\widetilde{\alpha}} \|u^s - u\|_U,$$

where the last relation follows on applying Jensen's inequality to (35). Notice that Lemma 3.10 can be employed since the $2s$-sparse RISP holds on the restricted domain Ω_{2s}. Combining all the inequalities yields the thesis. □

Finally, we provide a recovery estimate in probability. This is asymptotically optimal, but the constant grows like the inverse of the square root of the probability of failure.

Theorem 3.14 (Error estimate in probability)**.** *Given $\delta_M \in [0, 1)$, choose $M \in \mathbb{N}$ such that the truncation condition* (23) *is fulfilled. Then, for every $0 < \varepsilon < 1$ and $\delta_m \in [0, 1)$, provided*

$$m \geq 2\widetilde{C}_{2s}\, s\|\boldsymbol{\nu}^{N,M}\|_1[2s \log(eN/(2s)) + \log(2s/\varepsilon)],$$

with $\widetilde{C}_{2s}$ defined analogously to (26)*, with probability greater than or equal to $1 - 2\varepsilon$, the* **CORSING** *procedure computes a solution $\widehat{u}$ such that*

$$\|\widehat{u} - u\|_U < \left(1 + \frac{2\beta}{\widetilde{\alpha}\sqrt{\varepsilon}}\right) \|u^s - u\|_U$$

where $\widetilde{\alpha} := (1 - \delta_M)^{\frac{1}{2}}(1 - \delta_m)^{\frac{1}{2}}\alpha$ and β is the continuity constant of $a(\cdot, \cdot)$ defined in (3).

Proof. Define $e_s := \|u^s - u\|_U$ and the random variables $Z := \|\widehat{u} - u\|_U$ and $Z_s := \|\widehat{u} - u^s\|_U$. Moreover, consider the quantity

$$b_s := \left(1 + \frac{2\beta}{\widetilde{\alpha}\sqrt{\varepsilon}}\right) e_s. \tag{40}$$

The goal is to show that $\mathbb{P}\{Z \geq b_s\} \leq 2\varepsilon$. The triangle inequality implies $Z \leq Z_s + e_s$. Thus,

$$\mathbb{P}\{Z \geq b_s\} \leq \mathbb{P}\{Z_s \geq b_s - e_s\}.$$

Moreover, defining the event Ω_{2s} according to (27) and denoting by I_A the indicator function of a generic set A, we have

$$\begin{aligned}\mathbb{P}\{Z_s \geq b_s - e_s\} = \mathbb{E}[I_{\{Z_s \geq b_s - e_s\}}] &= \int_{\Omega_{2s}} I_{\{Z_s \geq b_s - e_s\}} \,\mathrm{d}\mathbb{P} + \int_{\Omega_{2s}^c} I_{\{Z_s \geq b_s - e_s\}} \,\mathrm{d}\mathbb{P} \\ &\leq \int_{\Omega_{2s}} I_{\{Z_s \geq b_s - e_s\}} \,\mathrm{d}\mathbb{P} + \mathbb{P}\{\Omega_{2s}^c\}.\end{aligned}$$

Theorem 3.9 implies $\mathbb{P}\{\Omega_{2s}^c\} \leq \varepsilon$. Moreover, employing Lemmas 3.10 and 3.11, we can bound the first integral as

$$\begin{aligned}\int_{\Omega_{2s}} I_{\{Z_s \geq b_s - e_s\}} \,\mathrm{d}\mathbb{P} &\leq \int_{\Omega_{2s}} I_{\{(2/\widetilde{\alpha})\mathcal{R}(u^s) > b_s - e_s\}} \,\mathrm{d}\mathbb{P} \\ &< \mathbb{E}\left[\frac{4\mathcal{R}(u^s)^2}{\widetilde{\alpha}^2 (b_s - e_s)^2}\right] \leq \frac{4\beta^2 e_s^2}{\widetilde{\alpha}^2 (b_s - e_s)^2} = \varepsilon,\end{aligned}$$

where the last equality follows from (40).

□

We conclude this section with a useful corollary dealing with a particular truncation condition. In practice, this corollary provides sufficient conditions for Theorem 3.13 to hold. We will apply this result to some examples in Section 4.

Corollary 3.15. *Suppose that there exist two positive constants $C_{\boldsymbol{\mu}}$ and γ_M such that*

$$\sum_{q > M} \mu_q^N \leq C_{\boldsymbol{\mu}} \left(\frac{N}{M}\right)^{1/\gamma_M}, \quad \forall M \in \mathbb{N}. \tag{41}$$

Then, for every $\varepsilon \in (0, 2^{-1/3}]$ and for $s \leq 2N/e$ there exist two positive constants C_M and C_m such that, for

$$M \geq C_M s^{\gamma_M} N \quad \text{and} \quad m \geq C_m s \|\boldsymbol{\nu}^{N,M}\|_1 [s \log(N/s) + \log(s/\varepsilon)], \tag{42}$$

the CORSING solution $\widehat{u}$ fulfills

$$\mathbb{E}[\|\mathcal{T}_{\mathcal{K}}\widehat{u} - u\|_U] < \left(1 + \frac{4\beta}{\alpha}\right)\|u^s - u\|_U + 2\mathcal{K}\varepsilon,$$

for every $\mathcal{K} > 0$ such that $\|u\|_U \leq \mathcal{K}$, with $\mathcal{T}_{\mathcal{K}}$ defined as in (38) *and where α and β are defined by* (2) *and* (3)*, respectively. In particular, two possible upper bounds for the constants C_M and C_m are*

$$C_M \leq \left(\frac{2C_{\boldsymbol{\mu}}}{\kappa_s\alpha^2}\right)^{\gamma_M} \quad \text{and} \quad C_m \leq \frac{105}{\alpha^2},$$

respectively, with κ_s defined in (24).

Proof. The idea is to choose $\delta_m = \delta_M = 1/2$ and, as anticipated, to apply Theorem 3.13. First, notice that assumption (41) is consistent with Hypothesis 2, on passing to the limit for $M \to +\infty$. In view of Theorem 3.13, we show that the second inequality in (42) implies (39) with a suitable choice of C_m. Moreover, the truncation condition (23), on which Theorem 3.13 relies on, is implied by

$$sC_{\boldsymbol{\mu}}\left(\frac{N}{M}\right)^{1/\gamma_M} \leq \frac{\alpha^2\kappa_s}{2},$$

that, in turn, is equivalent to

$$M \geq \left(\frac{2C_{\boldsymbol{\mu}}}{\kappa_s\alpha^2}\right)^{\gamma_M} s^{\gamma_M}N.$$

Moreover, thanks to the assumptions on ε and s, we have

$$\begin{aligned}
\varepsilon \leq 2^{-1/3} &\Longrightarrow \log(2s/\varepsilon) \leq 4\log(s/\varepsilon),\\
s \leq 2N/e &\Longrightarrow \log(eN/(2s)) \leq 2\log(N/s).
\end{aligned}$$

Thus, recalling the right-hand side of (39), we have

$$\begin{aligned}
2\,\widetilde{C}_{2s}\, s\|\boldsymbol{\nu}^{N,M}\|_1[2s\log(eN/(2s)) + \log(2s/\varepsilon)]&\\
\leq 8\,\widetilde{C}_{2s}\, s\|\boldsymbol{\nu}^{N,M}\|_1[s\log(N/s) + \log(s/\varepsilon)]&,
\end{aligned}$$

where $\widetilde{C}_{2s}$ is defined analogously to (26). In particular, if C_m in (42) is chosen such that

$$C_m \leq 8\,\widetilde{C}_{2s} = \frac{32}{(1-\log 2)\alpha^2} \leq \frac{105}{\alpha^2},$$

then (39) holds. Moreover, relation $\widetilde{\alpha} = (1-\delta_M)^{\frac{1}{2}}(1-\delta_m)^{\frac{1}{2}}\alpha$ yields $\widetilde{\alpha} = \frac{1}{2}\alpha$, so that the quantity $2\beta/\widetilde{\alpha}$ in Theorem 3.13 can be replaced by $4\beta/\alpha$. □

Remark 3.16. The assumptions $\varepsilon \leq 2^{-1/3} \approx 0.79$ and $s \leq 2N/e \approx 0.74N$ made in Corollary 3.15 are quite weak and they are chosen in such a way that the upper bounds to C_M and C_m are easy to derive. Of course, more restrictive hypotheses on ε and s would give sharper upper bounds for the asymptotic constants. Moreover, the parameters δ_M and δ_m could be chosen differently from $\delta_m = \delta_M = 1/2$ and this would lead to different values for the constant in the recovery error estimate.

Remark 3.17. If $\varepsilon \geq s^{s+1}/N^s$, then $s\log(N/s) + \log(s/\varepsilon) \leq 2s\log(N/s)$ and the term $\log(s/\varepsilon)$ disappears from the inequality on m by doubling the constant C_m, giving the trend

$$m \geq C_m \|\boldsymbol{\nu}^{N,M}\|_1 s^2 \log(N/s),$$

claimed in Algorithm 2.1. This assumption on ε is not restrictive, since $s \ll N$ guarantees $s^{s+1}/N^s \ll 1$.

Remark 3.18. A result analogous to Corollary 3.15 holds in probability by resorting to Theorem 3.14 instead of Theorem 3.13 in the proof.

4 Application to a 1D advection-diffusion-reaction equation

In this section, we apply the general theory presented in Section 3 to elliptic one-dimensional problems, such as the Poisson equation and an advection-diffusion-reaction (ADR) equation.

We adopt Corollary 3.15 as the main tool. In particular, we provide estimates for α, β, κ_s, $C_{\boldsymbol{\mu}}$, γ_M, $\boldsymbol{\nu}^N$ and $\|\boldsymbol{\nu}^{N,M}\|_1$, and then deduce suitable hypotheses on m and M such that the CORSING method recovers the best s-term approximation u^s to u. All the recovery results of the section are given in expectation, but they can be easily converted in probability (see Remark 3.18).

Let us first fix the notation. Consider $\Omega = (0,1)$, $U = V = H^1_0(\Omega)$ and

$$(u,v)_U = (u,v)_V = \int_\Omega u'(x)v'(x)\,\mathrm{d}x,$$

resulting in $\|\cdot\|_U = \|\cdot\|_V = |\cdot|_{H^1(\Omega)}$, the $H^1(\Omega)$-seminorm. Moreover, we introduce two Hilbert bases of $H^1_0(\Omega)$. The first one is the hierarchical multiscale basis [33, 13], defined as

$$\mathcal{H}_{\ell,k}(x) := 2^{-\frac{\ell}{2}}\mathcal{H}(2^\ell x - k), \quad \forall x \in [0,1],$$

for every $\ell \in \mathbb{N}_0$, $k = 0, \ldots, 2^\ell - 1$ and with $\mathcal{H}(x) := \max(0, \frac{1}{2} - |x - \frac{1}{2}|)$, for any $x \in [0,1]$, ordered according to the lexicographic mapping

$$j \mapsto (\ell(j), k(j)) := (\lfloor \log_2(j) \rfloor, j - 2^{\lfloor \log_2(j) \rfloor}).$$

The second one is the rescaled sine function basis

$$\mathcal{S}_r(x) := \frac{\sqrt{2}}{r\pi}\sin(r\pi x), \quad \forall x \in [0,1],\ \forall r \in \mathbb{N}.$$

For further details concerning these bases, see [6, Section 5]. It is easy to check that both bases are orthonormal with respect to $(\cdot,\cdot)_U$.

With reference to [6], when the following combination of trial and test functions is employed

$$\psi_j = \mathcal{H}_{\ell(j),k(j)}, \quad \varphi_q = \mathcal{S}_q,$$

we denote the approach by CORSING $\mathcal{HS}$. On the contrary, when the roles of the trial and test functions are switched, we denote it by CORSING $\mathcal{SH}$. In both cases, $\mathcal{HS}$ and $\mathcal{SH}$, we observe that Hypothesis 1 is fulfilled and that $\mathbf{K} = \mathbf{I}$. Thus, in particular, from (24), $\kappa_s = 1$.

As the reference problem, we consider the one-dimensional ADR equation over Ω, with Dirichlet boundary conditions

$$\begin{cases} -u'' + bu' + \eta u = f & \text{in } \Omega \\ u(0) = u(1) = 0, \end{cases} \tag{43}$$

with $b, \eta \in \mathbb{R}$ and $f : (0,1) \to \mathbb{R}$, corresponding to the weak problem

$$\text{find } u \in H_0^1(\Omega): \quad (u', v') + b(u', v) + \eta(u, v) = (f, v), \quad \forall v \in H_0^1(\Omega), \tag{44}$$

where $(\cdot,\cdot)$ denotes the standard inner product in $L^2(\Omega)$.

4.1 The Poisson equation ($\mathcal{HS}$).

First, we deal with the Poisson equation, corresponding to (43) with $b = \eta = 0$, whose weak formulation is

$$\text{find } u \in H_0^1(\Omega): \quad a_\Delta(u, v) = (f, v), \quad \forall v \in H_0^1(\Omega). \tag{45}$$

where $a_\Delta(u, v) := (u', v')$. In such a case, we denote the local a-coherence by $\boldsymbol{\mu}^N_\Delta$. The inf-sup and continuity constants of $a_\Delta(\cdot,\cdot)$ are $\alpha = \beta = 1$.

We can prove the following result for the CORSING $\mathcal{HS}$ procedure applied to (45).

Proposition 4.1. *Fix a maximum hierarchical level $L \in \mathbb{N}$, corresponding to $N = 2^{L+1} - 1$. Then, for every $\varepsilon \in (0, 2^{-1/3}]$ and $s \leq 2N/e$, provided that*

$$M \geq C_M s N, \quad m \geq C_m s \log M[s \log(N/s) + \log(s/\varepsilon)],$$

for suitable constants C_m and C_M, and chosen the upper bound $\boldsymbol{\nu}^N$ as

$$\nu_q^N := \frac{8}{\pi q}, \quad \forall q \in \mathbb{N},$$

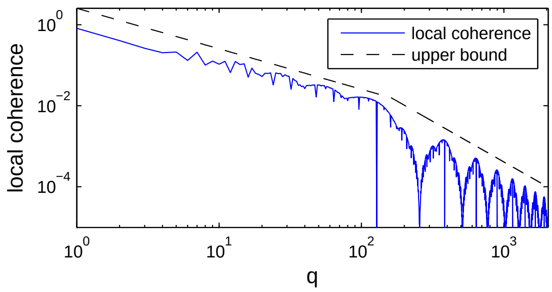

Figure 1: Sharpness of the upper bound (47) with $N = 127$ and $M = 2047$.

the CORSING $\mathcal{HS}$ *solution to* (45) *fulfills*

$$\mathbb{E}[|\mathcal{T}_{\mathcal{K}}\widehat{u} - u|_{H^1(\Omega)}] < 5|u^s - u|_{H^1(\Omega)} + 2\mathcal{K}\varepsilon,$$

for every $\mathcal{K} > 0$ *such that* $|u|_{H^1(\Omega)} \leq \mathcal{K}$, *with* $\mathcal{T}_{\mathcal{K}}$ *defined as in* (38). *In particular, two possible upper bounds for* C_M *and* C_m *are*

$$C_M \leq \frac{80}{3\pi^2} \approx 2.70 \quad \textit{and} \quad C_m \leq \frac{840}{\pi}\left(1 + \frac{1}{\log 3}\right) \approx 511.$$

Proof. An explicit computation yields the exact stiffness matrix entries (the dependence of ℓ and k on j is omitted)

$$a_\Delta(\mathcal{H}_{\ell,k}, \mathcal{S}_q) = \frac{4\sqrt{2}}{\pi}\frac{2^{\frac{\ell}{2}}}{q}\sin\left(\frac{\pi q}{2^\ell}\left(k + \frac{1}{2}\right)\right)\sin^2\left(\frac{\pi}{4}\frac{q}{2^\ell}\right). \tag{46}$$

Using Definition 2.3, employing the inequalities $\sin^2(x) \leq 1$ on the first sine and $\sin^4(x) \leq \min\{1, |x|\}$ on the second sine, for every $x \in \mathbb{R}$, we have

$$|a_\Delta(\mathcal{H}_{\ell,k}, \mathcal{S}_q)|^2 \leq \frac{32}{\pi^2}\frac{2^\ell}{q^2}\sin^4\left(\frac{\pi}{4}\frac{q}{2^\ell}\right) \leq \min\left\{\frac{32}{\pi^2}\frac{2^\ell}{q^2}, \frac{8}{\pi q}\right\},$$

and, thus, we obtain the upper bound

$$\mu_{\Delta,q}^N \leq \min\left\{\frac{32}{\pi^2}\frac{2^L}{q^2}, \frac{8}{\pi q}\right\}. \tag{47}$$

Figure 1 shows that this bound is sharp. Considering the first argument of the minimum in (47), on noticing that $2^L = (N+1)/2$, we obtain

$$\begin{aligned}\sum_{q>M}\mu_{\Delta,q}^N &\leq \frac{32}{\pi^2}\frac{N+1}{2}\sum_{q>M}\frac{1}{q^2} \leq \frac{16}{\pi^2}(N+1)\left[\frac{1}{(M+1)^2} + \int_{M+1}^{\infty}\frac{1}{q^2}\,\mathrm{d}q\right]\\ &= \frac{16}{\pi^2}\frac{N+1}{M+1}\left[\frac{1}{M+1} + 1\right] \leq \frac{20}{\pi^2}\frac{N+1}{M+1} \leq \frac{80}{3\pi^2}\frac{N}{M}.\end{aligned}$$

The fourth and fifth relations hinges on the assumption $L \geq 1$, that implies $N \geq 3$. Consequently, assuming $M \geq N$ we have also $M \geq 3$. This implies $1/(M+1) \leq 1/4$ (fourth relation) and $(N+1)/(M+1) \leq 4N/(3M)$ (fifth relation). Thus, in view of Corollary 3.15, we can pick

$$C_{\boldsymbol{\mu}} = \frac{80}{3\pi^2} \quad \text{and} \quad \gamma_M = 1.$$

Now, to bound $\|\boldsymbol{\nu}^{N,M}\|_1$, which is required by Corollary 3.15, we deal with the second argument of the minimum in (47) and set

$$\nu_q^N := \frac{8}{\pi q}.$$

This choice leads to the estimate

$$\|\boldsymbol{\nu}^{N,M}\|_1 = \frac{8}{\pi}\sum_{q=1}^{M}\frac{1}{q} \leq \frac{8}{\pi}\left[1+\int_1^M \frac{1}{q}\,\mathrm{d}q\right] = \frac{8}{\pi}(1+\log M) \leq \frac{8}{\pi}\left(1+\frac{1}{\log 3}\right)\log M, \tag{48}$$

since $M \geq 3$. Thus, combining the lower bound for m and M in Corollary 3.15 with (48), we conclude the proof. □

Remark 4.2. The upper bound $\sin^4(x) \leq \min\{1, |x|\}$ can be improved as $\sin^4(x) \leq \min\{1, 0.68\,|x|\}$. This change leads to rescale the value of C_m by a factor 0.68, i.e., $C_m \approx 347$.

Remark 4.3. The choice $\nu_q^N = 8/(\pi q)$ is suboptimal. If we choose the sharper upper bound

$$\nu_q^N = \min\left\{\frac{32}{\pi^2}\frac{2^L}{q^2}, \frac{8}{\pi q}\right\},$$

the term $\log M$ in the lower bound to m can be replaced by $\log N$. Indeed, in this case

$$\|\boldsymbol{\nu}^{N,M}\|_1 \lesssim \sum_{q=1}^{N}\frac{1}{q}+N\sum_{q=N+1}^{M}\frac{1}{q^2} \lesssim \log N+N\left(\frac{1}{N}-\frac{1}{M}\right) \lesssim \log N+1-\frac{1}{s} \lesssim \log N.$$

4.2 ADR equation $(\mathcal{HS})$

We consider problem (43) and state the following result.

Proposition 4.4. *Fix a maximum hierarchical level $L \in \mathbb{N}$, corresponding to $N = 2^{L+1}-1$. Then, for every $\varepsilon \in (0, 2^{-1/3}]$ and $s \leq 2N/e$, provided that*

$$M \gtrsim sN, \quad \frac{|b|}{M} \lesssim 1, \quad \frac{|\eta|}{M^2} \lesssim 1,$$

$$m \gtrsim s(\log M + |b|^2 + |\eta|^2)[s\log(N/s) + \log(s/\varepsilon)],$$

and chosen the upper bound $\boldsymbol{\nu}^N$ such that

$$\nu_q^N \sim \frac{1}{q} + \frac{|b|^2}{q^3} + \frac{|\eta|^2}{q^5}, \quad \forall q \in \mathbb{N},$$

the CORSING $\mathcal{HS}$ solution to (44), with $\eta > -2$, fulfills

$$\mathbb{E}[|\mathcal{T}_{\mathcal{K}}\widehat{u} - u|_{H^1(\Omega)}] < \left(1 + \frac{4 + 2\sqrt{2}|b| + 2|\eta|}{1 + \min(0, \eta/2)}\right) |u^s - u|_{H^1(\Omega)} + 2\mathcal{K}\varepsilon,$$

for every $\mathcal{K} > 0$ such that $|u|_{H^1(\Omega)} \leq \mathcal{K}$, with $\mathcal{T}_{\mathcal{K}}$ defined as in (38).

Proof. The argument is the same as in Proposition 4.1, thus we will just highlight the different parts. The precise values of the asymptotic constants will not be tracked during the proof.

First, a straightforward computation gives

$$a(\mathcal{H}_{\ell,k}, \mathcal{S}_q) = \frac{4\sqrt{2}}{\pi} \frac{2^{\frac{\ell}{2}}}{q} \sin^2\left(\frac{\pi}{4}\frac{q}{2^\ell}\right) \left[\left(1 + \frac{\eta}{(\pi q)^2}\right) \sin\left(\frac{\pi q}{2^\ell}(k + \tfrac{1}{2})\right) \right.$$
$$\left. - \frac{b}{\pi q} \cos\left(\frac{\pi q}{2^\ell}(k + \tfrac{1}{2})\right)\right].$$

Hence, using the same upper bounds as in Proposistion 4.1, we obtain

$$|a(\mathcal{H}_{\ell,k}, \mathcal{S}_q)|^2 \lesssim \min\left\{\frac{2^\ell}{q^2}, \frac{1}{q}\right\} \left(1 + \frac{|b|^2}{q^2} + \frac{|\eta|^2}{q^4}\right),$$

and, consequently,

$$\mu_q^N \lesssim \min\left\{\frac{N}{q^2}, \frac{1}{q}\right\} \left(1 + \frac{|b|^2}{q^2} + \frac{|\eta|^2}{q^4}\right). \tag{49}$$

Considering the first argument of the minimum in (49), yields

$$\sum_{q>M} \mu_q^N \lesssim N\left[\sum_{q>M} \frac{1}{q^2} + |b|^2 \sum_{q>M} \frac{1}{q^4} + |\eta|^2 \sum_{q>M} \frac{1}{q^6}\right]$$
$$\lesssim N\left[\frac{1}{M} + \frac{|b|^2}{M^3} + \frac{|\eta|^2}{M^5}\right] \lesssim \frac{N}{M}.$$

The second inequality hinges on estimates of the sums by suitable integrals, whereas the third one is implied by the hypotheses $|b|/M \lesssim 1$ and $|\eta|/M^2 \lesssim 1$.

Now, considering the second argument of the minimum in (49), we have the upper bound

$$\nu_q^N \sim \frac{1}{q} + \frac{|b|^2}{q^3} + \frac{|\eta|^2}{q^5}, \quad \forall q \in \mathbb{N},$$

and, consequently, the ℓ^1-norm of its truncation fulfills

$$\|\boldsymbol{\nu}^{N,M}\|_1 \sim \sum_{q=1}^{M} \frac{1}{q} + \sum_{q=1}^{M} \frac{|b|^2}{q^3} + \sum_{q=1}^{M} \frac{|\eta|^2}{q^5} \lesssim \log M + |b|^2 + |\eta|^2.$$

Finally, we notice that (2) and (3) hold with

$$\alpha = 1 + \min\left(0, \frac{\eta}{2}\right), \quad \beta = 1 + \frac{|b|}{\sqrt{2}} + \frac{|\eta|}{2},$$

thanks to the Poincaré inequality

$$\sqrt{2}\|v\|_{L^2(\Omega)} \leq |v|_{H^1(\Omega)}, \quad \forall v \in H_0^1(\Omega).$$

The thesis is now a direct consequence of Corollary 3.15. □

4.3 The Poisson equation ($\mathcal{SH}$)

We prove a recovery result for the CORSING $\mathcal{SH}$ method applied to the Poisson problem (45).

Proposition 4.5. *For every $\varepsilon \in (0, 2^{-1/3}]$ and $s \leq 2N/e$, there exist two positive constants C_m and C_M such that, provided*

$$M \geq C_M \sqrt{s} N, \quad m \geq C_m s \log(M)[s \log(N/s) + \log(s/\varepsilon)],$$

with M of the form $M = 2^{L+1} - 1$ for some $L \in \mathbb{N}$, and chosen the upper bound $\boldsymbol{\nu}^N$ as

$$\nu_q^N = \frac{1}{2^{\ell(q)-1}}, \quad \forall q \in \mathbb{N},$$

the CORSING $\mathcal{SH}$ solution to (45) fulfills

$$\mathbb{E}[|\mathcal{T}_{\mathcal{K}} \widehat{u} - u|_{H^1(\Omega)}] \leq 5|u^s - u|_{H^1(\Omega)} + 2\mathcal{K}\varepsilon,$$

for every $\mathcal{K} > 0$ such that $|u|_{H^1(\Omega)} \leq \mathcal{K}$, with $\mathcal{T}_{\mathcal{K}}$ defined as in (38) and where α and β are defined by (2) and (3), respectively. In particular, two possible upper bounds for C_M and C_m are

$$C_M \leq \frac{\pi}{\sqrt{3}} \approx 1.81 \quad \textit{and} \quad C_m \leq \frac{210 \log_2(e) \log(4)}{\log(3)} \approx 382.$$

Proof. The proof is analogous to that of Proposition 4.1. We highlight only the main differences. First, notice that

$$a_\Delta(\mathcal{S}_j, \mathcal{H}_{\ell(q),k(q)}) = a_\Delta(\mathcal{H}_{\ell(q),k(q)}, \mathcal{S}_j).$$

Moving from (46) and employing the inequality $\sin^4(x) \leq \min\{|x|^4, |x|^2\}$, for every $x \in \mathbb{R}$, we obtain

$$\mu_{\Delta,q}^N \leq \min\left\{\frac{\pi^2}{8} \frac{N^2}{2^{3\ell(q)}}, \frac{1}{2^{\ell(q)-1}}\right\}. \tag{50}$$

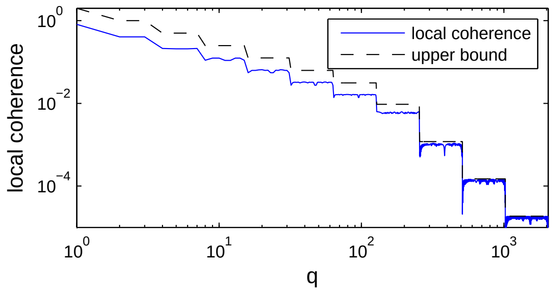

Figure 2: Sharpness of the upper bound (50) with $N = 127$ and $M = 2047$.

Figure 2 shows the sharpness of this bound.

Considering the first argument of the minimum in (50), and since $M = 2^{L+1} - 1$, we have that

$$\sum_{q>M} \mu^N_{\Delta,q} \leq \frac{\pi^2}{8} N^2 \sum_{\ell>L} \sum_{k=0}^{2^\ell - 1} \frac{1}{2^{3\ell}} = \frac{\pi^2}{8} N^2 \sum_{\ell>L} \frac{1}{2^{2\ell}} = \frac{\pi^2}{8} \frac{N^2}{2^{2(L+1)}} \sum_{\ell\geq 0} \frac{1}{2^{2\ell}} \leq \frac{\pi^2}{6} \left(\frac{N}{M}\right)^2$$

where the change of variable $q \mapsto (\ell, k)$ has been used. Thus, if follows that

$$C_{\boldsymbol{\mu}} = \frac{\pi^2}{6} \quad \text{and} \quad \gamma_M = \frac{1}{2}.$$

Now, by considering the second argument of the minimum in (50), we select

$$\nu^N_q := \frac{1}{2^{\ell-1}}$$

and conclude the proof by computing

$$\begin{aligned} \|\boldsymbol{\nu}^{N,M}\|_1 &= \sum_{\ell=0}^{L} \sum_{k=0}^{2^\ell - 1} \frac{1}{2^{\ell-1}} = 2(L+1) = 2\log_2(e)\log(M+1) \\ &\leq 2\log_2(e)\frac{\log(M+1)}{\log(M)}\log(M) \leq \frac{2\log_2(e)\log(4)}{\log(3)}\log(M), \end{aligned}$$

since $M \geq 3$, thanks to $L \geq 1$. □

Remark 4.6. The choice of $\mathbf{p}$ prompted by Proposition 4.5 (i.e., $p_q \sim 2^{-\ell(q)}$) coincides with that in [6], in the R-CORSING $\mathcal{SH}$ case, for the corresponding parameter $\mathbf{w}$, tuned via a trial-and-error procedure.

4.4 ADR equation ($\mathcal{SH}$)

Considerations analogous to those made in the $\mathcal{HS}$ case hold in the advective/reactive case. It suffices to notice that

$$(u', v') + b(u', v) + \eta(u, v) = (v', u') - b(v', u) + \eta(v, u), \quad \forall u, v \in H_0^1(\Omega).$$

Remark 4.7 (Application to more general cases). The main difficulty of the analysis of CORSING is the derivation of the upper bound $\boldsymbol{\nu}^N$ to the local a-coherence. For instance, in dealing with the ADR equation with non-constant coefficients, a highly oscillatory diffusion coefficient can considerably deteriorate $\boldsymbol{\nu}^N$. One possibility to tackle this issue is to expand the non-constant coefficient with respect to a suitable basis and then to exploit Propositions 4.1 and 4.5.

Considering the extension to higher-dimensional problems, first results are provided in [6, Section 6] where CORSING is applied to the ADR equation with constant coefficients, with hierarchical pyramids and tensor product of sine functions. Nevertheless, since the hierarchical pyramids are not orthonormal, they can only be used as trial functions in view of the theoretical setting of this work (Hypothesis 1 does not hold). In such a case, κ_s^{-1} grows at most logarithmically with respect to N [33]. A less trivial task is to provide a sharp upper bound $\boldsymbol{\nu}^N$ due to the involved expression of the stiffness matrix entries.

5 Numerical experiments

We validate the above theoretical results by both a qualitative and a quantitative analysis. For a more complete numerical assessment of CORSING, we refer to [6].

All the computations have been performed using MATLAB® R2013a 64-bit (version 8.1.0.604) on a MacBook Pro equipped with a 3GHz Intel Core i7 processor and 8GB of RAM.

5.1 Sensitivity analysis of the RISP constant

We investigate the sensitivity of $\widetilde{\alpha}$ to the constant C_m on the Poisson problem (45), in the setting $\mathcal{HS}$. We fix the hierarchical level to $L = 14$, corresponding to $N = 32767$. We consider the values $s = 1, 2, 3, 4, 5$ and choose $M = sN$, while selecting m according to one of the following rules

$$\begin{aligned} &\text{Rule 1:} \quad m = \lceil C_m s^2 \log M \log(N/s) \rceil, \\ &\text{Rule 2:} \quad m = \lceil C_m s \log M \log(N/s) \rceil, \\ &\text{Rule 3:} \quad m = \lceil C_m s \log(N/s) \rceil. \end{aligned} \tag{51}$$

Rule 1 is the one derived in this paper, corresponding to $\gamma_m = 2$. Rule 2 is associated with $\gamma_m = 1$, and Rule 3 is the asymptotically optimal lower bound that a general sparse recovery procedure requires to be stable (see [18, Proposition 10.7]). For each choice of M and m, we repeat the following experiment 50 times: first, extract $\tau_1, \ldots, \tau_m \in [M]$ i.i.d. with probability $p_q \sim 1/q$ and build the corresponding matrices $\mathbf{D}$ and $\mathbf{A}$; then, generate 1000 random subsets $\mathcal{S}_1, \ldots, \mathcal{S}_{1000} \subseteq [N]$ of cardinality s and compute the non-uniform RISP constant $\widetilde{\alpha}_{\mathcal{S}_k}$ for every $k \in [1000]$, corresponding to the minimum singular value of $\mathbf{DA}$, using the `svd` command; finally, approximate the uniform RISP constant as

$$\widetilde{\alpha} \approx \min_{k \in [1000]} \widetilde{\alpha}_{\mathcal{S}_k}.$$

We consider the three trends in (51) and $C_m = 2$ or 5. The corresponding six boxplots relative to the 50 different values of $\widetilde{\alpha}$, computed for each s, are shown in Figure 3, where the crosses represent the outliers.

For Rule 1 and 2, $\widetilde{\alpha}$ shows a similar behavior since both trends are approaching the value of the inf-sup constant, $\alpha = 1$, when s grows. We notice that the values computed for Rule 1 are more concentrated around the mean, implying that $\gamma_m = 2$ is a too conservative choice. For Rule 3, $\widetilde{\alpha}$ exhibits the lowest values, though the corresponding boxplots are quite aligned and have similar size, especially for $C_m = 5$, where $\widetilde{\alpha}$ seems to stabilize around the value $\alpha/2$. For $C_m = 2$, $\widetilde{\alpha}$ approaches the value $\alpha/4$, even though the presence of too many outliers suggests that the RISP is not being satisfied for a reasonable value of ε. However, since Rule 3 is quite satisfactory, especially for $C_m = 5$, the quantity $\log M$ does not seem to be really necessary in Rule 2. Moreover, Rule 1 is penalized by both the $\log M$ term and the extra s factor.

5.2 CORSING validation

We test CORSING $\mathcal{HS}$ on the one-dimensional Poisson equation (45), choosing the forcing term so that the exact solution be

$$u(x) := \widetilde{u}_{0.2,0.7,1000}(x) + 0.3 \cdot \widetilde{u}_{0.4,0.4005,2000}(x), \quad \forall x \in [0, 1] \tag{52}$$

with

$$\begin{aligned}
\widetilde{u}_{x_1,x_2,t}(x) &:= \overline{u}_{x_1,x_2,t}(x) - e_{x_1,x_2,t}(x), \\
e_{x_1,x_2,t}(x) &:= x\,\overline{u}_{x_1,x_2,t}(1) + (1 - x)\,\overline{u}_{x_1,x_2,t}(0), \\
\overline{u}_{x_1,x_2,t}(x) &:= \arctan(t(x - x_1)) - \arctan(t(x - x_2)),
\end{aligned}$$

for every $x \in [0, 1]$, $0 \leq x_1 < x_2 \leq 1$ and $t \in \mathbb{R}$. This particular solution is designed so as to exhibit two boundary layers at $x = 0.2$ and $x = 0.7$, and a small spike-shaped detail at $x = 0.4$ (see Figure 4). The hierarchical

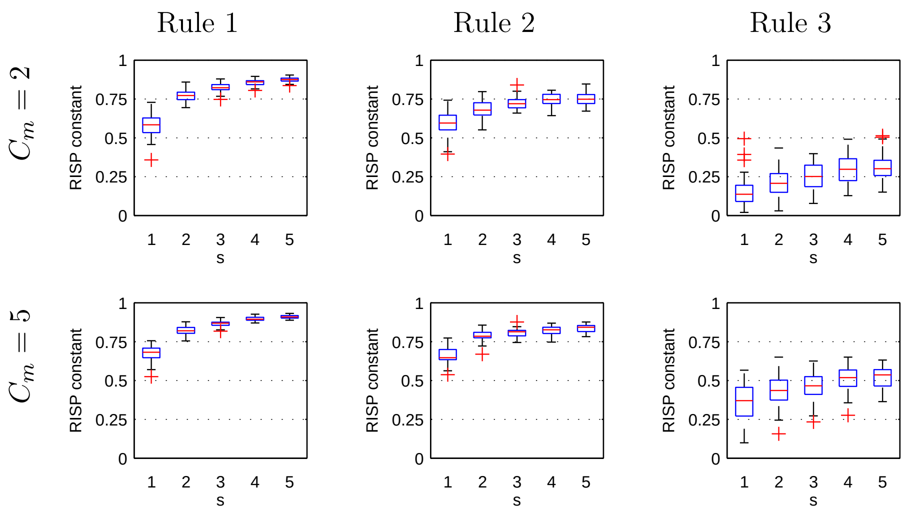

Figure 3: Sensitivity analysis of the RISP constant, with $M = sN$ and m defined according to (51).

multiscale basis is particularly suited to capture these sharp features. We fix $L = 12$, corresponding to $N = 8191$, $s = 50$, $M = sN$ and $m = 1200$.

In Figure 4, we compare u (dashed line) and $\widehat{u}$ (solid line). The exact solution is well recovered. Both boundary layers are correctly captured and also the spike-shaped feature is successfully detected. More quantitatively, the best 50-term relative error is $|u - u_{50}|_{H^1}/|u|_{H^1} \approx 0.092$ and the relative error of the CORSING solution is $|u - \widehat{u}|_{H^1}/|u|_{H^1} \approx 0.111$. Thus, via CORSING, we loose only the 21% of the best possible accuracy.

Figures 5 and 6, highlight that CORSING is able to find the most important coefficients of $\mathbf{u}$. In particular, in Figure 5, the coefficients of $\mathbf{u}$ and $\widehat{\mathbf{u}}$ are plotted according to the lexicographic ordering, whereas in Figure 6 they are shown in two dimensions: level ℓ is the vertical axis, and each level is divided horizontally into 2^ℓ parts, corresponding to $k = 0, \dots, 2^\ell - 1$, (left to right). The color plots refer to $|u_{\ell,k}|$ (left) and $|\widehat{u}_{\ell,k}|$ (right), in logarithmic scale. It is remarkable the capability of CORSING in detecting the localized features of the solution (see the isolated vertical line in Figure 6 (right)).

5.3 Convergence analysis

We now perform a convergence analysis of CORSING $\mathcal{HS}$ applied to (45), showing that the mean error shares the same trend as the best s-term approximation error, as predicted by the theoretical results. In particular, the forcing term f is chosen such that the exact solution be

$$u(x) := C_u(1 - x)(\exp(100x) - 1),$$

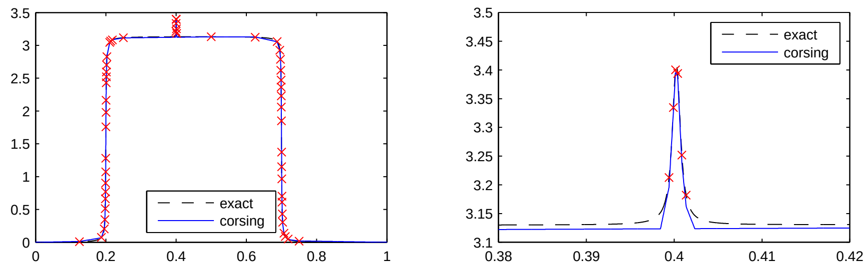

Figure 4: Left: comparison between u defined in (52) (dashed line) and $\widehat{u}$ (solid line). Right: a zoom in on the spike-shaped detail of u. Crosses correspond to the selected trial functions.

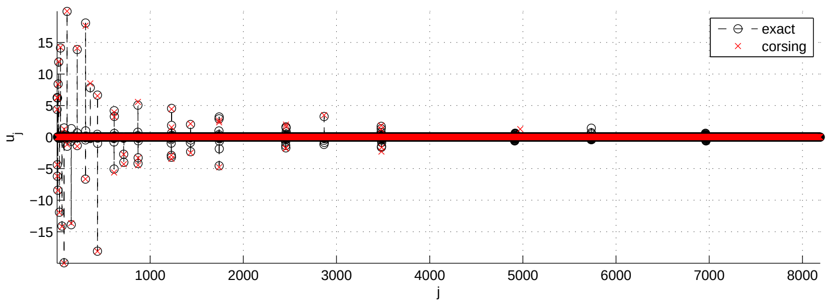

Figure 5: Comparison between $\mathbf{u}$ (circles) and $\widehat{\mathbf{u}}$ (crosses).

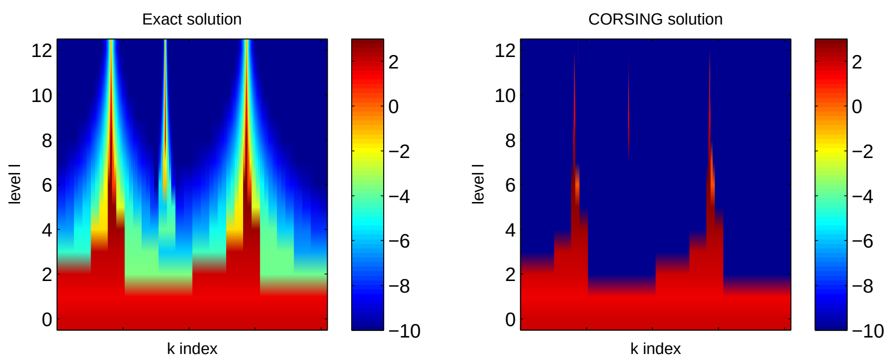

Figure 6: 2D color plot of $|u_{\ell,k}|$ and $|\widehat{u}_{\ell,k}|$ in logarithmic scale.

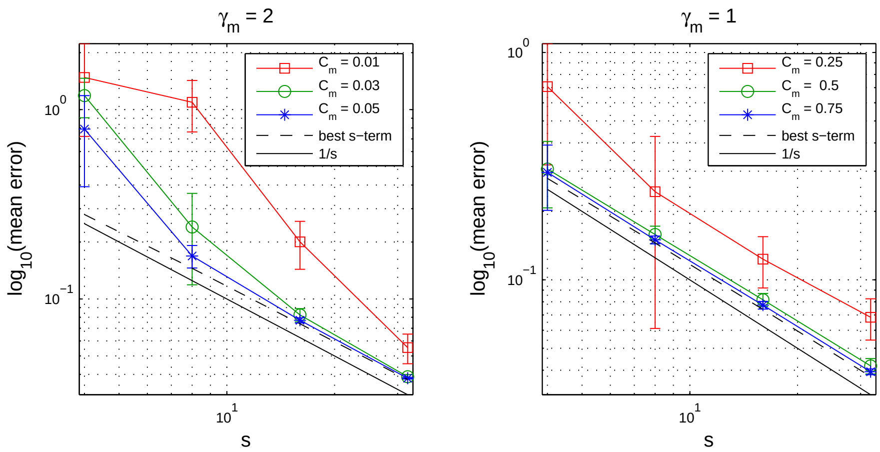

Figure 7: Convergence analysis: mean error $\pm$ standard deviation and best s-term approximation error. Case $\gamma_m = 2$ (right) and $\gamma_m = 1$ (left).

where C_u is chosen such that $|u|_{H^1} = 1$. We take $L = 11$, corresponding to $N = 4095$. For $s = 4, 8, 16, 32$, we define $M = sN$ and $m = \lceil C_m s^{\gamma_m} \log M \log(N/s) \rceil$ for $\gamma_m = 1, 2$, and for different values of C_m. For every combination of γ_m and C_m, we run 100 CORSING experiments and show the mean error obtained $\pm$ the standard deviation, computed using the unbiased estimator. In the case $\gamma_m = 1$, we select $C_m = 0.25, 0.5, 0.75$, whereas for $\gamma_m = 2$, we consider $C_m = 0.01, 0.03, 0.05$. The values of C_m are smaller for $\gamma_m = 2$, in order to ensure that $m < N$ for every s.

The results are shown in Figure 7. The mean error reaches the best s-term approximation rate, that is proportional to $1/s$.

6 Conclusions

We presented a rigorous formalization and provided a theoretical analysis of the CORSING (COmpRessed SolvING) method [6]. Our analysis essentially relies on the concepts of local a-coherence and restricted inf-sup property (RISP). In particular, we showed how suitable hypotheses on the local a-coherence are sufficient to guarantee the RISP. As a consequence, we provided estimates of the CORSING solution with respect to the best s-term approximation error in expectation (Theorem 3.13) and in probability (Theorem 3.14). This general theory has been applied to the case of the one-dimensional ADR equation with constant coefficients, and numerical experiments confirm the theoretical results.

Important issues are still open. For instance, the application of our theoretical results to more general cases, such as one-dimensional ADR equation with non-constant coefficients and the two- or three-dimensional case, is not a trivial extension of the results presented here (see Remark 4.7). Also the case of non-orthonormal test functions is an interesting open problem and the arguments employed here probably need to be substantially modified.

However, this first theoretical analysis of the method highlights the importance of the local a-coherence and the RISP as powerful picklocks, capable to cast the compressed sensing philosophy into the PDEs setting.

7 Acknowledgements

The first author acknowledges the INdAM research group "Gruppo Nazionale per il Calcolo Scientifico" for the economic support. The second author acknowledges the support from the Center for ADvanced MOdeling Science (CADMOS). The work of the third author was supported by the Project MIUR-PRIN 2010/2011 "Data-Centric Genomic Computing" (GenData 2020). The financial support of MIUR (Project "Innovative Methods for Water Resources under Hydro-Climatic Uncertainty Scenarios", PRIN 2010/2011) is gratefully acknowledged by the last author.

References

[1] B. Adcock and A.C. Hansen. Generalized Sampling and Infinite-Dimensional Compressed Sensing. *Found. Comput. Math.*, pages 1–61, 2015.

[2] B. Adcock, A.C. Hansen, C. Poon, and B. Roman. Breaking the coherence barrier: A new theory for compressed sensing. *arXiv preprint arXiv:1302.0561*, 2013.

[3] R. Ahlswede and A. Winter. Strong converse for identification via quantum channels. *IEEE Trans. Inform. Theory*, 48(3):569–579, 2002.

[4] A.K. Aziz and I. Babuška. *The Mathematical Foundations of the Finite Element Method with Applications to Partial Differential Equations*, chapter 1. Academic Press, New York, 1972.

[5] F. Brezzi and M. Fortin. *Mixed and Hybrid Finite Element Methods*, volume 15 of *Springer Ser. Comput. Math.* Springer-Verlag, New York, 1991.

[6] S. Brugiapaglia, S. Micheletti, and S. Perotto. Compressed solving: A numerical approximation technique for elliptic PDEs based on Compressed Sensing. *Comput. Math. Appl.*, 70(6):1306–1335, 2015.

[7] E.J. Candès, J.K. Romberg, and T. Tao. Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information. *IEEE Trans. Inform. Theory*, 52(2):489–509, 2006.

[8] S.S. Chen, D.L. Donoho, and M.A. Saunders. Atomic decomposition by basis pursuit. *SIAM J. Sci. Comput.*, 20:33–61, 1998.

[9] H. Chernoff. A measure of asymptotic efficiency for tests of a hypothesis based on the sums of observations. *Ann. Math. Stat.*, 23:409–507, 1952.

[10] A. Chkifa, A. Cohen, G. Migliorati, F. Nobile, and R. Tempone. Discrete least squares polynomial approximation with random evaluations - application to parametric and stochastic elliptic PDEs. *ESAIM: M2AN*, 49(3):815–837, 2015.

[11] A. Cohen, W. Dahmen, and R.A. DeVore. Orthogonal Matching Pursuit under the Restricted Isometry Property. *arXiv preprint arXiv:1506.04779*, 2015.

[12] A. Cohen, M.A. Davenport, and D. Leviatan. On the stability and accuracy of least squares approximations. *Found. Comput. Math.*, 13(5):819–834, 2013.

[13] W. Dahmen. Wavelet and multiscale methods for operator equations. *Acta Numer.*, 6:55–228, 1997.

[14] W. Dahmen, C. Huang, C. Schwab, and G. Welper. Adaptive Petrov–Galerkin Methods for First Order Transport Equations. *SIAM J. Num. Anal.*, 50(5):2420–2445, 2012.

[15] R.A. DeVore. Nonlinear approximation. *Acta Numer.*, 7:51–150, 1998.

[16] D.L. Donoho. Compressed sensing. *IEEE Trans. Inform. Theory*, 52:1289–1306, 2006.

[17] A. Ern and J.L. Guermond. *Theory and Practice of Finite Elements.* Number v. 159 in Appl. Math. Sci. Springer, 2004.

[18] S. Foucart and H. Rauhut. *A Mathematical Introduction to Compressive Sensing.* Appl. Numer. Harmon. Anal. Springer Science+Business Media, New York, 2013.

[19] F. Krahmer and R. Ward. Stable and robust sampling strategies for compressive imaging. *IEEE Trans. Image Process.*, 23(2):612–622, 2014.

[20] S. Mallat and Z. Zhang. Matching pursuits with time-frequency dictionaries. *IEEE Trans. Signal Process.*, 41(12):3397–3415, 1993.

[21] B.K. Natarajan. Sparse Approximate Solutions to Linear Systems. *SIAM J. Comput.*, 24(2):227–234, 1995.

[22] D. Needell and J.A. Tropp. CoSaMP: Iterative signal recovery from incomplete and inaccurate samples. *Appl. Comput. Harmon. Anal.*, 26(3):301 – 321, 2009.

[23] J. Nečas. Sur une méthode pour résoudre les équations aux dérivées partielles du type elliptique, voisine de la variationnelle. *Ann. Sc. Norm. Super. Pisa Cl. Sci.*, 16(4):305–326, 1962.

[24] R.H. Nochetto, K.G. Siebert, and A. Veeser. Theory of adaptive finite element methods: an introduction. In *Multiscale, nonlinear and adaptive approximation*, pages 409–542. Springer, 2009.

[25] Y.C. Pati, R. Rezaiifar, and P.S. Krishnaprasad. Orthogonal matching pursuit: Recursive function approximation with applications to wavelet decomposition. In *Proceedings of the 27th Annual Asilomar Conference on Signals, Systems, and Computers*, pages 40–44, 1993.

[26] A. Quarteroni and A. Valli. *Numerical Approximation of Partial Differential Equations*, volume 23 of *Springer Ser. Comput. Math.* Springer-Verlag, Berlin, 2008.

[27] H. Rauhut. Compressive sensing and structured random matrices. In M. Fornasier, editor, *Theoretical Foundations and Numerical Methods for Sparse Recovery*, volume 9 of *Radon Ser. Comput. Appl. Math.*, pages 1–92. deGruyter, 2010.

[28] R. Rubinstein. Omp-Box v10. `http://www.cs.technion.ac.il/~ronrubin/software.html`, 2009.

[29] R. Rubinstein, M. Zibulevsky, and M. Elad. Efficient implementation of the k-SVD algorithm using batch orthogonal matching pursuit. Technical Report CS-2008-08, Technion, Computer Science Department, 2008.

[30] V.N. Temlyakov. Nonlinear methods of approximation. *Found. Comput. Math.*, 3(1):33–107, 2003.

[31] J.A. Tropp. Greed is good: Algorithmic results for sparse approximation. *IEEE Trans. Inform. Theory*, 50:2231–2242, 2004.

[32] J.A. Tropp. User-friendly tail bounds for sums of random matrices. *Found. Comput. Math.*, 12(4):389–434, August 2012.

[33] H. Yserentant. On the multi-level splitting of finite element spaces. *Numer. Math.*, 49(4):379–412, 1986.

[34] T. Zhang. Sparse recovery with orthogonal matching pursuit under RIP. *IEEE Trans. Inform. Theory*, 57(9):6215–6221, 2011.

Recent publications:

MATHEMATICS INSTITUTE OF COMPUTATIONAL SCIENCE AND ENGINEERING
Section of Mathematics
Ecole Polytechnique Fédérale
CH-1015 Lausanne

11.2015 FEDERICO NEGRI:
A model order reduction framework for parametrized nonlinear PDE-constrained optimization

12.2015 ANNA TAGLIABUE, LUCA DEDÈ. ALFIO QUARTERONI:
Nitsche's method for parabolic partial differential equations with mixed time varying boundary conditions

13.2015 SIMONE DEPARIS, DAVIDE FORTI, GWENOL GRANDPERRIN, ALFIO QUARTERONI:
FaCSI: A block parallel preconditioner for fluid-structure interaction in hemodynamics

14.2015 ASSYR ABDULLE, TIMOTHÉE POUCHON:
A priori error analysis of the finite element heterogenenous multiscale method for the wave equation in heterogenenous media over long time

15.2015 ANDREA MANZONI, STEFANO PAGANI:
A certified reduced basis method for PDE-constrained parametric optimization problems by an adjoint-based approach

16.2015 SIMONE DEPARIS, DAVIDE FORTI, ALFIO QUARTERONI:
A fluid-structure interaction algorithm using radial basis function interpolation between non-conforming interfaces

17.2015 ASSYR ABDULLE, ONDREJ BUDAC:
A reduced basis finite element heterogeneous multiscale method for Stokes flow in porous media

18.2015 DANIEL KRESSNER, MICHAEL STEINLECHNER, BART VANDEREYCKEN:
Preconditioned low-rank Riemannian optimization for linear systems with tensor product structure

19.2015 ALESSANDRO S. PATELLI, LUCA DEDÈ, TONI LASSILA, ANDREA BARTEZZAGHI, ALFIO QUARTERONI:
Isogeometric approximation of cardiac electrophysiology models on surfaces: an accuracy study with application to the human left atrium

20.2015 MATTHIEU WILHELM, LUCA DEDÈ, LAURA M. SANGALLI, PIERRE WILHELM:
IGS: an IsoGeometric approach for Smoothing on surfaces

21.2015 SIMONE DEPARIS, DAVIDE FORTI, PAOLA GERVASIO, ALFIO QUARTERONI:
INTERNODES: an accurate interpolation-based method for coupling the Galerkin solutions of PDEs on subdomains featuring non-conforming interfaces

22.2015 ABDUL-LATEEF HAJI-ALI, FABIO NOBILE, LORENZO TAMELLINI, RAÙL TEMPONE:
Multi-index stochastic collocation for random PDEs

23.2015 SIMONE BRUGIAPAGLIA, FABIO NOBILE, STEFANO MICHELETTI, SIMONA PEROTTO:
A theoretical study of COmpRessed SolvING for advection-diffusion-reaction problems