IMA Journal of Numerical Analysis (2021) **41**, 2744–2781
https://doi.org/10.1093/imanum/draa036
Advance Access publication on 3 September 2020

Wavelet–Fourier CORSING techniques for multidimensional advection–diffusion–reaction equations

S. Brugiapaglia*
Department of Mathematics and Statistics, Concordia University, Montréal, QC H3G 1M8, Canada
*Corresponding author: simone.brugiapaglia@concordia.ca

S. Micheletti
MOX, Dipartimento di Matematica, Politecnico di Milano, 20133 Milano, Italy
stefano.micheletti@polimi.it

F. Nobile
MATHICSE-CSQI, École Polytechnique Fédérale de Lausanne, Lausanne, CH-1015, Switzerland
fabio.nobile@epfl.ch

and

S. Perotto
MOX, Dipartimento di Matematica, Politecnico di Milano, 20133 Milano, Italy
simona.perotto@polimi.it

[Received on 18 April 2019; revised on 2 July 2020]

We present and analyze a novel wavelet–Fourier technique for the numerical treatment of multidimensional advection–diffusion–reaction equations based on the COmpRessed SolvING (CORSING) paradigm. 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 to approximate the solution to the partial differential equation (PDE). In this paper we provide the first rigorous recovery error bounds and effective recipes for the implementation of the CORSING technique in the multidimensional setting. Our theoretical analysis relies on new estimates for the local *a*-coherence, which measures interferences between wavelet and Fourier basis functions with respect to the metric induced by the PDE operator. The stability and robustness of the proposed scheme are shown by numerical illustrations in the one-, two- and three-dimensional cases.

Keywords: compressed sensing; Petrov-Galerkin method; biorthogonal wavelets; advection-diffusion-reaction equation; local coherence.

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 Brugiapaglia (2016), Brugiapaglia *et al.* (2015, 2018) for the solution of linear PDEs set in Hilbert spaces and combines the Petrov–Galerkin discretization techniques with compressed sensing (Candès *et al.*, 2006;

Donoho, 2006). 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 Brugiapaglia *et al.* (2018) 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 multidimensional advection–diffusion–reaction (ADR) equations on a torus, and we employ biorthogonal wavelets as the 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 shown here can be extended to the nonperiodic case by considering biorthogonal wavelets on the interval (and on the hypercube) (see Dahmen, 1997; Urban, 2008; Pabel, 2015) as the trial functions basis and the sine basis $\{\sin(\pi kx)\}_{k\in\mathbb{N}}$ (or a tensorized version of it) as the test functions basis. However, we decided to stick to the periodic case to make the theoretical exposition free of an excessive quantity of technical detail 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. We also observe that the construction of wavelets over general domains is not a straightforward task. Hence, CORSING $\mathcal{WF}$ shares this difficulty with classic wavelet methods for PDEs. We refer to Brugiapaglia *et al.* (2020b) for an extension and implementation of the CORSING method to domains with more general geometries via the isogeometric analysis principle (Hughes *et al.*, 2005).

The choice of biorthogonal wavelet functions is motivated by the need to work with trial and test functions bases that satisfy the *Riesz basis* property (Brugiapaglia *et al.*, 2018). Roughly speaking this assumption guarantees control on the condition number of the Petrov–Galerkin discretization matrix, which is crucial for a successful application of the compressed sensing paradigm (see, e.g., Brugiapaglia, 2016, Theorem 1.21 or Adcock *et al.*, 2019a, Theorem 3.6). In Brugiapaglia *et al.* (2018) the authors considered the hierarchical basis of hat functions and the sine function basis as trial and test bases, respectively. While this choice ensures the Riesz basis property in the one-dimensional setting, tensorizing these two bases breaks the Riesz basis property in the multidimensional framework. Indeed, the tensorization of the sine functions basis is a Riesz basis (up to suitable diagonal rescaling) in any dimension, while the Riesz property is broken for the tensorized hierarchical basis of hat functions. However, thanks to the so-called *norm equivalence* property, tensorized biorthogonal wavelets form a Riesz basis in any dimension, hence providing a remedy for this issue.

Although compressed sensing is becoming a standard paradigm for signal processing applications, understanding its full potential and limitations in scientific computing is still the 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., Rauhut & Ward, 2012, Chkifa *et al.*, 2018, Adcock *et al.*, 2017a, 2019b) and of parametric PDEs, with special emphasis on uncertainty quantification (see, e.g., Doostan & Owhadi, 2011, Yang & Karniadakis, 2013, Bouchot *et al.*, 2017, Rauhut & Schwab, 2017). 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 Brugiapaglia (2018). Finally, we observe that wavelet–Fourier techniques are widely used in signal processing applications (see, e.g., Krahmer & Ward, 2014, Adcock & Hansen, 2016, Adcock *et al.*, 2017b, for theoretical contributions in this direction). Yet, to the best of our knowledge, this paper is the first comprehensive study of this type of technique in the context of numerical approximation of multidimensional PDEs.

1.1 *Main contributions*

The main contributions of this paper are 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 Brugiapaglia *et al.* (2018) 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 Brugiapaglia *et al.* (2018) 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 one-dimensional case with nonconstant coefficients in Theorem 4.2 and for the multidimensional 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 (one-dimensional case), Theorem 4.9 (multidimensional anisotropic case) and Theorem 4.12 (multidimensional 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 $\mathbb{N} := \{1, 2, 3, \ldots\}$, $\mathbb{N}_0 := \{0\} \cup \mathbb{N}$ 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, $\bar{z}$ is its complex conjugate and $\mathrm{Re}(z)$ and $\mathrm{Im}(z)$ denote its real and imaginary parts, respectively. Given a vector $\boldsymbol{x} \in \mathbb{C}^n$ and $1 \leq p < +\infty$ then $\|\boldsymbol{x}\|_p := (\sum_{j=1}^n |x_j|^p)^{1/p}$, $\mathrm{supp}(\boldsymbol{x}) := \{j \in [n] : x_j \neq 0\}$ and $\|\boldsymbol{x}\|_0 := |\mathrm{supp}(\boldsymbol{x})|$. 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 $\boldsymbol{x} \cdot \boldsymbol{y} := x_1\bar{y}_1 + \cdots + x_n\bar{y}_n$ is the standard inner product of $\mathbb{C}^n$. Given a matrix $A \in \mathbb{C}^{m\times n}$, A^* denotes its conjugate transpose. The set of sequences indexed by integer multi-indices is denoted by $\ell(\mathbb{Z}^n) := \{(x_{\boldsymbol{j}})_{\boldsymbol{j}\in\mathbb{Z}^n} : x_{\boldsymbol{j}} \in \mathbb{C}\ \forall \boldsymbol{j} \in \mathbb{Z}^n\}$.

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 $\|\cdot\|^2_{H^k(\mathcal{D})} = (\cdot, \cdot)_k$ and the $H^k(\mathcal{D})$-seminorm is given by

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

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

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

where

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

Notice that $H^0_{\text{per}}(\mathcal{D}) \equiv L^2(\mathcal{D})$ since $C^\infty_{\text{per}}(\mathcal{D})$ is dense in $L^2(\mathcal{D})$. Now let $u \in H^k_{\text{per}}(\mathcal{D})$ 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:

$$\|u\|^2_{H^k(\mathcal{D})} \sim \sum_{\boldsymbol{r}\in\mathbb{Z}^n} (1+|\boldsymbol{r}|^{2k})|c_{\boldsymbol{r}}|^2 \quad \forall\, u \in H^k_{\text{per}}(\mathcal{D}), \quad \forall\, k \in \mathbb{N}_0. \tag{2.1}$$

Moreover, we define $H^{-1}_{\text{per}}(\mathcal{D}) := [H^1_{\text{per}}(\mathcal{D})]^*$, 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 Adams & Fournier (2003), Taylor (2011), Temam (1995).

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

$$\text{find } u \in H^1_{\text{per}}(\mathcal{D}) : \quad a(u,v) = (f,v) \quad \forall\, v \in H^1_{\text{per}}(\mathcal{D}), \tag{2.2}$$

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 (2.2). A proof of this result is provided in Brugiapaglia *et al.* (2020a).

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 $f \in H^{-1}_{\mathrm{per}}(\mathcal{D})$, the weak problem (2.2) admits a unique solution u such that

$$\|u\|_{H^1(\mathcal{D})} \le \frac{1}{\alpha}\|f\|_{H^{-1}(\mathcal{D})}, \tag{2.3}$$

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

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

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

3. CORSING wavelet–Fourier

In order to solve problem (2.2) we describe the CORSING $\mathcal{WF}$ (wavelet–Fourier) method.

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 Brugiapaglia *et al.* (2020a) and to Dahmen (1997), Pabel (2015), Urban (2008) 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 a 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 Fig. 1, Setting 3.1 and Brugiapaglia *et al.*, 2020a).

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 Brugiapaglia *et al.*, 2020a). 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.,

$$\mathbb{Z}/(2^\ell\mathbb{Z}) \equiv \{0, 1, \dots, 2^\ell - 1\} \quad \forall\, \ell \in \mathbb{N}_0.$$

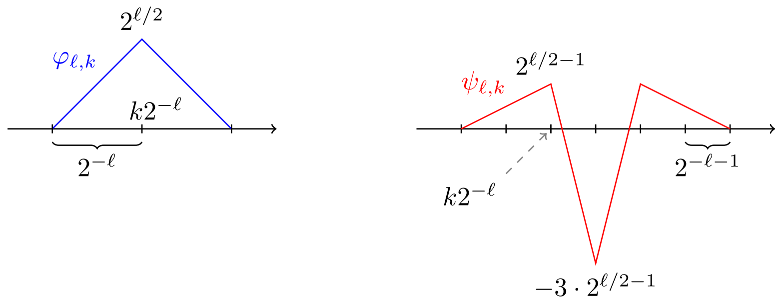

FIG. 1. The translated and rescaled scaling functions $\varphi_{\ell,k}$ (left) and wavelets $\psi_{\ell,k}$ (right) corresponding to the construction of biorthogonal B-spline wavelets of order $(d, \widetilde{d}) = (2, 2)$ on the real line (Section 3.1).

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

$$\psi_{\ell_0-1,k} \equiv \varphi_{\ell_0,k} \quad \forall\, k \in \mathbb{Z}. \tag{3.1}$$

SETTING 3.1 (One-dimensional biorthogonal B-spline wavelets; Cohen *et al.*, 1992). We consider one-dimensional 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 scaling functions Φ_ℓ^{per} and wavelets Ψ_ℓ^{per} to be normalized with respect to the $L^2(\mathcal{D})$-norm (i.e., $\|\varphi_{\ell,k}\|_{L^2(\mathcal{D})} \sim \|\psi_{\ell,k}\|_{L^2(\mathcal{D})} \sim 1$ for every $\ell \in \mathbb{N}_0$ and $k \in \mathbb{Z}/(2^\ell\mathbb{Z})$) and $\ell_0 \geq 2$.

When wavelets are built as in Setting 3.1 the corresponding Ψ^{per} is a basis for $H^1_{\text{per}}(\mathcal{D})$ when $L = \infty$. In fact, for $L \in \mathbb{N}$, $\text{span}(\Psi^{\text{per}})$ coincides with the space of functions in $H^1_{\text{per}}(\mathcal{D})$ 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]

$$\psi^{\text{per}}_{\ell,k}|_{\text{supp}(\psi_{\ell,k})} \equiv \psi_{\ell,k}|_{\text{supp}(\psi_{\ell,k})} \quad \forall\, \ell \geq \ell_0 - 1,\ \forall\, k \in \mathbb{Z}/(2^\ell\mathbb{Z}). \tag{3.2}$$

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

In order to generalize this construction to arbitrary dimension $n > 1$ we consider *anisotropic* and *isotropic* tensor product wavelets (see Fig. 2). Here we recall just the basic definitions and we refer the reader to Dahmen (1997), Pabel (2015), Urban (2008) for a more detailed discussion.

---

[1] In particular, the assumption $\ell_0 \geq 2$ ensures that the periodization $\psi^{\text{per}}_{\ell,k}$ of $\psi_{\ell,k}$ (defined as $\psi^{\text{per}}_{\ell,k}(x) = 2^{\ell/2} \sum_{j\in\mathbb{Z}} \psi(2^\ell(x + j) - k)$ for every $x \in \mathbb{R}$ – see also Brugiapaglia *et al.*, 2020a) is the sum of terms with disjoint support. On the contrary let us assume, for example, $\ell_0 = 1$. Then, since $\text{supp}(\psi_{1,1}) \subseteq (0, 3/2)$, the periodization $\psi^{\text{per}}_{1,1}$ is built by summing two overlapping terms on $[0, 1/2]$. In particular, $\psi^{\text{per}}_{1,1}$ is constant over $[0, 1/2]$, whereas $\psi_{1,1}$ is not. This shows that condition (3.2) is not satisfied for $\ell_0 = 1$.

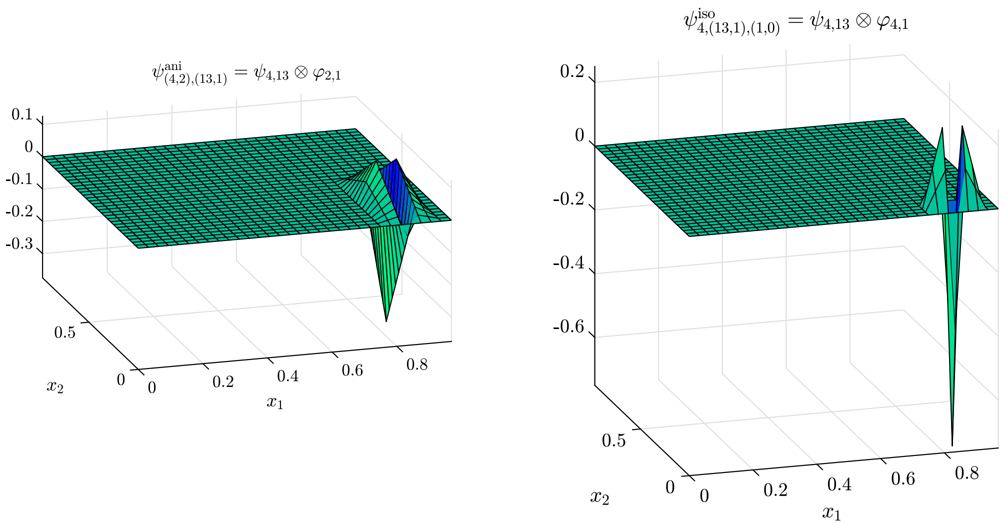

FIG. 2. Surface plot of the two-dimensional anisotropic (left) and isotropic (right) tensor product wavelets.

Anisotropic tensor product wavelets. Given an $\ell_0 \in \mathbb{N}$ 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 multidimensional basis is by tensorizing the one-dimensional wavelet basis Ψ with itself n times, namely,

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

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

$$\mathcal{J}^{\mathrm{ani}} := \{(\boldsymbol{\ell}, \boldsymbol{k}) : \boldsymbol{\ell} \in \mathbb{N}^n,\ \ell_0 - 1 \leq \boldsymbol{\ell} < L,\ \boldsymbol{k} \in \mathbb{Z}/(2^{\boldsymbol{\ell}}\mathbb{Z})\}, \tag{3.3}$$

we have

$$\Psi^{\mathrm{ani}} = \{\psi^{\mathrm{ani}}_{\boldsymbol{j}}\}_{\boldsymbol{j} \in \mathcal{J}^{\mathrm{ani}}}, \tag{3.4}$$

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

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

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

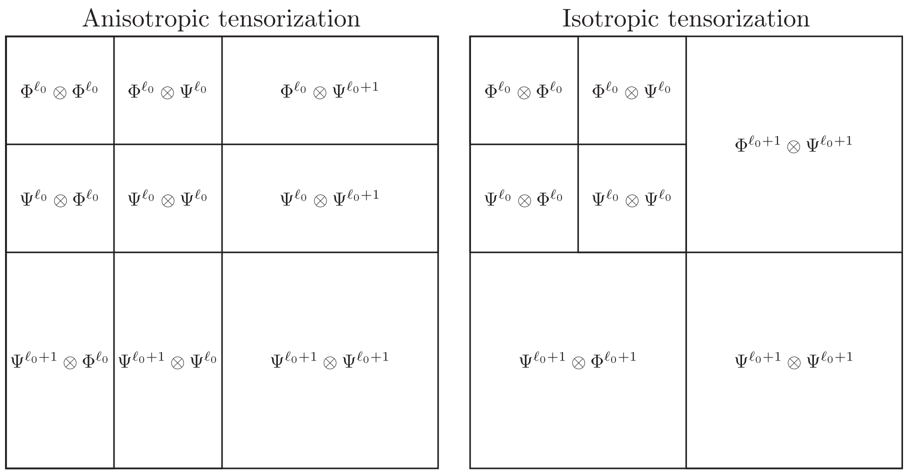

FIG. 3. Visualization of the two-dimensional tensorized anisotropic (left) and isotropic (right) wavelets for $L = \ell_0 + 1$.

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

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

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

$$\mathcal{J}^{\text{iso}} := \{(\ell,\boldsymbol{k},\boldsymbol{e}) : \ell \in \mathbb{N}, \ell_0 \leq \ell < L, \boldsymbol{k} \in (\mathbb{Z}/(2^\ell\mathbb{Z}))^n, \boldsymbol{e} \in \{0,1\}^n\}, \tag{3.5}$$

we have

$$\Psi^{\text{iso}} := \{\psi^{\text{iso}}_{\boldsymbol{j}}\}_{\boldsymbol{j}\in\mathcal{J}^{\text{iso}}}, \tag{3.6}$$

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

The main difference between anisotropic and isotropic tensor product structure is visualized in Figs 2 and 3 (in the two-dimensional case).

In particular, the anisotropic tensorization blends all the dyadic levels together, whereas the isotropic tensorization combines only 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 Brugiapaglia *et al.* (2020a) for the proof.

THEOREM 3.2 (Norm equivalence). Let $\Psi \in \{\Psi^{\text{ani}}, \Psi^{\text{iso}}\}$ be the tensor product wavelet basis defined as in (3.4) or (3.6) from periodized one-dimensional biorthogonal B-spline wavelets of order $(d,\widetilde{d})$. Then the following norm equivalence holds:

$$\left\| \sum_{\boldsymbol{j}\in\mathcal{J}^\star} c_{\boldsymbol{j}}\psi^\star_{\boldsymbol{j}} \right\|_{H^s(\mathcal{D})} \sim \|D^s_\star \boldsymbol{c}\|_2 \quad \forall \star \in \{\text{ani}, \text{iso}\}, \tag{3.7}$$

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

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

where $\delta_{\boldsymbol{j},\boldsymbol{j}'}$ is Kronecker's delta function and where $\mathcal{J}^{\text{ani}}$ and $\mathcal{J}^{\text{iso}}$ are defined as in (3.3) and (3.5), 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

$$\|\psi_{\boldsymbol{j}}^{\text{ani}}\|_{H^1(\mathcal{D})} \sim 2^{\|\boldsymbol{\ell}\|_\infty} \quad \forall \boldsymbol{j} \in \mathcal{J}^{\text{ani}}, \tag{3.8}$$

$$\|\psi_{\boldsymbol{j}}^{\text{iso}}\|_{H^1(\mathcal{D})} \sim 2^{\ell} \quad \forall \boldsymbol{j} \in \mathcal{J}^{\text{iso}}, \tag{3.9}$$

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 Brugiapaglia *et al.* (2020a).

3.2 *Test functions: Fourier basis*

Consider the one-periodic Fourier basis functions in one dimension 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

$$\|\xi_{\boldsymbol{q}}\|^2_{L^2(\mathcal{D})} = 1, \quad \|\xi_{\boldsymbol{q}}\|^2_{H^1(\mathcal{D})} = 1 + (2\pi\|\boldsymbol{q}\|_2)^2, \quad \forall \boldsymbol{q} \in \mathbb{Z}^n. \tag{3.10}$$

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

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

Then we define the Fourier basis as

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

In particular, $\varXi$ 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, $\Psi \in \{\Psi^{\text{ani}}, \Psi^{\text{iso}}\}$ and $\mathcal{J} \in \{\mathcal{J}^{\text{ani}}, \mathcal{J}^{\text{iso}}\}$. Let

$\varXi = \{\xi_{\boldsymbol{q}}\}_{\boldsymbol{q}\in\mathcal{Q}}$ 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,

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

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 (3.8), (3.9) and (3.10) this normalization can be realized by defining

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

Petrov–Galerkin discretization. We consider a Petrov–Galerkin discretization of (2.2) associated with the trial basis $\widehat{\varPsi}$ and test basis $\widehat{\varXi}$ (see Quarteroni & Valli, 2008, 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

$$B\boldsymbol{v} = \boldsymbol{g}, \tag{3.11}$$

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 (3.11) by random subsampling. Given a probability distribution $\boldsymbol{p} \in \mathbb{R}^M$ over $\mathcal{Q}$ we draw $m \ll M$ multi-indices $\boldsymbol{\tau}_1, \ldots, \boldsymbol{\tau}_m \in \mathcal{Q}$ 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

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

The CORSING reduced discretization corresponds to the underdetermined linear system

$$A\boldsymbol{z} = \boldsymbol{f}, \tag{3.12}$$

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

$$D_{i,k} = \delta_{i,k}/\sqrt{mp_{\tau_i}}, \quad i,k \in [m],$$

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:

$$\min_{z\in\mathbb{C}^N} \|D(Az - f)\|_2 \quad \text{s.t. } \|z\|_0 \leq s. \tag{3.13}$$

Although NP-hard, (3.13) can be approximated by the OMP algorithm. Finally, the CORSING solution is given by

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

where $\widehat{\boldsymbol{u}} = (\widehat{u}_j)_{j\in\mathcal{J}} \in \mathbb{C}^N$ is an approximate solution to (3.13) computed via OMP.

Note that the sampling complexity m can be reduced in order to avoid repeated indices among the τ_i. In that case the preconditioner D has to be slightly modified (see Brugiapaglia *et al.*, 2018, 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 we fix $s, N \in \mathbb{N}$ as parameters chosen by the user, 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 (3.11).

(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 (3.11) to (3.12)).

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

We observe that the choice of s and N theoretically depends on the best s-term approximation error decay rate for the target function class (containing u). This, in turn, could be studied using results from nonlinear approximation theory (see, e.g., DeVore, 1998 or Mallat, 1999, Chapter 9).

In order to solve issues (i), (ii) and (iii) we resort to the theoretical analysis carried out in Brugiapaglia *et al.* (2018), 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 $\boldsymbol{\mu} \in \ell(\mathbb{Z}^n)$ defined by

$$\mu_{\boldsymbol{q}} := \sup_{j\in\mathcal{J}} |a(\widehat{\psi}_j, \widehat{\xi}_{\boldsymbol{q}})|^2 \quad \forall \boldsymbol{q} \in \mathbb{Z}^n. \tag{3.14}$$

[2] The diagonal preconditioner D is chosen in such a way that $\mathbb{E}[(DA)^*DA] = B^*B$. For further details see Brugiapaglia *et al.* (2018).

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

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

We specialize Brugiapaglia *et al.*, 2018 (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 Brugiapaglia (2016), Brugiapaglia *et al.* (2015, 2018) does not give the Riesz basis assumption in the multidimensional case. It is also possible to state an analogous result in probability instead of expectation (see Brugiapaglia *et al.*, 2018, 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 (2.2), and assume we have an upper bound to the local a-coherence $\boldsymbol{\mu} \in \ell(\mathbb{Z}^n)$, i.e., there exists a sequence $\boldsymbol{\nu} \in \ell(\mathbb{Z}^n)$ such that

$$\boldsymbol{\mu} \lesssim \boldsymbol{\nu}. \tag{3.15}$$

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

$$\|\boldsymbol{\nu}|_{\mathcal{Q}^c}\|_1 \lesssim \frac{1}{s} \tag{3.16}$$

holds and where $\mathcal{Q}^c := \mathbb{Z}^n \setminus \mathcal{Q}$. Then, for every $0 < \varepsilon < 1$, the CORSING solution $\widehat{u} \in H^1_{\mathrm{per}}(\mathcal{D})$ exactly solving (3.13) satisfies

$$\mathbb{E} \left\| \min \left\{ 1, \frac{K}{\|\widehat{u}\|_{H^1(\mathcal{D})}} \right\} \widehat{u} - u \right\|_{H^1(\mathcal{D})} \lesssim \frac{\mathcal{A}}{\alpha} \sigma_s(u)_{H^1(\mathcal{D})} + K\varepsilon, \tag{3.17}$$

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

$$m \gtrsim s \|\boldsymbol{\nu}|_{\mathcal{Q}}\|_1 (s \ln(eN/s) + \ln(2s/\varepsilon)), \tag{3.18}$$

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

$$\boldsymbol{p} = \frac{\boldsymbol{\nu}|_{\mathcal{Q}}}{\|\boldsymbol{\nu}|_{\mathcal{Q}}\|_1}. \tag{3.19}$$

Some considerations are in order:

---

[3] Let us clarify a small difference between Brugiapaglia *et al.*, 2018, (Theorem 3.15) and Theorem 3.4. In Brugiapaglia *et al.* (2018, Theorem 3.15) $\boldsymbol{\nu}$ is an upper bound to $\boldsymbol{\mu}|_{\mathcal{Q}}$ and the truncation condition (3.16) involves $\boldsymbol{\mu}$ instead of $\boldsymbol{\nu}$. Therefore, the truncation condition of Theorem 3.4 implies the truncation condition of Brugiapaglia *et al.* (2018, Theorem 3.15). However, this does not make any difference since in practice the truncation condition of Brugiapaglia *et al.* (2018, Theorem 3.15) is verified using an upper bound to $\boldsymbol{\mu}$, and not $\boldsymbol{\mu}$ itself.

- Relation (3.18) 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 Brugiapaglia *et al.*, 2018, Section 5.4). A different theoretical analysis carried out in (Brugiapaglia, 2016, Section 3.2.5) based on the concept of a 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 (3.7) with $s = 1$. We have preferred to employ this slightly suboptimal result to avoid this technical rescaling issue.
- A necessary condition for (3.16) 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 (3.18).
- 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 (2.3).
- The hypothesis that $\widehat{\boldsymbol{u}}$ solves (3.13) 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.
- Condition (3.16) is sufficient to guarantee the so-called *s-sparse restricted inf-sup property* for the Petrov–Galerkin discretization (3.11). This is a variant of the classical inf-sup (or Ladyzhenskaya–Babuška–Brezzi) condition, adapted to the sparse case (see Brugiapaglia *et al.*, 2018, Section 3.3 for more details).

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 (3.16).

(ii) Choose $m = m(s, N)$ according to (3.18).

(iii) Choose $\boldsymbol{p} \in \mathbb{R}^M$ according to (3.19).

Our next goal is to find an upper bound $\boldsymbol{\nu}$ to the local a-coherence $\boldsymbol{\mu}$ as in (3.15) such that

(a) it is possible to find an explicit formula for $R = R(s, N)$ such that (3.16) 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 Brugiapaglia *et al.* (2018) by deriving upper bounds to the local a-coherence defined in (3.14) for ADR equations with nonconstant coefficients in one dimension (Section 4.1) and with constant coefficients in arbitrary dimension (Section 4.2).

4.1 *The one-dimensional 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 $\gamma \in [0,2]$:

$$\left|\left(D^{\alpha_1}\varphi_{\ell,k}, D^{\alpha_2}\xi_q\right)\right| \leq 2^{\alpha_1+\alpha_2}\, 2^{\left(\frac{3}{2}-\gamma\right)\ell}|\pi q|^{\gamma-2+\alpha_1+\alpha_2}, \tag{4.1}$$

$$\left|\left(D^{\alpha_1}\psi_{\ell,k}, D^{\alpha_2}\xi_q\right)\right| \leq 2^{\alpha_1+\alpha_2+1-\gamma}\|\boldsymbol{b}\|_2\|\boldsymbol{b}\|_0^{\frac{1}{2}}2^{\left(\frac{3}{2}-\gamma\right)\ell}|\pi q|^{\gamma-2+\alpha_1+\alpha_2}. \tag{4.2}$$

Moreover, for $q = 0$, we have

$$|(\varphi'_{\ell,k}, \xi'_0)| = |(\psi'_{\ell,k}, \xi'_0)| = 0, \tag{4.3}$$

$$|(\varphi'_{\ell,k}, \xi_0)| = |(\psi'_{\ell,k}, \xi_0)| = 0, \tag{4.4}$$

$$|(\varphi_{\ell,k}, \xi_0)| = 2^{-\ell/2}, \quad |(\psi_{\ell,k}, \xi_0)| = 0. \tag{4.5}$$

Proof. If $q = 0$ equations (4.3)–(4.5) 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 (3.2) we can directly compute

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

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

$$|(\varphi'_{\ell,k}, \xi'_q)| \leq 4\cdot 2^{(\frac{3}{2}-\gamma)\ell}|\pi q|^\gamma. \tag{4.6}$$

Moreover, thanks again to hypothesis (3.2), employing (4.6), the discrete Cauchy–Schwarz inequality and the definition of the mother wavelet (see also Brugiapaglia *et al.*, 2020a)

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

we obtain

$$\left|\left(\psi'_{\ell,k}, \xi'_q\right)\right| = \frac{1}{\sqrt{2}}\left|\sum_{j\in\mathbb{Z}/(2^{\ell+1}\mathbb{Z})} b_{j-2k}\left(\varphi'_{\ell+1,j}, \xi'_q\right)\right| \leq \frac{\|\boldsymbol{b}\|_2}{\sqrt{2}}\left[\sum_{j\in\mathbb{Z}/(2^{\ell+1}\mathbb{Z}),\, b_{j-2k}\neq 0}\left|\left(\varphi'_{\ell+1,j}, \xi'_q\right)\right|^2\right]^{\frac{1}{2}}$$

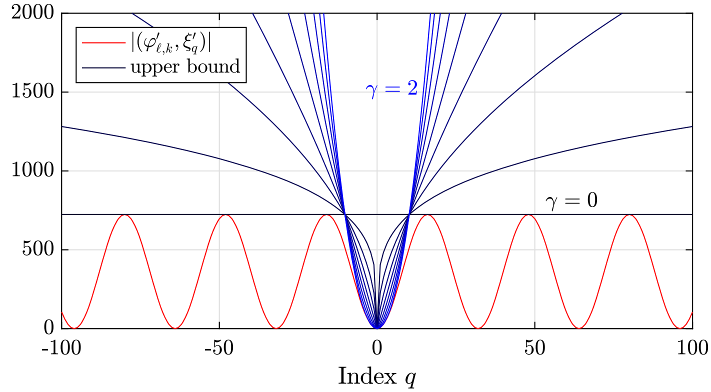

FIG. 4. Sharpness of the upper bound (4.6), with $\ell = 5$, $k = 4$ and different values of $\gamma \in [0, 2]$.

$$\leq 4 \cdot \frac{\|\boldsymbol{b}\|_2 \|\boldsymbol{b}\|_0^{\frac{1}{2}}}{\sqrt{2}} 2^{(\frac{3}{2}-\gamma)(\ell+1)} |\pi q|^{\gamma} = 2^{3-\gamma} \|\boldsymbol{b}\|_2 \|\boldsymbol{b}\|_0^{\frac{1}{2}} 2^{(\frac{3}{2}-\gamma)\ell} |\pi q|^{\gamma}.$$

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 one-dimensional ADR equations with nonconstant coefficients.

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 (3.14) can be bounded from above as

$$\mu_0 \lesssim 2^{-2\ell_0} \left(|\beta|^2_{H^1(\mathcal{D})} + \|\rho\|^2_{L^2(\mathcal{D})} \right),$$

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

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 (2.1). 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

$$\left|(\eta\psi'_{\ell,k}, \xi'_q)\right|^2 = \left|\sum_{r\in\mathbb{Z}} \left(\frac{1+r^2}{1+r^2}\right)^{\frac{1}{2}} \eta_r (\xi_r\psi'_{\ell,k}, \xi'_q)\right|^2 \leq \left(\sum_{r\in\mathbb{Z}}(1+r^2)|\eta_r|^2\right) \sum_{r\in\mathbb{Z}} \frac{\left|(\xi_r\psi'_{\ell,k}, \xi'_q)\right|^2}{1+r^2}. \quad (4.7)$$

Recalling (2.1) 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 (4.2) and (4.4), we obtain the following estimate for every $\gamma \in [0, 2]$:

$$\sum_{r\in\mathbb{Z}} \frac{\left|(\xi_r\psi'_{\ell,k}, \xi'_q)\right|^2}{1+r^2} \sim |q|^2 \sum_{r\in\mathbb{Z}} \frac{\left|(\psi'_{\ell,k}, \xi_{q-r})\right|^2}{1+r^2} \lesssim q^2 2^{(3-2\gamma)\ell} S_{2(1-\gamma)}(q), \quad (4.8)$$

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]

[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.

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,

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

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

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

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

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

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

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

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 (4.7) and (4.8) we have

$$|(\eta\psi'_{\ell,k}, \xi'_q)|^2 \lesssim \|\eta\|^2_{H^1(\mathcal{D})} q^2 \min\{2^{3\ell} S_2(q), 2^{2\ell} S_1(q)\} \lesssim \|\eta\|^2_{H^1(\mathcal{D})} 2^{2\ell} \min\{2^\ell, |q|\}. \tag{4.9}$$

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

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

Estimating the sum in the right-hand side as before we obtain

$$|(\beta\psi'_{\ell,k}, \xi_q)|^2 \lesssim \frac{\|\beta\|^2_{H^1(\mathcal{D})}}{q^2} 2^{2\ell} \min\{2^\ell, |q|\}. \tag{4.10}$$

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 (2.1) with $k = 0$ we have

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

[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.

for every $\gamma \in [0,2]$, where we have employed (4.1)–(4.2) with $(\alpha_1, \alpha_2) = (0,0)$ and ξ_{q-r} for $r \neq q$ and (4.5) with $\xi_{q-r} = \xi_0$ for $r = q$. Choosing $\gamma = 1/2$ we see that

$$|(\rho\psi_{\ell,k}, \xi_q)|^2 \lesssim \|\rho\|^2_{L^2(\mathcal{D})}(2^{2\ell}+1) \lesssim \|\rho\|^2_{L^2(\mathcal{D})} 2^{2\ell}, \tag{4.11}$$

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 (4.9), (4.10) and (4.11) 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 (3.8) and (3.10)), 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 $q \neq 0$:

$$\mu_q \lesssim \left(\|\eta\|^2_{H^1(\mathcal{D})} + \frac{\|\beta\|^2_{H^1(\mathcal{D})}}{q^2} + \frac{\|\rho\|^2_{H^1(\mathcal{D})}}{q^4} \right) \min\left\{ \frac{2^L}{q^2}, \frac{1}{|q|} \right\}. \tag{4.12}$$

Notice that (4.12) generalizes (Brugiapaglia *et al.*, 2018, Proposition 4.4) (in particular, we refer to Brugiapaglia *et al.*, 2018, equation (145)) for one-dimensional 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 one-dimensional 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 (3.17).

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 (3.16), 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 (3.16) 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 multidimensional case*

We consider issues (i), (ii) and (iii) in Section 3.4 for CORSING $\mathcal{WF}$ for multidimensional ADR equations with constant coefficients. In Section 4.2.1 we analyze the case of anisotropic tensor product wavelets, 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 notation 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]

$$\mathrm{zero}(\boldsymbol{q}) := [n] \setminus \mathrm{supp}(\boldsymbol{q}), \quad \mathrm{scal}(\boldsymbol{\ell}) := \{j \in [n] : \ell_j = \ell_0 - 1\}. \tag{4.13}$$

[6] Note that, according to (3.2), $\mathrm{scal}(\boldsymbol{\ell})$ is the set of indices j such that ψ_{ℓ_j, k_j} is a scaling function.

Then it follows that

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

$$\left|\left(\nabla \psi_{\boldsymbol{\ell},\boldsymbol{k}}^{\text{ani}}, \nabla \xi_{\boldsymbol{0}}\right)\right| = 0, \tag{4.14}$$

$$\left|\left(\boldsymbol{\beta} \cdot \nabla \psi_{\boldsymbol{\ell},\boldsymbol{k}}^{\text{ani}}, \xi_{\boldsymbol{0}}\right)\right| = 0 \quad \forall\, \boldsymbol{\beta} \in \mathbb{R}^n \tag{4.15}$$

$$\left|\left(\psi_{\boldsymbol{\ell},\boldsymbol{k}}^{\text{ani}}, \xi_{\boldsymbol{0}}\right)\right| = \begin{cases} 2^{-n\ell_0/2} & \text{if } \mathrm{scal}(\boldsymbol{\ell}) = [n], \\ 0 & \text{otherwise;} \end{cases} \tag{4.16}$$

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

$$|(\nabla \psi_{\boldsymbol{\ell},\boldsymbol{k}}^{\text{ani}}, \nabla \xi_{\boldsymbol{q}})| \lesssim 2^{-\frac{1}{2}(n-\|\boldsymbol{q}\|_0)\ell_0 + (\frac{3}{2}-\widehat{\boldsymbol{\gamma}})\cdot \widehat{\boldsymbol{\ell}}} |\widehat{\boldsymbol{q}}|^{\widehat{\boldsymbol{\gamma}}-\boldsymbol{2}} \|\boldsymbol{q}\|_2^2, \tag{4.17}$$

$$|(\boldsymbol{\beta} \cdot \nabla \psi_{\boldsymbol{\ell},\boldsymbol{k}}^{\text{ani}}, \xi_{\boldsymbol{q}})| \lesssim 2^{-\frac{1}{2}(n-\|\boldsymbol{q}\|_0)\ell_0 + (\frac{3}{2}-\widehat{\boldsymbol{\gamma}})\cdot \widehat{\boldsymbol{\ell}}} |\widehat{\boldsymbol{q}}|^{\widehat{\boldsymbol{\gamma}}-\boldsymbol{2}} \|\boldsymbol{\beta}\|_2 \|\boldsymbol{q}\|_2 \quad \forall\, \boldsymbol{\beta} \in \mathbb{R}^n, \tag{4.18}$$

$$|(\psi_{\boldsymbol{\ell},\boldsymbol{k}}^{\text{ani}}, \xi_{\boldsymbol{q}})| \lesssim 2^{-\frac{1}{2}(n-\|\boldsymbol{q}\|_0)\ell_0 + (\frac{3}{2}-\widehat{\boldsymbol{\gamma}})\cdot \widehat{\boldsymbol{\ell}}} |\widehat{\boldsymbol{q}}|^{\widehat{\boldsymbol{\gamma}}-\boldsymbol{2}}, \tag{4.19}$$

  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 $\boldsymbol{q} \neq \boldsymbol{0}$ and $\mathrm{zero}(\boldsymbol{q}) \not\subseteq \mathrm{scal}(\boldsymbol{\ell})$ then

$$\left|\left(\nabla \psi_{\boldsymbol{\ell},\boldsymbol{k}}^{\text{ani}}, \nabla \xi_{\boldsymbol{q}}\right)\right| = 0, \tag{4.20}$$

$$\left|\left(\boldsymbol{\beta} \cdot \nabla \psi_{\boldsymbol{\ell},\boldsymbol{k}}^{\text{ani}}, \xi_{\boldsymbol{q}}\right)\right| = 0 \quad \forall\, \boldsymbol{\beta} \in \mathbb{R}^n, \tag{4.21}$$

$$\left|\left(\psi_{\boldsymbol{\ell},\boldsymbol{k}}^{\text{ani}}, \xi_{\boldsymbol{q}}\right)\right| = 0. \tag{4.22}$$

Proof. In the case $\boldsymbol{q} = \boldsymbol{0}$ the equalities (4.14), (4.15) and (4.16) are a direct consequence of (4.3), (4.4) and (4.5). Let us consider $\boldsymbol{q} \neq \boldsymbol{0}$. 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:

$$\big(\nabla\psi^{\text{ani}}_{\boldsymbol{\ell},\boldsymbol{k}}, \nabla\xi_{\boldsymbol{q}}\big) = \sum_{j=1}^{n} \big(\psi'_{\ell_j,k_j}, \xi'_{q_j}\big) \prod_{i\neq j} \big(\psi_{\ell_i,k_i}, \xi_{q_i}\big), \tag{4.23}$$

$$\big(\boldsymbol{\beta}\cdot\nabla\psi^{\text{ani}}_{\boldsymbol{\ell},\boldsymbol{k}}, \xi_{\boldsymbol{q}}\big) = \sum_{j=1}^{n} \beta_j\big(\psi'_{\ell_j,k_j}, \xi_{q_j}\big) \prod_{i\neq j} \big(\psi_{\ell_i,k_i}, \xi_{q_i}\big) \quad \forall\,\boldsymbol{\beta}\in\mathbb{R}^n, \tag{4.24}$$

$$\big(\psi^{\text{ani}}_{\boldsymbol{\ell},\boldsymbol{k}}, \xi_{\boldsymbol{q}}\big) = \prod_{j=1}^{n} \big(\psi_{\ell_j,k_j}, \xi_{q_j}\big). \tag{4.25}$$

Plugging relations (4.1) and (4.2) into (4.23) we see that, for every $\boldsymbol{\gamma}\in[0,2]^n$, we have

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

Similarly, plugging (4.1) and (4.2) into (4.24), we obtain

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

Finally, plugging (4.1) and (4.2) into (4.25), it follows that

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

The relations above prove (4.17), (4.18) and (4.19).

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

First, assume that $\text{zero}(\boldsymbol{q}) \not\subseteq \text{scal}(\boldsymbol{\ell})$ (notice also that $\text{zero}(\boldsymbol{q})$ is nonempty since $\|\boldsymbol{q}\|_0 < n$). In this case we can pick an index $j_0 \in \text{zero}(\boldsymbol{q}) \setminus \text{scal}(\boldsymbol{\ell})$, 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 (4.23) with relations (4.3) and (4.5) yields

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

This proves (4.20) ((4.21) and (4.22) are shown analogously). As a consequence the only possibility for $(\nabla\psi^{\text{ani}}_{\boldsymbol{\ell},\boldsymbol{k}}, \nabla\xi_{\boldsymbol{q}})$ to be nonzero is to have $\text{zero}(\boldsymbol{q}) \subseteq \text{scal}(\boldsymbol{\ell})$. In this case, by splitting the sum above and employing (4.3), we obtain

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

Now splitting the product and using (4.5) yields

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

In order to prove (4.17) 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 (4.18) and (4.19). □

REMARK 4.6 (Curse of dimensionality). It is worth stressing that estimates (4.14)–(4.22) 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 always has 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 multidimensional 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:

$$\mu_{\mathbf{0}} \lesssim |\rho|^2 2^{-(2+n)\ell_0}, \tag{4.26}$$

$$\mu_{\boldsymbol{q}} \lesssim \left(|\eta|^2 + \frac{\|\boldsymbol{\beta}\|_2^2}{\|\boldsymbol{q}\|_2^2} + \frac{|\rho|^2}{\|\boldsymbol{q}\|_2^4}\right) 2^{-(n-\|\boldsymbol{q}\|_0)\ell_0} \min\left\{\frac{2^{(3\|\boldsymbol{q}\|_0-2)L}\|\boldsymbol{q}\|_2^2}{|\widehat{\boldsymbol{q}}|^{\mathbf{4}}}, \frac{\|\boldsymbol{q}\|_2^2}{\|\boldsymbol{q}\|_\infty^2 |\widehat{\boldsymbol{q}}|^{\mathbf{1}}}\right\} \quad \forall \boldsymbol{q} \neq \mathbf{0}, \tag{4.27}$$

where the inequalities hide constants depending exponentially on n.

Proof. Let us assume $\boldsymbol{q} = \mathbf{0}$. Recalling (4.14) and (4.15) we have

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

Employing (4.16) and recalling (3.8) and (3.10) 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 (4.26).

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 (4.17)–(4.19) (notice that in this case zero$(\boldsymbol{q}) = \emptyset \subseteq$ scal$(\boldsymbol{\ell})$) and recalling relations (3.8) and (3.10) on the $H^1(\mathcal{D})$-norm of $\psi^{\text{ani}}_{\boldsymbol{\ell},\boldsymbol{k}}$ and of $\xi_{\boldsymbol{q}}$, for every $\boldsymbol{\gamma} \in [0,2]^n$ and $\boldsymbol{q} \neq \mathbf{0}$ we have

$$|a(\widehat{\psi}^{\text{ani}}_{\boldsymbol{\ell},\boldsymbol{k}}, \widehat{\xi}_{\boldsymbol{q}})| = \frac{|a(\psi^{\text{ani}}_{\boldsymbol{\ell},\boldsymbol{k}}, \xi_{\boldsymbol{q}})|}{\|\psi^{\text{ani}}_{\boldsymbol{\ell},\boldsymbol{k}}\|_{H^1(\mathcal{D})}\|\xi_{\boldsymbol{q}}\|_{H^1(\mathcal{D})}} \lesssim 2^{(\frac{3}{2}-\boldsymbol{\gamma})\cdot\boldsymbol{\ell}-\|\boldsymbol{\ell}\|_\infty}|\boldsymbol{q}|^{\boldsymbol{\gamma}-\mathbf{2}}\|\boldsymbol{q}\|_2\left(|\eta| + \frac{\|\boldsymbol{\beta}\|_2}{\|\boldsymbol{q}\|_2} + \frac{|\rho|}{\|\boldsymbol{q}\|_2^2}\right). \quad (4.28)$$

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 (4.28) and considering the maximum over $\boldsymbol{j} \in \mathcal{J}$ we obtain

$$\mu_{\boldsymbol{q}} \lesssim \left(|\eta|^2 + \frac{\|\boldsymbol{\beta}\|_2^2}{\|\boldsymbol{q}\|_2^2} + \frac{|\rho|^2}{\|\boldsymbol{q}\|_2^4}\right)\frac{\|\boldsymbol{q}\|_2^2}{|\boldsymbol{q}|^{\mathbf{4}}}\max_{\boldsymbol{\ell}_0 \le \boldsymbol{\ell} < L} 2^{3\|\boldsymbol{\ell}\|_1 - 2\|\boldsymbol{\ell}\|_\infty}.$$

Finally, by noticing that

$$\max_{\boldsymbol{\ell}_0 \le \boldsymbol{\ell} < L} 2^{3\|\boldsymbol{\ell}\|_1 - 2\|\boldsymbol{\ell}\|_\infty} \le \max_{\boldsymbol{\ell}_0 \le \boldsymbol{\ell} < L} 2^{(3n-2)\|\boldsymbol{\ell}\|_\infty} \le 2^{(3n-2)L},$$

we obtain the upper bound corresponding to the first argument of the minimum in (4.27), 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

$$\widetilde{\gamma}_j := \begin{cases} 1/2 & \text{if } j = j_\infty, \\ 3/2 & \text{otherwise}, \end{cases} \quad (4.29)$$

and choose $\boldsymbol{\gamma} = \widetilde{\boldsymbol{\gamma}}$. Considering the set $\triangle := \{\boldsymbol{\ell} \in \mathbb{N}^n : \boldsymbol{\ell}_0 \le \boldsymbol{\ell} < L,\ \ell_1 \ge \ell_2 \ge \cdots \ge \ell_n\}$ and defining S_n as the permutation group of the set $[n]$ we have

$$\begin{aligned}\max_{\boldsymbol{\ell}_0 \le \boldsymbol{\ell} < L} 2^{\frac{3}{2}\|\boldsymbol{\ell}\|_1 - \|\boldsymbol{\ell}\|_\infty - \widetilde{\boldsymbol{\gamma}}\cdot\boldsymbol{\ell}} &= \max_{\boldsymbol{\ell}\in\triangle}\max_{\sigma\in S_n} 2^{\frac{3}{2}\|\sigma(\boldsymbol{\ell})\|_1 - \|\sigma(\boldsymbol{\ell})\|_\infty - \widetilde{\boldsymbol{\gamma}}\cdot\sigma(\boldsymbol{\ell})} \\ &= \max_{\boldsymbol{\ell}\in\triangle} 2^{\frac{3}{2}\|\boldsymbol{\ell}\|_1 - \|\boldsymbol{\ell}\|_\infty}\max_{\sigma\in S_n} 2^{-\widetilde{\boldsymbol{\gamma}}\cdot\sigma(\boldsymbol{\ell})} \\ &= \max_{\boldsymbol{\ell}\in\triangle} 2^{\frac{3}{2}\|\boldsymbol{\ell}\|_1 - \ell_1}2^{-\frac{1}{2}\ell_1 - \frac{3}{2}\sum_{j>1}\ell_j} = 1, \end{aligned} \quad (4.30)$$

where we have exploited the identity $\{\boldsymbol{\ell} \in \mathbb{N}^n : \boldsymbol{\ell}_0 \le \boldsymbol{\ell} < L\} = \bigcup_{\sigma\in S_n}\sigma(\triangle)$ and the fact that $\|\boldsymbol{\ell}\|_1$ and $\|\boldsymbol{\ell}\|_\infty$ are invariant with respect to permutations of the components of $\boldsymbol{\ell}$. Combining (4.28) with (4.30) (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

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

where the right-hand side does not depend on $\boldsymbol{\ell}$. This relation corresponds to the second argument of the minimum in (4.27) 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 $\|\widehat{\boldsymbol{\ell}}\|_\infty \leq \|\boldsymbol{\ell}\|_\infty$ (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 one-dimensional 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 immediately verified that (4.12) coincides with (4.27) when $n = 1$. Moreover, (4.26) 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 $|(\rho\psi_{\ell,k}, \xi_0)| \leq \|\rho\|_{L^2(\mathcal{D})}\|\psi_{\ell,k}\|_{L^2(\mathcal{D})}$ 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) $m \gtrsim C2^{n\ell_0}s(s\ln(eN/(2s)) + \ln(2s/\varepsilon))(\ln N + \ln s + \ln C)^n$,

(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 (3.17).

Proof. Let us consider the upper bound $\boldsymbol{\nu}$ defined according to (4.26) and (4.27). 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 (4.27). 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

$$\begin{aligned} \|\boldsymbol{\nu}|_{\mathcal{Q}^c}\|_1 &\leq \sum_{\boldsymbol{q}\in\mathbb{Z}^n:\|\boldsymbol{q}\|_\infty>Q} \nu_{\boldsymbol{q}} = \sum_{s=1}^{n} \sum_{\boldsymbol{q}\in\mathbb{Z}^n:\|\boldsymbol{q}\|_\infty>Q,\ \|\boldsymbol{q}\|_0=s} \nu_{\boldsymbol{q}} \\ &\leq C\sum_{s=1}^{n} \sum_{\boldsymbol{q}\in\mathbb{Z}^n:\|\boldsymbol{q}\|_\infty>Q,\ \|\boldsymbol{q}\|_0=s} 2^{-(n-s)\ell_0}\frac{2^{(3s-2)L}\|\boldsymbol{q}\|_2^2}{|\widehat{\boldsymbol{q}}|^{\boldsymbol{4}}} \\ &= C\sum_{s=1}^{n}\binom{n}{s}2^{-(n-s)\ell_0+(3s-2)L} \underbrace{\sum_{\boldsymbol{r}\in\mathbb{Z}^s:\|\boldsymbol{r}\|_\infty>Q,\ \|\boldsymbol{r}\|_0=s} \frac{\|\boldsymbol{r}\|_2^2}{|\boldsymbol{r}|^{\boldsymbol{4}}}}_{=:T(s)}. \end{aligned} \tag{4.31}$$

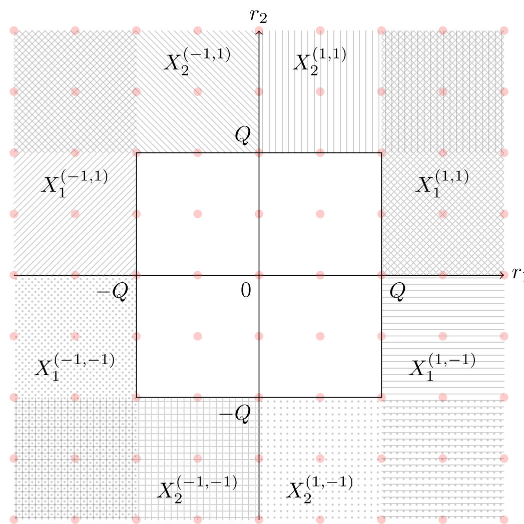

FIG. 5. The sets $X_k^{\boldsymbol{t}} \subseteq \mathbb{Z}^s$ in (4.32) for $s = 2$ and $Q = 2$, restricted to $[-4, 4]^2$. Different textures correspond to different sets.

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

$$\{\boldsymbol{r} \in \mathbb{Z}^s : \|\boldsymbol{r}\|_\infty > Q,\ \|\boldsymbol{r}\|_0 = s\} = \bigcup_{k=1}^{s} \bigcup_{\boldsymbol{t} \in \{-1,1\}^s} \underbrace{\{\boldsymbol{r} \in \mathbb{Z}^s : |r_k| > Q,\ \operatorname{sign}(\boldsymbol{r}) = \boldsymbol{t}\}}_{=:X_k^{\boldsymbol{t}}}, \tag{4.32}$$

where $X_k^{\boldsymbol{t}} \subseteq \mathbb{Z}^s$ is the set of multi-indices having the kth component larger than Q and with sign pattern $\boldsymbol{t}$ (see Fig. 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

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

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

(4.31) yields

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

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

Finally, in order to prove (ii), we estimate $\|\boldsymbol{v}|_{\mathcal{Q}}\|_1$. Employing the upper bound corresponding to (4.26) and to the second argument of the minimum in (4.27), and recalling that $\|\boldsymbol{q}\|_2^2 \leq \|\boldsymbol{q}\|_0\|\boldsymbol{q}\|_\infty^2$ for every $\boldsymbol{q} \in \mathbb{Z}^n$, we obtain

$$\|\boldsymbol{v}|_{\mathcal{Q}}\|_1 \leq C\left(2^{-(2+n)\ell_0} + \sum_{s=1}^{n}2^{-(n-s)\ell_0}\sum_{\boldsymbol{q}\in\mathbb{Z}^n:\|\boldsymbol{q}\|_\infty\leq Q,\ \|\boldsymbol{q}\|_0=s}\frac{\|\boldsymbol{q}\|_2^2}{\|\boldsymbol{q}\|_\infty^2|\widehat{\boldsymbol{q}}|^{\mathbf{1}}}\right)$$
$$\lesssim C\sum_{s=1}^{n}2^{-(n-s)\ell_0}\sum_{\boldsymbol{q}\in\mathbb{Z}^n:\|\boldsymbol{q}\|_\infty\leq Q,\ \|\boldsymbol{q}\|_0=s}\frac{s}{|\widehat{\boldsymbol{q}}|^{\mathbf{1}}} = C\sum_{s=1}^{n}\binom{n}{s}2^{-(n-s)\ell_0}s\sum_{\boldsymbol{r}\in\mathbb{Z}^s:\|\boldsymbol{r}\|_\infty\leq Q,\ \|\boldsymbol{r}\|_0=s}\frac{1}{|\boldsymbol{r}|^{\mathbf{1}}}$$
$$= C\sum_{s=1}^{n}\binom{n}{s}2^{-(n-s)\ell_0}s\left(\sum_{|q|\leq Q,\ q\neq 0}\frac{1}{|q|}\right)^s \lesssim 2^{-n\ell_0}C\sum_{s=1}^{n}\binom{n}{s}2^{s\ell_0}s(\log R)^s \lesssim C(\log R)^n,$$

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 zero($\boldsymbol{q}$) as in (4.13). Then the following inequalities hold:

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

$$|(\nabla\psi_{\ell,\boldsymbol{k},\boldsymbol{e}}^{\text{iso}}, \nabla\xi_{\mathbf{0}})| = 0,$$

$$|(\boldsymbol{\beta}\cdot\nabla\psi_{\ell,\boldsymbol{k},\boldsymbol{e}}^{\text{iso}}, \xi_{\mathbf{0}})| = 0 \quad \forall\,\boldsymbol{\beta}\in\mathbb{R}^n,$$

$$|(\psi_{\ell,\boldsymbol{k},\boldsymbol{e}}^{\text{iso}}, \xi_{\mathbf{0}})| = \begin{cases} 2^{-n\ell/2} & \text{if } \boldsymbol{e} = \mathbf{0},\\ 0 & \text{if } \boldsymbol{e} \neq \mathbf{0};\end{cases}$$

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

$$|(\nabla\psi^{\mathrm{iso}}_{\ell,\boldsymbol{k},\boldsymbol{e}}, \nabla\xi_{\boldsymbol{q}})| \lesssim 2^{(-\frac{n}{2}+2\|\boldsymbol{q}\|_0-\|\widehat{\boldsymbol{\gamma}}\|_1)\ell}|\widehat{\boldsymbol{q}}|^{\widehat{\boldsymbol{\gamma}}-\mathbf{2}}\|\boldsymbol{q}\|_2^2,$$

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

$$|(\psi^{\mathrm{iso}}_{\ell,\boldsymbol{k},\boldsymbol{e}}, \xi_{\boldsymbol{q}})| \lesssim 2^{(-\frac{n}{2}+2\|\boldsymbol{q}\|_0-\|\widehat{\boldsymbol{\gamma}}\|_1)\ell}|\widehat{\boldsymbol{q}}|^{\widehat{\boldsymbol{\gamma}}-\mathbf{2}},$$

  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 $\boldsymbol{q} \neq \mathbf{0}$ and $\mathrm{zero}(\boldsymbol{q}) \not\subseteq \mathrm{zero}(\boldsymbol{e})$ then

$$|(\nabla\psi^{\mathrm{iso}}_{\boldsymbol{\ell},\boldsymbol{k}}, \nabla\xi_{\boldsymbol{q}})| = 0,$$

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

$$|(\psi^{\mathrm{iso}}_{\boldsymbol{\ell},\boldsymbol{k}}, \xi_{\boldsymbol{q}})| = 0.$$

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:

$$\mu_{\mathbf{0}} \lesssim |\rho|^2 2^{-(2+n)\ell_0}, \tag{4.33}$$

$$\mu_{\boldsymbol{q}} \lesssim \left(|\eta|^2 + \frac{\|\boldsymbol{\beta}\|_2^2}{\|\boldsymbol{q}\|_2^2} + \frac{|\rho|^2}{\|\boldsymbol{q}\|_2^4}\right)\min\left\{\frac{(1+2^{2(-\frac{n}{2}+2\|\boldsymbol{q}\|_0-1)L})\|\boldsymbol{q}\|_2^2}{|\widehat{\boldsymbol{q}}|^{\mathbf{4}}}, \frac{2^{-(n-\|\boldsymbol{q}\|_0)\ell_0}\|\boldsymbol{q}\|_2^2}{\|\boldsymbol{q}\|_\infty^2|\widehat{\boldsymbol{q}}|^{\mathbf{1}}}\right\} \quad \forall\,\boldsymbol{q}\neq\mathbf{0}, \tag{4.34}$$

where the inequalities hide constants depending exponentially on n.

Proof. The upper bound (4.33) 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 $\boldsymbol{\gamma}\in[0,2]^n$,

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

Analogously to Theorem 4.7 we obtain the upper bound corresponding to the first argument of the minimum in (4.34) by choosing $\widehat{\boldsymbol{\gamma}} = \mathbf{0}$ in the inequality above. Notice in particular that

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

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 (4.34) is proved by letting $\boldsymbol{\gamma}=\widetilde{\boldsymbol{\gamma}}$, as in (4.29). 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) $m\gtrsim C2^{n\ell_0}s(s\ln(eN/(2s))+\ln(2s/\varepsilon))(\ln N+\ln s+\ln C)^n$,

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

where $C=|\eta|^2+\|\boldsymbol{\beta}\|_2^2+|\rho|^2$, the CORSING $\mathcal{WF}$ method with $\Psi=\Psi^{\text{iso}}$ recovers the best s-term approximation to u in expectation in the sense of estimate (3.17).

5. Numerical assessment

In this section we numerically investigate the reliability and robustness of the CORSING $\mathcal{WF}$ approach in different dimensions. In Section 5.1 we consider a one-dimensional ADR equation with constant and nonconstant coefficients. As predicted by the theory we show that CORSING $\mathcal{WF}$ is robust and reliable in the case of nonconstant coefficients also. In Section 5.2 we consider the two-dimensional 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 three-dimensional case.

All the numerical experiments have been performed in MATLAB with the aid of OMP-Box for OMP (Rubinstein *et al.*, 2008, Rubinstein, 2009). 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 *One-dimensional case with nonconstant diffusion*

We consider a one-dimensional 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 Fig. 6 (top row).

We set $L=9$, resulting in $N=512$, and choose $R=N$. We observe that the oscillations of the diffusion coefficient impact the stiffness matrix only ‘horizontally’. In particular this does not impact the decay properties of the local a-coherence $\boldsymbol{\mu}$ (see Fig. 6, bottom row). The wavelet–Fourier Petrov–Galerkin discretization of the ADR equation gives rise to a matrix with a comparable structure

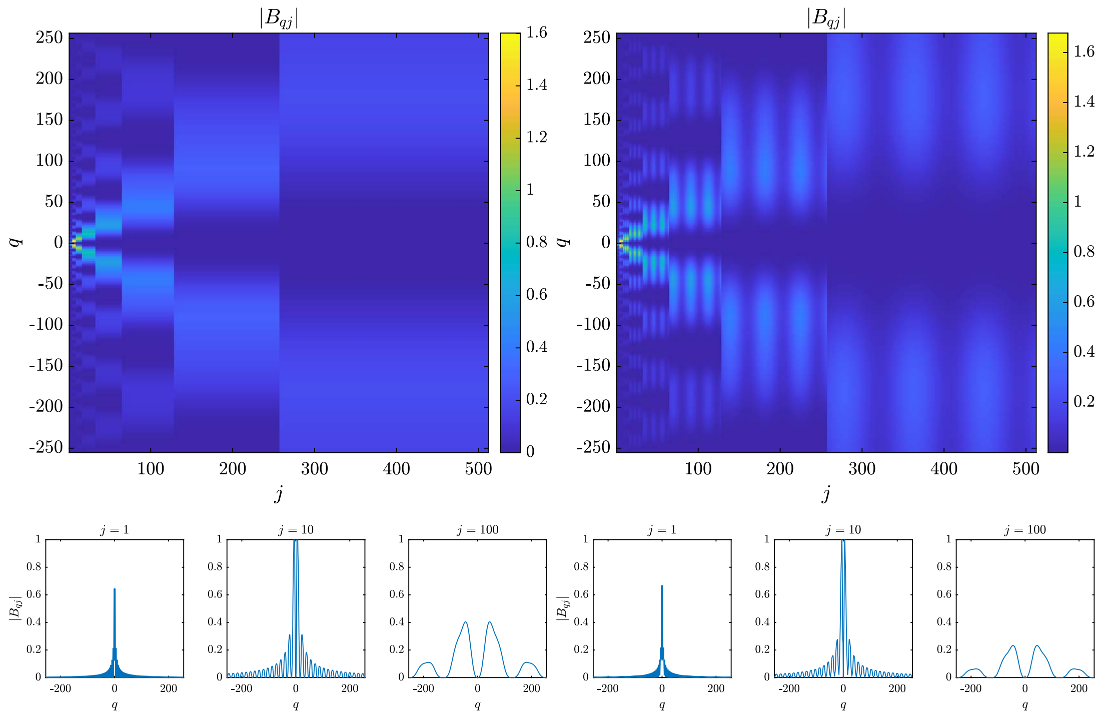

FIG. 6. (One-dimensional Diffusion Reaction (DR) problem) First row: Absolute value of the entries $|B_{q,j}|$ of the stiffness matrix associated with the Petrov–Galerkin discretization for $\eta \equiv 1$ (left) and $\eta = 1 + 0.5\sin(6\pi x)$ (right). Second row: Three vertical slices of the plots in the first row, i.e., plot of $|B_{q,j}|$ as a function of q, for $j = 1, 10, 100$.

with respect to the matrices in Adcock *et al.* (2017b, Fig. 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, 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 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 Fig. 7 (left)). In Fig. 7 (right) we show the wavelet coefficients and highlight in red the largest 50 in absolute value.

The resulting relative best 50-term approximation error with respect to the $H^1(\mathcal{D})$-norm is

$$\frac{\|u_1 - \widetilde{u}_1\|_{H^1(\mathcal{D})}}{\|u_1\|_{H^1(\mathcal{D})}} \sim \frac{\|\boldsymbol{u}_1 - \widetilde{\boldsymbol{u}}_1\|_2}{\|\boldsymbol{u}_1\|_2} = 1.68 \cdot 10^{-2}, \tag{5.1}$$

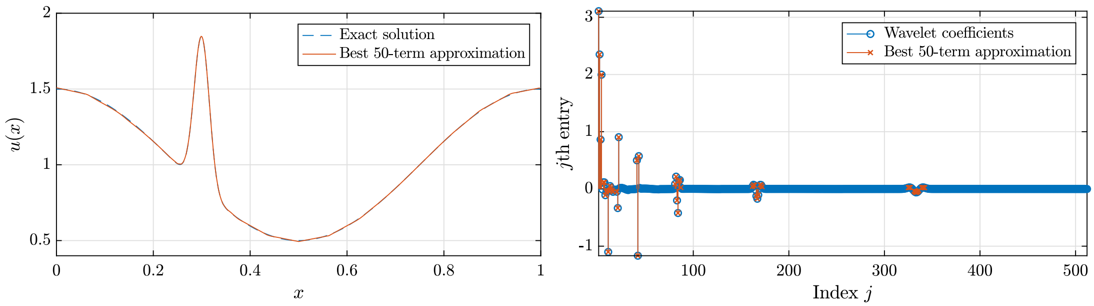

FIG. 7. (One-dimensional DR problem) Plot of u_1 and of its best 50-term approximation $\widetilde{u}_1$ (left). Wavelet coefficients of u with the 50 largest in magnitude highlighted (right).

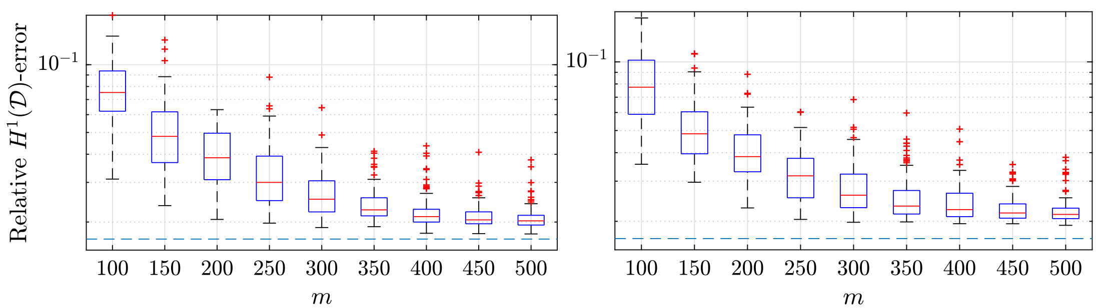

FIG. 8. (One-dimensional DR problem) Relative recovery error as a function of the number of random tests m for a constant (left) and nonconstant (right) diffusion term. The dashed line shows the relative best 50-term approximation error (5.2).

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 Fig. 8 we show box plots 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 (less than twice 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 observe more outliers only for the nonconstant diffusion.

Sensitivity of the recovery error to the sparsity. In Fig. 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 plots are 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 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.

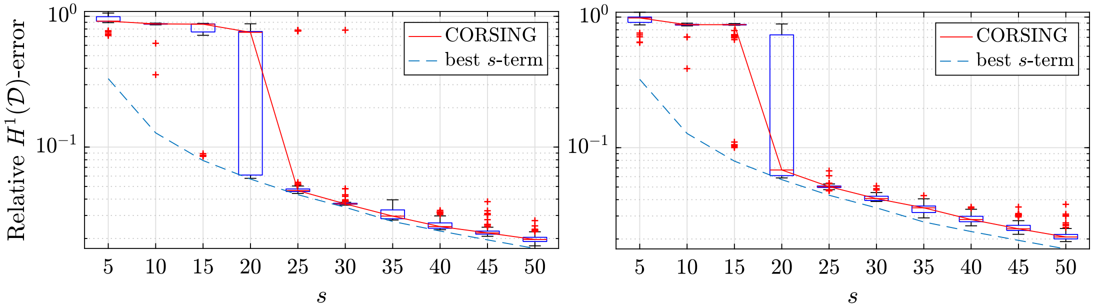

FIG. 9. (One-dimensional DR problem) Relative recovery and best s-term approximation errors as functions of the sparsity s with constant (left) and nonconstant (right) diffusion term.

5.2 *Two-dimensional case*

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

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 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 Fig. 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 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]

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

[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 (5.1).

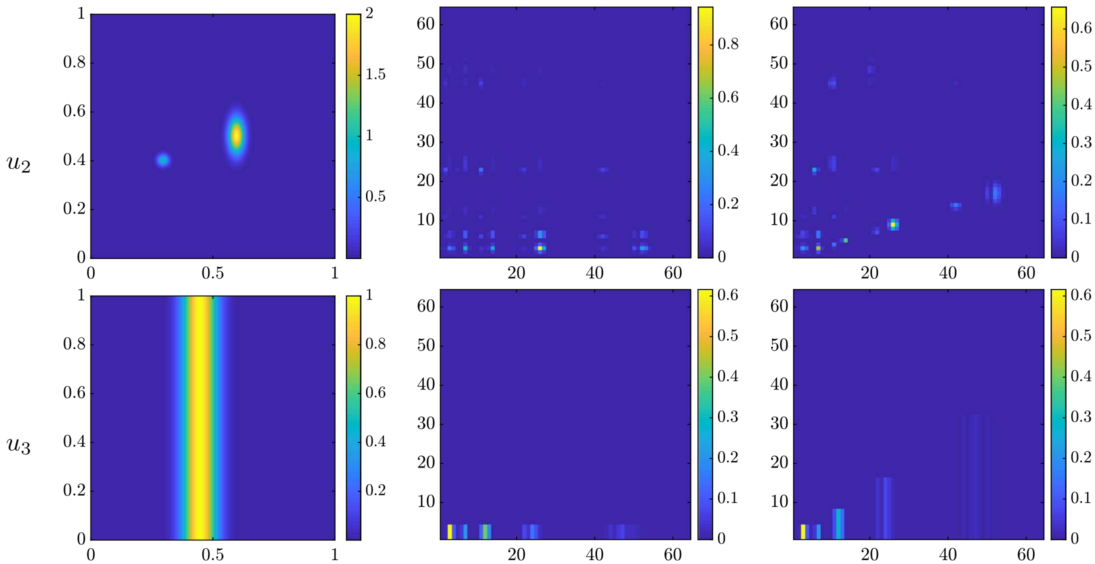

FIG. 10. (Two-dimensional ADR problem) Contour plots (left) of functions u_2 (top) and u_3 (bottom) and corresponding wavelet coefficients with respect to anisotropic (center) and isotropic (right) tensor product wavelets.

subsampling ($\boldsymbol{p} \propto \mathbf{1}$) with the nonuniform subsampling based on the local a-coherence upper bound

$$\nu_{\boldsymbol{q}} = \min\left\{1, \frac{\|\mathbf{q}\|_2^2}{\|\boldsymbol{q}\|_\infty^2 |\widehat{\boldsymbol{q}}|^{\mathbf{1}}}\right\},$$

which is obtained from the upper bounds in Theorems 4.7 and 4.11. In particular, in (4.27) and (4.34) 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 (4.26) and (4.33). 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 Figs 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 convergent behavior. This experiment confirms the key role played by the local a-coherence for successful implementation of the CORSING $\mathcal{WF}$ method.

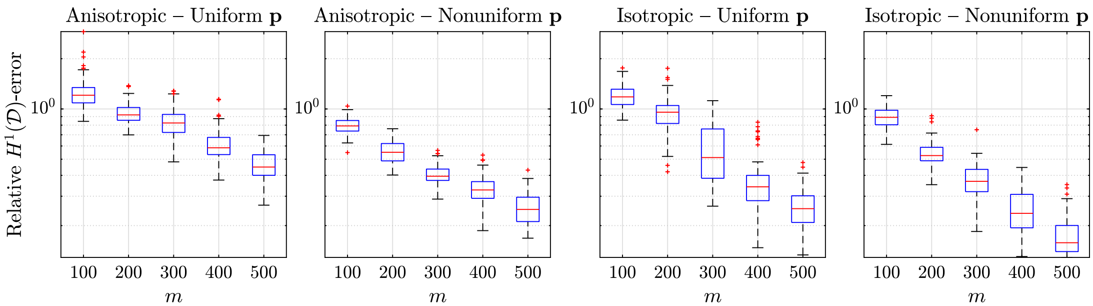

FIG. 11. (Two-dimensional ADR problem) Box plots of the relative recovery error for the function u_2 with respect to the $H^1(\mathcal{D})$-norm as a function of the number of tests m with anisotropic and isotropic wavelets and uniform and nonuniform subsampling.

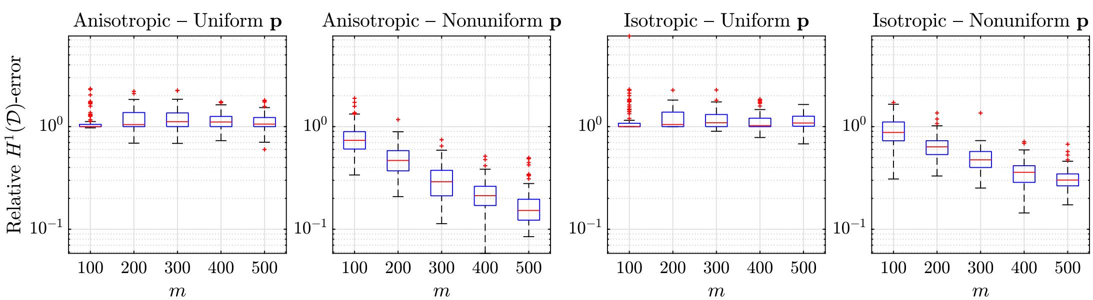

FIG. 12. (Two-dimensional ADR problem) Box plots of the relative recovery error for the function u_3 with respect to the $H^1(\mathcal{D})$-norm as a function of the number of tests m with anisotropic and isotropic wavelets and uniform and nonuniform subsampling.

5.3 *Three-dimensional case*

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

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

The 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 Fig. 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 Fig. 14.

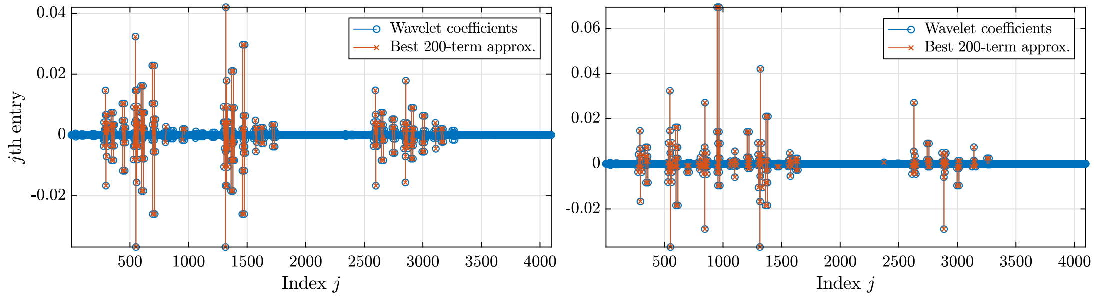

FIG. 13. (Three-dimensional ADR problem) Wavelet coefficients of the function u_4 and best 200-term approximation for anisotropic (left) and isotropic (right) wavelets.

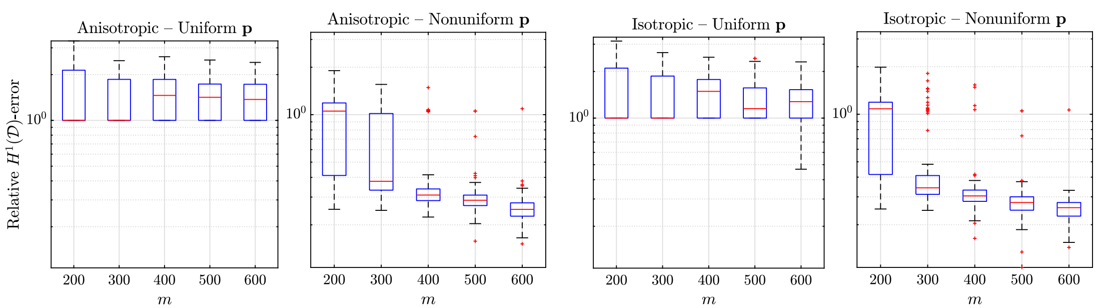

FIG. 14. (Three-dimensional ADR problem) Box plots of the relative recovery error with respect to the $H^1(\mathcal{D})$-norm as a function of the number of tests m with anisotropic and isotropic wavelets and for uniform and nonuniform subsampling.

The performance of anisotropic and isotropic wavelets is similar. However, uniform subsampling is not able to recover the solution at all. Comparing these results with those of the two-dimensional case (Figs 11 and 12) we note how using a bad probability measure (i.e., 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 a recovery error comparable to the best s-term approximation error and that the sampling measure based on the local a-coherence proposed here is able to successfully exploit the sparsity of the exact solution in the discretization (in contrast to other randomization strategies such as uniform random subsampling).

Several open issues still remain to be investigated. First, one needs to understand whether the sampling measure proposed in this paper is or is not the 'optimal' one (in some sense to be specified). On the practical and computational side there is still a lot of work to be done. Although we compared the accuracy of the CORSING $\mathcal{WF}$ solution with the best s-term approximation error (Fig. 9), the computational cost of the CORSING $\mathcal{WF}$ procedure with OMP reconstruction scales linearly in N (i.e., the dimension of the trial basis of wavelet functions). Yet adaptive wavelet methods can recover the best s-term approximation error accuracy with an optimal computational cost of $\mathcal{O}(s)$ flops. This is a crucial issue to address in order to understand what the real impact of CORSING $\mathcal{WF}$ is. In this direction, a line of research currently under investigation is the use of techniques for sublinear-time compressed sensing recently proposed in Choi *et al.* (2019). Finally, developing an effective and optimized implementation for CORSING $\mathcal{WF}$ that takes advantage of the wavelet transform, the Fourier transform and 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 thanks Ben Adcock, Wolfgang Dahmen and Holger Rauhut for very insightful discussions about approximation theory and compressed sensing. The authors would also like to thank the two anonymous reviewers for their helpful and constructive comments.

Funding

Postdoctoral Training Centre in Stochastics of the Pacifical Institute for the Mathematical Sciences (to S.B.); Centre for Advanced Modelling Science (to S.B.); Natural Sciences and Engineering Research Council of Canada (611675 to S.B.); Gruppo Nazionale Calcolo per il Scientifico (GNCS) - Istituto Nazionale di Alta Mathematica (INdAM) 2018 'Tecniche di Riduzione di Modello per le Applicazioni Mediche' (to S.P.).

References

Adams, R. A. & Fournier, J. F. (2003) *Sobolev Spaces*, vol. 140. Oxford, UK: Academic Press.

Adcock, B., Antun, V., & Hansen, A. C. (2019a) Uniform recovery in infinite-dimensional compressed sensing and applications to structured binary sampling. arXiv:1905.00126.

Adcock, B., Bao, A. & Brugiapaglia, S. (2019b) Correcting for unknown errors in sparse high-dimensional function approximation. *Numer. Math.*, **142**, 667–711.

Adcock, B., Brugiapaglia, S. & Webster, C. G. (2017a) Compressed sensing approaches for polynomial approximation of high-dimensional functions. *Compressed Sensing and Its Applications: Second International MATHEON Conference 2015* (H. Boche, G. Caire, R. Calderbank, M. März, G. Kutyniok & R. Mathar eds). Cham: Springer International Publishing, pp. 93–124.

Adcock, B. & Hansen, A. C. (2016) Generalized sampling and infinite-dimensional compressed sensing. *Found. Comput. Math.*, **16**, 1263–1323.

Adcock, B., Hansen, A. C., Poon, C. & Roman, B. (2017b) Breaking the coherence barrier: a new theory for compressed sensing. *Forum of Math., Sigma*, **5**, e4. Camdridge, UK

Bouchot, J.-L., Rauhut, H. & Schwab, C. (2017) Multi-level compressed sensing Petrov–Galerkin discretization of high-dimensional parametric PDEs. arXiv:1701.01671.

Brugiapaglia, S. (2016) *In Quantification of Uncertainty: Improving Efficiency and Technology*, Lecture Notes in Computational Science and Engineering, vol. 137. Springer International Publishing, 2020. arXiv:1807.06606. (in press).

Brugiapaglia, S. (2018) A compressive spectral collocation method for the diffusion equation under the restricted isometry property. arXiv:1807.06606.

Brugiapaglia, S., Micheletti, S., Nobile, F. & Perotto, S. (2020a) Supplementary material to "Wavelet-Fourier CORSING techniques for multidimensional advection–diffusion–reaction equations".

Brugiapaglia S., Micheletti S. & Perotto S. (2015) Compressed solving: a numerical approximation technique for elliptic PDEs based on compressed sensing. *Comput. Math. Appl.*, **70**. 1306–1335.

Brugiapaglia S., Nobile F., Micheletti S. & Perotto S. (2018) A theoretical study of compressed solving for advection–diffusion–reaction problems. *Math. Comp.*, **87**, 1–38.

Brugiapaglia, S., Tamellini, L. & Tani, M. (2020b) Compressive isogeometric analysis. arXiv:2003.06475. (in press).

Candès E. J., Romberg J. & Tao T. (2006) Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information. *IEEE Trans. Inf. Theory*, **52**, 489–509.

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

Choi, B., Iwen, M. & Volkmer, T. (2019) Sparse harmonic transforms II: best *s*-term approximation guarantees for bounded orthonormal product bases in sublinear-time. arXiv:1909.09564.

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

Dahmen, W. (1997) Wavelet and multiscale methods for operator equations. *Acta Numer.*, **6**, 55–228.

DeVore, R. A. (1998) Nonlinear approximation. *Acta Numer.*, **7**, 51–150.

Donoho, D. L. (2006) Compressed sensing. *IEEE Trans. Inf. Theory*, **52**, 1289–1306.

Doostan, A. & Owhadi, H. (2011) A non-adapted sparse approximation of PDEs with stochastic inputs. *J. Comput. Phys.*, **230**, 3015–3034.

Hughes, T. J. R., Cottrell J. A. & Bazilevs Y. (2005) Isogeometric analysis: CAD, finite elements, NURBS, exact geometry and mesh refinement. *Comput. Methods Appl. Mech. Engrg*, **194**, 4135–4195.

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

Mallat, S. (1999) *A Wavelet Tour of Signal Processing*. Burlington (MA), USA: Elsevier.

Pabel, R. (2015) *Adaptive Wavelet Methods for Variational Formulations of Nonlinear Elliptic PDEs on Tensor-Product Domains*. Logos Berlin, Germany: Verlag, GmbH.

Quarteroni, A. & Valli, A. *Numerical Approximation of Partial Differential Equations*, Springer Series in Computational Mathematics, vol. 23. Berlin: Springer, 2008.

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

Rauhut, H. & Ward, R. (2012) Sparse Legendre expansions via ℓ_1-minimization. *J. Approx. Theory*, **164**, 517–533.

Rubinstein, R. (2009) OMP-Box v10. Available at http://www.cs.technion.ac.il/~ronrubin/software.html.

Rubinstein, R., Zibulevsky, M. & Elad, M. (2008) Efficient implementation of the K-SVD algorithm using batch orthogonal matching pursuit. *Technical Report CS-2008-08*. Technion, Computer Science Department.

Taylor, M. E. (2011) *Partial Differential Equations I: Basic Theory*, 2nd edn. Appl. Math. Sci, vol. 115. New York: Springer.

Temam, R. (1995) *Navier–Stokes Equations and Nonlinear Functional Analysis*, 2nd edn. Philadelphia, PA: Society for Industrial and Applied Mathematics (SIAM).

Urban, K. (2008) *Wavelet Methods for Elliptic Partial Differential Equations*. Numerical Mathematics and Scientific Computation. Oxford University Press.

Yang, X. & Karniadakis, G. E. (2013) Reweighted ℓ^1 minimization method for stochastic elliptic differential equations. *J. Comput. Phys*, **248**, 87–108.