MOX-Report No. 63/2018

Wavelet-Fourier CORSING techniques for multi-dimensional advection-diffusion-reaction equations

Brugiapaglia, S.; Micheletti, S.; Nobile, F.; Perotto, S.

MOX, Dipartimento di Matematica
Politecnico di Milano, Via Bonardi 9 - 20133 Milano (Italy)

mox-dmat@polimi.it http://mox.polimi.it

Wavelet-Fourier CORSING techniques for multi-dimensional advection-diffusion-reaction equations

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

December 21, 2018

[♯] Department of Mathematics
Simon Fraser University
8888 University Drive, Burnaby, BC V5A 1S6, Canada

[♭] 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, local coherence, wavelets.

AMS Subject Classification (2010): Primary 65N30, 94A20, 65T60; Secondary 65T40, 65K10, 42A61

Abstract

We present and analyze a wavelet-Fourier technique for the numerical treatment of multi-dimensional advection-diffusion-reaction equations with periodic boundary conditions. Combining the Petrov-Galerkin technique with the compressed sensing approach, the proposed method is able to approximate the largest coefficients of the solution with respect to a biorthogonal wavelet basis. Namely, we assemble a compressed discretization based on randomized subsampling of the Fourier test space and we employ sparse recovery techniques. The proposed theoretical analysis is based on the local a-coherence and provides effective recipes for a practical implementation. The stability and robustness of the proposed scheme is shown by numerical illustrations in the one-, two-, and three-dimensional case.

[*]Corresponding author: simone_brugiapaglia@sfu.ca

1 Introduction

This paper deals with the theoretical analysis and the numerical implementation of a recently-introduced paradigm in numerical approximation of Partial Differential Equations (PDEs), named COmpRessed SolvING (in short, CORSING). The CORSING method has been proposed and studied in [7, 9, 10] for the solution of linear PDEs set in Hilbert spaces, and combines the Petrov-Galerkin discretization techniques with compressed sensing [11, 17]. Assuming the sparsity of the solution with respect to a suitable trial function basis, the idea is to build a reduced Petrov-Galerkin discretization of the weak formulation of the problem by considering a randomly subsampled test subspace, and then to recover a sparse approximation to the solution via sparse recovery techniques, such as ℓ^1-minimization or the greedy Orthogonal Matching Pursuit (OMP) algorithm. As discussed in [10], the main advantages of CORSING with respect to adaptive finite element techniques for PDEs are that no *a posteriori* error estimators are needed and that assembly and recovery via OMP are fully parallelizable.

In this paper, we focus on the scenario of multi-dimensional advection-diffusion-reaction (ADR) equations on a torus, and we employ biorthogonal wavelets as trial functions basis and a Fourier basis as test functions. In this way, we introduce a hybrid Wavelet-Fourier technique named CORSING $\mathcal{WF}$, which is able to approximate the largest wavelet coefficients of the solution to the PDE by sampling randomly the Fourier test space, using a suitable probability measure. It is worth noticing that the results showed here can be extended to the nonperiodic case by considering biorthogonal wavelets on the interval (and on the hypercube) (see [14, 23, 31]). However, we decided to stick to the periodic case to make the theoretical exposition free of an excessive quantity of technical details regarding the construction of wavelets at the boundary. In this respect, the present work should be considered as a first step towards the setup of CORSING in practical applications.

Although compressed sensing is becoming a standard paradigm for signal processing applications, understanding its full potential and limitations in scientific computing is still object of active work. This paper moves a step forward in this direction. In a fast-growing literature, it is worth mentioning here the applications of compressed sensing to numerical approximation of high-dimensional functions (see, e.g., [2, 3, 12, 26]) and of parametric PDEs, with special emphasis on uncertainty quantification (see, e.g., [6, 18, 25, 32]). In these cases, a smooth function, which can be the quantity of interest of a parametric PDE, is approximated with respect to a global sparsity basis like orthogonal polynomials by means of random pointwise observations. The PDE solver is a black box used to evaluate the quantity of interest for different values of the parameters, and compressed sensing is performed *outside* the black box. Our focus is different, since the CORSING method takes advantage of the compressed sensing paradigm *inside* the black box, i.e., to solve the PDE itself given a particular choice of the parameters.

The compressed sensing principle has also been recently employed for the efficient numerical approximation of diffusion equations via Sturm-Liouville spectral collocation in [8]. Finally, we observe that wavelet-Fourier techniques are widely used in signal processing applications (see, e.g., [4, 5, 20] for theoretical contributions in this direction). Yet, to the best of our knowledge, this paper is the first comprehensive study of this type of techniques in the context of numerical approximation of multidimensional PDEs.

1.1 Main contributions

The main contribution of this paper is threefold. First, we present and study the first hybrid wavelet-Fourier discretization for ADR equations in arbitrary dimension on the torus based on Petrov-Galerkin discretization and on compressed sensing, named CORSING $\mathcal{WF}$ (see Section 3). Moreover, in Section 4 we show the applicability of the theoretical analysis in [10] to n-dimensional ADR equations with constant coefficients for $n \geq 1$ and with nonconstant coefficients for $n = 1$. Finally, in Section 5 we provide a Matlab® implementation of the CORSING $\mathcal{WF}$ method for n-dimensional ADR equations, with $n = 1, 2, 3$.

In view of the aforementioned contributions, we will focus on the following three key technical issues necessary to implement and quantify the performance of the CORSING $\mathcal{WF}$ method (more details on these three issues are given in Section 3.4):

(i) Find a suitable truncation condition on the Fourier test space in order to guarantee stability of the resulting Petrov-Galerkin discretization.

(ii) Give lower bounds to the sampling complexity, i.e., the minimum number of randomly selected Fourier test functions needed to recover the s dominant coefficients of the solution in the wavelet expansion.

(iii) Provide explicit expressions for the probability distribution on the Fourier test space needed for the random selection of the basis functions.

In order to address issues (i), (ii), and (iii), we will take advantage of the general framework given in [10] for the analysis of CORSING, based on the so-called local a-coherence, which can be interpreted as a measure of the angle between the wavelet trial functions and the Fourier test functions with respect to the metric induced by the sesquilinear form associated with the ADR problem (see Definition 3.3). In view of this, our main efforts will be aimed at producing upper bounds to the local a-coherence in Section 4 for the specific ADR problems addressed. In particular, we provide local a-coherence upper bounds for the 1D case with nonconstant coefficients in Theorem 4.2, and for the multi-dimensional case with constant coefficients, when employing anisotropic and isotropic tensor product wavelets in Theorems 4.7 and 4.12, respectively. As a consequence, we derive explicit and computable answers to (i), (ii), and (iii) and provide recovery error guarantees for the CORSING $\mathcal{WF}$ method in Theorem 4.4 (1D case), Theorem 4.9 (multi-dimensional anisotropic case), and Theorem 4.12 (multi-dimensional isotropic case). Numerical results in Section 5 confirm the theoretical findings.

2 Problem setting

In this section, we recall some basics on periodic Sobolev spaces and discuss ADR problems in weak form with periodic boundary conditions.

Notation We denote by $\mathbb{N} := \{1, 2, 3, \ldots\}$, $\mathbb{N}_0 := \{0\} \cup \mathbb{N}$, and $\mathbb{Z} := \mathbb{N}_0 \cup (-\mathbb{N})$. We define $[k] := \{1, \ldots, k\}$ and $[k]_0 := \{0\} \cup [k]$. The notation $X \lesssim Y$ means $X \leq CY$, with $C > 0$ a constant independent of X and Y; $X \sim Y$ means that $X \lesssim Y$ and $X \gtrsim Y$ hold simultaneously. By $X \propto Y$, we understand that there exists a constant $C > 0$ independent of X and Y such that $X = CY$. Given a multi-index $\boldsymbol{r}$, we denote its 2-norm by $|\boldsymbol{r}|$. For every $z \in \mathbb{C}$, $|z|$ is its modulus, $\overline{z}$ is its complex conjugate, and $\mathrm{Re}(z)$ and $\mathrm{Im}(z)$ denote its real and imaginary

part, respectively. Given a vector $\boldsymbol{x} \in \mathbb{C}^n$ and $1 \leq p < +\infty$, then $\|\boldsymbol{x}\|_p := (\sum_{j=1}^n |x_j|^p)^{1/p}$, $\operatorname{supp}(\boldsymbol{x}) := \{j \in [n] : x_j \neq 0\}$, and $\|\boldsymbol{x}\|_0 := |\operatorname{supp}(\boldsymbol{x})|$. Inequalities between vectors in $\mathbb{R}^n$ have to be read componentwise; for example, $\boldsymbol{x} \leq \boldsymbol{y}$ means $x_i \leq y_i$ for every i. The vectors of the canonical basis of $\mathbb{C}^n$ are denoted by $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n$ and $\boldsymbol{x} \cdot \boldsymbol{y} := x_1\overline{y}_1 + \cdots + x_n\overline{y}_n$ is the standard inner product of $\mathbb{C}^n$. Given a matrix $A \in \mathbb{C}^{m\times n}$, A^* denotes its conjugate transpose.

2.1 Sobolev spaces

We start by recalling some standard notions about periodic Sobolev spaces. Let $n \in \mathbb{N}$ and consider the domain

$$\mathcal{D} = (0,1)^n \subseteq \mathbb{R}^n.$$

Given $k \in \mathbb{N}_0$, let $H^k(\mathcal{D}) = H^k(\mathcal{D};\mathbb{C})$ be the Sobolev space of order k of complex-valued functions over $\mathcal{D}$, being understood that $H^0(\mathcal{D}) = L^2(\mathcal{D})$. Moreover, we denote the $H^k(\mathcal{D})$-inner product by

$$(u,v)_k := \sum_{\substack{\boldsymbol{\alpha}\in[k]_0^n \\ \alpha_1+\cdots+\alpha_n\leq k}} \int_{\mathcal{D}} D^{\boldsymbol{\alpha}}u(\boldsymbol{x})\overline{D^{\boldsymbol{\alpha}}v(\boldsymbol{x})}\,\mathrm{d}\boldsymbol{x},$$

where $D^{\boldsymbol{\alpha}} := \frac{\partial^{\alpha_1+\cdots+\alpha_n}}{\partial x_1^{\alpha_1}\cdots\partial x_n^{\alpha_n}}$ is the derivative in the sense of distributions and $[k]_0^n := [k]_0 \times \cdots \times [k]_0$ (n times). For the sake of simplicity, we use the shorthand notation $(\cdot,\cdot) := (\cdot,\cdot)_0$ for the $L^2(\mathcal{D})$-inner product. The $H^k(\mathcal{D})$-norm is defined as $\|\cdot\|_{H^k(\mathcal{D})}^2 = (\cdot,\cdot)_k$ and the $H^k(\mathcal{D})$-seminorm is given by

$$|u|_{H^k(\mathcal{D})}^2 := \sum_{\substack{\boldsymbol{\alpha}\in[k]_0^n \\ \alpha_1+\cdots+\alpha_n=k}} \|D^{\boldsymbol{\alpha}}u\|_{L^2(\mathcal{D})}^2.$$

The periodic Sobolev space of order k, $H^k_{\mathrm{per}}(\mathcal{D}) \subseteq H^k(\mathcal{D})$, is then defined as

$$H^k_{\mathrm{per}}(\mathcal{D}) := \operatorname{clos}_{\|\cdot\|_{H^k(\mathcal{D})}}(C^\infty_{\mathrm{per}}(\mathcal{D})),$$

where

$$C^\infty_{\mathrm{per}}(\mathcal{D}) := \{v|_{\mathcal{D}} : v \in C^\infty(\mathbb{R}^n),\ v(\boldsymbol{x}+L\boldsymbol{e}_i) = v(\boldsymbol{x}),\ \forall L \in \mathbb{Z},\ \forall i \in [n]\}.$$

Notice that $H^0_{\mathrm{per}}(\mathcal{D}) \equiv L^2(\mathcal{D})$ since $C^\infty_{\mathrm{per}}(\mathcal{D})$ is dense in $L^2(\mathcal{D})$. Now, let $u \in H^k_{\mathrm{per}}(\mathcal{D})$ and consider its Fourier series expansion

$$u(\boldsymbol{x}) = \sum_{\boldsymbol{r}\in\mathbb{Z}^n} c_{\boldsymbol{r}} e^{2\pi\mathrm{i}\boldsymbol{r}\cdot\boldsymbol{x}}, \quad \text{with } c_{\boldsymbol{r}} := \int_{\mathcal{D}} u(\boldsymbol{x})e^{-2\pi\mathrm{i}\boldsymbol{r}\cdot\boldsymbol{x}}\,\mathrm{d}\boldsymbol{x}.$$

Then, the following norm equivalence holds:

$$\|u\|_{H^k(\mathcal{D})}^2 \sim \sum_{\boldsymbol{r}\in\mathbb{Z}^n} (1+|\boldsymbol{r}|^{2k})|c_{\boldsymbol{r}}|^2, \quad \forall u \in H^k_{\mathrm{per}}(\mathcal{D}), \quad \forall k \in \mathbb{N}_0. \tag{1}$$

Moreover, we define $H^{-1}_{\mathrm{per}}(\mathcal{D}) := [H^1_{\mathrm{per}}(\mathcal{D})]^*$, where the superscript * denotes the dual space. For the proofs of these results and for more details about periodic Sobolev spaces, we refer the reader to [1, 29, 30].

2.2 Advection-diffusion-reaction problems

Consider the sesquilinear form $a : H^1_{\text{per}}(\mathcal{D}) \times H^1_{\text{per}}(\mathcal{D}) \to \mathbb{C}$ defined as

$$a(u,v) := \int_{\mathcal{D}} \left[\eta(\boldsymbol{x}) \nabla u(\boldsymbol{x}) \cdot \overline{\nabla v(\boldsymbol{x})} + \boldsymbol{\beta}(\boldsymbol{x}) \cdot \nabla u(\boldsymbol{x}) \overline{v(\boldsymbol{x})} + \rho(\boldsymbol{x}) u(\boldsymbol{x}) \overline{v(\boldsymbol{x})} \right] \mathrm{d}\boldsymbol{x},$$

where $\eta, \rho : \mathcal{D} \to \mathbb{R}$ are the diffusion and reaction coefficients, respectively, and $\boldsymbol{\beta} : \mathcal{D} \to \mathbb{R}^n$ is the advective field. Then, the weak formulation of the periodic ADR equation reads

$$\text{find } u \in H^1_{\text{per}}(\mathcal{D}) : \quad a(u,v) = (f,v), \quad \forall v \in H^1_{\text{per}}(\mathcal{D}), \tag{2}$$

where $f : \mathcal{D} \to \mathbb{R}$ is a forcing term. Although we are interested in the case of real-valued coefficients η, ρ, $\boldsymbol{\beta}$, and f, the bilinear form $a(\cdot,\cdot)$ is defined over complex-valued Hilbert spaces to allow us the use of the Fourier basis in Section 3.

We recall that the coercivity constant of $a(\cdot,\cdot)$ is the largest $\alpha > 0$ such that

$$|a(u,u)| \geq \alpha \|u\|^2_{H^1(\mathcal{D})}, \quad \forall u \in H^1_{\text{per}}(\mathcal{D}),$$

and the continuity constant of $a(\cdot,\cdot)$ is the smallest constant $\mathcal{A} > 0$ such that

$$|a(u,v)| \leq \mathcal{A} \|u\|_{H^1(\mathcal{D})} \|v\|_{H^1(\mathcal{D})}, \quad \forall u, v \in H^1_{\text{per}}(\mathcal{D}).$$

In the following proposition, we provide conditions on the coefficients $\eta, \boldsymbol{\beta}, \rho$ sufficient to ensure the well-posedness of problem (2). A proof of this result is provided in Appendix A.

Proposition 2.1 (Well-posedness of the periodic ADR problem)**.** *Let $\eta, \rho \in L^\infty(\mathcal{D})$ and $\boldsymbol{\beta} \in [L^\infty(\mathcal{D})]^n$ be such that $\boldsymbol{\beta}$ is one-periodic with respect to each variable. Moreover, assume that there exist two constants $\eta_{\min}, \zeta > 0$ such that*

$$\eta \geq \eta_{\min}, \quad -\frac{1}{2} \nabla \cdot \boldsymbol{\beta} + \rho \geq \zeta, \quad \text{a.e. in } \mathcal{D}.$$

Then, for every $f \in H^{-1}_{\text{per}}(\mathcal{D})$, the weak problem (2) admits a unique solution u such that

$$\|u\|_{H^1(\mathcal{D})} \leq \frac{1}{\alpha} \|f\|_{H^{-1}(\mathcal{D})}, \tag{3}$$

with $\alpha \geq \min\{\eta_{\min}, \zeta\}$ the coercivity constant of $a(\cdot,\cdot)$. Moreover, the continuity constant of $a(\cdot,\cdot)$ satisfies

$$\mathcal{A} \leq \max \left\{ \|\eta\|_{L^\infty(\mathcal{D})}, \sup_{\boldsymbol{x} \in \mathcal{D}} \|\boldsymbol{\beta}(\boldsymbol{x})\|_2, \|\rho\|_{L^\infty(\mathcal{D})} \right\}.$$

Remark 2.2 (Diffusion equation). Notice that Proposition 2.1 does not encompass the purely diffusive case, i.e. $\boldsymbol{\beta} \equiv \mathbf{0}$ and $\rho \equiv 0$. Indeed, in this case we only have uniqueness of the solution up to constants. This issue can be fixed by assuming $u, v \in H^1_{\text{per}}(\mathcal{D})/\mathbb{R}$ in the weak formulation (2). ■

3 CORSING Wavelet-Fourier

In order to solve problem (2), we describe the CORSING $\mathcal{WF}$ (Wavelet-Fourier) method.

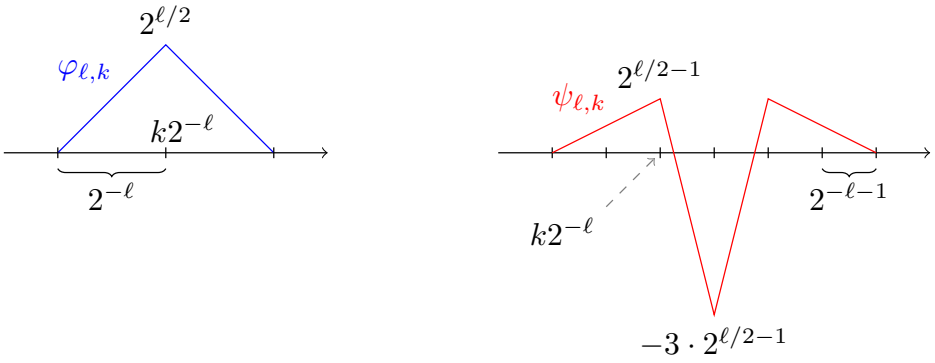

Figure 1: The translated and rescaled scaling functions $\varphi_{\ell,k}$ (left) and wavelets $\psi_{\ell,k}$ (right) corresponding to the construction of biorthogonal B-spline wavelets of order $(d, \widetilde{d}) = (2, 2)$ on the real line (Setting 3.1).

3.1 Trial functions: wavelets

In this section, we present the biorthogonal wavelets on the periodic interval and on the periodic hypercube that will be employed as trial functions of the Petrov-Galerkin discretization. We refer to Appendix B and to [14, 23, 31] for technical details about wavelet construction. A crucial property of the biorthogonal wavelets is that, after suitable normalization, they are a Riesz basis for $H^1_{\mathrm{per}}(\mathcal{D})$, thanks to the so-called norm equivalence property (see Theorem 3.2).

Biorthogonal wavelets on the periodic interval. Given $\ell_0, L \in \mathbb{N}_0$ with $\ell_0 < L$, we consider biorthogonal B-spline wavelet basis on the real line

$$\Psi := \Phi_{\ell_0} \cup \bigcup_{\ell=\ell_0}^{L-1} \Psi_\ell,$$

where $\Phi_\ell = \{\varphi_{\ell,k}\}_{k\in\mathbb{Z}}$ are scaling functions and $\Psi_\ell = \{\psi_{\ell,k}\}_{k\in\mathbb{Z}}$ are wavelet functions (for more details, see Figure 1, Setting 3.1 and Appendices B.1.1 and B.1.2). Notice that the dependence of Ψ on the levels ℓ_0 and L is understood. For $L = \infty$, Ψ is a basis for $L^2(\mathbb{R})$. In order to build a basis for $L^2(\mathcal{D})$, we resort to periodization (see Appendix B.1.3). Similarly, we denote by

$$\Psi^{\mathrm{per}} := \Phi^{\mathrm{per}}_{\ell_0} \cup \bigcup_{\ell=\ell_0}^{L-1} \Psi^{\mathrm{per}}_\ell,$$

where $\Phi^{\mathrm{per}}_\ell = \{\varphi^{\mathrm{per}}_{\ell,k}\}_{k\in\mathbb{Z}/(2^\ell\mathbb{Z})}$ are periodized scaling functions and $\Psi^{\mathrm{per}}_\ell = \{\psi^{\mathrm{per}}_{\ell,k}\}_{k\in\mathbb{Z}/(2^\ell\mathbb{Z})}$ are periodized wavelet functions. In particular, $\mathbb{Z}/(2^\ell\mathbb{Z})$ denotes the ring of integers modulo 2^ℓ that coincides with the set of canonical representatives, i.e.,

$$\mathbb{Z}/(2^\ell\mathbb{Z}) \equiv \{0, 1, \dots, 2^\ell - 1\}, \quad \forall \ell \in \mathbb{N}.$$

Assuming the coarsest level $\ell_0 \in \mathbb{N}_0$ to be fixed, we also introduce the following notation:

$$\psi_{\ell_0-1,k} \equiv \varphi_{\ell_0,k}, \quad \forall k \in \mathbb{Z}. \tag{4}$$

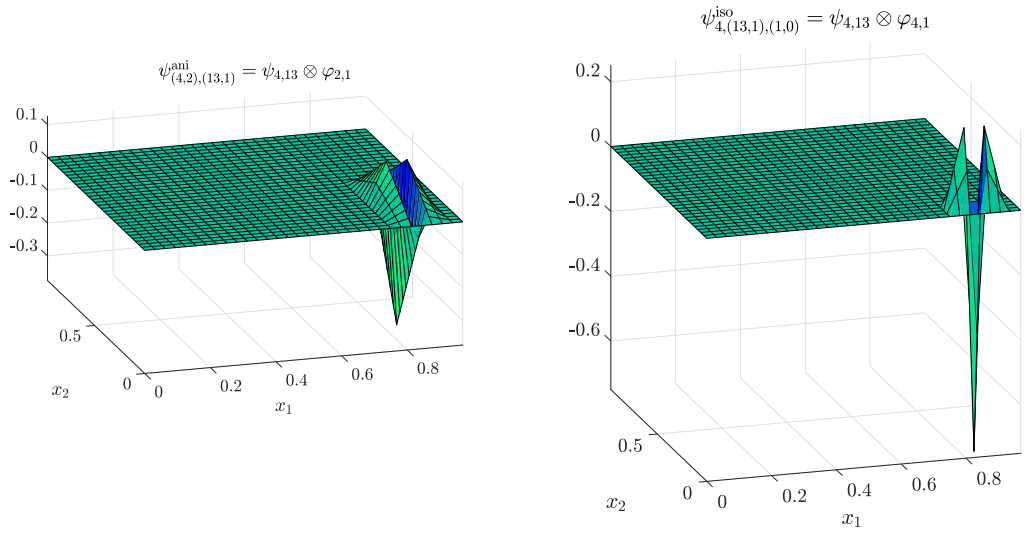

Figure 2: Surface plot of the 2D anisotropic (left) and isotropic (right) tensor product wavelets.

Setting 3.1 (1D biorthogonal B-spline wavelets)**.** *We consider 1D biorthogonal B-spline wavelets of order* $(d, \widetilde{d}) = (2, 2)$*, corresponding to primal and dual filters*

$$\boldsymbol{a}_{[-1:1]} = [\tfrac{1}{2}, 1, \tfrac{1}{2}], \qquad \widetilde{\boldsymbol{a}}_{[-2:2]} = [-\tfrac{1}{4}, \tfrac{1}{2}, \tfrac{3}{2}, \tfrac{1}{2}, -\tfrac{1}{4}],$$
$$\boldsymbol{b}_{[-1:3]} = [\tfrac{1}{4}, \tfrac{1}{2}, -\tfrac{3}{2}, \tfrac{1}{2}, \tfrac{1}{4}], \qquad \widetilde{\boldsymbol{b}}_{[0:2]} = [\tfrac{1}{2}, -1, \tfrac{1}{2}].$$

Moreover, we assume $\ell_0 \geq 2$.

When wavelets are built as in Setting 3.1, the corresponding Ψ^{per} is a basis for $H^1_{\mathrm{per}}(\mathcal{D})$ when $L = \infty$. In fact, for $L \in \mathbb{N}$, $\mathrm{span}(\Psi^{\mathrm{per}})$ coincides with the space of functions in $H^1_{\mathrm{per}}(\mathcal{D})$ that are continuous and piecewise linear on the uniform grid $2^{-L}\mathbb{Z} \cap [0, 1]$. Moreover, choosing $\ell_0 \geq 2$ guarantees that[1]

$$\psi^{\mathrm{per}}_{\ell,k}|_{\mathrm{supp}(\psi_{\ell,k})} \equiv \psi_{\ell,k}|_{\mathrm{supp}(\psi_{\ell,k})}, \quad \forall \ell \geq \ell_0 - 1, \ \forall k \in \mathbb{Z}/(2^\ell \mathbb{Z}). \tag{5}$$

For the sake of simplicity, the apex $^{\mathrm{per}}$ will be omitted in the subsequent developments. In fact, from now on we will always assume to be in the periodic setting.

In order to generalize this construction to arbitrary dimension $n > 1$, we consider *anisotropic* and *isotropic* tensor product wavelets (see Figure 2). Here we just recall the basic definitions and we refer the reader to [14, 23, 31] for a more detailed discussion.

Anisotropic tensor product wavelets. Given a $\ell_0 \in \mathbb{N}$ and multi-indices $\boldsymbol{\ell} \in \mathbb{N}^n$, with $\boldsymbol{\ell} \geq \ell_0 - 1$ and $\boldsymbol{k} \in \mathbb{Z}^n$, a first straightforward way to obtain a multi-dimensional basis is by tensorizing the 1D wavelet basis Ψ with itself n times, namely,

$$\psi^{\mathrm{ani}}_{\boldsymbol{\ell},\boldsymbol{k}} := \psi_{\ell_1,k_1} \otimes \cdots \otimes \psi_{\ell_n,k_n},$$

[1] In particular, the assumption $\ell_0 \geq 2$ ensures that the periodization $\psi^{\mathrm{per}}_{\ell,k}$ of $\psi_{\ell,k}$ (defined according to (61) in Appendix B.1.3) is the sum of terms with disjoint support. On the contrary, let us assume, for example, $\ell_0 = 1$. Then, since $\mathrm{supp}(\psi_{1,1}) \subseteq (0, 3/2)$ the periodization $\psi^{\mathrm{per}}_{1,1}$ is built by summing two overlapping terms on $[0, 1/2]$ (see (61)). In particular, $\psi^{\mathrm{per}}_{1,1}$ is constant over $[0, 1/2]$, whereas $\psi_{1,1}$ is not. This shows that condition (5) is not satisfied for $\ell_0 = 1$.

with $\psi_{\ell,k} = \psi^{\mathrm{per}}_{\ell,k}$. For every level multi-index $\boldsymbol{\ell} \in \mathbb{N}^n$, the spatial multi-index $\boldsymbol{k}$ takes values in $\mathbb{Z}/(2^{\boldsymbol{\ell}}\mathbb{Z})$, where

$$\mathbb{Z}/(2^{\boldsymbol{\ell}}\mathbb{Z}) := \mathbb{Z}/(2^{\ell_1}\mathbb{Z}) \times \cdots \times \mathbb{Z}/(2^{\ell_n}\mathbb{Z}) \equiv \prod_{j=1}^{n} \{0, 1, \ldots, 2^{\ell_j - 1}\}.$$

Therefore, fixing $L \in \mathbb{N}$ with $L > \ell_0$ and defining the multi-index set

$$\mathcal{J}^{\mathrm{ani}} := \{(\boldsymbol{\ell}, \boldsymbol{k}) : \boldsymbol{\ell} \in \mathbb{N}^n,\, \ell_0 - 1 \leq \boldsymbol{\ell} < L,\, \boldsymbol{k} \in \mathbb{Z}/(2^{\boldsymbol{\ell}}\mathbb{Z})\}, \tag{6}$$

we have

$$\Psi^{\mathrm{ani}} = \{\psi^{\mathrm{ani}}_{\boldsymbol{j}}\}_{\boldsymbol{j} \in \mathcal{J}^{\mathrm{ani}}}, \tag{7}$$

where we set $\boldsymbol{j} = (\boldsymbol{\ell}, \boldsymbol{k})$.

Isotropic tensor product wavelets. An alternative (and less straightforward) way to define a multi-dimensional wavelet basis is by isotropic tensorization. In this case, we need the following auxiliary notation:

$$\vartheta_{\ell,k,e} := \begin{cases} \varphi_{\ell,k}, & \text{if } e = 0, \\ \psi_{\ell,k}, & \text{if } e = 1. \end{cases}$$

Then, given $\ell \in \mathbb{N}$ with $\ell \geq \ell_0$, $\boldsymbol{k} \in \mathbb{Z}^n$, and $\boldsymbol{e} \in \{0,1\}^n$, we define

$$\psi^{\mathrm{iso}}_{\ell,\boldsymbol{k},\boldsymbol{e}} := \vartheta_{\ell,k_1,e_1} \otimes \cdots \otimes \vartheta_{\ell,k_n,e_n}.$$

Then, fixing $L \in \mathbb{N}$ with $L > \ell_0$ and defining the multi-index set

$$\mathcal{J}^{\mathrm{iso}} := \{(\ell, \boldsymbol{k}, \boldsymbol{e}) : \ell \in \mathbb{N}, \ell_0 \leq \ell < L, \boldsymbol{k} \in (\mathbb{Z}/(2^{\ell}\mathbb{Z}))^n, \boldsymbol{e} \in \{0,1\}^n\}, \tag{8}$$

we have

$$\Psi^{\mathrm{iso}} := \{\psi^{\mathrm{iso}}_{\boldsymbol{j}}\}_{\boldsymbol{j} \in \mathcal{J}^{\mathrm{iso}}}, \tag{9}$$

where we set $\boldsymbol{j} = (\ell, \boldsymbol{k}, \boldsymbol{e})$.

The main difference between anisotropic and isotropic tensor product structure is visualized in Figures 2 and 3 (in the 2D case). In particular, the anisotropic tensorization blends all the dyadic levels together, whereas the isotropic tensorization only combines basis functions of the same level.

Riesz basis property. The biorthogonal wavelets employed here satisfy the norm equivalence property, presented in the following theorem. We refer to Appendix B.2 for the proof.

Theorem 3.2 (Norm equivalence)**.** *Let $\Psi \in \{\Psi^{\mathrm{ani}}, \Psi^{\mathrm{iso}}\}$ be the tensor product wavelet basis defined as in* (7) *or* (9) *from periodized 1D biorthogonal B-spline wavelets of order $(d, \widetilde{d})$. Then, the following norm equivalence holds:*

$$\Big\| \sum_{\boldsymbol{j} \in \mathcal{J}^{\star}} c_{\boldsymbol{j}} \psi^{\star}_{\boldsymbol{j}} \Big\|_{H^s(\mathcal{D})} \sim \|D^s_{\star} \boldsymbol{c}\|_2, \quad \forall \star \in \{\mathrm{ani}, \mathrm{iso}\}, \tag{10}$$

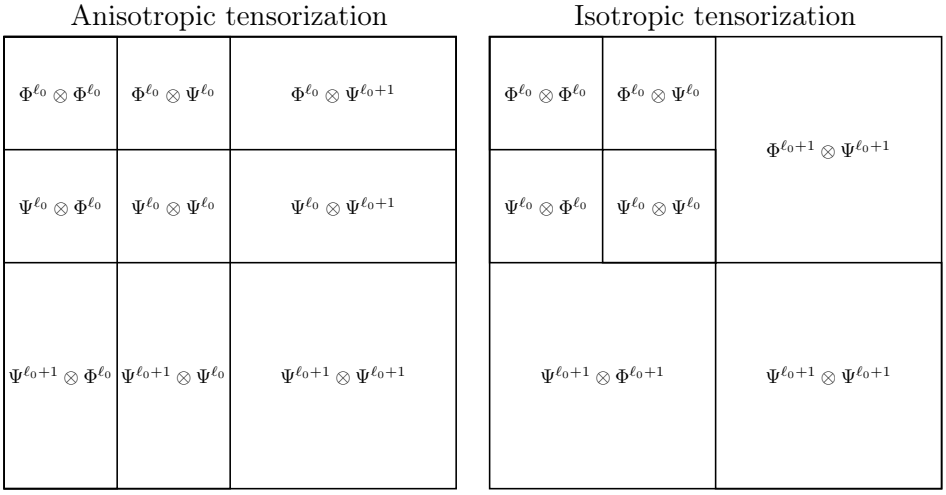

Figure 3: Visualization of the 2D tensorized anisotropic (left) and isotropic (right) wavelets for $L = \ell_0 + 1$.

for every $s \in [0, d-1]$*, where* $\boldsymbol{c} = (\boldsymbol{c})_{\boldsymbol{\ell},\boldsymbol{k}}$ *and* $D_\star \in \{D_{\mathrm{ani}}, D_{\mathrm{iso}}\}$ *is the diagonal matrix defined by*

$$(D_\star)_{\boldsymbol{j},\boldsymbol{j}'} = \begin{cases} 2^{\|\boldsymbol{\ell}\|_\infty}\delta_{\boldsymbol{j},\boldsymbol{j}'}, & \text{if } \star = \mathrm{ani}, \\ 2^{\ell}\delta_{\boldsymbol{j},\boldsymbol{j}'}, & \text{if } \star = \mathrm{iso}, \end{cases} \quad \forall \boldsymbol{j}, \boldsymbol{j}' \in \mathcal{J}^\star,$$

where $\delta_{\boldsymbol{j},\boldsymbol{j}'}$ *is Kronecker's delta function and where* $\mathcal{J}^{\mathrm{ani}}$ *and* $\mathcal{J}^{\mathrm{iso}}$ *are defined as in* (6) *and* (8)*, respectively.*

In Setting 3.1, Theorem 3.2 implies that a suitable weighted ℓ^2-norm of the wavelet coefficients is equivalent to the $H^1(\mathcal{D})$-norm. In particular, using the fact that

$$\|\psi^{\mathrm{ani}}_{\boldsymbol{j}}\|_{H^1(\mathcal{D})} \sim 2^{\|\boldsymbol{\ell}\|_\infty}, \quad \forall \boldsymbol{j} \in \mathcal{J}^{\mathrm{ani}}, \tag{11}$$

$$\|\psi^{\mathrm{iso}}_{\boldsymbol{j}}\|_{H^1(\mathcal{D})} \sim 2^{\ell}, \quad \forall \boldsymbol{j} \in \mathcal{J}^{\mathrm{iso}}. \tag{12}$$

we can rescale the basis functions of Ψ^{ani} and Ψ^{iso} and normalize them with respect to the $H^1(\mathcal{D})$-norm and obtain two Riesz bases with respect to the $H^1(\mathcal{D})$-norm. More details on the norm equivalence property are provided in Appendix B.2.

3.2 Test functions: Fourier basis

Consider the one-periodic Fourier basis functions in dimension one and their tensorized version in the n-dimensional case

$$\xi_q(x) := \exp(2\pi \mathrm{i} q x), \quad \forall q \in \mathbb{Z}, \quad \xi_{\boldsymbol{q}} := \xi_{q_1} \otimes \cdots \otimes \xi_{q_n}, \quad \forall \boldsymbol{q} \in \mathbb{Z}^n.$$

It is easy to verify that $\{\xi_{\boldsymbol{q}} : \boldsymbol{q} \in \mathbb{Z}^n\}$ is an orthonormal basis for $L^2(\mathcal{D})$ and that its elements are orthogonal with respect to the $H^1(\mathcal{D})$-inner product. Moreover, we have

$$\|\xi_{\boldsymbol{q}}\|^2_{L^2(\mathcal{D})} = 1, \quad \|\xi_{\boldsymbol{q}}\|^2_{H^1(\mathcal{D})} = 1 + (2\pi\|\boldsymbol{q}\|_2)^2, \quad \forall \boldsymbol{q} \in \mathbb{Z}^n. \tag{13}$$

Given $R \in \mathbb{N}$, let us consider the following finite multi-index set:

$$\mathcal{Q} := \{\boldsymbol{q} \in \mathbb{Z}^n : -\lfloor R/2 \rfloor + 1 \leq q_i \leq \lfloor R/2 \rfloor,\ \forall i \in [n]\}.$$

Then, we define the Fourier basis as

$$\Xi := \{\xi_{\boldsymbol{q}}\}_{\boldsymbol{q} \in \mathcal{Q}}.$$

In particular, Ξ is a Riesz basis with respect to the $H^1(\mathcal{D})$-norm.

3.3 The CORSING Wavelet-Fourier method

We are now in a position to introduce the CORSING $\mathcal{WF}$ (Wavelet-Fourier) method.

Normalization with respect to the $H^1(\mathcal{D})$-norm. Let $\Psi = \{\psi_{\boldsymbol{j}}\}_{\boldsymbol{j} \in \mathcal{J}}$ be a tensor product of periodized biorthogonal B-spline wavelets (or simply a family of periodized biorthogonal B-spline wavelets for $n = 1$) defined as in Section 3.1. In particular, $\Psi \in \{\Psi^{\mathrm{ani}}, \Psi^{\mathrm{iso}}\}$ and $\mathcal{J} \in \{\mathcal{J}^{\mathrm{ani}}, \mathcal{J}^{\mathrm{iso}}\}$. Let $\Xi = \{\xi_{\boldsymbol{q}}\}_{\boldsymbol{q} \in \mathcal{Q}}$ be the Fourier basis in Section 3.2. Then, we normalize both trial and test functions with respect to the $H^1(\mathcal{D})$-norm, namely,

$$\widehat{\Psi} = \{\widehat{\psi}_{\boldsymbol{j}}\}_{\boldsymbol{j} \in \mathcal{J}}, \quad \widehat{\Xi} = \{\widehat{\xi}_{\boldsymbol{q}}\}_{\boldsymbol{q} \in \mathcal{Q}},$$

such that

$$\|\widehat{\psi}_{\boldsymbol{j}}\|_{H^1(\mathcal{D})} \sim 1, \quad \|\widehat{\xi}_{\boldsymbol{q}}\|_{H^1(\mathcal{D})} = 1, \quad \forall \boldsymbol{j} \in \mathcal{J},\ \forall \boldsymbol{q} \in \mathcal{Q}.$$

In view of (11), (12), and (13), this normalization can be realized by defining

$$\widehat{\psi}^{\mathrm{ani}}_{\boldsymbol{j}} := 2^{-\|\boldsymbol{\ell}\|_\infty} \psi^{\mathrm{ani}}_{\boldsymbol{j}}, \quad \widehat{\psi}^{\mathrm{iso}}_{\boldsymbol{j}} := 2^{-\ell} \psi^{\mathrm{iso}}_{\boldsymbol{j}}, \quad \widehat{\xi}_{\boldsymbol{q}} := (1 + (2\pi\|\boldsymbol{q}\|_2)^2)^{-\frac{1}{2}} \xi_{\boldsymbol{q}}.$$

Petrov-Galerkin discretization. We consider a Petrov-Galerkin discretization of (2) associated with the trial basis $\widehat{\Psi}$ and test basis $\widehat{\Xi}$ (see [24] for an introduction to the Petrov-Galerkin method). The stiffness matrix $B \in \mathbb{C}^{M \times N}$, with $M, N \in \mathbb{N}$ and $M \geq N$, and the load vector $\boldsymbol{g} \in \mathbb{C}^M$ are defined as

$$B_{\boldsymbol{q},\boldsymbol{j}} := a(\widehat{\psi}_{\boldsymbol{j}}, \widehat{\xi}_{\boldsymbol{q}}), \quad g_{\boldsymbol{q}} := (f, \widehat{\xi}_{\boldsymbol{q}}), \quad \forall \boldsymbol{j} \in \mathcal{J},\ \forall \boldsymbol{q} \in \mathcal{Q},$$

where

$$N := |\mathcal{J}| = 2^{nL}, \quad M := |\mathcal{Q}| = R^n,$$

and the resulting Petrov-Galerkin discretization is given by the linear system

$$B\boldsymbol{v} = \boldsymbol{g}, \tag{14}$$

with $\boldsymbol{v} \in \mathbb{C}^N$ the vector of the unknowns.

The CORSING $\mathcal{WF}$ method. The next step is to reduce the dimensionality of the Petrov-Galerkin discretization (14) by random subsampling. Given a probability distribution $\boldsymbol{p} \in \mathbb{R}^M$ over $\mathcal{Q}$, we draw $m \ll M$ multi-indices $\boldsymbol{\tau}_1, \dots, \boldsymbol{\tau}_m \in \mathcal{Q}$ i.i.d. randomly according to

$$\mathbb{P}\{\boldsymbol{\tau}_i = \boldsymbol{q}\} = p_{\boldsymbol{q}}, \quad \forall \boldsymbol{q} \in \mathcal{Q},\ \forall i \in [m].$$

Then, we define the CORSING stiffness matrix $A \in \mathbb{C}^{m \times N}$ and load vector $\boldsymbol{f} \in \mathbb{C}^m$ as

$$A_{i,\boldsymbol{j}} := a(\widehat{\psi}_{\boldsymbol{j}}, \widehat{\xi}_{\boldsymbol{\tau}_i}), \quad f_i := (f, \widehat{\xi}_{\boldsymbol{\tau}_i}), \quad \boldsymbol{j} \in \mathcal{J},\ i \in [m].$$

The CORSING reduced discretization corresponds to the underdetermined linear system

$$A\boldsymbol{z} = \boldsymbol{f}, \tag{15}$$

with $\boldsymbol{z} \in \mathbb{C}^N$ the vector of the unknowns. Then, we consider the diagonal preconditioner $D \in \mathbb{C}^{m \times m}$, defined as[2]

$$D_{i,k} = \delta_{i,k} / \sqrt{m p_{\boldsymbol{\tau}_i}}, \quad i, k \in [m],$$

where $\delta_{i,k}$ is Kronecker's delta function. Given a target sparsity level $s \in \mathbb{N}$, with $s \ll N$, we consider the following optimization problem:

$$\min_{\boldsymbol{z} \in \mathbb{C}^N} \|D(A\boldsymbol{z} - \boldsymbol{f})\|_2, \quad \text{s.t. } \|\boldsymbol{z}\|_0 \leq s. \tag{16}$$

Although NP-hard, (16) can be approximated by the orthogonal matching pursuit (OMP) algorithm. Finally, the CORSING solution is given by

$$\widehat{u} := \sum_{\boldsymbol{j} \in \mathcal{J}} \widehat{u}_{\boldsymbol{j}} \widehat{\psi}_{\boldsymbol{j}},$$

where $\widehat{\boldsymbol{u}} = (\widehat{u}_{\boldsymbol{j}})_{\boldsymbol{j} \in \mathcal{J}} \in \mathbb{C}^N$ is an approximate solution to (16) computed via OMP.

Note that the sampling complexity m can be reduced in order to avoid repeated indices among the τ_i's. In that case, the preconditioner D has to be slightly modified (see [10, Remark 3.9]).

3.4 Towards a recovery error analysis

We are now able to restate issues (i), (ii), and (iii) in Section 1.1 in a rigorous way. Assuming to fix $s, N \in \mathbb{N}$ with $s \ll N$, the CORSING procedure outlined above depends on the following three choices:

(i) Choose $R = R(s, N)$, defining the size of $\mathcal{Q}$ and, consequently, the test space dimension $M = R^n$ of the Petrov-Galerkin discretization (14).

(ii) Choose the number $m = m(s, N, M)$ of random samples, depending sublinearly on N and M (in order to have dimensionality reduction from (14) to (15));

(iii) Define the probability distribution $\boldsymbol{p} \in \mathbb{R}^M$ on the test multi-index space $\mathcal{Q}$.

[2] The diagonal preconditioner D is chosen in such a way that $\mathbb{E}[(DA)^* DA] = B^* B$. For further details, see [10].

In order to solve issues (i), (ii), and (iii), we resort to the theoretical analysis carried out in [10], based on the following

Definition 3.3 (Local a-coherence)**.** The local a-coherence of $\widehat{\Psi}$ with respect to $\widehat{\Xi}$ is a sequence $\boldsymbol{\mu} \in \ell(\mathbb{N})$ defined by

$$\mu_{\boldsymbol{q}} := \sup_{\boldsymbol{j} \in \mathcal{J}} |a(\widehat{\psi}_{\boldsymbol{j}}, \widehat{\xi}_{\boldsymbol{q}})|^2, \quad \forall \boldsymbol{q} \in \mathcal{Q}. \tag{17}$$

Since the CORSING solution $\widehat{u}$ is s-sparse with respect to Ψ (or, equivalently, to $\widehat{\Psi}$), the corresponding best possible accuracy is the best s-term approximation error

$$\sigma_s(u)_{H^1(\mathcal{D})} := \inf \left\{ \|u - w\|_{H^1(\mathcal{D})} : w = \sum_{\boldsymbol{j} \in \mathcal{J}} c_{\boldsymbol{j}} \psi_{\boldsymbol{j}},\ \|\boldsymbol{c}\|_0 \leq s \right\}.$$

We specialize [10, Theorem 3.15] to the CORSING $\mathcal{WF}$ setting, providing a recovery error estimate in expectation. The fact that $\widehat{\Psi}$ and $\widehat{\Xi}$ are Riesz bases with respect to the $H^1(\mathcal{D})$-norm (guaranteed by Theorem 3.2) is required to apply this result, and this justifies the use of biorthogonal wavelets when $n > 1$. Indeed, tensorizing the hierachical basis of hat functions as in [7, 9, 10] does not guarantee the Riesz basis assumption in the multi-dimensional case. It is also possible to state an analogous result in probability instead of expectation (see [10, Theorems 3.16 and 3.18]).

Theorem 3.4 (CORSING $\mathcal{WF}$ recovery in expectation)**.** *Let $n, s, L \in \mathbb{N}$, with $s \ll N = 2^{nL}$, $K > 0$ be such that $\|u\|_{H^1(\mathcal{D})} \leq K$, where u is the unique solution to (2), and assume to have an upper bound to the local a-coherence $\boldsymbol{\mu} \in \ell(\mathbb{Z})$, i.e., there exists a sequence $\boldsymbol{\nu} \in \ell(\mathbb{Z})$ such that*

$$\boldsymbol{\mu} \lesssim \boldsymbol{\nu}. \tag{18}$$

Choose $R \in \mathbb{N}$ (or, equivalently, $\mathcal{Q}$) such that the truncation condition[3]

$$s \|\boldsymbol{\nu}|_{\mathcal{Q}^c}\|_1 \lesssim 1, \tag{19}$$

holds. Then, for every $0 < \varepsilon < 1$, the CORSING solution $\widehat{u} \in H^1_{\mathrm{per}}(\mathcal{D})$ exactly solving (16) satisfies

$$\mathbb{E} \left\| \min \left\{ 1, \frac{K}{\|\widehat{u}\|_{H^1(\mathcal{D})}} \right\} \widehat{u} - u \right\|_{H^1(\mathcal{D})} \lesssim \frac{\mathcal{A}}{\alpha} \sigma_s(u)_{H^1(\mathcal{D})} + K\varepsilon, \tag{20}$$

where α and $\mathcal{A}$ are the inf-sup and the continuity constants respectively associated with $a(\cdot, \cdot)$, provided that

$$m \gtrsim s \|\boldsymbol{\nu}|_{\mathcal{Q}}\|_1 (s \ln(eN/s) + \ln(2s/\varepsilon)), \tag{21}$$

and that the drawings $\tau_1, \ldots, \tau_m \in \mathcal{Q}$ are i.i.d. according to the probability distribution

$$\boldsymbol{p} = \frac{\boldsymbol{\nu}|_{\mathcal{Q}}}{\|\boldsymbol{\nu}|_{\mathcal{Q}}\|_1}. \tag{22}$$

[3] Let us clarify a small difference between [10, Theorem 3.15] and Theorem 3.4. In [10, Theorem 3.15], $\boldsymbol{\nu}$ is an upper bound to $\boldsymbol{\mu}|_{\mathcal{Q}}$ and the truncation condition (19) involves $\boldsymbol{\mu}$ instead of $\boldsymbol{\nu}$. Therefore, the truncation condition of Theorem 3.4 implies the truncation condition of [10, Theorem 3.15]. However, this does not make any difference since in practice the truncation condition of [10, Theorem 3.15] is verified using an upper bound to $\boldsymbol{\mu}$, and not $\boldsymbol{\mu}$ itself.

Some considerations are in order:

- Relation (21) corresponds to a quadratic scaling of m with respect to s (up to logarithmic factors and up to the quantity $\|\boldsymbol{\nu}|_{\mathcal{Q}}\|_1$). In practice, a linear dependence of m on s (up to logarithmic factors) seems to be sufficient (see the [10, Section 5.4]). A different theoretical analysis carried out in [7, Section 3.2.5] based on the concept of restricted isometry property seems to confirm this conjecture up to rescaling DA by a suitable factor depending on $\mathcal{A}$ and on the true Riesz constant of Ψ, hidden in the norm equivalence (10) with $s = 1$. We have preferred to employ this slightly suboptimal result to avoid this technical rescaling issue.
- A necessary condition for (19) is $\|\boldsymbol{\nu}\|_1 < \infty$. This will always be the case in the applications discussed in this paper.
- In order to have an actual compression of the Petrov-Galerkin discretization, $\|\boldsymbol{\nu}|_{\mathcal{Q}}\|_1$ has to depend sublinearly on N and M in (21).
- Notice that knowing an upper bound K to $\|u\|_{H^1(\mathcal{D})}$ is not a restrictive hypothesis in view of the *a priori* estimate (3).
- The hypothesis that $\widehat{\boldsymbol{u}}$ solves (16) exactly does not take into account the approximation error due to the OMP algorithm, which can be included by resorting to the restricted isometry property analysis.

Starting from Theorem 3.4, we can tackle issues (i), (ii), and (iii) in Section 1.1:

(i) Choose $R = R(s, N)$ large enough to satisfy (19).

(ii) Choose $m = m(s, N)$ according to (21).

(iii) Choose $\boldsymbol{p} \in \mathbb{R}^M$ according to (22).

Our next goal is to find an upper bound $\boldsymbol{\nu}$ to the local a-coherence $\boldsymbol{\mu}$ as in (18) such that:

(a) It is possible to find an explicit formula for $R = R(s, N)$ such that (19) is satisfied.

(b) The quantity $\|\boldsymbol{\nu}|_{\mathcal{Q}}\|_1$ depends sublinearly on M and N.

4 Local a-coherence estimates

This section is the technical core of the article. We extend the results in [10] by deriving upper bounds to the local a-coherence defined in (17) for ADR equations with nonconstant coefficients in dimension one (Section 4.1) and with constant coefficients in arbitrary dimension (Section 4.2).

4.1 The 1D ADR problem with nonconstant coefficients

We start by showing some auxiliary inequalities involving inner products between biorthogonal B-spline wavelets and the Fourier basis functions, and their respective derivatives. Note that in the following statement, the basis functions are normalized with respect to the $L^2(\mathcal{D})$-norm.

Lemma 4.1 (Auxiliary inequalities, $n = 1$)**.** *In Setting 3.1, the following inequalities hold for every $\ell \geq \ell_0$, $k \in \mathbb{Z}/(2^\ell\mathbb{Z})$, $q \in \mathbb{Z} \setminus \{0\}$, $(\alpha_1, \alpha_2) \in \{0,1\}^2$, and $\gamma \in [0,2]$:*

$$|(D^{\alpha_1}\varphi_{\ell,k}, D^{\alpha_2}\xi_q)| \leq 2^{\alpha_1+\alpha_2}\, 2^{(\frac{3}{2}-\gamma)\ell}|\pi q|^{\gamma-2+\alpha_1+\alpha_2}, \tag{23}$$

$$|(D^{\alpha_1}\psi_{\ell,k}, D^{\alpha_2}\xi_q)| \leq 2^{\alpha_1+\alpha_2+1-\gamma}\|\boldsymbol{b}\|_2\|\boldsymbol{b}\|_0^{\frac{1}{2}} 2^{(\frac{3}{2}-\gamma)\ell}|\pi q|^{\gamma-2+\alpha_1+\alpha_2}. \tag{24}$$

Moreover, for $q = 0$, we have

$$|(\varphi'_{\ell,k}, \xi'_0)| = |(\psi'_{\ell,k}, \xi'_0)| = 0, \tag{25}$$

$$|(\varphi'_{\ell,k}, \xi_0)| = |(\psi'_{\ell,k}, \xi_0)| = 0, \tag{26}$$

$$|(\varphi_{\ell,k}, \xi_0)| = 2^{-\ell/2}, \quad |(\psi_{\ell,k}, \xi_0)| = 0. \tag{27}$$

Proof. If $q = 0$, estimates (25)-(27) can be verified via direct computation, being $\xi_0 \equiv 1$. Now, we analyze the case $q \neq 0$. Considering the case $(\alpha_1, \alpha_2) = (1,1)$, thanks to hypothesis (5), we can directly compute

$$\begin{aligned}
(\varphi'_{\ell,k}, \xi'_q) &= 2^{3\ell/2}\bigg(\int_{(k-1)2^{-\ell}}^{k2^{-\ell}} \overline{\xi'_q(x)}\,\mathrm{d}x - \int_{k2^{-\ell}}^{(k+1)2^{-\ell}} \overline{\xi'_q(x)}\,\mathrm{d}x\bigg) \\
&= 2^{3\ell/2}(2\mathrm{e}^{-2\mathrm{i}\pi qk2^{-\ell}} - \mathrm{e}^{-2\mathrm{i}\pi q(k-1)2^{-\ell}} - \mathrm{e}^{-2\mathrm{i}\pi q(k+1)2^{-\ell}}) \\
&= 2^{3\ell/2}\mathrm{e}^{-2\mathrm{i}\pi qk2^{-\ell}}(2 - \mathrm{e}^{2\mathrm{i}\pi q2^{-\ell}} - \mathrm{e}^{-2\mathrm{i}\pi q2^{-\ell}}) \\
&= 2^{3\ell/2}\mathrm{e}^{-2\mathrm{i}\pi qk2^{-\ell}}2(1 - \cos(2\pi q2^{-\ell})) = 4 \cdot 2^{3\ell/2}\mathrm{e}^{-2\mathrm{i}\pi qk2^{-\ell}}\sin^2(\pi q2^{-\ell}).
\end{aligned}$$

The inequality $\sin^2(x) \leq |x|^\gamma$, which holds for every $x \in \mathbb{R}$ and $\gamma \in [0,2]$ (see Figure 4), yields

$$|(\varphi'_{\ell,k}, \xi'_q)| \leq 4 \cdot 2^{(\frac{3}{2}-\gamma)\ell}|\pi q|^\gamma \tag{28}$$

Moreover, thanks again to hypothesis (5), employing (59) and (28) and the discrete Cauchy-Schwarz inequality, we obtain

$$\begin{aligned}
|(\psi'_{\ell,k}, \xi'_q)| &= \frac{1}{\sqrt{2}}\bigg|\sum_{j\in\mathbb{Z}/(2^{\ell+1}\mathbb{Z})} b_{j-2k}(\varphi'_{\ell+1,j}, \xi'_q)\bigg| \leq \frac{\|\boldsymbol{b}\|_2}{\sqrt{2}}\bigg[\sum_{j\in\mathbb{Z}/(2^{\ell+1}\mathbb{Z}),\, b_{j-2k}\neq 0} |(\varphi'_{\ell+1,j}, \xi'_q)|^2\bigg]^{\frac{1}{2}} \\
&\leq 4 \cdot \frac{\|\boldsymbol{b}\|_2\|\boldsymbol{b}\|_0^{\frac{1}{2}}}{\sqrt{2}} 2^{(\frac{3}{2}-\gamma)(\ell+1)}|\pi q|^\gamma = 2^{3-\gamma}\|\boldsymbol{b}\|_2\|\boldsymbol{b}\|_0^{\frac{1}{2}} 2^{(\frac{3}{2}-\gamma)\ell}|\pi q|^\gamma.
\end{aligned}$$

This concludes the case $(\alpha_1, \alpha_2) = (1,1)$. The case $(\alpha_1, \alpha_2) \neq (1,1)$ can be addressed using integration by parts since $\xi'_q = (2\pi\mathrm{i}q)\xi_q$.

□

We are now in a position to estimate the local a-coherence of the wavelet basis with respect to the Fourier basis for 1D ADR equations with nonconstant coefficients.

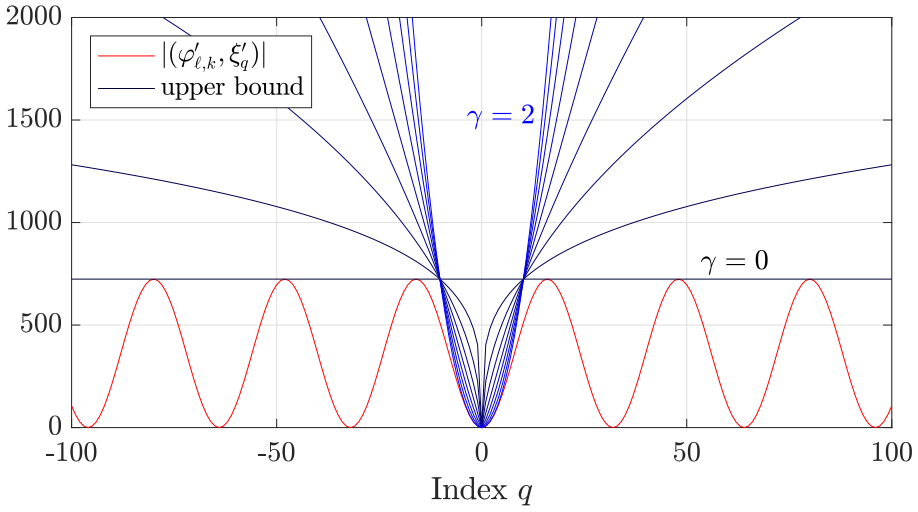

Figure 4: Sharpness of the upper bound (28), with $\ell = 5$, $k = 4$, and different values of $\gamma \in [0, 2]$.

Theorem 4.2 (Local a-coherence upper bound, $n = 1$)**.** *In Setting 3.1, for $\eta, \beta \in H^1_{\mathrm{per}}(\mathcal{D})$ and $\rho \in L^2(\mathcal{D})$, the local a-coherence in* (17) *can be bounded from above as*

$$\mu_0 \lesssim 2^{-2\ell_0}(|\beta|^2_{H^1(\mathcal{D})} + \|\rho\|^2_{L^2(\mathcal{D})})$$
$$\mu_q \lesssim \left(\|\eta\|^2_{H^1(\mathcal{D})} + \frac{\|\beta\|^2_{H^1(\mathcal{D})}}{q^2} + \|\rho\|^2_{L^2(\mathcal{D})}\right) \min\left\{\frac{2^L}{q^2}, \frac{1}{|q|}\right\}, \quad \forall q \in \mathbb{Z} \setminus \{0\}.$$

Proof. First, let us consider the case $q = 0$. We have $|(\eta\psi'_{\ell,k}, \xi'_0)| = |(\eta\psi'_{\ell,k}, 0)| = 0$,

$$|(\beta\psi'_{\ell,k}, \xi_0)| = \left|\int_{\mathcal{D}} \beta(x)\psi'_{\ell,k}(x)\,\mathrm{d}x\right| = \left|\int_{\mathcal{D}} \beta'(x)\psi_{\ell,k}(x)\,\mathrm{d}x\right| \leq |\beta|_{H^1(\mathcal{D})}\|\psi_{\ell,k}\|_{L^2(\mathcal{D})},$$

and

$$|(\rho\psi_{\ell,k}, \xi_0)| \leq \|\rho\|_{L^2(\mathcal{D})}\|\psi_{\ell,k}\|_{L^2(\mathcal{D})}.$$

Therefore,

$$|a(\widehat{\psi}_{\ell,k}, \widehat{\xi}_0)|^2 \leq \left(\frac{|\beta|_{H^1(\mathcal{D})}\|\psi_{\ell,k}\|_{L^2(\mathcal{D})} + \|\rho\|_{L^2(\mathcal{D})}\|\psi_{\ell,k}\|_{L^2(\mathcal{D})}}{\|\psi_{\ell,k}\|_{H^1(\mathcal{D})}\|\xi_0\|_{H^1(\mathcal{D})}}\right)^2$$
$$= \left(\frac{\|\psi_{\ell,k}\|_{L^2(\mathcal{D})}}{\|\psi_{\ell,k}\|_{H^1(\mathcal{D})}}\right)^2 (|\beta|_{H^1(\mathcal{D})} + \|\rho\|_{L^2(\mathcal{D})})^2 \lesssim 2^{-2\ell}(|\beta|^2_{H^1(\mathcal{D})} + \|\rho\|^2_{L^2(\mathcal{D})}),$$

which, by maximization over ℓ and k, implies the estimate for μ_0.

Now, let $q \neq 0$. The idea is to expand the diffusion, advection, and reaction terms with respect to the Fourier basis

$$\eta = \sum_{r\in\mathbb{Z}} \eta_r \xi_r, \quad \beta = \sum_{r\in\mathbb{Z}} \beta_r \xi_r, \quad \rho = \sum_{r\in\mathbb{Z}} \rho_r \xi_r,$$

with $\eta_r := (\eta, \xi_r)$, $\beta_r := (\beta, \xi_r)$, and $\rho_r := (\rho, \xi_r)$, for every $r \in \mathbb{Z}$. The decay of the coefficients $(\eta_r)_{r\in\mathbb{Z}}$, $(\beta_r)_{r\in\mathbb{Z}}$, and $(\rho_r)_{r\in\mathbb{Z}}$ is strictly linked with the Sobolev regularity of η, β, and ρ,

respectively, thanks to the norm equivalence (1). The estimate of μ_q is divided into four parts. In the first three parts, we assess the impact of the terms μ, β, and ρ on the final estimate separately. Then, we combine these estimates in the fourth part.

Part I: diffusion term η ($q \neq 0$). We start by considering the diffusion term

$$|(\eta\psi'_{\ell,k}, \xi'_q)|^2 = \left| \sum_{r\in\mathbb{Z}} \left(\frac{1+r^2}{1+r^2}\right)^{\frac{1}{2}} \eta_r (\xi_r \psi'_{\ell,k}, \xi'_q) \right|^2 \leq \left(\sum_{r\in\mathbb{Z}} (1+r^2)|\eta_r|^2 \right) \sum_{r\in\mathbb{Z}} \frac{|(\xi_r\psi'_{\ell,k}, \xi'_q)|^2}{1+r^2}. \quad (29)$$

Recalling (1), the first factor is bounded from above by $\|\eta\|^2_{H^1(\mathcal{D})}$. Now, exploiting standard algebraic properties of the Fourier basis, performing integration by parts, and using (24) and (26), we obtain the following estimate for every $\gamma \in [0, 2]$:

$$\sum_{r\in\mathbb{Z}} \frac{|(\xi_r\psi'_{\ell,k}, \xi'_q)|^2}{1+r^2} \sim |q|^2 \sum_{r\in\mathbb{Z}} \frac{|(\psi'_{\ell,k}, \xi_{q-r})|^2}{1+r^2} \lesssim q^2 2^{(3-2\gamma)\ell} S_{2(1-\gamma)}(q), \quad (30)$$

where

$$S_y(q) := \sum_{r\in\mathbb{Z}\setminus\{q\}} \frac{1}{|q-r|^y(1+r^2)}, \quad \forall y \in \mathbb{R}.$$

Now, we study the asymptotic behavior of $S_y(q)$ for $y = 1$, and $y = 2$ (corresponding to $\gamma = 1/2$, and $\gamma = 0$, respectively).[4]

We start by considering $S_1(q)$, when $q > 0$. We have the splitting

$$S_1(q) = \underbrace{\sum_{r=-\infty}^{-1} \frac{1}{(q-r)(1+r^2)}}_{S_{1,1}(q)} + \underbrace{\sum_{r=0}^{q-1} \frac{1}{(q-r)(1+r^2)}}_{S_{1,2}(q)} + \underbrace{\sum_{r=q+1}^{+\infty} \frac{1}{(r-q)(1+r^2)}}_{S_{1,3}(q)}.$$

We study the three sums separately, to show that each term can be bounded from above by $1/q$, up to a constant. Indeed,

$$\begin{aligned}
S_{1,1}(q) &= \sum_{r=1}^{+\infty} \frac{1}{(q+r)(1+r^2)} \leq \frac{1}{q} \sum_{r=1}^{+\infty} \frac{1}{1+r^2} \lesssim \frac{1}{q}, \\
S_{1,2}(q) &\leq \frac{1}{q} + \frac{1}{1+(q-1)^2} + \int_0^{q-1} \frac{1}{(q-r)(1+r^2)}\,\mathrm{d}r \\
&= \frac{1}{q} + \frac{1}{1+(q-1)^2} + \left[\frac{2q\arctan(r) - 2\log(q-r) + \log(1+r^2)}{2(1+q^2)} \right]\Bigg|_{r=0}^{r=q-1} \\
&= \frac{1}{q} + \frac{1}{1+(q-1)^2} + \frac{2q\arctan(q-1) + 2\log(q) + \log(1+(q-1)^2)}{2(1+q^2)} \lesssim \frac{1}{q}, \\
S_{1,3}(q) &\lesssim \frac{1}{q} \sum_{r=q+1}^{+\infty} \frac{1}{(r-q)(r+1)} \lesssim \frac{1}{q} \sum_{r=1}^{+\infty} \frac{1}{r^2} \lesssim \frac{1}{q}.
\end{aligned}$$

[4] Choosing $\gamma = 1/2$ and $\gamma = 0$ leads to the estimates $S_1(q) \lesssim 1/|q|$ and $S_2(q) \lesssim 1/|q|^2$, respectively. These, in turn, imply two upper bounds to μ_q: the first one independent of N and decaying linearly with respect to q, the second one linear in N but decaying quadratically with respect to q. These two properties will be crucial to answer issues (i) and (ii) in Theorem 4.4.

The first inequality employed to bound $S_{1,2}(q)$ relies on a property of the function $g(r) := 1/[(q-r)(1+r^2)]$ such that for every $q \geq 2$, g admits only one stationary point $r^* = \frac{1}{3}(q - \sqrt{q^2-3})$ in the open interval $(0, q-1)$ such that $g(r^*) \leq g(q-1) \leq g(0)$, and g decreases monotonically in $[0, r^*]$ and increases monotonically in $[r^*, q-1]$. In particular, this implies

$$\sum_{r=0}^{q-1} g(r) = \sum_{r=0}^{\lfloor r^* \rfloor} g(r) + \sum_{r=\lfloor r^* \rfloor + 1}^{q-1} g(r) \leq g(0) + \int_0^{\lfloor r^* \rfloor} g(r)\,\mathrm{d}r + g(q-1) + \int_{\lfloor r^* \rfloor + 1}^{q-1} g(r)\,\mathrm{d}r$$
$$\leq g(0) + g(q-1) + \int_0^{q-1} g(r)\,\mathrm{d}r.$$

Also, notice that when $q = 1$ the term $S_{1,2}(q)$ is equal to 1.

Observing that $S_y(q)$ is even with respect to q, we conclude that $S_1(q) \lesssim 1/|q|$, for every $q \neq 0$.

We carry out a similar analysis for $S_2(q)$. For $q > 0$, we have

$$S_2(q) = \underbrace{\sum_{r=-\infty}^{-1} \frac{1}{(q-r)^2(1+r^2)}}_{S_{2,1}(q)} + \underbrace{\sum_{r=0}^{q-1} \frac{1}{(q-r)^2(1+r^2)}}_{S_{2,2}(q)} + \underbrace{\sum_{r=q+1}^{+\infty} \frac{1}{(r-q)^2(1+r^2)}}_{S_{2,3}(q)}.$$

Using arguments similar to those employed for $S_1(q)$, we obtain

$$S_{2,1}(q) \leq \frac{1}{(q+1)^2} \sum_{r=1}^{+\infty} \frac{1}{1+r^2} \lesssim \frac{1}{q^2}$$

$$S_{2,2}(q) \leq \frac{1}{q^2} + \frac{1}{1+(q-1)^2} + \int_0^{q-1} \frac{1}{(q-r)^2(1+r^2)}\,\mathrm{d}r$$
$$\lesssim \frac{1}{q^2} + \frac{(q-1)(1+q^2) + (q^3-q)\arctan(q-1) + q^2(2\log(q) + \log(1+(q-1)^2))}{q(1+q^2)^2} \lesssim \frac{1}{q^2},$$

$$S_{2,3}(q) \leq \frac{1}{1+(q+1)^2} \sum_{r=q+1}^{+\infty} \frac{1}{(r-q)^2} \lesssim \frac{1}{q^2}.$$

Using again that $S_y(q)$ is even with respect to q, we conclude that $S_2(q) \lesssim 1/q^2$, for every $q \neq 0$.[5]

Recalling relations (29) and (30), we have

$$|(\eta\psi'_{\ell,k}, \xi'_q)|^2 \lesssim \|\eta\|^2_{H^1(\mathcal{D})} q^2 \min\{2^{3\ell} S_2(q), 2^{2\ell} S_1(q)\} \lesssim \|\eta\|^2_{H^1(\mathcal{D})} 2^{2\ell} \min\{2^\ell, |q|\}. \tag{31}$$

Part II: advection term β ($q \neq 0$). Analogously to the diffusion case, we have

$$|(\beta\psi'_{\ell,k}, \xi_q)|^2 \leq \left(\sum_{r\in\mathbb{Z}} |\beta_r|^2(1+r^2)\right) \sum_{r\in\mathbb{Z}} \frac{|(\xi_r\psi'_{\ell,k}, \xi_q)|^2}{1+r^2} \sim \|\beta\|^2_{H^1(\mathcal{D})} \sum_{r\in\mathbb{Z}} \frac{|(\psi'_{\ell,k}, \xi_{q-r})|^2}{1+r^2}.$$

[5] In view of Lemma 4.1, we conjecture that $S_y(q) \lesssim 1/|q|^y$ for every $y \in [-4, 2]$ (corresponding to $\gamma \in [0, 2]$ and $y = 2(1-\gamma)$). Nevertheless, proving this rigorously is not straightforward, since the terms corresponding to $S_{1,2}(q)$ and $S_{2,2}(q)$ become very difficult to analyze.

Estimate the sum in the right-hand side as before, we obtain

$$|(\beta\psi'_{\ell,k},\xi_q)|^2 \lesssim \frac{\|\beta\|^2_{H^1(\mathcal{D})}}{q^2}2^{2\ell}\min\{2^\ell,|q|\}. \tag{32}$$

Part III: reaction term ρ ($q\neq 0$). We deal with the nonconstant reaction term in an analogous way. Recalling Lemma 4.1 and the norm equivalence (1) with $k=0$, we have

$$|(\rho\psi_{\ell,k},\xi_q)|^2 \le \left(\sum_{r\in\mathbb{Z}}|\rho_r|^2\right)\sum_{r\in\mathbb{Z}}|(\psi_{\ell,k},\xi_{q-r})|^2 \lesssim \|\rho\|^2_{L^2(\mathcal{D})}\left(2^{(3-2\gamma)\ell}\left(\sum_{r\in\mathbb{Z}\setminus\{q\}}|q-r|^{2(\gamma-2)}\right)+2^{-\ell/2}\right),$$

for every $\gamma\in[0,2]$, where we have employed (23)-(24) with $(\alpha_1,\alpha_2)=(0,0)$ and ξ_{q-r} for $r\neq q$ and (27) with $\xi_{q-r}=\xi_0$ for $r=q$. Choosing $\gamma=1/2$, we see that

$$|(\rho\psi_{\ell,k},\xi_q)|^2 \lesssim \|\rho\|^2_{L^2(\mathcal{D})}(2^{2\ell}+1) \lesssim \|\rho\|^2_{L^2(\mathcal{D})}2^{2\ell}, \tag{33}$$

since $\sum_{r\in\mathbb{Z}\setminus\{q\}}|q-r|^{-3}\lesssim 1$, for every $q\neq 0$.

Part IV: conclusion ($q\neq 0$). Combining (31), (32), and (33) finally yields

$$|a(\psi_{\ell,k},\xi_q)|^2 \lesssim \left(\|\eta\|^2_{H^1(\mathcal{D})}+\frac{\|\beta\|^2_{H^1(\mathcal{D})}}{q^2}+\|\rho\|^2_{L^2(\mathcal{D})}\right)2^{2\ell}\min\{2^\ell,|q|\}.$$

As a consequence, normalizing the trial and test functions with respect to the $H^1(\mathcal{D})$-norm and using that $\|\psi_{\ell,k}\|_{H^1(\mathcal{D})}\sim 2^\ell$ and that $\|\xi_q\|_{H^1(\mathcal{D})}\sim|q|$ (recall (11) and (13)), we obtain

$$\mu_q \lesssim \left(\|\eta\|^2_{H^1(\mathcal{D})}+\frac{\|\beta\|^2_{H^1(\mathcal{D})}}{q^2}+\|\rho\|^2_{L^2(\mathcal{D})}\right)\min\left\{\frac{2^L}{q^2},\frac{1}{|q|}\right\}.$$

This completes the proof. □

Remark 4.3 (Sharper upper bound for $\rho\in H^1_{\mathrm{per}}(\mathcal{D})$). It is not difficult to show that an upper bound for μ_q as in Theorem 4.2 holds when $\rho\in H^1_{\mathrm{per}}(\mathcal{D})$, with the following estimate for the local a-coherence when $q\neq 0$:

$$\mu_q \lesssim \left(\|\eta\|^2_{H^1(\mathcal{D})}+\frac{\|\beta\|^2_{H^1(\mathcal{D})}}{q^2}+\frac{\|\rho\|^2_{H^1(\mathcal{D})}}{q^4}\right)\min\left\{\frac{2^L}{q^2},\frac{1}{|q|}\right\}. \tag{34}$$

Notice that (34) generalizes [10, Proposition 4.4] (in particular, we refer to [10, equation (145)]) for 1D ADR equations with constant coefficients and nonperiodic boundary conditions. ■

Finally, we have the CORSING $\mathcal{WF}$ recovery theorem that provides an answer to the three items (i), (ii), and (iii) in Section 3.4 for the 1D case.

Theorem 4.4 (CORSING $\mathcal{WF}$ recovery)**.** *In Setting 3.1 and under the same hypotheses as in Theorem 3.4, let $n=1$, $\eta,\beta\in H^1_{\mathrm{per}}(\mathcal{D})$ and $\rho\in L^2(\mathcal{D})$. Then, provided*

(i) $R\sim CsN$,

(ii) $m \gtrsim Cs(s\ln(eN/(2s)) + \ln(2s/\varepsilon))(\ln N + \ln s + \ln C)$,

(iii) $p_q \propto \begin{cases} 1, & q = 0, \\ \min\left\{\frac{N}{q^2}, \frac{1}{|q|}\right\}, & q \neq 0, \end{cases}$

where

$$C := \|\eta\|^2_{H^1(\mathcal{D})} + \|\beta\|^2_{H^1(\mathcal{D})} + \|\rho\|^2_{L^2(\mathcal{D})},$$

the CORSING $\mathcal{WF}$ *method recovers the best* s*-term approximation to* u *in expectation, in the sense of estimate* (20).

Proof. Employing Theorem 4.2, we choose

$$\nu_q = \begin{cases} C, & q = 0, \\ C\min\left\{\frac{N}{q^2}, \frac{1}{|q|}\right\}, & q \neq 0. \end{cases}$$

As a consequence, we have $\|\boldsymbol{\nu}\|_1 < +\infty$. Then, to ensure condition (19), using that $\nu_q \leq CN/q^2$ for $q \neq 0$ we estimate

$$s\|\boldsymbol{\nu}|_{\mathcal{Q}^c}\|_1 \leq CsN \sum_{|q|\geq\lfloor R/2\rfloor-1} \frac{1}{q^2} \lesssim \frac{CsN}{R}.$$

Therefore, to ensure (19) we let $R = R(s,N) \sim CsN$. Moreover, using that $\nu_q \leq C/|q|$ for $q \neq 0$, we see that

$$\|\boldsymbol{\nu}|_{\mathcal{Q}}\|_1 \leq C\left(1 + \sum_{0<|q|\leq\lfloor R/2\rfloor} \frac{1}{|q|}\right) \lesssim C\ln R = C\ln M \sim C(\ln N + \ln s + \ln C),$$

which depends sublinearly on M and N, as desired. □

4.2 The multi-dimensional case

We consider issues (i), (ii), and (iii) in Section 3.4 for the CORSING $\mathcal{WF}$ for multi-dimensional ADR equations with constant coefficients. In Section 4.2.1 we analyze the case of anisotropic tensor product wavelet, while in Section 4.2.2 we deal with the isotropic case.

4.2.1 Anisotropic tensor product wavelets

We provide local a-coherence upper bounds (Theorem 4.7) and a recovery result (Theorem 4.9) for the CORSING $\mathcal{WF}$ method with anisotropic tensor product wavelets.

We start by proving a technical result analogous to Lemma 4.1. To shorten notations, we introduce

$$|\boldsymbol{x}|^{\boldsymbol{y}} := \prod_{j=1}^{k} |x_j|^{y_j}, \quad \forall \boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^k,\ \forall k \in \mathbb{N}.$$

Moreover, we denote $\mathbf{1} = (1, 1, \ldots, 1)$, $\mathbf{2} = (2, 2, \ldots, 2)$, and so on.

Lemma 4.5 (Auxiliary inequalities, anisotropic wavelets)**.** *In Setting 3.1, let $n > 1$, $\boldsymbol{\ell} \in \mathbb{N}^n$, with $\boldsymbol{\ell} \geq \ell_0 - 1$, $\boldsymbol{k} \in \mathbb{Z}/(2^{\boldsymbol{\ell}}\mathbb{Z})$, and $\boldsymbol{q} \in \mathbb{Z}^n$. Moreover, define*[6]

$$\mathrm{zero}(\boldsymbol{q}) := [n] \setminus \mathrm{supp}(\boldsymbol{q}), \quad \mathrm{scal}(\boldsymbol{\ell}) := \{j \in [n] : \ell_j = \ell_0 - 1\}. \tag{35}$$

Then, it follows

- *if $\boldsymbol{q} = \mathbf{0}$, we have*

$$|(\nabla \psi^{\mathrm{ani}}_{\boldsymbol{\ell},\boldsymbol{k}}, \nabla \xi_{\mathbf{0}})| = 0, \tag{36}$$
$$|(\boldsymbol{\beta} \cdot \nabla \psi^{\mathrm{ani}}_{\boldsymbol{\ell},\boldsymbol{k}}, \xi_{\mathbf{0}})| = 0, \quad \forall \boldsymbol{\beta} \in \mathbb{R}^n \tag{37}$$
$$|(\psi^{\mathrm{ani}}_{\boldsymbol{\ell},\boldsymbol{k}}, \xi_{\mathbf{0}})| = \begin{cases} 2^{-n\ell_0/2}, & \text{if } \mathrm{scal}(\boldsymbol{\ell}) = [n], \\ 0, & \text{otherwise;} \end{cases} \tag{38}$$

- *if $\boldsymbol{q} \neq \mathbf{0}$ and $\mathrm{zero}(\boldsymbol{q}) \subseteq \mathrm{scal}(\boldsymbol{\ell})$, then, for every $\boldsymbol{\gamma} \in [0,2]^{\|\boldsymbol{q}\|_0}$, it holds*

$$|(\nabla \psi^{\mathrm{ani}}_{\boldsymbol{\ell},\boldsymbol{k}}, \nabla \xi_{\boldsymbol{q}})| \lesssim 2^{-\frac{1}{2}(n-\|\boldsymbol{q}\|_0)\ell_0 + (\frac{3}{2} - \widehat{\boldsymbol{\gamma}}) \cdot \widehat{\boldsymbol{\ell}}} |\widehat{\boldsymbol{q}}|^{\widehat{\boldsymbol{\gamma}} - \mathbf{2}} \|\boldsymbol{q}\|_2^2, \tag{39}$$
$$|(\boldsymbol{\beta} \cdot \nabla \psi^{\mathrm{ani}}_{\boldsymbol{\ell},\boldsymbol{k}}, \xi_{\boldsymbol{q}})| \lesssim 2^{-\frac{1}{2}(n-\|\boldsymbol{q}\|_0)\ell_0 + (\frac{3}{2} - \widehat{\boldsymbol{\gamma}}) \cdot \widehat{\boldsymbol{\ell}}} |\widehat{\boldsymbol{q}}|^{\widehat{\boldsymbol{\gamma}} - \mathbf{2}} \|\boldsymbol{\beta}\|_2 \|\boldsymbol{q}\|_2, \quad \forall \boldsymbol{\beta} \in \mathbb{R}^n, \tag{40}$$
$$|(\psi^{\mathrm{ani}}_{\boldsymbol{\ell},\boldsymbol{k}}, \xi_{\boldsymbol{q}})| \lesssim 2^{-\frac{1}{2}(n-\|\boldsymbol{q}\|_0)\ell_0 + (\frac{3}{2} - \widehat{\boldsymbol{\gamma}}) \cdot \widehat{\boldsymbol{\ell}}} |\widehat{\boldsymbol{q}}|^{\widehat{\boldsymbol{\gamma}} - \mathbf{2}}, \tag{41}$$

  where $\widehat{\boldsymbol{x}} := \boldsymbol{x}|_{\mathrm{supp}(\boldsymbol{q})} \in \mathbb{R}^{\|\boldsymbol{q}\|_0}$ for every $\boldsymbol{x} \in \mathbb{R}^n$ and the inequalities hide constants that depend exponentially on n;

- *if $\boldsymbol{q} \neq \mathbf{0}$ and $\mathrm{zero}(\boldsymbol{q}) \not\subseteq \mathrm{scal}(\boldsymbol{\ell})$, it holds*

$$|(\nabla \psi^{\mathrm{ani}}_{\boldsymbol{\ell},\boldsymbol{k}}, \nabla \xi_{\boldsymbol{q}})| = 0, \tag{42}$$
$$|(\boldsymbol{\beta} \cdot \nabla \psi^{\mathrm{ani}}_{\boldsymbol{\ell},\boldsymbol{k}}, \xi_{\boldsymbol{q}})| = 0, \quad \forall \boldsymbol{\beta} \in \mathbb{R}^n \tag{43}$$
$$|(\psi^{\mathrm{ani}}_{\boldsymbol{\ell},\boldsymbol{k}}, \xi_{\boldsymbol{q}})| = 0. \tag{44}$$

Proof. In the case $\boldsymbol{q} = \mathbf{0}$, the equalities (36), (37), and (38) are a direct consequence of (25), (26), and (27). Let us consider $\boldsymbol{q} \neq \mathbf{0}$. Then, we organize the proof discussing two cases: $\|\boldsymbol{q}\|_0 = n$ and $\|\boldsymbol{q}\|_0 < n$.

Case $\|\boldsymbol{q}\|_0 = n$. Due to the tensorized form of the trial and the test basis functions, the following relations hold:

$$(\nabla \psi^{\mathrm{ani}}_{\boldsymbol{\ell},\boldsymbol{k}}, \nabla \xi_{\boldsymbol{q}}) = \sum_{j=1}^{n} (\psi'_{\ell_j,k_j}, \xi'_{q_j}) \prod_{i \neq j} (\psi_{\ell_i,k_i}, \xi_{q_i}), \tag{45}$$
$$(\boldsymbol{\beta} \cdot \nabla \psi^{\mathrm{ani}}_{\boldsymbol{\ell},\boldsymbol{k}}, \xi_{\boldsymbol{q}}) = \sum_{j=1}^{n} \beta_j (\psi'_{\ell_j,k_j}, \xi_{q_j}) \prod_{i \neq j} (\psi_{\ell_i,k_i}, \xi_{q_i}), \quad \forall \boldsymbol{\beta} \in \mathbb{R}^n, \tag{46}$$
$$(\psi^{\mathrm{ani}}_{\boldsymbol{\ell},\boldsymbol{k}}, \xi_{\boldsymbol{q}}) = \prod_{j=1}^{n} (\psi_{\ell_j,k_j}, \xi_{q_j}). \tag{47}$$

[6]Note that, according to (4), $\mathrm{scal}(\boldsymbol{\ell})$ is the set of indices j such that ψ_{ℓ_j,k_j} is a scaling function.

Plugging relations (23) and (24) into (45) we see that, for every $\boldsymbol{\gamma} \in [0,2]^n$, we have

$$|(\nabla\psi^{\mathrm{ani}}_{\boldsymbol{\ell},\boldsymbol{k}}, \nabla\xi_{\boldsymbol{q}})| \lesssim \sum_{j=1}^{n} 2^{(\frac{3}{2}-\gamma_j)\ell_j}|q_j|^{\gamma_j} \prod_{i\neq j} 2^{(\frac{3}{2}-\gamma_i)\ell_i}|q_i|^{\gamma_i-2} = 2^{\frac{3}{2}\|\boldsymbol{\ell}\|_1-\boldsymbol{\gamma}\cdot\boldsymbol{\ell}}|\boldsymbol{q}|^{\boldsymbol{\gamma}-\mathbf{2}}\|\boldsymbol{q}\|_2^2.$$

Similarly, plugging (23) and (24) into (46), we obtain

$$|(\boldsymbol{\beta}\cdot\nabla\psi^{\mathrm{ani}}_{\boldsymbol{\ell},\boldsymbol{k}}, \xi_{\boldsymbol{q}})| \lesssim \sum_{j=1}^{n} |\beta_j| 2^{(\frac{3}{2}-\gamma_j)\ell_j}|q_j|^{\gamma_j-1} \prod_{i\neq j} 2^{(\frac{3}{2}-\gamma_i)\ell_i}|q_i|^{\gamma_i-2} \leq 2^{\frac{3}{2}\|\boldsymbol{\ell}\|_1-\boldsymbol{\gamma}\cdot\boldsymbol{\ell}}|\boldsymbol{q}|^{\boldsymbol{\gamma}-\mathbf{2}}\|\boldsymbol{\beta}\|_2\|\boldsymbol{q}\|_2.$$

Finally, plugging (23) and (24) into (47), it follows

$$|(\psi^{\mathrm{ani}}_{\boldsymbol{\ell},\boldsymbol{k}}, \xi_{\boldsymbol{q}})| \lesssim \prod_{j=1}^{n} 2^{(\frac{3}{2}-\gamma_j)\ell_j}|q_j|^{\gamma_j-2} = 2^{\frac{3}{2}\|\boldsymbol{\ell}\|_1-\boldsymbol{\gamma}\cdot\boldsymbol{\ell}}|\boldsymbol{q}|^{\boldsymbol{\gamma}-\mathbf{2}}.$$

The relations above prove (39), (40), and (41).

Case $\|\boldsymbol{q}\|_0 < n$. Let us consider the diffusion term $(\nabla\psi^{\mathrm{ani}}_{\boldsymbol{\ell},\boldsymbol{k}}, \nabla\xi_{\boldsymbol{q}})$.

First, assume that $\mathrm{zero}(\boldsymbol{q}) \not\subseteq \mathrm{scal}(\boldsymbol{\ell})$ (notice also that $\mathrm{zero}(\boldsymbol{q})$ is nonempty since $\|\boldsymbol{q}\|_0 < n$). In this case, we can pick an index $j_0 \in \mathrm{zero}(\boldsymbol{q}) \setminus \mathrm{scal}(\boldsymbol{\ell})$, i.e., such that $q_{j_0} = 0$ and that $\ell_{j_0} \geq \ell_0$ (that is such that $\psi_{\ell_{j_0},k_{j_0}}$ is a wavelet function and not a scaling function). Combining (45) with relations (25) and (27) yields

$$(\nabla\psi^{\mathrm{ani}}_{\boldsymbol{\ell},\boldsymbol{k}}, \nabla\xi_{\boldsymbol{q}}) = \underbrace{(\psi'_{\ell_{j_0},k_{j_0}}, \xi'_0)}_{=0} \prod_{i\neq j_0} (\psi_{\ell_i,k_i}, \xi_{q_i}) + \sum_{j\neq j_0} (\psi'_{\ell_j,k_j}, \xi'_{q_j}) \underbrace{(\psi_{\ell_{j_0},k_{j_0}}, \xi_0)}_{=0} \prod_{i\notin\{j,j_0\}} (\psi_{\ell_i,k_i}, \xi_{q_i}) = 0.$$

This proves (42) ((43) and (44) are shown analogously). As a consequence, the only possibility for $(\nabla\psi^{\mathrm{ani}}_{\boldsymbol{\ell},\boldsymbol{k}}, \nabla\xi_{\boldsymbol{q}})$ to be nonzero is to have $\mathrm{zero}(\boldsymbol{q}) \subseteq \mathrm{scal}(\boldsymbol{\ell})$. In this case, by splitting the sum above and employing (25), we obtain

$$|(\nabla\psi^{\mathrm{ani}}_{\boldsymbol{\ell},\boldsymbol{k}}, \nabla\xi_{\boldsymbol{q}})| \leq \sum_{j\in\mathrm{zero}(\boldsymbol{q})} \underbrace{|(\varphi'_{\ell_0,k_j}, \xi'_0)|}_{=0} \prod_{i\neq j} |(\psi_{\ell_i,k_i}, \xi_{q_i})| + \sum_{j\in\mathrm{supp}(\boldsymbol{q})} |(\psi'_{\ell_j,k_j}, \xi'_{q_j})| \prod_{i\neq j} |(\psi_{\ell_i,k_i}, \xi_{q_i})|.$$

Now, splitting the product and using (27) yields

$$\begin{aligned}|(\nabla\psi^{\mathrm{ani}}_{\boldsymbol{\ell},\boldsymbol{k}}, \nabla\xi_{\boldsymbol{q}})| &\leq \sum_{j\in\mathrm{supp}(\boldsymbol{q})} |(\psi'_{\ell_j,k_j}, \xi'_{q_j})| \prod_{i\in\mathrm{zero}(\boldsymbol{q})} \underbrace{|(\varphi_{\ell_0,k_i}, \xi_0)|}_{=2^{-\ell_0/2}} \prod_{i\in\mathrm{supp}(\boldsymbol{q})\setminus\{j\}} |(\psi_{\ell_i,k_i}, \xi_{q_i})| \\ &= 2^{-(n-\|\boldsymbol{q}\|_0)\ell_0/2} \sum_{j\in\mathrm{supp}(\boldsymbol{q})} |(\psi'_{\ell_j,k_j}, \xi'_{q_j})| \prod_{i\in\mathrm{supp}(\boldsymbol{q})\setminus\{j\}} |(\psi_{\ell_i,k_i}, \xi_{q_i})|.\end{aligned}$$

In order to prove (39), it is sufficient to apply an argument analogous to the case $\|\boldsymbol{q}\|_0 = n$, where the set $[n]$ is replaced with $\mathrm{supp}(\boldsymbol{q})$, $\boldsymbol{q}$ with $\widehat{\boldsymbol{q}}$ and $\boldsymbol{\ell}$ with $\widehat{\boldsymbol{\ell}}$.

Similar arguments lead to (40), (41). □

Remark 4.6 (Curse of dimensionality). It is worth stressing that estimates (36)–(44) are affected by the curse of dimensionality, since they hide constants that blow up exponentially with n. This makes the CORSING $\mathcal{WF}$ approach applicable only for moderate values of n. This is not a major problem for fluid-dynamics applications, where the physical domain has always dimension $n \leq 3$. ■

Equipped with the auxiliary inequalities of Lemma 4.5, we are now in a position to provide upper bounds to the local a-coherence in the multi-dimensional case for anisotropic tensor product wavelets.

Theorem 4.7 (Local a-coherence upper bound, anisotropic wavelets)**.** *In Setting 3.1, let $n > 1$. Then, for every $\eta, \rho \in \mathbb{R}$ and $\boldsymbol{\beta} \in \mathbb{R}^n$, the following upper bounds hold:*

$$\mu_{\mathbf{0}} \lesssim |\rho|^2 2^{-(2+n)\ell_0}. \tag{48}$$

$$\mu_{\boldsymbol{q}} \lesssim \left(|\eta|^2 + \frac{\|\boldsymbol{\beta}\|_2^2}{\|\boldsymbol{q}\|_2^2} + \frac{|\rho|^2}{\|\boldsymbol{q}\|_2^4}\right) 2^{-(n-\|\boldsymbol{q}\|_0)\ell_0} \min\left\{ \frac{2^{(3\|\boldsymbol{q}\|_0-2)L}\|\boldsymbol{q}\|_2^2}{|\widehat{\boldsymbol{q}}|^{\mathbf{4}}}, \frac{\|\boldsymbol{q}\|_2^2}{\|\boldsymbol{q}\|_\infty^2 |\widehat{\boldsymbol{q}}|^{\mathbf{1}}} \right\}, \quad \forall \boldsymbol{q} \neq \mathbf{0}, \tag{49}$$

where the inequalities hide constants depending exponentially on n.

Proof. Let us assume $\boldsymbol{q} = \mathbf{0}$. Recalling (36) and (37), we have

$$|a(\psi^{\mathrm{ani}}_{\boldsymbol{\ell},\boldsymbol{k}}, \xi_{\mathbf{0}})| \leq |\eta| \underbrace{|(\nabla \psi^{\mathrm{ani}}_{\boldsymbol{\ell},\boldsymbol{k}}, \nabla \xi_{\mathbf{0}})|}_{=0} + \underbrace{|(\boldsymbol{\beta} \cdot \nabla \psi^{\mathrm{ani}}_{\boldsymbol{\ell},\boldsymbol{k}}, \xi_{\mathbf{0}})|}_{=0} + |\rho| |(\psi^{\mathrm{ani}}_{\boldsymbol{\ell},\boldsymbol{k}}, \xi_{\mathbf{0}})|.$$

Employing (38) and recalling (11) and (13), we obtain

$$|a(\widehat{\psi}^{\mathrm{ani}}_{\boldsymbol{\ell},\boldsymbol{k}}, \widehat{\xi}_{\mathbf{0}})|^2 \leq \begin{cases} |\rho|^2 2^{-n\ell_0}/2^{2\ell_0} = 2^{-(2+n)\ell_0}|\rho|^2, & \text{if } \mathrm{scal}(\boldsymbol{\ell}) = [n], \\ 0, & \text{otherwise,} \end{cases}$$

which, in turn, implies (48).

When $\boldsymbol{q} \neq \mathbf{0}$, we consider the cases $\|\boldsymbol{q}\|_0 = n$ and $\|\boldsymbol{q}\|_0 < n$.

Case $\|\boldsymbol{q}\|_0 = n$. Employing the auxiliary inequalities (39)–(41) (notice that in this case $\mathrm{zero}(\boldsymbol{q}) = \emptyset \subseteq \mathrm{scal}(\boldsymbol{\ell})$) and recalling relations (11) and (13) on the $H^1(\mathcal{D})$-norm of $\psi^{\mathrm{ani}}_{\boldsymbol{\ell},\boldsymbol{k}}$ and of $\xi_{\boldsymbol{q}}$, for every $\boldsymbol{\gamma} \in [0,2]^n$ and $\boldsymbol{q} \neq \mathbf{0}$, we have

$$|a(\widehat{\psi}^{\mathrm{ani}}_{\boldsymbol{\ell},\boldsymbol{k}}, \widehat{\xi}_{\boldsymbol{q}})| = \frac{|a(\psi^{\mathrm{ani}}_{\boldsymbol{\ell},\boldsymbol{k}}, \xi_{\boldsymbol{q}})|}{\|\psi^{\mathrm{ani}}_{\boldsymbol{\ell},\boldsymbol{k}}\|_{H^1(\mathcal{D})} \|\xi_{\boldsymbol{q}}\|_{H^1(\mathcal{D})}} \lesssim 2^{(\frac{3}{2}-\boldsymbol{\gamma})\cdot\boldsymbol{\ell} - \|\boldsymbol{\ell}\|_\infty} |\boldsymbol{q}|^{\boldsymbol{\gamma}-\mathbf{2}} \|\boldsymbol{q}\|_2 \left(|\eta| + \frac{\|\boldsymbol{\beta}\|_2}{\|\boldsymbol{q}\|_2} + \frac{|\rho|}{\|\boldsymbol{q}\|_2^2}\right). \tag{50}$$

For any fixed $\boldsymbol{q} \in \mathbb{Z}^n \setminus \{\mathbf{0}\}$, we make two different choices of $\boldsymbol{\gamma} = \boldsymbol{\gamma}(\boldsymbol{q})$ that, in turn, generate two upper bounds to the local a-coherence.

We start by choosing $\boldsymbol{\gamma} = \mathbf{0}$, in order to let the upper bound decay as fast as possible in $\boldsymbol{q}$. Squaring (50) and considering the maximum over $\boldsymbol{j} \in \mathcal{J}$, we obtain

$$\mu_{\boldsymbol{q}} \lesssim \left(|\eta|^2 + \frac{\|\boldsymbol{\beta}\|_2^2}{\|\boldsymbol{q}\|_2^2} + \frac{|\rho|^2}{\|\boldsymbol{q}\|_2^4}\right) \frac{\|\boldsymbol{q}\|_2^2}{|\boldsymbol{q}|^{\mathbf{4}}} \max_{\ell_0 \leq \boldsymbol{\ell} < L} 2^{3\|\boldsymbol{\ell}\|_1 - 2\|\boldsymbol{\ell}\|_\infty}.$$

Finally, by noticing that

$$\max_{\ell_0 \leq \boldsymbol{\ell} < L} 2^{3\|\boldsymbol{\ell}\|_1 - 2\|\boldsymbol{\ell}\|_\infty} \leq \max_{\ell_0 \leq \boldsymbol{\ell} < L} 2^{(3n-2)\|\boldsymbol{\ell}\|_\infty} \leq 2^{(3n-2)L},$$

we obtain the upper bound corresponding to the first argument of the minimum in (49), for $\|\boldsymbol{q}\|_0 = n$.

Now, we find a second upper bound to $\mu_{\boldsymbol{q}}$ that is independent of L, but decays more slowly with respect to $\boldsymbol{q}$. Consider an index $j_\infty \in [n]$ such that $|q_{j_\infty}| = \|\boldsymbol{q}\|_\infty$. We define $\widetilde{\boldsymbol{\gamma}} = \widetilde{\boldsymbol{\gamma}}(\boldsymbol{q})$ componentwise as

$$\widetilde{\gamma}_j := \begin{cases} 1/2 & \text{if } j = j_\infty, \\ 3/2 & \text{otherwise,} \end{cases} \tag{51}$$

and choose $\boldsymbol{\gamma} = \widetilde{\boldsymbol{\gamma}}$. Considering the set $\triangle := \{\boldsymbol{\ell} \in \mathbb{N}^n : \ell_0 \le \boldsymbol{\ell} < L,\, \ell_1 \ge \ell_2 \ge \cdots \ge \ell_n\}$ and defining S_n as the permutation group of the set $[n]$, we have

$$\begin{aligned} \max_{\ell_0 \le \boldsymbol{\ell} < L} 2^{\frac{3}{2}\|\boldsymbol{\ell}\|_1 - \|\boldsymbol{\ell}\|_\infty - \widetilde{\boldsymbol{\gamma}}\cdot\boldsymbol{\ell}} &= \max_{\boldsymbol{\ell}\in\triangle} \max_{\sigma\in S_n} 2^{\frac{3}{2}\|\sigma(\boldsymbol{\ell})\|_1 - \|\sigma(\boldsymbol{\ell})\|_\infty - \widetilde{\boldsymbol{\gamma}}\cdot\sigma(\boldsymbol{\ell})} \\ &= \max_{\boldsymbol{\ell}\in\triangle} 2^{\frac{3}{2}\|\boldsymbol{\ell}\|_1 - \|\boldsymbol{\ell}\|_\infty} \max_{\sigma\in S_n} 2^{-\widetilde{\boldsymbol{\gamma}}\cdot\sigma(\boldsymbol{\ell})} \\ &= \max_{\boldsymbol{\ell}\in\triangle} 2^{\frac{3}{2}\|\boldsymbol{\ell}\|_1 - \ell_1} 2^{-\frac{1}{2}\ell_1 - \frac{3}{2}\sum_{j>1}\ell_j} = 1, \end{aligned} \tag{52}$$

where we have exploited the identity $\{\boldsymbol{\ell} \in \mathbb{N}^n : \ell_0 \le \boldsymbol{\ell} < L\} = \bigcup_{\sigma\in S_n} \sigma(\triangle)$ and the fact that $\|\boldsymbol{\ell}\|_1$ and $\|\boldsymbol{\ell}\|_\infty$ are invariant with respect to permutations of the components of $\boldsymbol{\ell}$. Combining (50) with (52) (with $\boldsymbol{\gamma} = \widetilde{\boldsymbol{\gamma}}$), and observing that

$$|\boldsymbol{q}|^{2\widetilde{\boldsymbol{\gamma}}(\boldsymbol{q})-\mathbf{4}} = |q_{j_\infty}|^{-3} \prod_{j\neq j_\infty} |q_j|^{-1} = \frac{1}{\|\boldsymbol{q}\|_\infty^2 |\boldsymbol{q}|^{\mathbf{1}}},$$

we obtain

$$\mu_{\boldsymbol{q}} \lesssim |\boldsymbol{q}|^{2\widetilde{\boldsymbol{\gamma}}(\boldsymbol{q})-\mathbf{4}} \|\boldsymbol{q}\|_2^2 \left(|\eta|^2 + \frac{\|\boldsymbol{\beta}\|_2^2}{\|\boldsymbol{q}\|_2^2} + \frac{|\rho|^2}{\|\boldsymbol{q}\|_2^4} \right) = \frac{\|\boldsymbol{q}\|_2^2}{\|\boldsymbol{q}\|_\infty^2 |\boldsymbol{q}|^{\mathbf{1}}} \left(|\eta|^2 + \frac{\|\boldsymbol{\beta}\|_2^2}{\|\boldsymbol{q}\|_2^2} + \frac{|\rho|^2}{\|\boldsymbol{q}\|_2^4} \right),$$

where the right hand side does not depend on $\boldsymbol{\ell}$. This relation corresponds to the second argument of the minimum in (49) when $\|\boldsymbol{q}\|_0 = n$. The case $\|\boldsymbol{q}\|_0 = n$ is hence concluded.

Case $0 < \|\boldsymbol{q}\|_0 < n$. The argument is analogous to the case $\|\boldsymbol{q}\|_0 = n$. We just need to replace $[n]$ with $\mathrm{supp}(\boldsymbol{q})$, n with $\|\boldsymbol{q}\|_0$, $\boldsymbol{q}$ with $\widehat{\boldsymbol{q}}$, $\boldsymbol{\ell}$ with $\widehat{\boldsymbol{\ell}}$, and $\boldsymbol{\gamma}$ with $\widehat{\boldsymbol{\gamma}}$. Moreover, notice that $\|\widehat{\boldsymbol{q}}\|_2 = \|\boldsymbol{q}\|_2$, $\|\widehat{\boldsymbol{q}}\|_\infty = \|\boldsymbol{q}\|_\infty$, and $\|\widehat{\boldsymbol{\ell}}\|_\infty \le \|\boldsymbol{\ell}\|_\infty$ (note that the last relation is not an equality since $\widehat{\boldsymbol{\ell}} = \boldsymbol{\ell}|_{\mathrm{supp}(\boldsymbol{q})}$). This concludes the proof. □

Remark 4.8 (Consistency with the 1D case.). The upper bounds of Theorem 4.2 are compatible with those in Theorem 4.7. Indeed, when the coefficients are constant, they all belong to $H^1_{\mathrm{per}}(\mathcal{D})$ and it is immediate to verify that (34) coincides with (49) when $n = 1$. Moreover, (48) with $n = 1$ yields $\mu_0 \lesssim |\rho|^2 2^{-3\ell_0}$, which is sharper than the upper bound $\mu_0 \lesssim |\rho|^2 2^{-2\ell_0}$ implied by Theorem 4.2 when the coefficients are constant. This small discrepancy is due to the upper bound $|(\rho\psi_{\ell,k}, \xi_0)| \le \|\rho\|_{L^2(\mathcal{D})} \|\psi_{\ell,k}\|_{L^2(\mathcal{D})}$ employed in the proof of Theorem 4.2, which is not sharp when ρ is constant. ■

The local a-coherence estimates of Theorem 4.7 answer issues (i), (ii), and (iii) in Section 3.4 and lead to the following

Theorem 4.9 (CORSING $\mathcal{WF}$ recovery, anisotropic wavelets)**.** *In Setting 3.1 and under the same assumptions as in Theorem 3.4, let $n > 1$ and $\eta, \rho \in \mathbb{R}$ and $\boldsymbol{\beta} \in \mathbb{R}^n$. Then, provided*

(i) $R \sim CsN^{3-\frac{2}{n}}$,

(ii) $m \gtrsim C 2^{n\ell_0} s(s\ln(eN/(2s)) + \ln(2s/\varepsilon))(\ln N + \ln s + \ln C)^n$,

(iii) $p_{\boldsymbol{q}} \propto \begin{cases} 2^{-(2+n)\ell_0}, & \boldsymbol{q} = \boldsymbol{0}, \\ 2^{-(n-\|\boldsymbol{q}\|_0)\ell_0} \min\left\{ \frac{2^{(3\|\boldsymbol{q}\|_0 - 2)L}\|\boldsymbol{q}\|_2^2}{|\widehat{\boldsymbol{q}}|^{\boldsymbol{4}}}, \frac{\|\boldsymbol{q}\|_2^2}{\|\boldsymbol{q}\|_\infty^2 |\widehat{\boldsymbol{q}}|^{\boldsymbol{1}}} \right\}, & \boldsymbol{q} \neq \boldsymbol{0}, \end{cases}$

where $N = 2^{nL}$ and $C = |\eta|^2 + \|\boldsymbol{\beta}\|_2^2 + |\rho|^2$, the CORSING $\mathcal{WF}$ method with $\Psi = \Psi^{\mathrm{ani}}$ recovers the best s-term approximation to u in expectation in the sense of estimate (20).

Proof. Let us consider the upper bound $\boldsymbol{\nu}$ defined according to (48) and (49). The definition of $p_{\boldsymbol{q}}$ in (iii) directly follows from the definition of $\nu_{\boldsymbol{q}}$. Now, we derive condition (i) by estimating the tail $\|\boldsymbol{\nu}|_{\mathcal{Q}^c}\|_1$ and using the upper bound corresponding to the first argument of the minimum in (49). Letting $Q := \lfloor R/2 \rfloor$ and splitting the sum involved in the 1-norm with respect to the sparsity levels of $\boldsymbol{q}$, we obtain

$$\begin{aligned} \|\boldsymbol{\nu}|_{\mathcal{Q}^c}\|_1 &\leq \sum_{\boldsymbol{q} \in \mathbb{Z}^n : \|\boldsymbol{q}\|_\infty > Q} \nu_{\boldsymbol{q}} = \sum_{s=1}^{n} \sum_{\boldsymbol{q} \in \mathbb{Z}^n : \|\boldsymbol{q}\|_\infty > Q,\ \|\boldsymbol{q}\|_0 = s} \nu_{\boldsymbol{q}} \\ &\leq C \sum_{s=1}^{n} \sum_{\boldsymbol{q} \in \mathbb{Z}^n : \|\boldsymbol{q}\|_\infty > Q,\ \|\boldsymbol{q}\|_0 = s} 2^{-(n-s)\ell_0} \frac{2^{(3s-2)L}\|\boldsymbol{q}\|_2^2}{|\widehat{\boldsymbol{q}}|^{\boldsymbol{4}}} \\ &= C \sum_{s=1}^{n} \binom{n}{s} 2^{-(n-s)\ell_0 + (3s-2)L} \underbrace{\sum_{\boldsymbol{r} \in \mathbb{Z}^s : \|\boldsymbol{r}\|_\infty > Q,\ \|\boldsymbol{r}\|_0 = s} \frac{\|\boldsymbol{r}\|_2^2}{|\boldsymbol{r}|^{\boldsymbol{4}}}}_{=:T(s)}. \end{aligned} \tag{53}$$

Now, we analyze $T(s)$. It is easy to verify that, for every $s \in [n]$, it holds

$$\{\boldsymbol{r} \in \mathbb{Z}^s : \|\boldsymbol{r}\|_\infty > Q,\ \|\boldsymbol{r}\|_0 = s\} = \bigcup_{k=1}^{s} \bigcup_{\boldsymbol{t} \in \{-1,1\}^s} \underbrace{\{\boldsymbol{r} \in \mathbb{Z}^s : |r_k| > Q,\ \operatorname{sign}(\boldsymbol{r}) = \boldsymbol{t}\}}_{=:X_k^{\boldsymbol{t}}}, \tag{54}$$

where $X_k^{\boldsymbol{t}} \subseteq \mathbb{Z}^s$ is the set of multi-indices having the k^{th} component larger than Q and with sign pattern $\boldsymbol{t}$ (see Figure 5 for $s = 2$ and $Q = 2$). Since the function $\boldsymbol{r} \mapsto \|\boldsymbol{r}\|_2^2/|\boldsymbol{r}|^{\boldsymbol{4}}$ is invariant with respect to $\operatorname{sign}(\boldsymbol{r})$ and to permutations of the components of $\boldsymbol{r}$, we can just consider the set $X_1^{\boldsymbol{1}}$. Moreover, the number of possible sets $X_k^{\boldsymbol{t}}$ in $\mathbb{Z}^s$ is $2^s s$. Therefore, we estimate

$$\begin{aligned} T(s) &\leq 2^s s \sum_{\boldsymbol{r} \in X_1^{\boldsymbol{1}}} \frac{\|\boldsymbol{r}\|_2^2}{|\boldsymbol{r}|^4} = 2^s s \sum_{r_1 > Q} \sum_{r_2 > 0} \cdots \sum_{r_s > 0} \sum_{k=1}^{s} \frac{1}{|\boldsymbol{r}|^2} \prod_{i \neq k} \frac{1}{r_i^2} \\ &= 2^s s \bigg(\Big(\sum_{r_1 > Q} \frac{1}{r_1^2} \Big) \prod_{i \neq 1} \Big(\sum_{r_i > 0} \frac{1}{r_i^4} \Big) + \sum_{k=2}^{s} \Big(\sum_{r_1 > Q} \frac{1}{r_1^4} \Big) \Big(\sum_{r_k > 0} \frac{1}{r_k^2} \Big) \prod_{i \notin \{1,k\}} \Big(\sum_{r_i > 0} \frac{1}{r_i^4} \Big) \bigg) \\ &\lesssim 2^s s \bigg(\frac{c^{s-1}}{Q} + s \frac{c^{s-2}}{Q^3} \bigg) \lesssim \frac{1}{Q}, \end{aligned}$$

where $c := \sum_{k \in \mathbb{N}} 1/k^4 < \infty$ and where the last inequality involves a constant depending exponentially on s and hence on n (since $s \leq n$). Plugging the estimate for $T(s)$ above into

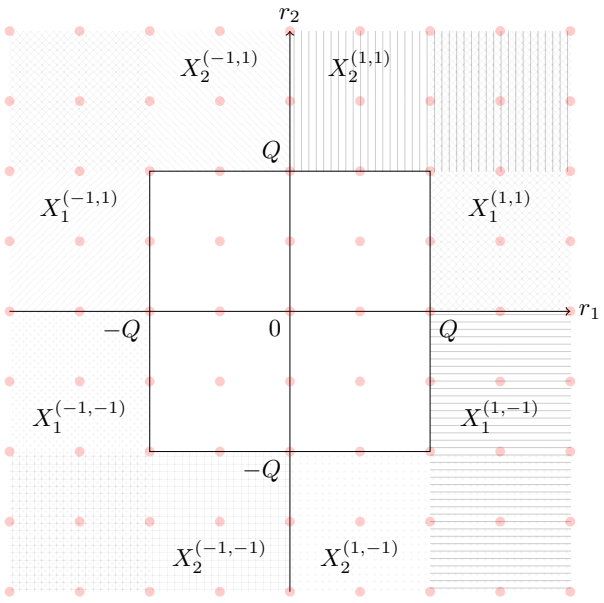

Figure 5: The sets $X_k^{\boldsymbol{t}} \subseteq \mathbb{Z}^s$ in (54) for $s = 2$ and $Q = 2$, restricted to $[-4, 4]^2$. Different textures correspond to different sets.

(53) yields

$$\begin{aligned}\|\boldsymbol{\nu}|_{\mathcal{Q}^c}\|_1 &\lesssim \frac{C}{Q}\sum_{s=1}^{n}\binom{n}{s}2^{-(n-s)\ell_0+(3s-2)L} = C\frac{2^{-n\ell_0-2L}}{\lfloor R/2\rfloor}\sum_{s=1}^{n}\binom{n}{s}2^{(\ell_0+3L)s}\\ &\lesssim C\frac{2^{-n\ell_0-2L}}{R}2^{n(\ell_0+3L)} = C\frac{2^{nL(3-\frac{2}{n})}}{R}.\end{aligned}$$

Condition (i) is obtained from $s\|\boldsymbol{\nu}|_{\mathcal{Q}^c}\|_1 \lesssim 1$. Of course, the above inequality also shows implicitly that $\|\boldsymbol{\nu}\|_1 < +\infty$.

Finally, in order to prove (ii), we estimate $\|\boldsymbol{\nu}|_{\mathcal{Q}}\|_1$. Employing the upper bound corresponding to (48) and to the second argument of the minimum in (49), and recalling that $\|\boldsymbol{q}\|_2^2 \le \|\boldsymbol{q}\|_0\|\boldsymbol{q}\|_\infty^2$ for every $\boldsymbol{q} \in \mathbb{Z}^n$, we obtain

$$\begin{aligned}\|\boldsymbol{\nu}|_{\mathcal{Q}}\|_1 &\le C\bigg(2^{-(2+n)\ell_0} + \sum_{s=1}^{n}2^{-(n-s)\ell_0}\sum_{\boldsymbol{q}\in\mathbb{Z}^n:\|\boldsymbol{q}\|_\infty\le Q,\ \|\boldsymbol{q}\|_0=s}\frac{\|\boldsymbol{q}\|_2^2}{\|\boldsymbol{q}\|_\infty^2|\widehat{\boldsymbol{q}}|^{\mathbf{1}}}\bigg)\\ &\lesssim C\sum_{s=1}^{n}2^{-(n-s)\ell_0}\sum_{\boldsymbol{q}\in\mathbb{Z}^n:\|\boldsymbol{q}\|_\infty\le Q,\ \|\boldsymbol{q}\|_0=s}\frac{s}{|\widehat{\boldsymbol{q}}|^{\mathbf{1}}} = C\sum_{s=1}^{n}\binom{n}{s}2^{-(n-s)\ell_0}s\sum_{\boldsymbol{r}\in\mathbb{Z}^s:\|\boldsymbol{r}\|_\infty\le Q,\ \|\boldsymbol{r}\|_0=s}\frac{1}{|\boldsymbol{r}|^{\mathbf{1}}}\\ &= C\sum_{s=1}^{n}\binom{n}{s}2^{-(n-s)\ell_0}s\bigg(\sum_{|q|\le Q,\ q\ne 0}\frac{1}{|q|}\bigg)^s \lesssim 2^{-n\ell_0}C\sum_{s=1}^{n}\binom{n}{s}2^{s\ell_0}s(\log R)^s \lesssim C(\log R)^n,\end{aligned}$$

which proves the theorem. □

4.2.2 Isotropic tensor product wavelets

As in the previous sections, we provide local a-coherence upper bounds (Theorem 4.11) and a recovery result (Theorem 4.12) for the CORSING $\mathcal{WF}$ method.

We start by proving auxiliary inequalities analogous to those in Lemma 4.5. We skip the proof, which follows the same arguments as in Lemma 4.5.

Lemma 4.10 (Auxiliary inequalities, isotropic wavelets)**.** *In Setting 3.1, let $n > 1$, $\ell \in \mathbb{N}$, with $\ell \geq \ell_0$, $\boldsymbol{k} \in (\mathbb{Z}/(2^\ell\mathbb{Z}))^n$, $\boldsymbol{e} \in \{0,1\}^n$, and $\boldsymbol{q} \in \mathbb{Z}^n$. Moreover, define* $\mathrm{zero}(\boldsymbol{q})$ *as in* (35)*. Then, the following inequalities hold:*

- *If $\boldsymbol{q} = \boldsymbol{0}$, we have*

$$\begin{aligned} |(\nabla\psi^{\mathrm{iso}}_{\ell,\boldsymbol{k},\boldsymbol{e}}, \nabla\xi_{\boldsymbol{0}})| &= 0, \\ |(\boldsymbol{\beta}\cdot\nabla\psi^{\mathrm{iso}}_{\ell,\boldsymbol{k},\boldsymbol{e}}, \xi_{\boldsymbol{0}})| &= 0, \quad \forall \boldsymbol{\beta}\in\mathbb{R}^n \\ |(\psi^{\mathrm{iso}}_{\ell,\boldsymbol{k},\boldsymbol{e}}, \xi_{\boldsymbol{0}})| &= \begin{cases} 2^{-n\ell/2}, & \text{if } \boldsymbol{e} = \boldsymbol{0}, \\ 0, & \text{if } \boldsymbol{e} \neq \boldsymbol{0}; \end{cases} \end{aligned}$$

- *if $\boldsymbol{q} \neq \boldsymbol{0}$ and $\mathrm{zero}(\boldsymbol{q}) \subseteq \mathrm{zero}(\boldsymbol{e})$, then, for every $\boldsymbol{\gamma} \in [0,2]^{\|\boldsymbol{q}\|_0}$, it holds*

$$\begin{aligned} |(\nabla\psi^{\mathrm{iso}}_{\ell,\boldsymbol{k},\boldsymbol{e}}, \nabla\xi_{\boldsymbol{q}})| &\lesssim 2^{(-\frac{n}{2}+2\|\boldsymbol{q}\|_0-\|\widehat{\boldsymbol{\gamma}}\|_1)\ell}|\widehat{\boldsymbol{q}}|^{\widehat{\boldsymbol{\gamma}}-\boldsymbol{2}}\|\boldsymbol{q}\|_2^2, \\ |(\boldsymbol{\beta}\cdot\nabla\psi^{\mathrm{iso}}_{\ell,\boldsymbol{k},\boldsymbol{e}}, \xi_{\boldsymbol{q}})| &\lesssim 2^{(-\frac{n}{2}+2\|\boldsymbol{q}\|_0-\|\widehat{\boldsymbol{\gamma}}\|_1)\ell}|\widehat{\boldsymbol{q}}|^{\widehat{\boldsymbol{\gamma}}-\boldsymbol{2}}\|\boldsymbol{\beta}\|_2\|\boldsymbol{q}\|_2, \quad \forall\boldsymbol{\beta}\in\mathbb{R}^n, \\ |(\psi^{\mathrm{iso}}_{\ell,\boldsymbol{k},\boldsymbol{e}}, \xi_{\boldsymbol{q}})| &\lesssim 2^{(-\frac{n}{2}+2\|\boldsymbol{q}\|_0-\|\widehat{\boldsymbol{\gamma}}\|_1)\ell}|\widehat{\boldsymbol{q}}|^{\widehat{\boldsymbol{\gamma}}-\boldsymbol{2}}, \end{aligned}$$

  where $\widehat{\boldsymbol{x}} := \boldsymbol{x}|_{\mathrm{supp}(\boldsymbol{q})} \in \mathbb{R}^{\|\boldsymbol{q}\|_0}$ for every $\boldsymbol{x} \in \mathbb{R}^n$ and the above inequalities hide constants depending exponentially on n;

- *if $\boldsymbol{q} \neq \boldsymbol{0}$ and $\mathrm{zero}(\boldsymbol{q}) \not\subseteq \mathrm{zero}(\boldsymbol{e})$, it holds*

$$\begin{aligned} |(\nabla\psi^{\mathrm{iso}}_{\boldsymbol{\ell},\boldsymbol{k}}, \nabla\xi_{\boldsymbol{q}})| &= 0, \\ |(\boldsymbol{\beta}\cdot\nabla\psi^{\mathrm{iso}}_{\boldsymbol{\ell},\boldsymbol{k}}, \xi_{\boldsymbol{q}})| &= 0, \quad \forall\boldsymbol{\beta}\in\mathbb{R}^n \\ |(\psi^{\mathrm{iso}}_{\boldsymbol{\ell},\boldsymbol{k}}, \xi_{\boldsymbol{q}})| &= 0. \end{aligned}$$

In the following theorem, we provide upper bounds to the local a-coherence for isotropic tensor product wavelets. We only outline a sketch of its proof, which is analogous to that of Theorem 4.7.

Theorem 4.11 (Local a-coherence upper bound, $n > 1$, isotropic wavelets)**.** *In Setting 3.1, let $n > 1$. Then, for every $\eta, \rho \in \mathbb{R}$ and $\boldsymbol{\beta} \in \mathbb{R}^n$, the following upper bounds hold:*

$$\mu_{\boldsymbol{0}} \lesssim |\rho|^2 2^{-(2+n)\ell_0}, \tag{55}$$

$$\mu_{\boldsymbol{q}} \lesssim \left(|\eta|^2 + \frac{\|\boldsymbol{\beta}\|_2^2}{\|\boldsymbol{q}\|_2^2} + \frac{|\rho|^2}{\|\boldsymbol{q}\|_2^4}\right) \min\left\{\frac{(1+2^{2(-\frac{n}{2}+2\|\boldsymbol{q}\|_0-1)L})\|\boldsymbol{q}\|_2^2}{|\widehat{\boldsymbol{q}}|^{\boldsymbol{4}}}, \frac{2^{-(n-\|\boldsymbol{q}\|_0)\ell_0}\|\boldsymbol{q}\|_2^2}{\|\boldsymbol{q}\|_\infty^2|\widehat{\boldsymbol{q}}|^{\boldsymbol{1}}}\right\}, \quad \forall \boldsymbol{q} \neq \boldsymbol{0}, \tag{56}$$

where the inequalities hide constants depending exponentially on n.

Proof. The upper bound (55) to $\mu_{\mathbf{0}}$ is easy to verify. Let us consider $\boldsymbol{q} \neq \mathbf{0}$. Employing the auxiliary inequalities of Lemma 4.10 and using arguments analogous to those in Theorem 4.7, we obtain, for every $\boldsymbol{\gamma} \in [0,2]^n$,

$$|a(\widehat{\psi}^{\mathrm{iso}}_{\boldsymbol{j}}, \widehat{\xi}_{\boldsymbol{q}})|^2 \lesssim \begin{cases} (|\eta|^2 + \frac{\|\boldsymbol{\beta}\|_2^2}{\|\boldsymbol{q}\|_2^2} + \frac{|\rho|^2}{\|\boldsymbol{q}\|_2^4}) 2^{2(-\frac{n}{2}+2\|\boldsymbol{q}\|_0 - \|\widehat{\boldsymbol{\gamma}}\|_1 - 1)\ell} |\widehat{\boldsymbol{q}}|^{2\widehat{\boldsymbol{\gamma}}-\mathbf{4}} \|\boldsymbol{q}\|_2^2, & \text{if } \mathrm{zero}(\boldsymbol{q}) \subseteq \mathrm{zero}(\boldsymbol{e}), \\ 0 & \text{otherwise.} \end{cases}$$

Analogously to Theorem 4.7, we obtain the upper bound corresponding to the first argument of the minimum in (56) by choosing $\widehat{\boldsymbol{\gamma}} = \mathbf{0}$ in the inequality above. Notice in particular that

$$\max_{\ell_0 \leq \ell < L} 2^{2(-\frac{n}{2}+2\|\boldsymbol{q}\|_0 - 1)\ell} \leq \max\{1, 2^{2(-\frac{n}{2}+2\|\boldsymbol{q}\|_0 - 1)L}\} \leq 1 + 2^{2(-\frac{n}{2}+2\|\boldsymbol{q}\|_0 - 1)L},$$

where, in the second expression, we have used that $2^{2(-\frac{n}{2}+\|\boldsymbol{q}\|_0 - 1)\ell} \leq 1$ when $-\frac{n}{2} + \|\boldsymbol{q}\|_0 - 1$ is negative. The upper bound corresponding to the second argument of the minimum in (56) is proved by letting $\boldsymbol{\gamma} = \widetilde{\boldsymbol{\gamma}}$, as in (51). Note that, in this case, we have $\|\widehat{\boldsymbol{\gamma}}\|_1 = \frac{3}{2}\|\boldsymbol{q}\|_0 - 1$ and, consequently,

$$\max_{\ell_0 \leq \ell < L} 2^{2(-\frac{n}{2}+2\|\boldsymbol{q}\|_0 - \|\widehat{\boldsymbol{\gamma}}\|_1 - 1)\ell} = \max_{\ell_0 \leq \ell < L} 2^{-(n-\|\boldsymbol{q}\|_0)\ell} \leq 2^{-(n-\|\boldsymbol{q}\|_0)\ell_0}.$$

This concludes the proof. □

Finally, we obtain the CORSING $\mathcal{WF}$ recovery theorem for isotropic tensor product wavelets solving issues (i), (ii), and (iii) in Section 3.4. The proof is analogous to that of Theorem 4.9 and therefore will be omitted.

Theorem 4.12 (CORSING $\mathcal{WF}$ recovery, isotropic wavelets)**.** *In Setting 3.1 and under the same assumptions as in Theorem 3.4, let $n > 1$ and let $\eta, \rho \in \mathbb{R}$ and $\boldsymbol{\beta} \in \mathbb{R}^n$. Then, provided*

(i) $R \sim CsN^{3-\frac{2}{n}}$,

(ii) $m \gtrsim C 2^{n\ell_0} s(s\ln(eN/(2s)) + \ln(2s/\varepsilon))(\ln N + \ln s + \ln C)^n$,

(iii) $p_{\boldsymbol{q}} \propto \begin{cases} 2^{-(2+n)\ell_0}, & \boldsymbol{q} = \mathbf{0}, \\ \min\left\{ \frac{(1+2^{2(-\frac{n}{2}+2\|\boldsymbol{q}\|_0-1)L})\|\boldsymbol{q}\|_2^2}{|\widehat{\boldsymbol{q}}|^{\mathbf{4}}}, \frac{2^{-(n-\|\boldsymbol{q}\|_0)\ell_0}\|\boldsymbol{q}\|_2^2}{\|\boldsymbol{q}\|_\infty^2 |\widehat{\boldsymbol{q}}|^{\mathbf{1}}} \right\}, & \boldsymbol{q} \neq \mathbf{0}, \end{cases}$

where $C = |\eta|^2 + \|\boldsymbol{\beta}\|_2^2 + |\rho|^2$, the CORSING $\mathcal{WF}$ method with $\Psi = \Psi^{\mathrm{iso}}$ recovers the best s-term approximation to u in expectation in the sense of estimate (20).

5 Numerical assessment

In this section, numerically investigate the reliability and robustness of the CORSING $\mathcal{WF}$ approach in different dimensions. In Section 5.1, we consider a 1D ADR equation with constant and nonconstant coefficients. As predicted by the theory, we show that CORSING $\mathcal{WF}$ is robust and reliable also in the case of nonconstant coefficients. In Section 5.2, we consider the 2D case and compare the performance of isotropic and anisotropic wavelets in different case studies. Moreover, it turns out that the nonuniform subsampling strategy based on the local a-coherence significantly outperforms uniform random subsampling. Finally, in Section 5.3, we consider a 3D case.

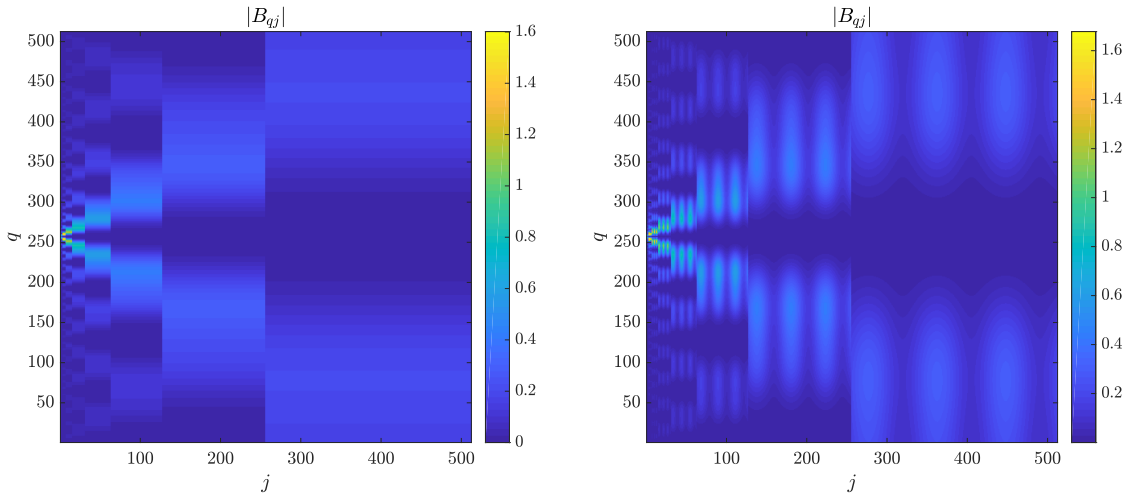

Figure 6: Absolute value of the entries $|B_{q,j}|$ of the stiffness matrix associated with the Petrov-Galerkin discretization for $\eta \equiv 1$ (left) and $\eta = 1 + 0.5\sin(6\pi x)$ (right).

All the numerical experiments have been performed in MATLAB® with the aid of OMP-Box for OMP [27, 28]. We have used MATLAB® R2017b version 9.3 64-bit on a MacBook Pro equipped with a 3 GHz Intel Core i7 processor and with 8 GB DDR3 RAM.

5.1 1D case with nonconstant diffusion

We consider a 1D equation with $\beta = 0$, $\rho = 1$ and let the diffusion coefficient vary. In particular, we consider $\eta(x_1) \equiv 1$ and $\eta(x_1) = 1 + 0.5\sin(6\pi x_1)$.

The Petrov-Galerkin stiffness matrix B. We compute the stiffness matrix B associated with the Petrov-Galerkin discretization and show the absolute value of the entries in Figure 6. We set $L = 9$, resulting in $N = 512$, and choose $R = N$. We observe that the oscillations of the diffusion coefficient only impact the stiffness matrix "horizontally". In particular, this does not impact the decay properties of the local a-coherence $\boldsymbol{\mu}$. The Wavelet-Fourier Petrov-Galerkin discretization of the ADR equation gives rise to a matrix with a comparable structure with respect to the matrices in [5, Figure 4]. This qualitatively confirms that the proposed discretization is suitable for compressed sensing. We also point out that the condition number of $B \in \mathbb{C}^{512\times 512}$ is very small, being 20.4 for the constant diffusion case and 28.2 for the nonconstant diffusion case, to be compared with 10^6 and $1.6 \cdot 10^6$, respectively, when the trial and test functions are not normalized with respect to the $H^1(\mathcal{D})$-norm.

Best s-term approximation. We consider the synthetic solution

$$u_1(x_1) = 1 + \exp\left(-\frac{(x_1 - 0.3)^2}{0.0005}\right) + \frac{1}{2}\cos(2\pi x_1), \quad 0 \le x_1 < 1.$$

This solution is smooth, with a global support in $[0, 1)$, and exhibits a bump close to the point $x = 0.3$. Moreover, it is periodic up to machine precision (see Figure 7 (left)). In Figure 7 (right), we show the wavelet coefficients and highlight in red the largest 50 ones in absolute value. The resulting relative best 50-term approximation error with respect to the $H^1(\mathcal{D})$-norm

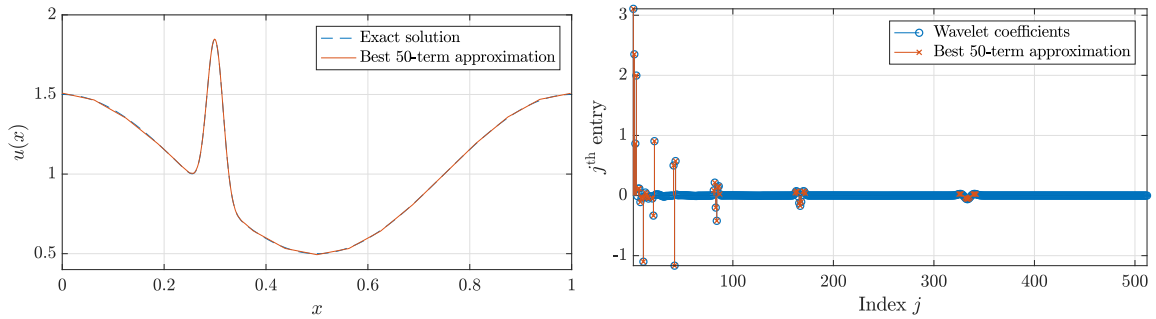

Figure 7: Plot of u_1 and of its best 50-term approximation $\widetilde{u}_1$ (left). Wavelet coefficients of u with the 50 largest in magnitude highlighted (right).

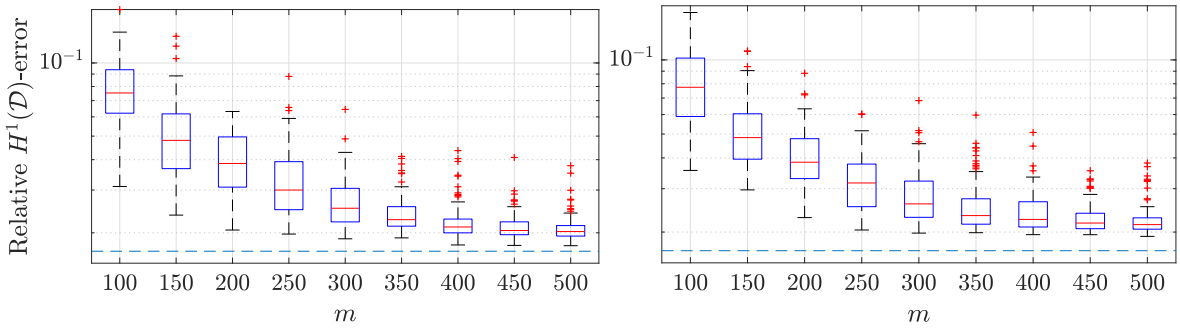

Figure 8: Relative recovery error as a function of the number of random tests, m, for a constant (left) and nonconstant (right) diffusion term. In dashed line, the relative best 50-term approximation error (57).

is

$$\frac{\|u_1 - \widetilde{u}_1\|_{H^1(\mathcal{D})}}{\|u_1\|_{H^1(\mathcal{D})}} \sim \frac{\|\boldsymbol{u}_1 - \widetilde{\boldsymbol{u}}_1\|_2}{\|\boldsymbol{u}_1\|_2} = 1.68 \cdot 10^{-2}, \tag{57}$$

where $\boldsymbol{u}_1$ is the vector of coefficients of u_1 with respect to the biorthogonal wavelet basis $\widehat{\Psi}$ (normalized with respect to the $H^1(\mathcal{D})$-norm), $\widetilde{\boldsymbol{u}}_1$ is the best 50-term approximation of $\boldsymbol{u}_1$, and $\widetilde{u}_1$ is the function corresponding to the wavelet coefficients in $\widetilde{\boldsymbol{u}}_1$.

Sensitivity of the recovery error to the number of test functions. In Figure 8, we show the box plot of the relative error between u_1 and the CORSING approximation, $\widehat{u}_1$, with respect to the $H^1(\mathcal{D})$-norm as a function of the number m of test functions. We fix $s = 50$ and let m vary from 100 to 500 ($N = 512$). The data are relative to 100 random runs of the CORSING procedure. We can appreciate that for both choices of η CORSING is able to reach a good accuracy (comparable to the best 50-term approximation error) for $m \geq 250$. The presence of a nonconstant diffusion term does not impact the performance of the method to a substantial extent. We only observe more outliers for the nonconstant diffusion.

Sensitivity of the recovery error to the sparsity. In Figure 9, we show the relative CORSING error with respect to the $H^1(\mathcal{D})$-norm as a function of s and compare it with the best s-term approximation error. The box plot is relative to 100 runs of CORSING. For each value of s varying between 5 and 50, we set $m = \lceil 2s \log(N) \rceil$. We remark that, for s large

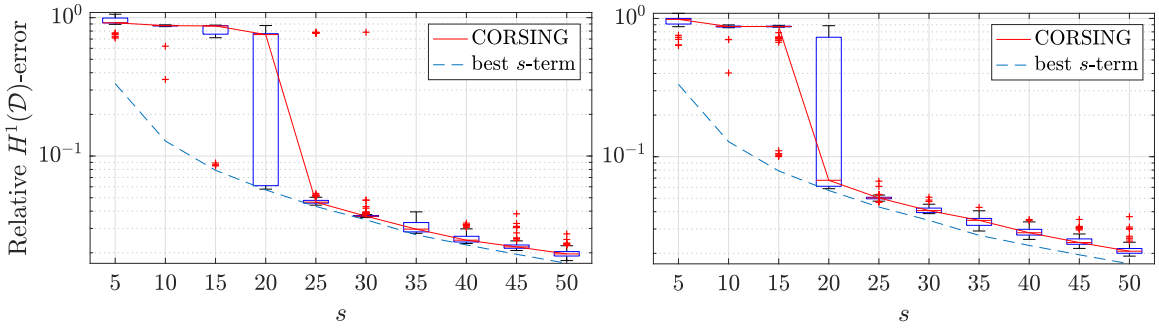

Figure 9: Relative recovery and best s-term approximation errors as functions of the sparsity s with constant (left) and nonconstant (right) diffusion term.

enough, the recovery error exhibits the same decay rate as the best s-term approximation error. No striking difference can be detected varying the diffusion.

5.2 2D case

We consider a 2D ADR problem over $\mathcal{D} = (0,1)^2$ with constant coefficients $\mu = \rho = 1$, and $\boldsymbol{b} = [1,1]^T$.

On the one hand, we compare the performance of anisotropic and isotropic wavelets on solutions that exhibit different spatial features. On the other hand, we show that nonuniform sampling strategy based on local a-coherence outperforms the uniform random subsampling.

Wavelet coefficients and best s-term approximation error. We consider the following solutions:

$$u_2(x_1,x_2) = \exp\left(-\frac{(x_1-0.3)^2}{0.0005}\right)\exp\left(-\frac{(x_2-0.4)^2}{0.0005}\right) + 2\exp\left(-\frac{(x_1-0.6).^2}{0.001}\right)\exp\left(-\frac{(x_2-0.5)^2}{0.005}\right),$$

$$u_3(x_1,x_2) = \exp\left(-\frac{(x_1-0.45)^2}{0.005}\right),$$

both periodic up to machine precision. The function u_2 exhibits two local Gaussian-shaped features, one isotropic around the point $(0.3, 0.4)$ and the other anisotropic, around $(0.6, 0.5)$. The function u_3 is purely anisotropic, having a Gaussian behavior along the x_1-direction and being constant along the x_2-direction. The functions u_2 and u_3 are shown in Figure 10 along with the corresponding anisotropic and isotropic 64×64 wavelet coefficients (with respect to $H^1(\mathcal{D})$-normalized wavelets). Letting $s = 100$, the relative best s-term approximation error with respect to the $H^1(\mathcal{D})$-norm of is 10^{-1} (anisotropic wavelets) and $6.6 \cdot 10^{-2}$ (isotropic wavelets) for u_2, and $8.2 \cdot 10^{-3}$ (anisotropic wavelets) and $1.2 \cdot 10^{-1}$ (isotropic wavelets) for u_3. As expected, anisotropic wavelets generate a very sparse representation of u_3. For u_2, the compression achieved by anisotropic and isotropic wavelets is comparable, in slight favor of isotropic wavelets.[7]

[7] Taking advantage of the norm equivalence property, the $H^1(\mathcal{D})$-norm is approximated using the ℓ^2-norm of the wavelet coefficients as in (57).

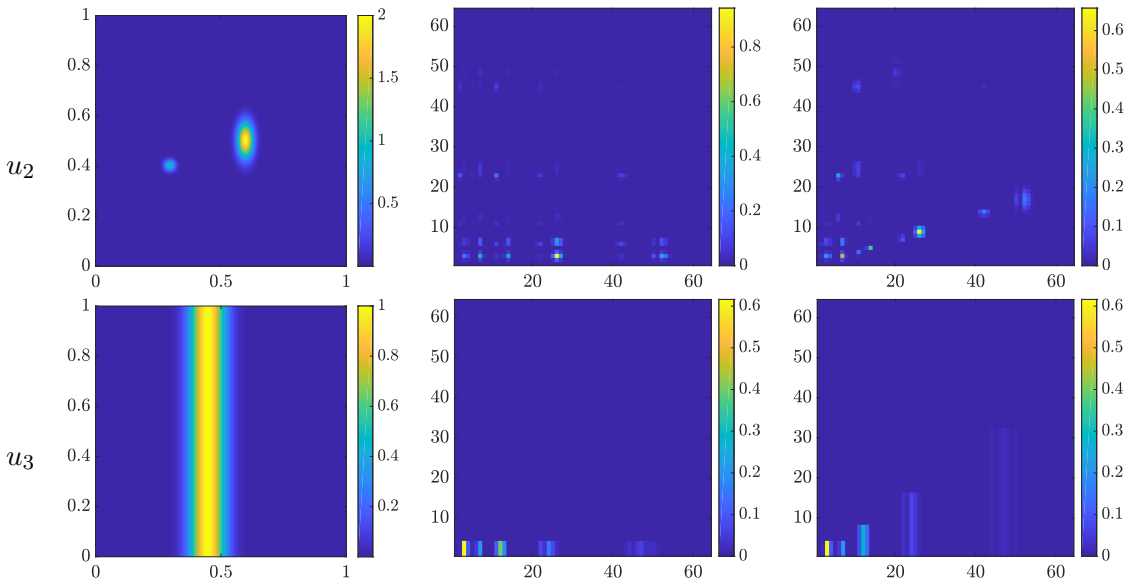

Figure 10: Contour plots (left) of functions u_2 (top) and u_3 (bottom) and corresponding wavelet coefficients with respect to anisotropic (center) and isotropic (right) tensor product wavelets.

Sensitivity of the recovery error to the number of test functions. We assess the performance of CORSING in the case of anisotropic and isotropic wavelets and compare uniform random subsampling ($\boldsymbol{p} \propto \mathbf{1}$) with the nonuniform subsampling based on the local a-coherence upper bound

$$\nu_{\boldsymbol{q}} = \min\left\{1, \frac{\|\mathbf{q}\|_{\mathbf{2}}^{\mathbf{2}}}{\|\boldsymbol{q}\|_\infty^2 |\widehat{\boldsymbol{q}}|^{\mathbf{1}}}\right\},$$

which is obtained from the upper bounds in Theorems 4.7 and 4.11. In particular, in (49) and (56), we consider the second argument of the minimum and use that $2^{-(n-\|\boldsymbol{q}\|_0)\ell_0} \leq 1$, while we use that $2^{-(2+n)\ell_0} \leq 1$ in (48) and (55). We set $\ell_0 = 2$, $L = 6$ (corresponding to $N = 2^{2L} = 4096$). As for the test space, we fix $R = 2^L$, corresponding to $M = N$. Although issue (i) in Theorems 4.9 and 4.12 suggests choosing $R \sim sN^{3-\frac{2}{n}}$, the choice $R = N$ turns out to be sufficient to have a well-conditioned Petrov-Galerkin discretization matrix B in practice. We set $s = 100$ and let $m = 100, 200, 300, 400, 500$. For each value of m, we run 100 tests of CORSING with uniform and nonuniform subsampling. We plot the relative recovery error measured with respect to the $H^1(\mathcal{D})$-norm as a function of the number of tests m in Figures 11 and 12 for u_2 and u_3, respectively. In the case of u_2, isotropic wavelets slightly outperform anisotropic wavelets. The benefit of nonuniform subsampling over uniform subsampling is evident both in terms of the probability of success (smaller boxes) and of accuracy. For u_3, anisotropic wavelets significantly outperform isotropic wavelets, thanks to the better compressibility of the solution. Moreover, uniform sampling fails to recover the solutions in both cases, whereas nonuniform sampling exhibits a convergent behavior. This experiment confirms the key role played by the local a-coherence for a successful implementation of the CORSING $\mathcal{WF}$ method.

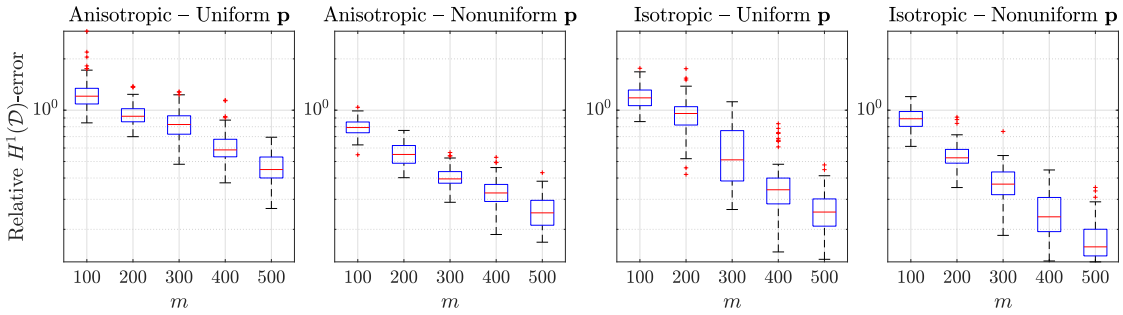

Figure 11: Box plot of the relative recovery error for the function u_2 with respect to the $H^1(\mathcal{D})$-norm as a function of the number of tests m with anisotropic and isotropic wavelets and uniform and nonuniform subsampling.

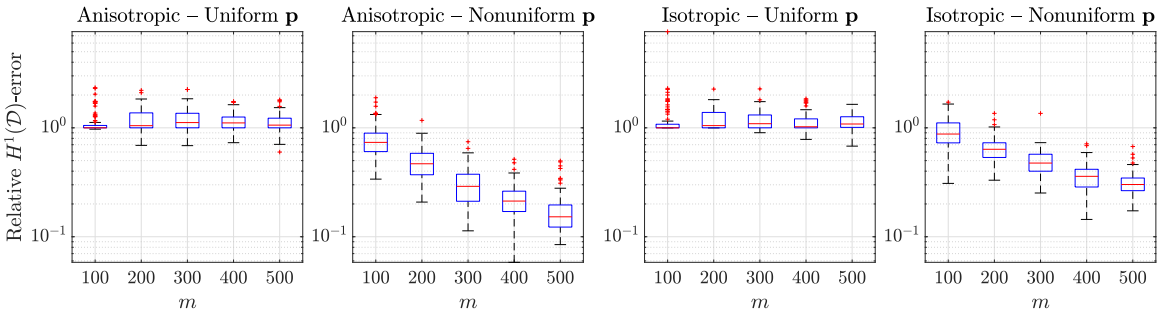

Figure 12: Box plot of the relative recovery error for the function u_3 with respect to the $H^1(\mathcal{D})$-norm as a function of the number of tests m with anisotropic and isotropic wavelets and uniform and nonuniform subsampling.

5.3 The 3D case

We validate the CORSING $\mathcal{WF}$ method on a 3D ADR problem on $\mathcal{D} = (0,1)^3$ with constant coefficients $\mu = \rho = 1$, $\boldsymbol{b} = [1,1,1]^T$. We consider exact solution

$$u_4(x_1,x_2,x_3) = \exp\left(-\frac{(x_1-0.4)^2}{0.005}\right)\exp\left(-\frac{(x_2-0.5)^2}{0.0005}\right)\exp\left(-\frac{(x_3-0.6)^2}{0.005}\right).$$

Function u_4 exhibits an anisotropic Gaussian-shaped feature centered at the point $(0.4, 0.5, 0.6)$. We compare anisotropic and isotropic wavelets.

Wavelet coefficients and best s-term approximation. We fix $\ell_0 = 2$, $L = 4$ (corresponding to a trial space of dimension $N = 2^{3L} = 4096$) and $s = 200$. The wavelet coefficients and the best s-term approximation are shown in Figure 13. The relative best s-term approximation error with respect to the $H^1(\mathcal{D})$-norm is $9.3 \cdot 10^{-2}$ for the anisotropic wavelets and $2.5 \cdot 10^{-2}$ for the isotropic wavelets. In this case, the isotropic tensorization is able to sparsify the function to a slightly better extent.

Sensitivity of the recovery error to the number of test functions. We compare anisotropic and isotropic wavelets and uniform and nonuniform subsampling. We set $m = 200, 300, 400, 500, 600$. The box plots corresponding to 100 runs of CORSING are shown in

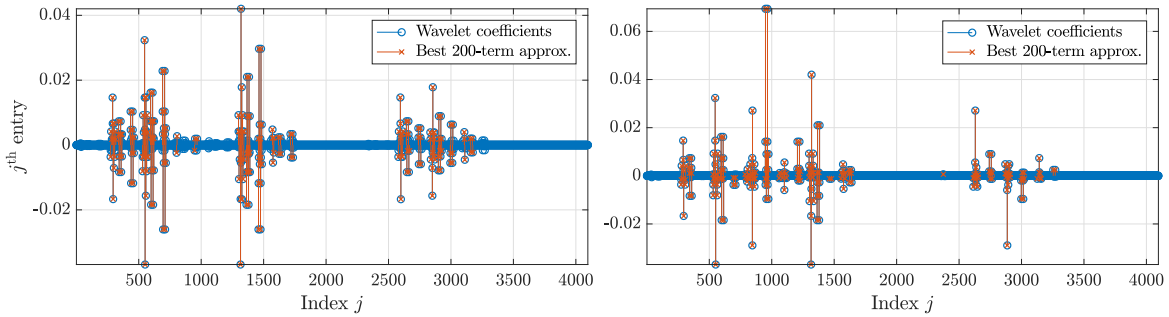

Figure 13: Wavelet coefficients of the function u_4 and best 200-term approximation for anisotropic (left) and isotropic (right) wavelets.

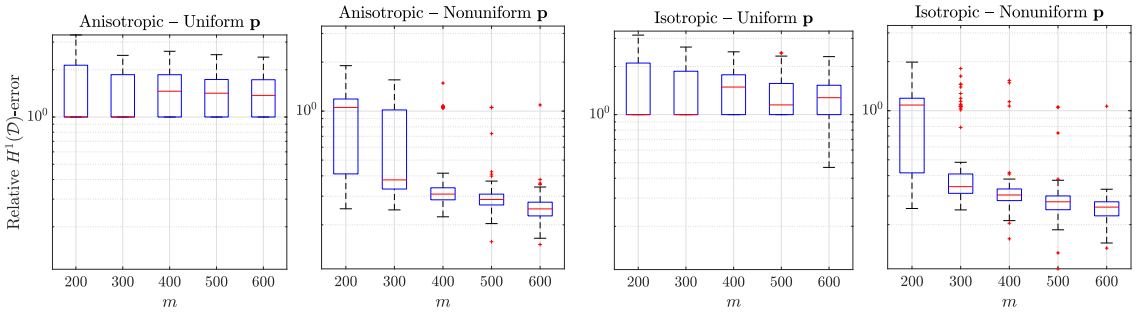

Figure 14: Box plot of the relative recovery error with respect to the $H^1(\mathcal{D})$-norm as a function of the number of tests m with anisotropic and isotropic wavelets and for uniform and nonuniform subsampling.

Figure 14. The performance of anisotropic and isotropic wavelets is similar. However, uniform subsampling is not able to recover the solution at all. Comparing this results with those of the 2D case (Figures 11 and 12), we note how using a bad probability measure (i.e., the uniform) deteriorates the performance of the method more heavily as the dimension of the domain increases. Finally, this experiment confirms that the theoretical analysis carried out in Section 4 turns out to be a useful tool for an effective implementation of the CORSING $\mathcal{WF}$ method.

6 Conclusions

We presented a wavelet-Fourier discretization technique for ADR equations based on the Petrov-Galerkin method and on the compressed sensing paradigm, called CORSING $\mathcal{WF}$. We carried out a theoretical analysis of the method, which hinges on the concept of local a-coherence and provides practical recipes for a successful implementation of the method. Numerical experiments confirm the robustness and reliability of the CORSING $\mathcal{WF}$ approach for n-dimensional ADR equations with $n = 1, 2, 3$. In particular, we showed that the method achieves recovery error comparable to the best s-term approximation error and that the sampling measure here proposed and based on the local a-coherence is able to successfully exploit the sparsity of the exact solution the discretization (in contrast to other randomization strategies such as uniform random subsampling).

Several open issues still remain to be investigated. First, understanding whether the sampling measure proposed in this paper is or not the "optimal" one (in some sense to be specified).

On the practical and computational viewpoint there is still a lot of work to be done. Developing an effective and optimized implementation for CORSING $\mathcal{WF}$ that takes advantage of the wavelet transform, the Fourier transform, and of the tensor product structure of the basis functions is still an open issue that has to be tackled to implement CORSING $\mathcal{WF}$ in dimension $n > 3$.

Acknowledgements

The first author acknowledges the support of the Postdoctoral Training Centre in Stochastics of the Pacifical Institute for the Mathematical Sciences (PIMS) and the Centre for Advanced Modelling Science (CADMOS) for the financial support. The first thanks Ben Adcock, Wolfgang Dahmen, and Holger Rauhut for very insightful discussions about approximation theory and compressed sensing. The fourth author acknowledges the research project GNCS-INdAM 2018 "Tecniche di Riduzione di Modello per le Applicazioni Mediche", which partially supported this research.

A Proof of Proposition 2.1

Employing the Cauchy-Schwarz inequality, it follows that

$$|a(u,v)| \leq \max\left\{\|\eta\|_{L^\infty(\mathcal{D})}, \sup_{\boldsymbol{x}\in\mathcal{D}}\|\boldsymbol{\beta}(\boldsymbol{x})\|_2, \|\rho\|_{L^\infty(\mathcal{D})}\right\}\|u\|_{H^1(\mathcal{D})}\|v\|_{H^1(\mathcal{D})}, \quad \forall u, v \in H^1_{\mathrm{per}}(\mathcal{D}),$$

which, in turn, implies the continuity of $a(\cdot,\cdot)$.

Let us now estimate the coercivity constant α. Observing that $\mathrm{Re}(\overline{u}\,\nabla u) = \frac{1}{2}\nabla(|u|^2)$ and using Green's formula, we have

$$\mathrm{Re}\left(\int_{\mathcal{D}}(\boldsymbol{\beta}\cdot\nabla u)\overline{u}\,\mathrm{d}\boldsymbol{x}\right) = \frac{1}{2}\int_{\mathcal{D}}\boldsymbol{\beta}\cdot\nabla(|u|^2)\,\mathrm{d}\boldsymbol{x} = -\frac{1}{2}\int_{\mathcal{D}}(\nabla\cdot\boldsymbol{\beta})|u|^2\,\mathrm{d}\boldsymbol{x} + \underbrace{\frac{1}{2}\int_{\partial\mathcal{D}}(\mathbf{n}\cdot\boldsymbol{\beta})|u|^2\,\mathrm{d}\boldsymbol{x}}_{=0},$$

where the boundary term vanishes thanks to the periodicity of $\boldsymbol{\beta}$ and u. As a consequence,

$$\mathrm{Re}\left(\int_{\mathcal{D}}(\boldsymbol{\beta}\cdot\nabla u + \rho u)\overline{u}\,\mathrm{d}\boldsymbol{x}\right) = \int_{\mathcal{D}}\left(-\frac{1}{2}\nabla\cdot\boldsymbol{\beta} + \rho\right)|u|^2\,\mathrm{d}\boldsymbol{x} \geq \zeta\|u\|^2_{L^2(\mathcal{D})}.$$

Therefore, we have the desired coercivity

$$|a(u,u)| \geq \mathrm{Re}(a(u,u)) \geq \eta_{\min}\|\nabla u\|^2_{L^2(\mathcal{D})} + \zeta\|u\|^2_{L^2(\mathcal{D})} \geq \min\{\eta_{\min},\zeta\}\|u\|^2_{H^1(\mathcal{D})}.$$

The thesis is directly implied by the Lax-Milgram lemma (see, for example, [22, Lemma 3.1]). □

B A wavelet digest

In this appendix, we show how to construct biorthogonal B-spline wavelets on the periodic interval (Section B.1) and provide more details on the norm equivalence property and the proof of Theorem 3.2 (Section B.2). We refer the reader to [14, 23, 31] for a more thorough presentation.

B.1 Biorthogonal B-spline wavelets on periodic intervals

The construction of biorthogonal wavelets on the periodic interval is presented in three steps. First, we illustrate the construction of biorthogonal wavelets on the real line. Then, we provide details about the special case of biorthogonal B-spline wavelets on the real line. Finally, we describe the periodization technique.

B.1.1 Biorthogonal wavelets on the real line

Consider a *scaling function* $\varphi \in L^2(\mathbb{R})$ refinable with *refinement filter* $\boldsymbol{a} \in \mathbb{R}^{\mathbb{Z}}$ such that

$$\varphi(x) = \sum_{k\in\mathbb{Z}} a_k \varphi(2x-k), \quad \forall x \in \mathbb{R}. \tag{58}$$

Notice that equation (58) defines φ up to a scaling factor. Here, we assume φ to be normalized such that $\|\varphi\|_{L^2(\mathbb{R})} \sim 1$. We also define a *mother wavelet* $\psi \in L^2(\mathbb{R})$ associated with a *wavelet filter* $\boldsymbol{b} \in \mathbb{R}^{\mathbb{Z}}$ such that

$$\psi(x) = \sum_{k\in\mathbb{Z}} b_k \varphi(2x-k), \quad \forall x \in \mathbb{R}. \tag{59}$$

Then, we define the associated diadically rescaled and translated versions

$$\varphi_{\ell,k}(x) = 2^{\ell/2}\varphi(2^\ell x - k), \quad \psi_{\ell,k}(x) = 2^{\ell/2}\psi(2^\ell x - k), \quad \forall x \in \mathbb{R}, \quad \forall \ell, k \in \mathbb{Z}. \tag{60}$$

The parameter ℓ denotes the dyadic level of the basis function and the index k encodes information about the spatial location. Moreover, definition (60) implies

$$\|\varphi_{\ell,k}\|_{L^2(\mathbb{R})} = \|\varphi\|_{L^2(\mathbb{R})} \sim 1, \quad \|\psi_{\ell,k}\|_{L^2(\mathbb{R})} = \|\psi\|_{L^2(\mathbb{R})} \sim 1.$$

We now illustrate the construction of a biorthogonal wavelet basis. Given $\ell_0 \in \mathbb{N}_0$, we introduce the sets

$$\Phi_\ell := \{\varphi_{\ell,k} : k \in \mathbb{Z}\}, \quad \widetilde{\Phi}_\ell := \{\widetilde{\varphi}_{\ell,k} : k \in \mathbb{Z}\}, \quad \forall \ell \in \mathbb{Z},\ \ell \geq \ell_0.$$

The *dual scaling function* $\widetilde{\varphi}$ is chosen such that $\{\Phi_\ell\}_{\ell\geq\ell_0}$ and $\{\widetilde{\Phi}_\ell\}_{\ell\geq\ell_0}$ generate two biorthogonal *Multiresolution Analyses* (MRAs) (see [21]), i.e., two collections of nested sets $\mathcal{S} = \{S_\ell\}_{\ell\geq\ell_0}$ and $\widetilde{\mathcal{S}} = \{\widetilde{S}_\ell\}_{\ell\geq\ell_0}$ in $L^2(\mathbb{R})$, such that

$$S_{\ell_0} \subseteq S_{\ell_0+1} \subseteq \cdots \subseteq L^2(\mathbb{R}), \quad \mathrm{clos}_{\|\cdot\|_{L^2(\mathbb{R})}}\left(\bigcup_{\ell\geq\ell_0} S_\ell\right) = L^2(\mathbb{R}), \quad S_\ell := \mathrm{clos}_{\|\cdot\|_{L^2(\mathbb{R})}}(\mathrm{span}(\Phi_\ell)),$$

$$\widetilde{S}_{\ell_0} \subseteq \widetilde{S}_{\ell_0+1} \subseteq \cdots \subseteq L^2(\mathbb{R}), \quad \mathrm{clos}_{\|\cdot\|_{L^2(\mathbb{R})}}\left(\bigcup_{\ell\geq\ell_0} \widetilde{S}_\ell\right) = L^2(\mathbb{R}), \quad \widetilde{S}_\ell := \mathrm{clos}_{\|\cdot\|_{L^2(\mathbb{R})}}(\mathrm{span}(\widetilde{\Phi}_\ell)),$$

where $\mathrm{clos}_{\|\cdot\|_{L^2(\mathbb{R})}}(X)$ denotes the closure of a set X in $L^2(\mathbb{R})$ and such that the following *biorthogonality* condition holds

$$(\varphi_{\ell,k}, \widetilde{\varphi}_{\ell,j})_{L^2(\mathbb{R})} = \delta_{(\ell,k),(\ell,j)}, \quad \forall \ell \in \mathbb{Z},\ \ell \geq \ell_0,\ \forall j,k \in \mathbb{Z}.$$

The wavelet family is then built by considering the so-called *detail spaces* $W_\ell, \widetilde{W}_\ell \subseteq L^2(\mathbb{R})$. These satisfy the relations

$$S_{\ell+1} = S_\ell \oplus W_\ell, \quad W_\ell \perp \widetilde{S}_\ell, \qquad \widetilde{S}_{\ell+1} = \widetilde{S}_\ell \oplus \widetilde{W}_\ell, \quad \widetilde{W}_\ell \perp S_\ell,$$

and are spanned by the sets

$$\Psi_\ell = \{\psi_{\ell,k} : k \in \mathbb{Z}\}, \quad \widetilde{\Psi}_\ell = \{\widetilde{\psi}_{\ell,k} : k \in \mathbb{Z}\}, \quad \forall \ell \in \mathbb{Z},\ \ell \geq \ell_0,$$

where $\widetilde{\psi}$ is the *dual mother wavelet*. Namely, $W_\ell = \mathrm{clos}_{\|\cdot\|_{L^2(\mathbb{R})}}(\mathrm{span}(\Psi_\ell))$, and $\widetilde{W}_\ell = \mathrm{clos}_{\|\cdot\|_{L^2(\mathbb{R})}}(\mathrm{span}(\widetilde{\Psi}_\ell))$. These bases are said to be biorthogonal if

$$(\varphi_{\ell,k}, \widetilde{\psi}_{\ell,j})_{L^2(\mathbb{R})} = (\widetilde{\varphi}_{\ell,k}, \psi_{\ell,j})_{L^2(\mathbb{R})} = 0, \quad (\psi_{\ell,k}, \widetilde{\psi}_{\ell,j}) = \delta_{(\ell,k),(\ell,j)}, \quad \forall k, j \in \mathbb{Z}.$$

Then, fixing the coarsest level ℓ_0 and the maximum level $L \in \mathbb{N}$ with $L > \ell_0$, the corresponding wavelet basis is defined as

$$\Psi := \Phi_{\ell_0} \cup \bigcup_{\ell=\ell_0}^{L-1} \Psi_\ell.$$

Notice that $\mathrm{span}(\Psi) = S_L$. In the above notation, the dependence of Ψ on ℓ_0 and L is hidden for the sake of simplicity.

B.1.2 B-spline biorthogonal wavelets on the real line

We now detail how to build an explicit example of biorthogonal wavelets on the real line, i.e., the biorthogonal B-spline wavelets [13].

The set of *cardinal B-splines* $N_d : \mathbb{R} \to \mathbb{R}$ is recursively defined as

$$N_d(x) := \int_0^1 N_{d-1}(x-t)\,\mathrm{d}t, \quad N_1(x) := \chi_{[0,1)}(x), \quad \forall x \in \mathbb{R}, \quad \forall d \in \mathbb{N}, d \geq 2.$$

The *centralized B-splines* ${}_d\varphi : \mathbb{R} \to \mathbb{R}$ are then defined as

$${}_d\varphi(x) := N_d\Big(x + \Big\lfloor \frac{d}{2} \Big\rfloor\Big), \quad \forall x \in \mathbb{R},$$

with support in $(-\lfloor \frac{d}{2} \rfloor, \lceil \frac{d}{2} \rceil)$. The function ${}_d\varphi$ has *polynomial exactness* of order d (i.e., all polynomials of degree less than d can be represented as linear combinations of the shifted functions $\{{}_d\varphi(\cdot - k) : k \in \mathbb{Z}\}$) and fulfills the refinement equation (58) with scaling filter $\boldsymbol{a}$ defined as

$$a_k = 2^{1-d}\binom{d}{k + \lfloor \frac{d}{2} \rfloor}, \quad \forall k = -\lfloor \tfrac{d}{2} \rfloor, \dots, \lceil \tfrac{d}{2} \rceil.$$

Following the construction in [13], we consider the dual scaling function $\widetilde{\varphi}$, also denoted by ${}_{d,\widetilde{d}}\widetilde{\varphi}$, of polynomial exactness $\widetilde{d}$, compactly supported on $[-\lfloor \frac{d}{2} \rfloor - \widetilde{d} + 1, \lceil \frac{d}{2} \rceil + \widetilde{d} - 1]$, refinable with some scaling filter $\widetilde{\boldsymbol{a}}$ (the filters $\boldsymbol{a}$ and $\widetilde{\boldsymbol{a}}$ corresponding to different values of d and $\widetilde{d}$ can be found in [31, Table 2.6 and 2.7]). Moreover, choosing the wavelet filter $\boldsymbol{b}$ as

$$b_k = (-1)^k \widetilde{a}_{1-k}, \quad \widetilde{b}_k = (-1)^k a_{1-k},$$

yields a biorthogonal wavelet basis Ψ (see [31, Proposition 5.5]).

B.1.3 Periodization

Finally, we make use of *periodization* in order to adapt the biorthogonal wavelet construction on the real line to the case of a periodic interval (see also [14, Section 4.3]). The idea is to replace all the expressions of the form $g_{\ell,k} = 2^{\ell/2} g(2^\ell \cdot - k)$ (where $g : \mathbb{R} \to \mathbb{R}$ is arbitrary) with

$$g^{\mathrm{per}}_{\ell,k}(x) = 2^{\ell/2} \sum_{j \in \mathbb{Z}} g(2^\ell (x + j) - k), \quad \forall x \in \mathbb{R}. \tag{61}$$

The main difference is that at level ℓ the spatial index is $\mathbb{Z}/(2^\ell \mathbb{Z})$ (the ring of integers modulo 2^ℓ). For the sake of convenience, we also identify $\mathbb{Z}/(2^\ell \mathbb{Z})$ with the set of canonical representatives, i.e.,

$$\mathbb{Z}/(2^\ell \mathbb{Z}) \equiv \{0, 1, \dots, 2^\ell - 1\}, \quad \forall \ell \in \mathbb{N}.$$

With this convention, $\mathbb{Z}/(2^\ell \mathbb{Z}) \subseteq \mathbb{Z}$. We have

$$\Phi^{\mathrm{per}}_\ell := \{\varphi^{\mathrm{per}}_{\ell,k} : k \in \mathbb{Z}/(2^\ell \mathbb{Z})\}, \quad \Psi^{\mathrm{per}}_\ell := \{\psi^{\mathrm{per}}_{\ell,k} : k \in \mathbb{Z}/(2^\ell \mathbb{Z})\},$$

and, analogously to the case of the real line, we define

$$\Psi^{\mathrm{per}} := \Phi^{\mathrm{per}}_{\ell_0} \cup \bigcup_{\ell=\ell_0}^{L-1} \Psi^{\mathrm{per}}_\ell.$$

The same holds for the dual bases $\widetilde{\Phi}_\ell$ and $\widetilde{\Psi}_\ell$.

B.2 Norm equivalence

For the purposes of this paper, the most important feature of biorthogonal wavelets is the norm equivalence property. Theorem 3.2 states that, up to a diagonal rescaling, the anisotropic and isotropic tensor product wavelets introduced in Section 3.1 form a Riesz basis with respect to Sobolev norms of suitable orders. The aim of this section is to prove this theorem.

Before proving Theorem 3.2, we need some auxiliary results. The first technical lemma gives sufficient conditions for Jackson's inequality (see [31, Proposition 2.20] for the proof). Notice that Jackson's and Bernstein's inequalities are properties of the MRA (see Section 3.1), and not of the wavelet spaces. In the following, $\mathcal{P}_d$ will denote the space of polynomials of one variable of degree less than or equal to d.

Lemma B.1 (Jackson's inequality)**.** *Let $I_{\ell,k} := \operatorname{supp}(\varphi_{\ell,k}) \cup \operatorname{supp}(\widetilde{\varphi}_{\ell,k})$ and assume that the following conditions hold for every $\ell \geq \ell_0$ and for every $k \in \mathbb{Z}$:*

$$|\{j \in \mathbb{Z} : I_{\ell,k} \cap I_{\ell,j} \neq \emptyset\}| \lesssim 1, \tag{62}$$

$$|I_{\ell,k}| \lesssim 2^{-\ell}, \tag{63}$$

$$\|\varphi_{\ell,k}\|_{L^2(\mathbb{R})} \lesssim 1, \quad \|\widetilde{\varphi}_{\ell,k}\|_{L^2(\mathbb{R})} \lesssim 1. \tag{64}$$

(Note that the upper bounds in (62) and (64) are independent of ℓ). Moreover, suppose that there exists $d \in \mathbb{N}$ such that

$$\mathcal{P}_{d-1} \subseteq \operatorname{span}(\Phi_0) = \left\{ \sum_{k \in \mathbb{Z}} c_k \varphi_{0,k} : \mathbf{c} \in \ell(\mathbb{Z}) \right\}. \tag{65}$$

Then, Jackson's inequality holds for every $s \leq d$ and $\ell \geq \ell_0$, i.e.

$$\inf_{v_\ell \in S_\ell} \|g - v_\ell\|_{L^2(\mathbb{R})} \lesssim 2^{-s\ell} |g|_{H^s(\mathbb{R})}, \quad \forall g \in H^s(\mathbb{R}). \tag{66}$$

The second technical result provides sufficient conditions for Bernstein's inequality to hold (the result is proved in [31, Lemma 5.11]).

Lemma B.2 (Bernstein's inequality)**.** *Let $\varphi \in L^2(\mathbb{R}^n)$ be a compactly supported refinable function and define*

$$\gamma := \sup\{s \in \mathbb{R} : \varphi \in H^s(\mathbb{R}^n)\}.$$

Then, the following inequality

$$\|v_\ell\|_{H^s(\mathbb{R}^n)} \lesssim 2^{s\ell} \|v_\ell\|_{L^2(\mathbb{R}^n)}, \quad \forall v_\ell \in S_\ell, \tag{67}$$

holds for every $0 \leq s \leq \gamma$.

In the following result, we see that biorthogonal wavelet families corresponding to an MRAs that verify Jackson's and Bernstein's inequalities have the norm equivalence property. This result is proved in [31, Theorem 5.12].

Lemma B.3 (Jackson + Bernstein + Biorthogonality $\Rightarrow$ Norm equivalence)**.** *Let $\mathcal{S} = \{S_\ell\}$ and $\widetilde{\mathcal{S}} = \{\widetilde{S}_\ell\}$ be dual MRAs with associated biorthogonal wavelets Ψ, $\widetilde{\Psi}$, respectively, such that Jackson's inequality* (66) *and Bernstein's inequality* (67) *hold for $\mathcal{S}$ with $d_{\mathcal{S}}$, $\gamma_{\mathcal{S}}$ and for $\widetilde{\mathcal{S}}$ with $d_{\widetilde{\mathcal{S}}}$, $\gamma_{\widetilde{\mathcal{S}}}$, respectively. Then, the following norm equivalence holds:*

$$\left\| \sum_{\ell,k} c_{\ell,k} \psi_{\ell,k} \right\|^2_{H^s(\mathbb{R})} \sim \sum_{\ell,k} 2^{2s\ell} |c_{\ell,k}|^2,$$

for every $s \in [-\min\{d_{\widetilde{\mathcal{S}}}, \gamma_{\widetilde{\mathcal{S}}}\}, \min\{d_{\mathcal{S}}, \gamma_{\mathcal{S}}\}]$.

We are now in a position to prove Theorem 3.2.

Proof of Theorem 3.2. We start by showing that the 1D norm equivalence[8]

$$\left\| \sum_{\ell,k} c_{\ell,k} \psi_{\ell,k} \right\|^2_{H^s(\mathcal{D})} \sim \sum_{\ell,k} 2^{2s\ell} |c_{\ell,k}|^2, \tag{68}$$

holds for every $0 \leq s \leq d-1$, by applying Lemmas B.1, B.2, and B.3.

First, hypotheses (62),(63), and (64) of Lemma B.1 are satisfied by construction. Regarding hypothesis (65), we notice that, thanks to the Curry-Schoenberg Theorem (see, e.g., [16, Theorem 3.1] or [15, Theorem 44]), linear combination of integer translates of ${}_d\varphi$ generates the space $\mathcal{P}_{d-1}$. Therefore we can apply Lemma B.1 for the primal MRA with $d_{\mathcal{S}} = d$. The same considerations hold for ${}_{d,\widetilde{d}}\varphi$ (see [31, Proposition 2.13]). In particular, ${}_{d,\widetilde{d}}\varphi$ is exact of order $\widetilde{d}$, thus Lemma B.1 applies to the dual MRA with $d_{\widetilde{\mathcal{S}}} = \widetilde{d} \geq 0$.[9]

[8] The norm equivalence (68) holds also for negative values of the exponent s, since the regularity of ${}_{d,\widetilde{d}}\varphi$ increases with $\widetilde{d}$ (see [31, Proposition 2.13]). However, quantifying this statement precisely is not trivial.

[9] Notice that the coefficients $\mathbf{c}$ in expression (65) need be in $\ell(\mathbb{Z})$, since, e.g., the constant function is given by the sum $\sum_{k\in\mathbb{Z}} {}_d\varphi(\cdot - k)$ and $\mathbf{c} = \mathbf{1} \in \ell(\mathbb{Z}) \setminus \ell^2(\mathbb{Z})$.

Moreover, we have that ${}_{d,\widetilde{d}}\varphi$ is compactly supported and belongs to $L^2(\mathbb{R})$, hence Bernstein's inequality holds for the dual MRA with $\gamma_{\widetilde{\mathcal{S}}} = 0$, thanks to Lemma B.2. Moreover, since ${}_{d}\varphi \in H^{d-\frac{1}{2}-\varepsilon}(\mathbb{R})$, for every $\varepsilon > 0$, Bernstein's inequality holds for the primal MRA with $\gamma_{\mathcal{S}} = d - 1/2 - \varepsilon \geq d - 1$ (for $\varepsilon \leq 1/2$).

Now, since Jackson's and Bernstein's inequalities hold, we apply Lemma B.3 and employ periodization to obtain (68). To conclude, the multi-dimensional norm equivalence (10) is guaranteed thanks to [19, Theorem 3.3] for the anisotropic case and to [14, Remark 5.9] for the isotropic case.

□

References

[1] R. A. Adams and J. F. Fournier. *Sobolev spaces*, volume 140. Academic press, 2003.

[2] B. Adcock, A. Bao, and S. Brugiapaglia. Correcting for unknown errors in sparse high-dimensional function approximation. *arXiv preprint arXiv:1711.07622*, 2017.

[3] B. Adcock, S. Brugiapaglia, and C. G. Webster. Compressed sensing approaches for polynomial approximation of high-dimensional functions. In H. Boche, G. Caire, R. Calderbank, M. März, G. Kutyniok, and R. Mathar, editors, *Compressed Sensing and its Applications: Second International MATHEON Conference 2015*, pages 93–124. Springer International Publishing, Cham, 2017.

[4] B. Adcock and A. C. Hansen. Generalized Sampling and Infinite-Dimensional Compressed Sensing. *Found. Comput. Math.*, 16(5):1263–1323, 2016.

[5] B. Adcock, A. C. Hansen, C. Poon, and B. Roman. Breaking the coherence barrier: A new theory for compressed sensing. In *Forum of Mathematics, Sigma*, volume 5. Cambridge University Press, 2017.

[6] J.-L. Bouchot, H. Rauhut, and C. Schwab. Multi-level Compressed Sensing Petrov-Galerkin discretization of high-dimensional parametric PDEs. *arXiv preprint arXiv:1701.01671*, 2017.

[7] S. Brugiapaglia. *COmpRessed SolvING: sparse approximation of PDEs based on compressed sensing*. PhD thesis, MOX - Politecnico di Milano, 2016.

[8] S. Brugiapaglia. A compressive spectral collocation method for the diffusion equation under the restricted isometry property. *arXiv preprint arXiv:1807.06606*, 2018.

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

[10] S. Brugiapaglia, F. Nobile, S. Micheletti, and S. Perotto. A theoretical study of compressed solving for advection-diffusion-reaction problems. *Math. Comp.*, 87(309):1–38, 2018.

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

[12] A. Chkifa, N. Dexter, H. Tran, and C.G. Webster. Polynomial approximation via compressed sensing of high-dimensional functions on lower sets. *Math. Comp.*, 87(311):1415–1450, 2018.

[13] A. Cohen, I. Daubechies, and J.-C. Feauveau. Biorthogonal bases of compactly supported wavelets. *Comm. Pure Appl. Math.*, 45(5):485–560, 1992.

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

[15] C. de Boor. *A Practical Guide to Splines*. Appl. Math. Sci. Springer New York, 2001.

[16] R. A. DeVore and G G Lorentz. *Constructive Approximation*. Die Grundlehren der mathematischen Wissenschaften in Einzeldarstellungen. Springer, 1993.

[17] D. L. Donoho. Compressed Sensing. *IEEE Trans. Inf. Theory*, 52(4):1289–1306, 2006.

[18] A. Doostan and H. Owhadi. A non-adapted sparse approximation of PDEs with stochastic inputs. *J. Comput. Phys.*, 230(8):3015–3034, 2011.

[19] M. Griebel and S. Knapek. Optimized general sparse grid approximation spaces for operator equations. *Math. Comp.*, 78(268):2223–2257, 2009.

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

[21] S. G. Mallat. Multiresolution approximations and wavelet orthonormal bases of L^2(R). *Trans. Amer. Math. Soc.*, 315(1):69–87, 1989.

[22] J. Nečas. *Direct methods in the theory of elliptic equations*. Springer Science & Business Media, 2011.

[23] R. Pabel. *Adaptive Wavelet Methods for Variational Formulations of Nonlinear Elliptic PDEs on Tensor-Product Domains*. Logos Verlag Berlin GmbH, 2015.

[24] A. Quarteroni and A. Valli. *Numerical Approximation of Partial Differential Equations*, volume 23 of *Springer Series in Computational Mathematics*. Springer-Verlag, Berlin, 2008.

[25] H. Rauhut and C. Schwab. Compressive sensing Petrov-Galerkin approximation of high-dimensional parametric operator equations. *Math. Comp.*, 86(304):661–700, 2017.

[26] H. Rauhut and R. Ward. Sparse legendre expansions via ℓ_1-minimization. *J. Approx. Theory*, 164(5):517–533, 2012.

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

[28] 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.

[29] M. E. Taylor. *Partial Differential Equations I: Basic Theory*. Appl. Math. Sci. 115. Springer-Verlag New York, 2nd edition, 2011.

[30] R. Temam. *Navier-Stokes equations and nonlinear functional analysis*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, second edition, 1995.

[31] K. Urban. *Wavelet Methods for Elliptic Partial Differential Equations*. Numer. Math. Sci. Comput. OUP Oxford, 2008.

[32] X. Yang and G.E. Karniadakis. Reweighted ℓ^1 minimization method for stochastic elliptic differential equations. *J. Comput. Phys.*, 248:87–108, 2013.

MOX Technical Reports, last issues

Dipartimento di Matematica
Politecnico di Milano, Via Bonardi 9 - 20133 Milano (Italy)

62/2018 Perotto, S.; Carlino, M.G.; Ballarin, F.
Model reduction by separation of variables: a comparison between Hierarchical Model reduction and Proper Generalized Decomposition

61/2018 Zakerzadeh, R.; Zunino P.
A Computational Framework for Fluid-Porous Structure Interaction with Large Structural Deformation

59/2018 Martino, A.; Guatteri, G.; Paganoni, A. M.
Multivariate Hidden Markov Models for disease progression

58/2018 Ferro, N.; Micheletti, S.; Perotto, S.
A sequential coupling of shape and topology optimization for structural design

56/2018 Antonietti, P.F.; Manzini, G.; Verani, M.
The conforming virtual element method for polyharmonic problems

57/2018 Ferro, N.; Micheletti, S.; Perotto, S.
POD-assisted strategies for structural topology optimization

53/2018 Giantesio, G.; Musesti, A.; Riccobelli, D.
A comparison between active strain and active stress in transversely isotropic hyperelastic materials

55/2018 Cerroni, D.; Laurino, F.; Zunino, P.
Mathematical analysis, finite element approximation and numerical solvers for the interaction of 3D reservoirs with 1D wells

52/2018 Possenti, L.; di Gregorio, S.; Gerosa, F.M.; Raimondi, G.; Casagrande, G.; Costantino, M.L.; Z
A computational model for microcirculation including Fahraeus-Lindqvist effect, plasma skimming and fluid exchange with the tissue interstitium

54/2018 Dal Santo, N.; Deparis, S.; Manzoni, A.; Quarteroni, A.
Multi space reduced basis preconditioners for parametrized Stokes equations