IMA Journal of Numerical Analysis (2005) Page 1 of 27
doi: 10.1093/imanum/dri017

Computation of Differential Operators in Aggregated Wavelet Frame Coordinates *

ROB STEVENSON†
Department of Mathematics, Utrecht University
P.O. Box 80.010 NL–3508 TA Utrecht, The Netherlands

AND

MANUEL WERNER‡
Philipps-Universität Marburg, FB 12 Mathematik und Informatik
Hans-Meerwein Straße, Lahnberge, D–35032 Marburg, Germany

[Submitted in January 2007]

Adaptive wavelet algorithms for solving operator equations have been shown to converge with the best possible rates in linear complexity. For the latter statement all costs are taken into account, i.e., also the cost of approximating entries from the infinite stiffness matrix with respect to the wavelet basis using suitable quadrature. A difficulty is the construction of a suitable wavelet basis on the generally non-trivially shaped domain on which the equation is posed. In view of this, recently corresponding algorithms have been proposed that require only a wavelet frame instead of a basis. By employing an overlapping decomposition of the domain, where each subdomain is the smooth parametric image of the unit cube, and by lifting a wavelet basis on this cube to each of the subdomains, the union of these collections defines such a frame. A potential bottleneck within this approach is the efficient approximation of entries corresponding to pairs of wavelets from different collections. Indeed, such wavelets are piecewise smooth with respect to mutually non-nested partitions. In this paper, considering partial differential operators and spline wavelets on the subdomains, we propose an easy implementable quadrature scheme to approximate the required entries, which allows the fully discrete adaptive frame algorithm to converge with the optimal rate in linear complexity.

Keywords: Adaptive algorithms, boundary value problems, optimal computational complexity, frames, wavelets, splines, matrix compression, numerical integration

1. Motivation and Background

For some separable Hilbert space H, a boundedly invertible operator $L : H \to H'$, and a $g \in H'$, we consider the problem of finding $u \in H$ such that

$$Lu = g.$$

We assume that we are given a *frame* $\Psi = \{\psi_\lambda : \lambda \in \Lambda\}$ for H, i.e., a countable collection in H such that for some constants $A_\Psi, B_\Psi > 0$,

$$A_\Psi \|f\|_{H'}^2 \leqslant \|[f(\psi_\lambda)]_{\lambda\in\Lambda}\|_{\ell_2(\Lambda)}^2 \leqslant B_\Psi \|f\|_{H'}^2, \quad (f \in H') \tag{1.1}$$

*This work was supported in part by the Netherlands Organization for Scientific Research, the Deutsche Forschungsgemeinschaft DFG, Grants Da 360/4-2, Da 360/4-3, and the EC-IHP project "Breaking Complexity".
†Email: stevenson@math.uu.nl
‡Email: werner@mathematik.uni-marburg.de

or, equivalently, $\operatorname{clos}\operatorname{span}\Psi = H$ and

$$B_\Psi^{-1}\|v\|_H^2 \leqslant \inf_{\{\mathbf{v}\in\ell_2(\Lambda):\sum_{\lambda\in\Lambda}\mathbf{v}_\lambda\psi_\lambda=v\}}\|\mathbf{v}\|_{\ell_2(\Lambda)}^2 \leqslant A_\Psi^{-1}\|v\|_H^2, \qquad (v\in H). \tag{1.2}$$

The frame operator $F: H'\to\ell_2(\Lambda): f\mapsto[f(\psi_\lambda)]_{\lambda\in\Lambda}$ has dual $F':\ell_2(\Lambda)\to H:\mathbf{v}\mapsto\sum_{\lambda\in\Lambda}\mathbf{v}_\lambda\psi_\lambda$. We have $\ell_2(\Lambda)=\operatorname{Ran}F\oplus^\perp\operatorname{Ker}F'$, and Ψ is called a *(Riesz) basis* for H when $\operatorname{Ker}F'=\{0\}$, or, equivalently, $\operatorname{Ran}F=\ell_2(\Lambda)$.

Writing the solution of (1.1) as $u=\sum_{\lambda\in\Lambda}\mathbf{u}_\lambda\psi_\lambda$ for some $\mathbf{u}\in\ell_2(\Lambda)$, $\mathbf{u}$ is a solution of

$$\mathbf{Mu}=\mathbf{g}, \tag{1.3}$$

where $\mathbf{M}=[(L\psi_{\lambda'})(\psi_\lambda)]_{\lambda,\lambda'\in\Lambda}$, and $\mathbf{g}=[g(\psi_\lambda)]_{\lambda\in\Lambda}$. The, generally, bi-infinite stiffness matrix $\mathbf{M}$ is bounded with $\|\mathbf{M}\|_{\ell_2(\Lambda)\to\ell_2(\Lambda)}\leqslant B_\Psi\|L\|_{H\to H'}$, $\operatorname{Ker}\mathbf{M}=\operatorname{Ker}F'$, $\mathbf{M}|_{\operatorname{Ran}F}:\operatorname{Ran}F\to\operatorname{Ran}F$ is boundedly invertible with $\|\mathbf{M}|_{\operatorname{Ran}F}^{-1}\|_{\ell_2(\Lambda)\to\ell_2(\Lambda)}\leqslant A_\Psi^{-1}\|L^{-1}\|_{H'\to H}$, and for $\mathbf{g}$, that is in $\operatorname{Ran}F$, we have $\|\mathbf{g}\|_{\ell_2(\Lambda)}\leqslant B_\Psi^{\frac12}\|g\|_{H'}$.

In Cohen et al. (2001, 2002); Gantumur et al. (2007) or Stevenson (2003); Dahlke et al. (2004, 2006) for the basis or (true) frame case, respectively, adaptive iterative schemes have been proposed for solving (1.3). Under some conditions, these schemes were shown to be *optimal* in the following sense: Let for some $s>0$, some solution $\mathbf{u}$ of (1.3) be in

$$\mathscr{A}^s=\mathscr{A}^s_\infty=\{\mathbf{v}\in\ell_2(\Lambda):|\mathbf{v}|_{\mathscr{A}^s}:=\sup_N N^s\|\mathbf{v}-\mathbf{v}_N\|_{\ell_2(\Lambda)}<\infty\},$$

where $\mathbf{v}_N$ denotes a *best N-term approximation* for $\mathbf{v}$, i.e., a vector with $\#\operatorname{supp}\mathbf{v}_N\leqslant N$, that has distance to $\mathbf{v}$ not larger than any vector with this support length. Note that the *positions* of the non-zero coefficients of $\mathbf{v}_N$ generally depend on $\mathbf{v}$, meaning that here we are dealing with *nonlinear approximation*. Membership of $\mathbf{u}\in\mathscr{A}^s$ means that for any $\varepsilon>0$, there *exists* a $\mathbf{u}_\varepsilon$ with $\#\operatorname{supp}\mathbf{u}_\varepsilon\leqslant\lceil\varepsilon^{-1/s}|\mathbf{u}|_{\mathscr{A}^s}^{1/s}\rceil$ and $\|\mathbf{u}-\mathbf{u}_\varepsilon\|_{\ell_2(\Lambda)}\leqslant\varepsilon$. For bases or frames that are commonly used in this setting, this membership is related to smoothness of u in a scale of *Besov spaces*, being a much weaker notion of smoothness than that in the standard scale of Sobolev spaces. This is the motivation to consider nonlinear approximation and adaptive schemes. Now *suppose* that $\mathbf{M}$ can be sufficiently well approximated by computable sparse matrices, in the sense that for some $s^*>s$ it is *s^*-computable*. This means that

> *for each $j\in\mathbb{N}_0$, one can construct a matrix $\mathbf{M}_j^*$ having in each column $\mathscr{O}(2^j)$ non-zero entries,* whose joint computation takes $\mathscr{O}(2^j)$ operations, *such that for any $\tilde{s}<s^*$, $\|\mathbf{M}-\mathbf{M}_j^*\|_{\ell_2(\Lambda)\to\ell_2(\Lambda)}\lesssim 2^{-j\tilde{s}}$.*

A consequence of $\mathbf{u}\in\mathscr{A}^s$, and $\mathbf{M}$ being s^*-computable for some $s^*>s$ is that $\mathbf{g}\in\mathscr{A}^s$ with $|\mathbf{g}|_{\mathscr{A}^s}\lesssim|\mathbf{u}|_{\mathscr{A}^s}$. Let us secondly *assume* that given any $\varepsilon>0$, one knows how to produce an approximation $\mathbf{g}_\varepsilon$ in $\mathscr{O}(\varepsilon^{-1/s}|\mathbf{g}|_{\mathscr{A}^s}^{1/s})$ operations, and thus in particular with $\#\operatorname{supp}\mathbf{g}_\varepsilon\lesssim\varepsilon^{-1/s}|\mathbf{g}|_{\mathscr{A}^s}^{1/s}$, with $\|\mathbf{g}-\mathbf{g}_\varepsilon\|_{\ell_2(\Lambda)}\leqslant\varepsilon$. Then, given any $\varepsilon>0$, the aforementioned algorithms are proven to *produce* an approximation $\mathbf{u}_\varepsilon$ in $\mathscr{O}(\varepsilon^{-1/s}|\mathbf{u}|_{\mathscr{A}^s}^{1/s})$ operations, and so $\#\operatorname{supp}\mathbf{u}_\varepsilon\lesssim\varepsilon^{-1/s}|\mathbf{u}|_{\mathscr{A}^s}^{1/s}$, with $\|\mathbf{u}-\mathbf{u}_\varepsilon\|_{\ell_2(\Lambda)}\leqslant\varepsilon$. In view of the assumption $\mathbf{u}\in\mathscr{A}^s$, these bounds on the work and the support length are the best possible modulo a constant factor. Note that $\|u-\sum_{\lambda\in\Lambda}(\mathbf{u}_\varepsilon)_\lambda\psi_\lambda\|_H\leqslant B_\Psi^{\frac12}\varepsilon$.

REMARK 1.1 Actually to arrive at this result in the frame case, an additional technical third assumption was made concerning the $\ell_2(\Lambda)$-orthogonal projector onto $\operatorname{Ran}F$. Although we expect it to hold much more generally, so far it was verified rigorously only in a special situation, see (Stevenson; 2003, §4.3). In Stevenson (2003), we therefore introduced an alternative algorithm that does not require this assumption, which however we expect to have worse quantitative properties.

REMARK 1.2 Although the algorithms from Cohen et al. (2001); Gantumur et al. (2007) are of somewhat different type, one may think of the adaptive algorithms to consist of the application of a simple iterative scheme to (1.3), as the damped Richardson iteration or the Steepest Descent scheme, where in each iteration the application of $\mathbf{M}$ to the current finitely supported iterant, as well as the vector $\mathbf{g}$ are approximated. Such schemes are convergent when $\mathbf{M} = \mathbf{M}^T \geqslant 0$, that we silently assumed above. If L is symmetric and positive definite, i.e., $L' = L$ and $\inf_{0 \neq v \in H} (Lv)(v)/\|v\|_H^2 > 0$, then $\mathbf{M}^T = \mathbf{M} \geqslant 0$ (> 0 in the basis case). Otherwise, one can apply the algorithms to the normal equations $\mathbf{M}^T\mathbf{M}\mathbf{u} = \mathbf{M}^T\mathbf{g}$, although, depending on L, quantitative better options may be possible (cf. Dahmen et al. (2002); Gantumur (2006)).

The validity of the assumption on the approximatibility of $\mathbf{g}$ depends on the right-hand side at hand. In any case, it is satisfied when g is sufficiently smooth. The value of s^* for which $\mathbf{M}$ is s^*-computable depends on the frame or basis Ψ and the operator L. Let us consider L to be a partial differential or integral operator of *order* $2t$, so that typically H is a Sobolev space with smoothness index t, on an *n-dimensional* domain or manifold. Then for Ψ being a *wavelet basis* of *order* d, even for a smooth solution u the largest s for which membership $\mathbf{u} \in \mathscr{A}^s$ can be expected is $s = \frac{d-t}{n}$. For biorthogonal *spline* wavelets that have $\tilde{d} > d - 2t$ *vanishing moments*, in Gantumur and Stevenson (2006a) or Gantumur and Stevenson (2006b) for differential or singular integral operators, respectively, s^*-computability for $s^* > \frac{d-t}{n}$ was shown, being thus sufficient for optimality of the adaptive algorithms. The argument was first, using the smoothness and vanishing moments of the wavelets, to show that the corresponding $\mathbf{M}$ is s^*-*compressible*. This means that

> *for each* $j \in \mathbb{N}_0$, *there exists an infinite matrix* $\mathbf{M}_j$, constructed by dropping entries from $\mathbf{M}$, *such that in each column it has* $\mathscr{O}(2^j)$ *non-zero entries, and such that for any* $\tilde{s} < s^*$, $\|\mathbf{M} - \mathbf{M}_j\|_{\ell_2(\Lambda) \to \ell_2(\Lambda)} \lesssim 2^{-j\tilde{s}}$.

Secondly, by applying suitable *quadrature*, it was shown that each column of $\mathbf{M}_j$ can be approximately *computed*, taking *on average* $\mathscr{O}(1)$ operations per entry, while the order of approximation with respect to $\mathbf{M}$ is maintained.

The bottleneck for the application of the adaptive wavelet basis algorithms is the construction of suitable biorthogonal wavelet bases on the generally non-trivially shaped domains or manifolds on which the equations are posed. The common construction principle is that via a *non-overlapping* domain decomposition, where each subdomain is a smooth parametric image of the n-dimensional unit cube (Dahmen and Schneider (1999a); Canuto et al. (1999); Cohen and Masson (2000)). Biorthogonal multiresolution analyses on this cube are lifted to the subdomains, and continuously connected to biorthogonal multiresolution analyses on the whole domain or manifold, giving rise to biorthogonal wavelets, called *composite wavelets*. In view of obtaining s^*-computability for a sufficiently large value of s^*, difficulties are that wavelets with supports that intersect interfaces between subdomains generally have no vanishing moments, and that their smoothness is restricted to continuity. Proposals to circumvent these problems have been made in Harbrecht and Stevenson (2006); Stevenson (2007), however resulting in wavelets with larger supports, or requiring a more complicated construction. Another difficulty is that continuous "gluing" of the multiresolution analyses over the interfaces requires some matching condition on the parametrizations, that in practical situations might be difficult to fulfill. An elegant construction that

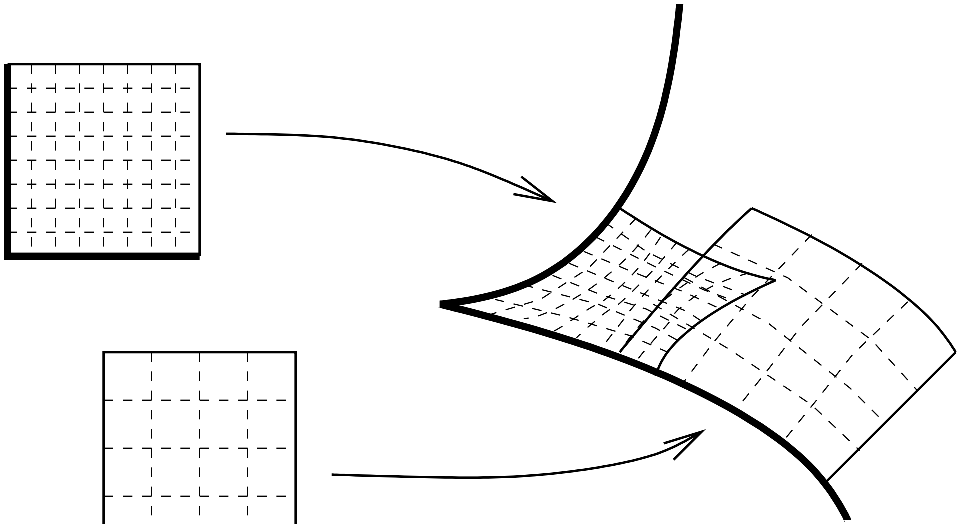

FIG. 1. Construction of aggregated wavelet frame based on an overlapping domain decomposition.

does not require this matching, and yields wavelets that satisfy all requirements concerning smoothness and vanishing moments was proposed in Dahmen and Schneider (1999b). Unfortunately, so far with this approach it seems not easy to obtain wavelets that have competitive quantitative properties. A recent investigation of this approach was made in Kunoth and Sahner (2006).

Above problems with the construction of wavelet *bases* led us to consider adaptive algorithms based on *frames*. A special kind of frame for a Sobolev space on a domain or manifold, called *aggregated wavelet frame* in Dahlke et al. (2004), can be easily constructed. Since in this paper we will consider L to be a *partial differential operator* of some *order* $2t$, appended with *homogeneous Dirichlet boundary conditions*, we describe the construction for the case that for some $t \in \mathbb{N}_0$, $H = H_0^t(\Omega)$ and $\Omega \subset \mathbb{R}^n$ is a domain. Consider an *overlapping domain decomposition* of the domain into a finite number of subdomains, each of them being a smooth parametric image of the n-dimensional unit cube. Then, with $\Psi^\square$ being a wavelet basis for $H_0^t(0,1)^n$, the union of the lifted bases is a frame for $H = H_0^t(\Omega)$, see Figure 1. This construction is simple, and can be applied on any domain having a piecewise smooth boundary. It can be expected that the effective condition number of the aggregated frame, i.e., the condition number without taking the zero eigenvalues into account, is (much) smaller than that of the corresponding composite wavelet basis.

In view of obtaining a sufficiently compressible stiffness matrix, we will consider $\Psi^\square$ to be a biorthogonal spline wavelet basis of *order* d, having $\tilde{d}$ *vanishing moments*. Moreover, to obtain frame elements that are globally sufficiently smooth, i.e., that are in $C^{d-2}(\Omega)$, on those faces of $(0,1)^n$ that are mapped into the interior of Ω we incorporate homogeneous Dirichlet boundary conditions into $\Psi^\square$ of order $d-2$. The resulting collection is then still a frame for $H_0^t(\Omega)$. Note that generally $\Psi^\square$ now depends on the subdomain. An alternative to obtain globally smooth frame elements, which has some advantages in view of Remark 1.1, is to multiply the lifted basis on subdomain Ω_i with χ_i, where $\{\chi_i\}_i$ is a collection of smooth non-negative functions on Ω with χ_i vanishing outside Ω_i and $\sum_i \chi_i \eqsim 1$. Although in the following we will not consider this option, all results from this paper trivially extend to this construction.

As in the basis case, the largest s for which membership of some solution $\mathbf{u} \in \mathscr{A}^s$ can be expected is $s = \frac{d-t}{n}$. For $\tilde{d} > d - 2t$, the proof given in Stevenson (2004) of s^*-compressibility of the stiffness matrix with $s^* > \frac{d-t}{n}$ for the basis case carries directly over to the aggregated frame case. A potential problem, however, is the quadrature, i.e., the question whether $\mathbf{M}$ is also s^**-computable* for some $s^* > \frac{d-t}{n}$. In the overlapping region of two subdomains, there are two collections of lifted wavelets whose elements are piecewise smooth with respect to images of square meshes on $(0,1)^n$ under *different* smooth mappings, see Figure 1. The question is whether entries involving pairs of wavelets from different collections can be approximated within the required tolerance at sufficiently low cost. By carefully distributing computational cost over the entries, in this paper we will show that indeed $\mathbf{M}$ is s^*-computable for $s^* > \frac{d-t}{n}$, at least when $\frac{d-t}{n} > \frac{1}{2}$, $t > 0$ and $d - 1 > t$. For $\frac{d-t}{n} > \frac{1}{2}$, and $t = 0$ or $d - 1 = t$, the suboptimal result $s^* = \frac{d-t}{n}$ will be shown. Although quantitatively better schemes might be possible, in doing so we exclusively use simple composite quadrature rules of fixed order and variable rank in the parameter space of the wavelet that has the highest level of the two involved in an entry.

As follows from the preceding discussion, our result on s^*-computability of the stiffness matrix $\mathbf{M}$ of the boundary value problem with respect to an aggregated wavelet frame is a key ingredient in the proof of optimality of adaptive frame algorithms. In a forthcoming paper, we will study *overlapping domain decomposition (Schwarz) algorithms* applied on the continuous level for solving boundary value problems, where the subdomain solves are approximated by adaptive wavelet basis algorithms. Our first experiments show much better performance of these algorithms compared to the adaptive frame method. The approximate application of $\mathbf{M}$ enters these domain decomposition algorithms for the transport of information between the subdomains. By the s^*-computability of $\mathbf{M}$ shown here, also these algorithms can be shown to be optimal.

This paper is organized as follows: In Sect. 2, a result is proven concerning s^*-compressibility of partial differential operators in (aggregated) frame coordinates, which slightly improves upon the corresponding result from Stevenson (2004) (cf. Remark 2.1). Another reason to include it is that for both the present result and that from Stevenson (2004), homogeneous Dirichlet boundary conditions are essential, whereas this restriction was overlooked in Stevenson (2004). In Sect. 3, we develop quadrature schemes to approximate the remaining entries after compression, and show the required s^*-computability. In Sect. 4, in a slightly specialized setting we give much sharper estimates for certain quadrature errors, which may help improving the quantitative behaviour of adaptive wavelet and frame algorithms. Finally, in Sect. 5, we report on numerical tests to verify the sharpness of bounds on the sizes of the entries in the stiffness matrix, and on those on quadrature errors.

In this paper, by $C \lesssim D$ we will mean that C can be bounded by a multiple of D, independently of parameters which C and D may depend on. Obviously, $C \gtrsim D$ is defined as $D \lesssim C$, and $C \eqsim D$ as $C \lesssim D$ and $C \gtrsim D$.

For any countable set Σ, we will use $\|\cdot\|$ to denote $\|\cdot\|_{\ell_2(\Sigma)}$ or $\|\cdot\|_{\ell_2(\Sigma)\to\ell_2(\Sigma)}$.

2. Compressibility of Partial Differential Operators in (Aggregated) Frame Coordinates

For some domain $\Omega \subset \mathbb{R}^n$ and $t \in \mathbb{N}_0$, let $L : H_0^t(\Omega) \to H^{-t}(\Omega)$ be defined by

$$(Lw)(v) = \sum_{|\alpha|,|\beta| \leqslant t} \int_\Omega a_{\alpha,\beta} \partial^\alpha w \partial^\beta v, \quad (w, v \in H_0^t(\Omega)),$$

where the *coefficients $a_{\alpha,\beta}$ are sufficiently smooth*.

Let

$$\Psi = \{\psi_\lambda : \lambda \in \Lambda\}$$

be a countable collection of functions in $H_0^t(\Omega)$, where we have in mind Ψ to be an aggregated wavelet frame. The index λ encodes both the *level*, denoted by $|\lambda| \in \mathbb{N}_0$, and the location of the wavelet ψ_λ.

We assume that the wavelets are *local* in the sense that

$$\operatorname{diam}(\operatorname{supp}\psi_\lambda) \lesssim 2^{-|\lambda|} \quad \text{and} \quad \sup_{x\in\Omega,\ell\in\mathbb{N}_0} \#\{|\lambda| = \ell : B(x;2^{-\ell}) \cap \operatorname{supp}\psi_\lambda \neq \emptyset\} < \infty,$$

and that they are *piecewise smooth*, with which we mean that $\operatorname{supp}\psi_\lambda \backslash \operatorname{sing}\operatorname{supp}\psi_\lambda$ is the disjoint union of m domains $\Xi_{\lambda,1},\ldots,\Xi_{\lambda,m}$, with $\cup_{i=1}^m \overline{\Xi_{\lambda,i}} = \operatorname{supp}\psi_\lambda$, where $\psi_\lambda|_{\Xi_{\lambda,i}}$ is smooth with, for any $\gamma \in \mathbb{N}_0^n$,

$$\sup_{x\in\Xi_{\lambda,i}} |\partial^\gamma \psi_\lambda(x)| \lesssim 2^{|\lambda|(\frac{n}{2}+|\gamma|-t)}. \tag{2.1}$$

We assume that there is a smooth, regular mapping $\kappa_\lambda : \mathbb{R}^n \to \mathbb{R}^n$, for which each derivative is bounded, uniformly in λ, such that $\kappa_\lambda^{-1}(\Xi_{\lambda,i})$ is an n-cube aligned with the Cartesian coordinates, and

$$(\psi_\lambda \circ \kappa_\lambda)|_{\kappa_\lambda^{-1}(\Xi_{\lambda,i})} \in Q_{d-1},$$

with Q_{d-1} being the $n-$fold tensor product of the space of univariate polynomials of degree $d-1$. Thinking of an aggregated wavelet frame, κ_λ is just the parametric mapping used to lift the wavelet on $(0,1)^n$ to the subdomain. For some

$$\mathbb{N} \ni d \geqslant t+1,$$

we assume that, when $d \geqslant 2$,

$$\psi_\lambda \in C^{d-2}(\Omega).$$

By (2.1), this shows that for $k \in [0, d-1]$,

$$\|\psi_\lambda\|_{W_\infty^k(\Omega)} \lesssim 2^{|\lambda|(\frac{n}{2}+k-t)}. \tag{2.2}$$

In view of these assumptions, recall from Sect. 1 that even for smooth u, the largest s for which any $\mathbf{u} \in \ell_2(\Lambda)$ with $u = \sum_{\lambda\in\Lambda} \mathbf{u}_\lambda \psi_\lambda$ can be expected to be in $\mathscr{A}^s$ is $s = \frac{d-t}{n}$. Our task is therefore to prove s^*-computability of $\mathbf{M} = [(L\psi_{\lambda'})(\psi_\lambda)]_{\lambda,\lambda'\in\Lambda}$ for some $s^* > \frac{d-t}{n}$.

First of all, we show that $\mathbf{M}$ is sufficiently compressible. To this end, we split $\mathbf{M}$ into $\mathbf{M}^{(\mathrm{r})} + \mathbf{M}^{(\mathrm{s})}$, with $\mathbf{M}^{(\mathrm{r})}$ containing those entries $(L\psi_{\lambda'})(\psi_\lambda)$ of $\mathbf{M}$ with

$$\begin{cases} \operatorname{supp}\psi_\lambda \subset \overline{\Xi_{\lambda',i'}}, \text{ for some } 1 \leqslant i' \leqslant m, & \text{when } |\lambda| > |\lambda'|, \\ \operatorname{supp}\psi_{\lambda'} \subset \overline{\Xi_{\lambda,i}}, \text{ for some } 1 \leqslant i \leqslant m, & \text{when } |\lambda| < |\lambda'|, \end{cases}$$

and zeros at the remaining locations in $\Lambda \times \Lambda$, and thus with $\mathbf{M}^{(\mathrm{s})}$ being the matrix containing the remaining entries of $\mathbf{M}$, and zeros otherwise (see Figure 2). *The indices "r" and "s" refer to regular and singular, respectively.*

The collection of wavelets Ψ is said to have $\tilde{d} \in \mathbb{N}_0$ *vanishing moments* when, if $\tilde{d} > 0$,

$$\psi_\lambda \circ \kappa_\lambda \perp P_{\tilde{d}-1},$$

possibly with the exception of the λ with $|\lambda| = 0$.

In order not to be forced to handle $n = 1$ as an exceptional, although easy case, unless explicitly stated otherwise, in the following we will always assume that

$$n > 1.$$

THEOREM 2.1 For $j \in \mathbb{N}_0$, we define the infinite matrices $\mathbf{M}_j^{(\mathrm{r})}$ and $\mathbf{M}_j^{(\mathrm{s})}$ by dropping the entries $\mathbf{M}_{\lambda,\lambda'} = (L\psi_{\lambda'})(\psi_\lambda)$ from $\mathbf{M}^{(\mathrm{r})}$ or $\mathbf{M}^{(\mathrm{s})}$ when

$$\big||\lambda| - |\lambda'|\big| > \tfrac{j}{n} \quad \text{or} \quad \big||\lambda| - |\lambda'|\big| > \tfrac{j}{n-1}, \quad \text{respectively.}$$

Then the number of nonzero entries in each row and column of $\mathbf{M}_j^{(\mathrm{r})}$ and $\mathbf{M}_j^{(\mathrm{s})}$ is of order 2^j, and

$$\|\mathbf{M}^{(\mathrm{r})} - \mathbf{M}_j^{(\mathrm{r})}\| \lesssim 2^{-j(\frac{t+\tilde{d}}{n})}, \quad \|\mathbf{M}^{(\mathrm{s})} - \mathbf{M}_j^{(\mathrm{s})}\| \lesssim 2^{-j(\frac{d-1/2-t}{n-1})}, \tag{2.3}$$

for the latter estimate assuming that $\tilde{d} \geqslant d - 2t - 1$.

REMARK 2.1 The corresponding result from Stevenson (2004) gives the same bound for $\|\mathbf{M}^{(\mathrm{r})} - \mathbf{M}_j^{(\mathrm{r})}\|$, whereas it shows that for any $s < \frac{d-1/2-t}{n-1}$, $\|\mathbf{M}^{(\mathrm{s})} - \mathbf{M}_j^{(\mathrm{s})}\| \lesssim 2^{-js}$.

So Theorem 2.1 shows that $\mathbf{M}^{(\mathrm{r})}$ is s^*-compressible with $s^* \geqslant \frac{d-t}{n}$ or $s^* > \frac{d-t}{n}$ when $\tilde{d} \geqslant d - 2t$ or $\tilde{d} > d - 2t$, and that $\mathbf{M}^{(\mathrm{s})}$ is s^*-compressible with $s^* \geqslant \frac{d-t}{n}$ or $s^* > \frac{d-t}{n}$ when $\frac{d-t}{n} \geqslant \frac{1}{2}$ or $\frac{d-t}{n} > \frac{1}{2}$, and $\tilde{d} \geqslant d - 2t - 1$.

In order to prove it, we start with bounding the individual entries of $\mathbf{M}$.

LEMMA 2.1 We have

$$|\mathbf{M}_{\lambda,\lambda'}^{(\mathrm{r})}| \lesssim 2^{-\big||\lambda|-|\lambda'|\big|(\frac{n}{2}+t+\tilde{d})}, \quad |\mathbf{M}_{\lambda,\lambda'}^{(\mathrm{s})}| \lesssim 2^{-\big||\lambda|-|\lambda'|\big|(\frac{n}{2}+d-1-t)},$$

for the latter estimate assuming that $\tilde{d} \geqslant d - 2t - 1$.

Proof. Let $|\lambda| \geqslant |\lambda'|$, $|\lambda| > 0$. By a transformation of coordinates, we can write

$$(L\psi_{\lambda'})(\psi_\lambda) = \sum_{|\alpha|,|\beta| \leqslant t} \int_{\kappa_\lambda^{-1}(\operatorname{supp}\psi_\lambda)} \tilde{a}_{\alpha,\beta} \partial^\alpha(\psi_{\lambda'} \circ \kappa_\lambda) \partial^\beta(\psi_\lambda \circ \kappa_\lambda), \tag{2.4}$$

for some smooth $\tilde{a}_{\alpha,\beta}$ depending on the coefficients $a_{\alpha,\beta}$ and κ_λ. Since bounding the lower order terms is easier, we consider a term of the right-hand side of (2.4) for arbitrary $|\alpha| = |\beta| = t$.

When $d - 1 \leqslant 2t$, select a $\gamma \leqslant \beta$ with $|\alpha + \gamma| = d - 1$ and so $|\beta - \gamma| = 2t - (d-1)$. Using the homogeneous Dirichlet boundary conditions for the case that $\operatorname{supp}\psi_\lambda \cap \operatorname{supp}\psi_{\lambda'} \cap \partial\Omega \neq \emptyset$, integration by parts, $\operatorname{vol}(\operatorname{supp}\psi_\lambda) \lesssim 2^{-|\lambda|n}$, and (2.2) show that

$$\begin{aligned}
&\Big|\int_{\kappa_\lambda^{-1}(\operatorname{supp}\psi_\lambda)} \tilde{a}_{\alpha,\beta} \partial^\alpha(\psi_{\lambda'} \circ \kappa_\lambda) \partial^\beta(\psi_\lambda \circ \kappa_\lambda)\Big| \\
&\qquad = \Big|\int_{\kappa_\lambda^{-1}(\operatorname{supp}\psi_\lambda)} (-1)^{|\gamma|} \partial^\gamma(\tilde{a}_{\alpha,\beta} \partial^\alpha(\psi_{\lambda'} \circ \kappa_\lambda)) \partial^{\beta-\gamma}(\psi_\lambda \circ \kappa_\lambda)\Big| \\
&\qquad \lesssim 2^{-|\lambda|n} 2^{|\lambda'|(\frac{n}{2}+d-1-t)} 2^{|\lambda|(\frac{n}{2}+2t-(d-1)-t)} = 2^{-(|\lambda|-|\lambda'|)(\frac{n}{2}+d-1-t)}.
\end{aligned}$$

When $d-1>2t$, by additionally using that the wavelets have $\tilde{d} \geqslant d-2t-1$ vanishing moments, we obtain

$$\begin{aligned}
|\int_{\kappa_\lambda^{-1}(\operatorname{supp}\psi_\lambda)} \cdots| &= |\int_{\kappa_\lambda^{-1}(\operatorname{supp}\psi_\lambda)} (-1)^{|\beta|}\partial^\beta(\tilde{a}_{\alpha,\beta}\partial^\alpha(\psi_{\lambda'}\circ\kappa_\lambda))(\psi_\lambda\circ\kappa_\lambda)| \\
&\lesssim 2^{-|\lambda|n} \inf_{p\in P_{d-2t-2}} \|(-1)^{|\beta|}\partial^\beta(\tilde{a}_{\alpha,\beta}\partial^\alpha(\psi_{\lambda'}\circ\kappa_\lambda)) - p\|_{L_\infty(\kappa_\lambda^{-1}(\operatorname{supp}\psi_\lambda))} 2^{|\lambda|(\frac{n}{2}-t)} \\
&\lesssim 2^{-|\lambda|n}\operatorname{diam}(\kappa_\lambda^{-1}(\operatorname{supp}\psi_\lambda))^{d-2t-1}\|(-1)^{|\beta|}\partial^\beta(\tilde{a}_{\alpha,\beta}\partial^\alpha(\psi_{\lambda'}\circ\kappa_\lambda))\|_{W_\infty^{d-2t-1}} 2^{|\lambda|(\frac{n}{2}-t)} \\
&\lesssim 2^{-|\lambda|n}2^{-|\lambda|(d-1-2t)}2^{|\lambda'|(\frac{n}{2}+d-1-t)}2^{|\lambda|(\frac{n}{2}-t)} = 2^{-(|\lambda|-|\lambda'|)(\frac{n}{2}+d-1-t)},
\end{aligned}$$

which completes the proof of the second estimate.

Finally, when $\operatorname{supp}\psi_\lambda \subset \overline{\Xi_{\lambda',i'}}$ for some $1 \leqslant i' \leqslant m$, i.e., when $(L\psi_{\lambda'})(\psi_\lambda) = \mathbf{M}^{(\mathrm{r})}_{\lambda,\lambda'}$, using (2.1) we find

$$\begin{aligned}
|\int_{\kappa_\lambda^{-1}(\operatorname{supp}\psi_\lambda)} \cdots| &\lesssim 2^{-|\lambda|n} \inf_{p\in P_{\tilde{d}-1}} \|(-1)^{|\beta|}\partial^\beta(\tilde{a}_{\alpha,\beta}\partial^\alpha(\psi_{\lambda'}\circ\kappa_\lambda)) - p\|_{L_\infty(\kappa_\lambda^{-1}(\operatorname{supp}\psi_\lambda))} 2^{|\lambda|(\frac{n}{2}-t)} \\
&\lesssim 2^{-|\lambda|n}2^{-|\lambda|\tilde{d}}2^{|\lambda'|(\frac{n}{2}+\tilde{d}+2t-t)}2^{|\lambda|(\frac{n}{2}-t)} = 2^{-(|\lambda|-|\lambda'|)(\frac{n}{2}+t+\tilde{d})}.
\end{aligned}$$

□

REMARK 2.2 Since $\psi_{\lambda'}\circ\kappa_{\lambda'} \in Q_{d-1}$, in the exponent of the upper bound $\eqsim 2^{|\lambda'|(\frac{n}{2}+|\gamma|-t)}$ for $\sup_{x\in\kappa_\lambda^{-1}(\overline{\Xi_{\lambda',i}})} |\partial^\gamma(\psi_{\lambda'}\circ\kappa_\lambda)(x)|$ the term $|\gamma|$ can be replaced by $\max(n(d-1),|\gamma|)$. As a consequence, for sufficiently large $\tilde{d}$, the last estimate from above proof, and so the upper bound for $\mathbf{M}^{(\mathrm{r})}_{\lambda,\lambda'}$ is not sharp for $|\lambda'|>0$, in the sense that the bound is still valid when multiplied with a certain positive power of $2^{-|\lambda'|}$. So the case $|\lambda'|=0$ is the most difficult one. This observation is valid for many estimates that will be derived in this paper.

Proof of Theorem 2.1. The locality of the wavelets shows that the number of nonzero entries in each row of $\mathbf{M}^{(\mathrm{r})}_{\ell,\ell'} := [\mathbf{M}^{(\mathrm{r})}_{\lambda,\lambda'}]_{|\lambda|=\ell,|\lambda'|=\ell'}$ and column of $\mathbf{M}^{(\mathrm{r})}_{\ell',\ell}$ is $\mathscr{O}(\max\{1,2^{(\ell'-\ell)n}\})$. The piecewise smoothness of the wavelets shows that the number of nonzero entries in each row of $\mathbf{M}^{(\mathrm{s})}_{\ell,\ell'} := [\mathbf{M}^{(\mathrm{s})}_{\lambda,\lambda'}]_{|\lambda|=\ell,|\lambda'|=\ell'}$ and column of $\mathbf{M}^{(\mathrm{s})}_{\ell',\ell}$ is $\mathscr{O}(\max\{1,2^{(\ell'-\ell)(n-1)}\})$. The definition of $\mathbf{M}^{(\mathrm{r})}_j$ and $\mathbf{M}^{(\mathrm{s})}_j$ shows that in each row and column they have $\mathscr{O}(2^j)$ nonzero entries.

Estimating $\|\mathbf{M}^{(\mathrm{r})}_{\ell,\ell'}\|^2$ and $\|\mathbf{M}^{(\mathrm{s})}_{\ell,\ell'}\|^2$ on the products of their maximal absolute row- and column sums, taking into account Lemma 2.1, we find

$$\|\mathbf{M}^{(\mathrm{r})}_{\ell,\ell'}\|^2 \lesssim 2^{|\ell'-\ell|n}2^{-|\ell'-\ell|(\frac{n}{2}+t+\tilde{d})2}, \quad \|\mathbf{M}^{(\mathrm{s})}_{\ell,\ell'}\|^2 \lesssim 2^{|\ell'-\ell|(n-1)}2^{-|\ell'-\ell|(\frac{n}{2}+d-1-t)2}.$$

By applying $\|\mathbf{M}^{(\mathrm{r})}-\mathbf{M}^{(\mathrm{r})}_j\|^2 \leqslant \max_{\ell'}\sum_{\{\ell:|\ell-\ell'|>\frac{j}{n}\}}\|\mathbf{M}^{(\mathrm{r})}_{\ell,\ell'}\| \times \max_\ell \sum_{\{\ell':|\ell-\ell'|>\frac{j}{n}\}}\|\mathbf{M}^{(\mathrm{r})}_{\ell,\ell'}\|$ and $\|\mathbf{M}^{(\mathrm{s})}-\mathbf{M}^{(\mathrm{s})}_j\|^2 \leqslant \max_{\ell'}\sum_{\{\ell:|\ell-\ell'|>\frac{j}{n-1}\}}\|\mathbf{M}^{(\mathrm{r})}_{\ell,\ell'}\| \times \max_\ell \sum_{\{\ell':|\ell-\ell'|>\frac{j}{n-1}\}}\|\mathbf{M}^{(\mathrm{r})}_{\ell,\ell'}\|$,
the proof is completed. □

3. Computability

The following lemma will be applied for $\mathbf{B}_j = \mathbf{M}^{(\mathrm{r})}_j$ with $k=n$, and for $\mathbf{B}_j = \mathbf{M}^{(\mathrm{s})}_j$ with $k=n-1$.

LEMMA 3.1 For some fixed $k \in \mathbb{N}$, and all $j \in \mathbb{N}_0$, let $\mathbf{B}_j = ((\mathbf{B}_j)_{\lambda,\lambda'})_{\lambda,\lambda' \in \Lambda}$ be a matrix such that the number of possible nonzero entries in each row of $(\mathbf{B}_j)_{\ell,\ell'} := [(\mathbf{B}_j)_{\lambda,\lambda'}]_{|\lambda|=\ell,|\lambda'|=\ell'}$ and column of $(\mathbf{B}_j)_{\ell',\ell}$ is $\mathscr{O}(\max\{1, 2^{(\ell'-\ell)k}\})$, and

$$(\mathbf{B}_j)_{\lambda,\lambda'} = 0 \quad \text{when } \big||\lambda| - |\lambda'|\big| > \tfrac{j}{k}.$$

Let $\mathbf{B}_j^*$ be an approximation for $\mathbf{B}_j$, zero on positions where $\mathbf{B}_j$ is known to be zero, and for which the computation of $(\mathbf{B}_j^*)_{\lambda,\lambda'}$ takes $\mathscr{O}(N_{j,\lambda,\lambda'})$ operations otherwise, where, for some absolute constants $r, q \geqslant 0$, $r \neq q$,

$$|(\mathbf{B}_j)_{\lambda,\lambda'} - (\mathbf{B}_j^*)_{\lambda,\lambda'}| \lesssim N_{j,\lambda,\lambda'}^{-q} 2^{-\big||\lambda|-|\lambda'|\big|(k/2+rk)}. \tag{3.1}$$

For some $\rho \in (1, r/q)$ when $r > q$, and $\rho \in (r/q, 1)$ when $r < q$, and $\theta \leqslant \min\{1,\rho\}$, select

$$N_{j,\lambda,\lambda'} \eqsim \max\Big\{1, 2^{j\theta - \big||\lambda|-|\lambda'|\big|\rho k}\Big\}.$$

Then the work for computing each column of $\mathbf{B}_j^*$ is $\mathscr{O}(2^j)$, and

$$\|\mathbf{B}_j - \mathbf{B}_j^*\| \lesssim \begin{cases} 2^{-jq\theta} & \text{when } r > q, \\ 2^{-j(r+(\theta-\rho)q)} & \text{when } r < q. \end{cases} \tag{3.2}$$

In particular, taking $\theta = \min\{1,\rho\}$, we have $\|\mathbf{B}_j - \mathbf{B}_j^*\| \lesssim 2^{-j\min\{q,r\}}$.

Proof. The work per column (or row) is less than or equal to an absolute multiple of

$$\sum_{m=0}^{j/k} 2^{mk} \max\Big\{1, 2^{j\theta - m\rho k}\Big\} \eqsim 2^j + 2^{j\theta} \sum_{m=0}^{j/k} 2^{mk(1-\rho)} \eqsim 2^j + 2^{j\theta} \max\{1, 2^{j(1-\rho)}\} \eqsim 2^j,$$

because of $\theta \leqslant \min\{1,\rho\}$.

By bounding the squared norm of $(\mathscr{E}_j)_{\ell,\ell'} = (\mathbf{B}_j)_{\ell,\ell'} - [(\mathbf{B}_j^*)_{\lambda,\lambda'}]_{|\lambda|=\ell,|\lambda'|=\ell'}$ on the product of its maximal absolute row- and column sum, taking into account the number of non-zero entries in each row and column, (3.1), and the selection of $N_{j,\lambda,\lambda'}$, we find that

$$\|\mathscr{E}_{j,\ell,\ell'}\| \lesssim (1 \cdot 2^{|\ell-\ell'|k})^{\frac{1}{2}} (2^{j\theta - |\ell-\ell'|\rho k})^{-q} 2^{-|\ell-\ell'|(k/2+rk)} = 2^{-j\theta q} 2^{-|\ell-\ell'|k(r-\rho q)}.$$

By bounding $\|\mathbf{B}_j - \mathbf{B}_j^*\|^2$ on $\max_{\ell'} \sum_{\{\ell : |\ell-\ell'| \leqslant \frac{j}{k}\}} \|(\mathscr{E}_j)_{\ell,\ell'}\| \times \max_{\ell} \sum_{\{\ell' : |\ell-\ell'| \leqslant \frac{j}{k}\}} \|(\mathscr{E}_j)_{\ell,\ell'}\|$, we arrive at (3.2). $\square$

REMARK 3.1 For $r = q$, one easily infers that one can compute a $\mathbf{B}_j^*$ taking $\mathscr{O}(2^j)$ operations per column, with for any $\varepsilon > 0$, $\|\mathbf{B}_j - \mathbf{B}_j^*\| \lesssim 2^{-j(\min\{q,r\}-\varepsilon)}$.

Now we come to the task of approximately computing the entries of $\mathbf{M}_j^{(\mathrm{r})}$ and $\mathbf{M}_j^{(\mathrm{s})}$. We will exclusively apply composite quadrature rules of *variable rank* $N \in \mathbb{N}$, depending on j and $\big||\lambda| - |\lambda'|\big|$, but *fixed order* $p \in \mathbb{N}$, on n-cubes aligned with the Cartesian coordinates. That is, we subdivide the n-cube under consideration into N equal subcubes, and, on each of these subcubes $\square$, we apply a quadrature rule that is *exact on* $Q_{p-1}(\square)$. We assume that this rule is internal, i.e., that all abscissae are in the closure of the subcube, and that it is uniformly stable, i.e., that the sum of the absolute values of the

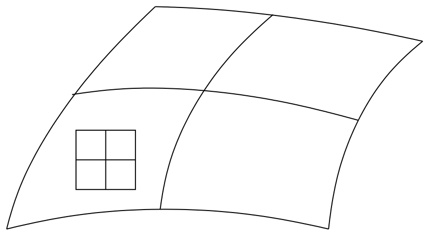

FIG. 2. $\operatorname{supp}\psi_\lambda \subset \overline{\Xi_{\lambda',i'}}$, i.e., an entry of $\mathbf{M}^{(\mathrm{r})}$.

weights can be bounded by an absolute multiple of the volume of the subcube. Finally, having a fixed p, we can assume that the total number of abscissae is $\mathscr{O}(N)$.

Without loss of generality, *in the remainder of this section and in the next section* we assume that

$$\boxed{|\lambda| \geqslant |\lambda'|}.$$

For notational convenience, and, in view of (2.4), without loss of generality, we will assume that $\boxed{\kappa_\lambda = \mathrm{id}}$, so that

$$\mathbf{M}_{\lambda,\lambda'} = \sum_{i=1}^{m} \sum_{|\alpha|,|\beta|\leqslant t} \int_{\Xi_{\lambda,i}} a_{\alpha,\beta}\partial^\alpha \psi_{\lambda'}\partial^\beta \psi_\lambda, \tag{3.3}$$

each $\Xi_{\lambda,i}$ being an n-cube aligned with the Cartesian coordinates, and $\boxed{\psi_\lambda|_{\Xi_{\lambda,i}} \in Q_{d-1}}$.

We first consider the task of approximating the entries of $\mathbf{M}_j^{(\mathrm{r})}$.

PROPOSITION 3.1 Let $\operatorname{supp}\psi_\lambda \subset \overline{\Xi_{\lambda',i'}}$ for some $1 \leqslant i' \leqslant m$, see Figure 2 for an illustration. Let $\mathbf{M}^{(\mathrm{r}),*}_{\lambda,\lambda'}$ be the result of the application of a composite rule of rank N and order p applied to each of the integrals from (3.3). Then

$$|\mathbf{M}^{(\mathrm{r})}_{\lambda,\lambda'} - \mathbf{M}^{(\mathrm{r}),*}_{\lambda,\lambda'}| \lesssim N^{-p/n} 2^{-(|\lambda|-|\lambda'|)(n/2+p-d+1)}.$$

Proof. Although this proof can already be found in Gantumur and Stevenson (2006a), for convenience we recall it here. It is sufficient to consider one integral $\int_{\Xi_{\lambda,i}} a_{\alpha,\beta}\partial^\alpha \psi_{\lambda'}\partial^\beta \psi_\lambda$. Using that each of the N subcubes of $\Xi_{\lambda,i}$ has diameter $\lesssim 2^{-|\lambda|}N^{-1/n}$, and thus volume $\lesssim 2^{-|\lambda|n}N^{-1}$, standard estimates show that the quadrature error can be bounded by an absolute multiple of

$$\sum_{\square} 2^{-|\lambda|n} N^{-1} (2^{-|\lambda|}N^{-1/n})^p \max_{1\leqslant k\leqslant n} \|\partial_k^p (a_{\alpha,\beta}\partial^\alpha \psi_{\lambda'}\partial^\beta \psi_\lambda)\|_{L_\infty(\square)},$$

where $\square$ runs over the N subcubes. In order to bound $\|\partial_k^u a_{\alpha,\beta}\partial_k^v\partial^\alpha \psi_{\lambda'}\partial_k^w\partial^\beta \psi_\lambda\|_{L_\infty(\square)}$ for any $u+v+w=p$, since $a_{\alpha,\beta}$ is smooth, $|\lambda| \geqslant |\lambda'|$, and $\partial_k^w\partial^\beta\psi_\lambda$ vanishes when $\beta_k + w \geqslant d$, by invoking the bound (2.1) on the partial derivatives of $\psi_{\lambda'}$ and ψ_λ we see that the worst case occurs when $u = 0$, $\beta_k + w = r_k := \min\{d-1, \beta_k + p\}$, and thus $v = p - r_k + \beta_k$, yielding

$$\begin{aligned} \|\partial_k^p \left(a_{\alpha,\beta}\partial^\alpha \psi_{\lambda'}\partial^\beta \psi_\lambda\right)\|_{L_\infty(\square)} &\lesssim 2^{|\lambda'|(|\alpha|+p-r_k+\beta_k+n/2-t)} 2^{|\lambda|(|\beta|+r_k-\beta_k+n/2-t)} \\ &\leqslant 2^{|\lambda'|(p-d+1+|\alpha|+n/2-t)} 2^{|\lambda|(|\beta|+d-1+n/2-t)}, \end{aligned} \tag{3.4}$$

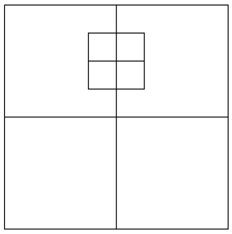

FIG. 3. $\operatorname{supp}\psi_\lambda \not\subset \overline{\Xi_{\lambda',i'}}$ for any $1 \leqslant i' \leqslant m$, i.e., an entry of $\mathbf{M}^{(\mathrm{s})}$, but $\Xi_{\lambda,i} \subset \Xi_{\lambda',i'(i)}$ for any i.

where we used that $2^{(|\lambda|-|\lambda'|)(r_k-\beta_k)} \leqslant 2^{(|\lambda|-|\lambda'|)(d-1)}$, that is sharp when $\beta_k = 0$ and $p \geqslant d-1$. Upon applying that $|\alpha|, |\beta| \leqslant t$, and using that the number of the subcubes is N, the proof is completed. $\square$

REMARK 3.2 In above proof, the worst case occurs when $\beta_k = 0$ and $|\beta| = t$. For $n = 1$ and $t > 0$, both equalities cannot hold simultaneously. One may verify that for $n = 1$, the upper bound from Proposition 3.1 can be sharpened to $N^{-p}2^{-(|\lambda|-|\lambda'|)(3/2+p+t-d)}$.

COROLLARY 3.1 Let $\tilde{d} > d - 2t$ and $p > 2d - t - 1$, then $\mathbf{M}^{(\mathrm{r})}$ is s^*-computable with $s^* > \frac{d-t}{n}$.

Proof. Recall that the number of nonzero entries in each row of $\mathbf{M}^{(\mathrm{r})}_{\ell,\ell'}$ and column of $\mathbf{M}^{(\mathrm{r})}_{\ell',\ell}$ is $\mathscr{O}(\max\{1, 2^{(\ell'-\ell)n}\})$. Using Proposition 3.1, an application of Lemma 3.1 for $k = n$ yields a matrix $\mathbf{M}^{(\mathrm{r}),*}_j$, for which the computation of each column takes $\mathscr{O}(2^j)$ operations, and for which $\|\mathbf{M}^{(\mathrm{r})}_j - \mathbf{M}^{(\mathrm{r}),*}_j\| \lesssim 2^{-j\left(\frac{p-d+1}{n}\right)}$ when $d > 1$, and $\|\mathbf{M}^{(\mathrm{r})}_j - \mathbf{M}^{(\mathrm{r}),*}_j\| \lesssim 2^{-j\left(\frac{p-d+1}{n}-\varepsilon\right)}$ for any $\varepsilon > 0$ otherwise (cf. Remark 3.1). Using that $\|\mathbf{M}^{(\mathrm{r})} - \mathbf{M}^{(\mathrm{r})}_j\| \lesssim 2^{-j\left(\frac{t+\tilde{d}}{n}\right)}$ by Theorem 2.1, the proof is completed. $\square$

Next we consider the approximation of the non-zero entries $\mathbf{M}_{\lambda,\lambda'}$ of $\mathbf{M}^{(\mathrm{s})}_j$. Note that for these entries, $\operatorname{supp}\psi_\lambda$ will have a non-empty intersection with the singular support of $\psi_{\lambda'}$. As a consequence, for p not too small, generally the decay of the quadrature error will not be as fast as function of the rank $N \to \infty$ or $\big||\lambda| - |\lambda'|\big| \to \infty$ as with the entries of $\mathbf{M}^{(\mathrm{r})}_j$. However, since the number of non-zero entries in $\mathbf{M}^{(\mathrm{s})}_{\ell,\ell'}$ increases less fast as function of $|\ell - \ell'| \to \infty$ as that in $\mathbf{M}^{(\mathrm{r})}_{\ell,\ell'}$, as shown in Lemma 3.1, this effect can be compensated by investing some more work in their computation without increasing the overall complexity.

REMARK 3.3 Let (λ, λ') correspond to a non-zero entry of $\mathbf{M}^{(\mathrm{s})}$, i.e., $\operatorname{supp}\psi_\lambda \not\subset \overline{\Xi_{\lambda',i'}}$ for any $1 \leqslant i' \leqslant m$, but such that for all $1 \leqslant i \leqslant m$ there exists an $1 \leqslant i'(i) \leqslant m$ with $\Xi_{\lambda,i} \subset \Xi_{\lambda',i'(i)}$, meaning that $\operatorname{sing\,supp}\psi_{\lambda'} \cap \operatorname{supp}\psi_\lambda \subset \operatorname{sing\,supp}\psi_\lambda$, see Figure 3. Then it is obvious that the bound of the quadrature error from Proposition 3.1 is still valid. This situation occurs when $\psi_{\lambda'}$ and ψ_λ are piecewise smooth with respect to partitions that are *nested* as function of the level, e.g., in the case of an aggregated wavelet frame, when $\psi_{\lambda'}$ and ψ_λ are wavelets lifted by the same parametric mapping. In order not to complicate our exposition, in the following we will ignore this fact, and use also for such entries the less favourable bound on the quadrature error from the following proposition.

PROPOSITION 3.2 Let $\mathbf{M}^*_{\lambda,\lambda'}$ be the result of the application of a composite rule of rank N and order p

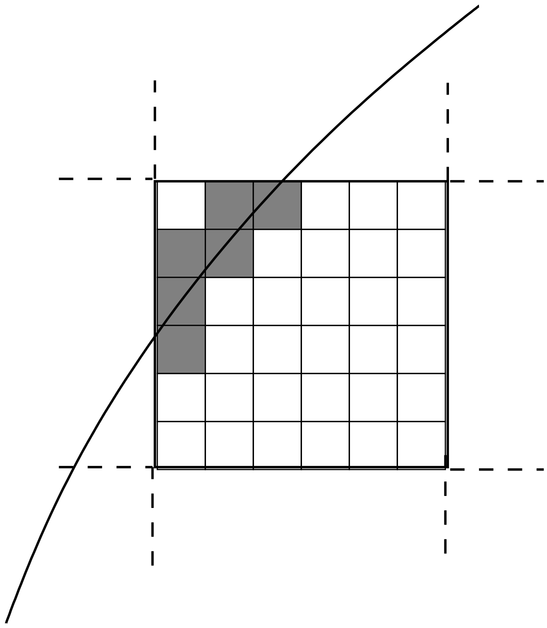

FIG. 4. Set $V_{\lambda,i,\lambda'}$ of subcubes in the quadrature mesh on $\Xi_{\lambda,i}$ on which $\psi_{\lambda'}$ is not arbitrarily smooth.

applied to each of the integrals from (3.3). Then

$$|\mathbf{M}_{\lambda,\lambda'} - \mathbf{M}^*_{\lambda,\lambda'}| \lesssim N^{-p/n} 2^{-(|\lambda|-|\lambda'|)(n/2+p-d+1)} + N^{-(d-t)/n} 2^{-(|\lambda|-|\lambda'|)(n/2+d-1-t)},$$

which is valid without any assumption on the location of $\operatorname{sing\,supp} \psi_\lambda$.

Note that the first term in the upper bound is equal to the bound given in Proposition 3.1.
Proof. As in the proof of Proposition 3.1, we have to consider only one integral $\int_{\Xi_{\lambda,i}} a_{\alpha,\beta} \partial^\alpha \psi_{\lambda'} \partial^\beta \psi_\lambda$, where it is now sufficient to consider the hard case that $\Xi_{\lambda,i} \cap \operatorname{sing\,supp} \psi_{\lambda'} \neq \emptyset$. Let us denote with $V_{\lambda,i,\lambda'}$ the set of n-cubes $\square$ in the quadrature mesh on $\Xi_{\lambda,i}$ on which $\psi_{\lambda'}$ is not arbitrarily smooth, see Figure 4. Using that $\partial^\alpha \psi_{\lambda'} \in W^{d-1-|\alpha|}_\infty(\Omega)$, in particular using the bound (2.2), for any $\square \in V_{\lambda,i,\lambda'}$ its Taylor polynomial $q \in P_{d-2-|\alpha|}$ of order $d-1-|\alpha|$ at some point $y \in \square$, with $q := 0$ when $d-1-|\alpha| = 0$, satisfies

$$\begin{aligned}\|\partial^\alpha \psi_{\lambda'} - q\|_{L_\infty(\square)} &\lesssim (N^{-1/n} 2^{-|\lambda|})^{d-1-|\alpha|} \|\partial^\alpha \psi_{\lambda'}\|_{W^{d-1-|\alpha|}_\infty(\square)} \\ &\lesssim N^{-(d-1-|\alpha|)/n} 2^{-|\lambda|(d-1-|\alpha|)} 2^{|\lambda'|(d-1+n/2-t)}.\end{aligned}$$

From $q \in P_{d-2-|\alpha|}$, for $|\eta| \leqslant d-2-|\alpha|$ we have $\partial^\eta q \in P_{d-2-|\alpha|-|\eta|}$, and so

$$\begin{aligned}\partial^\eta q(x) &= \sum_{j=0}^{d-2-|\alpha|-|\eta|} \frac{1}{j!}\left\{ \Big(\sum_{i=1}^{n}(x_i-y_i)\frac{\partial}{\partial x_i'}\Big)^j(\partial^\eta q)(x')\right\}\Big|_{x'=y} \\ &= \sum_{j=0}^{d-2-|\alpha|-|\eta|} \frac{1}{j!}\left\{ \Big(\sum_{i=1}^{n}(x_i-y_i)\frac{\partial}{\partial x_i'}\Big)^j(\partial^{\eta+\alpha} \psi_{\lambda'})(x')\right\}\Big|_{x'=y}.\end{aligned}$$

Invoking the bounds for $\|\partial^\zeta \psi_{\lambda'}\|_{L_\infty(\Omega)}$ for $|\zeta| \leqslant d-2$, we infer that for $|x-y| \lesssim 2^{-|\lambda'|}$, so in particular for $x \in \square$, we have

$$|\partial^\eta q(x)| \lesssim 2^{|\lambda'|(|\alpha|+|\eta|+n/2-t)}, \tag{3.5}$$

obviously being also valid for $|\eta| > d-2-|\alpha|$. Note that this bound on $\|\partial^\eta q\|_{L_\infty(\square)}$ is equal to that on $\|\partial^{\eta+\alpha}\psi_{\lambda'}\|_{L_\infty(\square)}$ from (2.1) in case $\square \notin V_{\lambda,i,\lambda'}$.

Let us now think of the composite quadrature rule as being applied to the integrand $a_{\alpha,\beta}\partial^\alpha\psi_{\lambda'}\partial^\beta\psi_\lambda$ with, on any $\square \in V_{\lambda,i,\lambda'}$, the factor $\partial^\alpha\psi_{\lambda'}$ being replaced by the corresponding q. Since the fact that generally the modified integrand is discontinuous over interfaces between different n-cubes in the quadrature mesh does not effect the error of the composite quadrature rule, using (3.5) the proof of Proposition 3.1 shows that this quadrature error can be bounded by an absolute multiple of

$$N^{-p/n} 2^{-(|\lambda|-|\lambda'|)(n/2+p-d+1)}. \tag{3.6}$$

To bound the total error we have to add the sum over $\square \in V_{\lambda,i,\lambda'}$ of bounds for the error of the quadrature rule on $\square$ with integrand $a_{\alpha,\beta}(\partial^\alpha\psi_{\lambda'} - q)\partial^\beta\psi_\lambda$. On each of such a $\square$, this error can be bounded by an absolute multiple of

$$\begin{aligned}&\mathrm{vol}(\square)\|a_{\alpha,\beta}(\partial^\alpha\psi_{\lambda'}-q)\partial^\beta\psi_\lambda\|_{L_\infty(\square)} \\ &\lesssim (N^{-1/n}2^{-|\lambda|})^n N^{-(d-1-|\alpha|)/n} 2^{-|\lambda|(d-1-|\alpha|)} 2^{|\lambda'|(d-1+n/2-t)} 2^{|\lambda|(|\beta|+n/2-t)} \\ &\lesssim N^{-1-(d-1-t)/n} 2^{-(|\lambda|-|\lambda'|)(n/2+d-1-t)},\end{aligned} \tag{3.7}$$

the last inequality being sharp for $|\alpha| = |\beta| = t$. Since $\#V_{\lambda,i,\lambda'} \lesssim N^{(n-1)/n}$, we find that the additional error can be bounded by an absolute multiple of

$$N^{-(d-t)/n} 2^{-(|\lambda|-|\lambda'|)(n/2+d-1-t)}.$$

which completes the proof. $\square$

REMARK 3.4 One might wonder why for $\square \in V_{\lambda,i,\lambda'}$ we wrote $a_{\alpha,\beta}\partial^\alpha\psi_{\lambda'}\partial^\beta\psi_\lambda = a_{\alpha,\beta}q\partial^\beta\psi_\lambda + a_{\alpha,\beta}(\partial^\alpha\psi_{\lambda'} - q)\partial^\beta\psi_\lambda$, and estimated the quadrature error for both terms separately. Indeed, using $a_{\alpha,\beta}\partial^\alpha\psi_{\lambda'}\partial^\beta\psi_\lambda \in W_\infty^{d-1-|\alpha|}(\square)$, alternatively one can apply a standard error estimate for a quadrature rule of order $\min\{d-1-|\alpha|, p\}$ applied to this integrand. Invoking the bounds on the partial derivatives of ψ_λ and $\psi_{\lambda'}$, however, one would end up with a bound on the quadrature error as in Proposition 3.2 with the second term $N^{-(d-t)/n}2^{-(|\lambda|-|\lambda'|)(n/2+d-1-t)}$ replaced by $N^{-(d-t)/n}2^{-(|\lambda|-|\lambda'|)n/2}$.

COROLLARY 3.2 Let $p \geqslant \max\{d-t, 2d-2-t\}$, $\frac{d-t}{n} > \frac{1}{2}$ and $\tilde{d} \geqslant d-2t-1$. Then $\mathbf{M}^{(s)}$ is s^*-computable with $s^* = \frac{d-t}{n}$.

Proof. Recall that the number of nonzero entries in each row of $\mathbf{M}^{(\mathrm{s})}_{\ell,\ell'}$ and column of $\mathbf{M}^{(\mathrm{s})}_{\ell',\ell}$ is $\mathscr{O}(\max\{1,2^{(\ell'-\ell)(n-1)}\})$. The condition $p \geqslant \max\{d-t, 2d-2-t\}$ shows that in the bound from Proposition 3.2 the first term is never larger than the second one, that can be written as $N^{-(d-t)/n}2^{-(|\lambda|-|\lambda'|)((n-1)/2+\frac{d-1/2-t}{n-1}(n-1))}$. The condition $\frac{d-t}{n} > \frac{1}{2}$ shows that $\frac{d-1/2-t}{n-1} > \frac{(d-t)}{n}$, and so an application of Lemma 3.1 for $k=n-1$ yields a matrix $\mathbf{M}^{(\mathrm{s}),*}_j$, for which the computation of each column takes $\mathscr{O}(2^j)$ operations, and for which $\|\mathbf{M}^{(\mathrm{s})}_j - \mathbf{M}^{(\mathrm{s}),*}_j\| \lesssim 2^{-j\left(\frac{d-t}{n}\right)}$. Using that by $\frac{d-t}{n} \geqslant \frac{1}{2}$ and $\tilde{d} \geqslant d-2t-1$, $\|\mathbf{M}^{(\mathrm{s})} - \mathbf{M}^{(\mathrm{s})}_j\| \lesssim 2^{-j\left(\frac{d-t}{n}\right)}$ by Theorem 2.1, the proof is completed. $\square$

Above result is not fully satisfactory, since actually we need $s^* > \frac{d-t}{n}$. For $s^* = \frac{d-t}{n}$, generally the adaptive frame algorithms can only be shown to be optimal up to some log factors. Below, we reconsider the task of approximately computing the entries of $\mathbf{M}^{(\mathrm{s})}_j$. For $\eta \leqslant \alpha$, $|\beta|+|\eta| \leqslant d-1$, *integration by parts* shows that

$$\int_\Omega a_{\alpha\beta}\partial^\alpha\psi_{\lambda'}\partial^\beta\psi_\lambda = (-1)^{|\eta|}\sum_{i=1}^m \int_{\Xi_{\lambda,i}} \partial^{\alpha-\eta}\psi_{\lambda'}\partial^\eta(a_{\alpha,\beta}\partial^\beta\psi_\lambda), \tag{3.8}$$

so that alternatively one can apply the composite quadrature rule to each term on the right-hand side, with the advantage that $\partial^{\alpha-\eta}\psi_{\lambda'}$ is smoother than $\partial^\alpha\psi_{\lambda'}$. We will consider this approach for the *largest possible* η, i.e., $|\eta| = \min\{|\alpha|, d-1-|\beta|\}$. Formula (3.4) now reads as

$$\|\partial_k^p\left(\partial^{\alpha-\eta}\psi_{\lambda'}\partial^\eta(a_{\alpha,\beta}\partial^\beta\psi_\lambda)\right)\|_{L_\infty(\square)} \lesssim 2^{|\lambda'|(p-d+1+|\alpha|-|\eta|+n/2-t)}2^{|\lambda|(|\beta|+|\eta|+d-1+n/2-t)},$$

which is sharp when $\beta_k = \eta_k = 0$ and $p \geqslant d-1$. As in the proof of Proposition 3.2, for each $\square \in V_{\lambda,i,\lambda'}$, let us think of $\partial^{\alpha-\eta}\psi_{\lambda'}$ being replaced by a Taylor polynomial q of order $d-1-|\alpha-\eta|$, with $q=0$ when $d-1-|\alpha-\eta|=0$. Then the quadrature error can be bounded by some absolute multiple of

$$\begin{aligned}&\sum_\square 2^{-|\lambda|n}N^{-1}(2^{-|\lambda|}N^{-1/n})^p 2^{|\lambda'|(p-d+1+|\alpha|-|\eta|+n/2-t)}2^{|\lambda|(|\beta|+|\eta|+d-1+n/2-t)}\\ &\leqslant N^{-p/n}2^{-(|\lambda|-|\lambda'|)(n/2+p-d+1-\min\{t,d-1-t\})},\end{aligned}$$

with the last inequality being sharp when $|\alpha| = |\beta| = t$.

Instead of (3.7), for each $\square \in V_{\lambda,i,\lambda'}$ we get

$$\begin{aligned}&\mathrm{vol}(\square)\|(\partial^{\alpha-\eta}\psi_{\lambda'}-q)\partial^\eta(a_{\alpha,\beta}\partial^\beta\psi_\lambda)\|_{L_\infty(\square)}\\ &\lesssim (N^{-1/n}2^{-|\lambda|})^n N^{-(d-1-|\alpha|+|\eta|)/n}2^{-|\lambda|(d-1-|\alpha|+|\eta|)}2^{|\lambda'|(d-1+n/2-t)}2^{|\lambda|(|\beta|+|\eta|+n/2-t)}\\ &\lesssim N^{-1-(d-1-t+\min\{t,d-1-t\})/n}2^{-(|\lambda|-|\lambda'|)(d-1+n/2-t)},\end{aligned}$$

with the last inequality being sharp for $|\alpha| = |\beta| = t$. So, using $\#V_{\lambda,i,\lambda'} \lesssim N^{(n-1)/n}$, the additional error can be bounded by an absolute multiple of

$$N^{-(d-t+\min\{t,d-1-t\})/n}2^{-(|\lambda|-|\lambda'|)(n/2+d-1-t)},$$

and we have arrived at the following result:

PROPOSITION 3.3 Let $\mathbf{M}^*_{\lambda,\lambda'}$ be the result of the application of a composite rule of rank N and order p applied to each of the integrals from (3.8) for each $|\alpha|,|\beta| \leqslant t$, where for each of these integrals the largest possible η is taken. Then

$$|\mathbf{M}_{\lambda,\lambda'} - \mathbf{M}^*_{\lambda,\lambda'}| \lesssim N^{-p/n} 2^{-(|\lambda|-|\lambda'|)(n/2+p-d+1-\boxed{\min\{t,d-1-t\}})} + N^{-(d-t+\boxed{\min\{t,d-1-t\}})/n} 2^{-(|\lambda|-|\lambda'|)(n/2+d-1-t)},$$

which is valid without any assumption on the location of $\operatorname{sing\,supp}\psi_\lambda$. (The terms in the exponents within frames indicate the differences to the bound from Proposition 3.2.)

COROLLARY 3.3 Let $p \geqslant 2d-2-t+\min\{t,d-1-t\}$, $\frac{d-t}{n} > \frac{1}{2}$, $d-1>t>0$ and $\tilde{d} \geqslant d-2t-1$. Then $\mathbf{M}^{(\mathrm{s})}$ is s^*-computable for some $s^* > \frac{d-t}{n}$.

Proof. The number of nonzero entries in each row of $\mathbf{M}^{(\mathrm{s})}_{\ell,\ell'}$ and column of $\mathbf{M}^{(\mathrm{s})}_{\ell',\ell}$ is $\mathscr{O}(\max\{1, 2^{(\ell'-\ell)(n-1)}\})$. The conditions $p \geqslant 2d-2-t+\min\{t,d-1-t\}$, and $d-1>t>0$ that implies $d \geqslant 2$, show that in the bound from Proposition 3.3 the first term is never larger than the second one, that can be written as $N^{-(d-t+\min\{t,d-1-t\})/n} 2^{-(|\lambda|-|\lambda'|)((n-1)/2+\frac{d-1/2-t}{n-1}(n-1))}$. The condition $\frac{d-t}{n} > \frac{1}{2}$ shows that $r := \frac{d-1/2-t}{n-1} > \frac{d-t}{n}$, and the conditions $t>0$ and $d-1>t$ show that $\min\{t,d-1-t\}>0$ and thus that $q := \frac{d-t+\min\{t,d-1-t\}}{n} > \frac{d-t}{n}$. So an application of Lemma 3.1 for $k=n-1$ yields a matrix $\mathbf{M}^{(\mathrm{s}),*}_j$, for which the computation of each column takes $\mathscr{O}(2^j)$ operations, and for which $\|\mathbf{M}^{(\mathrm{s})}_j - \mathbf{M}^{(\mathrm{s}),*}_j\| \lesssim 2^{-j\min\{q,r\}}$, or $\|\mathbf{M}^{(\mathrm{s})}_j - \mathbf{M}^{(\mathrm{s}),*}_j\| \lesssim 2^{-j(\min\{q,r\})-\varepsilon}$ for any $\varepsilon>0$ in case $q=r$ (see Remark 3.1). Using that by $\tilde{d} \geqslant d-2t-1$, $\|\mathbf{M}^{(\mathrm{s})} - \mathbf{M}^{(\mathrm{s})}_j\| \lesssim 2^{-j\left(\frac{d-1/2-t}{n-1}\right)}$ by Theorem 2.1, and $\frac{d-1/2-t}{n-1} > \frac{d-t}{n}$ by the assumption that $\frac{d-t}{n} > \frac{1}{2}$, the proof is completed. □

Concluding we can say that under the conditions of Corollary 3.1 and 3.2, $\mathbf{M}$ is s^*-computable with $s^* \geqslant \frac{d-t}{n}$, and that under the conditions of Corollary 3.1 and 3.3, it is is s^*-computable with $s^* > \frac{d-t}{n}$. Corollary 3.3 requires that $t>0$ and $d-1>t$, the latter meaning that lowest possible order spline wavelets are not covered.

The underlying quadrature scheme consists of the application of a simple composite quadrature rule of fixed order p, and a suitably chosen rank N on the integral in the parameter space of the wavelet that has the highest level of the two involved in an entry. The rank N depends on j, i.e., on the total number of operations one is prepared to spend on the computation of each column of the approximate stiffness matrix, on the difference in levels of the wavelets involved, and on whether the support of the wavelet on the higher level intersects the singular support of the wavelet on the lower level or not. Fully satisfactory results are only obtained, when in the first, singular case, quadrature is applied after first applying an integration by parts.

For that case, even more favourable bounds, possibly leading to better quantitative behaviour of the adaptive frame scheme, could be obtained by applying adaptive quadrature, in the sense that subcubes that intersect the singular support could be more refined than those that do not (cf. Figure 4). On the latter subcubes, instead of h-refinement also p-refinement could be considered. These modifications would require more programming efforts though.

We apply fixed order composite quadrature rules mainly because of the approximation of the singular entries of $\mathbf{M}$, i.e, the entries of $\mathbf{M}^{(\mathrm{s})}$, and because of the ease with which they can be adjusted to give approximations within any prescribed tolerance. Numerical experiments learned us that when applied to regular entries their error decreases much faster as function of the difference in levels of the wavelets involved than predicted by the bound of Proposition 3.1. Although for our goal of proving optimal computational complexity of adaptive frame algorithms this bound was satisfactory, in view of the quantitative behaviour it is interesting to derive a more accurate bound. This will be the topic of the next section.

4. The Regular Case Revisited

All quadrature error bounds derived so far were obtained by summing over $1 \leqslant i \leqslant m$ bounds for quadrature errors in approximating integrals over the individual $\Xi_{\lambda,i}$, being the regions restricted to which ψ_λ is in Q_{d-1}. The fact that ψ_λ is a wavelet, i.e., that it has vanishing moments did not play any role. In this section, we will see that also for bounding quadrature errors the vanishing moments can be exploited. Although for relatively large d and $\tilde{d}$ compared to p, also for entries of $\mathbf{M}^{(\mathrm{s})}$ better bounds can be obtained, for simplicity here *we consider only entries of* $\mathbf{M}^{(\mathrm{r})}$ (actually, we will even exclude some of these entries from our considerations, see below).

In order to simplify our analysis we make the following observation. By a scaling of the integration domain, $\int a_{\alpha,\beta}(x)\partial_x^\alpha \psi_{\lambda'}(x)\partial_x^\beta \psi_\lambda(x)dx$ can be written as $\int \hat{a}_{\alpha,\beta}(y)\partial_y^\alpha \hat{\psi}_{\lambda'}(y)\partial_y^\beta \hat{\psi}_\lambda(y)dy$, where $\hat{a}_{\alpha,\beta}(y) = a_{\alpha,\beta}(2^{-|\lambda'|}y)$, and where $\hat{\psi}_{\lambda'}$ and $\hat{\psi}_\lambda$ have all the properties of a wavelet on level 0 and $|\lambda|-|\lambda'|$, respectively. In the same way, a composite quadrature rule of rank N and order p to approximate $\int_{\Xi_{\lambda,i}} a_{\alpha,\beta}(x)\partial_x^\alpha \psi_{\lambda'}(x)\partial_x^\beta \psi_\lambda(x)dx$ can be transformed to such a rule to approximate $\int_{2^{|\lambda'|}\Xi_{\lambda,i}} \hat{a}_{\alpha,\beta}(y)\partial_y^\alpha \hat{\psi}_{\lambda'}(y)\partial_y^\beta \hat{\psi}_\lambda(y)dy$. In other words, if for $|\lambda'|=0$ we can prove an upper bound for the quadrature error in approximating $\mathbf{M}^{(\mathrm{r})}_{\lambda,\lambda'}$ of type $\eqsim N^{-q}2^{-|\lambda|r}$, then we have shown an upper bound of type $\eqsim N^{-q}2^{-(|\lambda|-|\lambda'|)r}$ for general $|\lambda| \geqslant |\lambda'|$. (Since qualitatively, $\hat{a}_{\alpha,\beta}$ becomes increasingly smooth with increasing $|\lambda'|$, as in Remark 2.2 we see that the case $|\lambda'|=0$ is actually the most demanding one). Since for any non-zero entry $\mathbf{M}^{(\mathrm{r})}_{\lambda,\lambda'}$, $\psi_{\lambda'}$ is infinitely smooth on supp ψ_λ, it is even sufficient to prove an upper bound of type $\eqsim N^{-q}2^{-|\lambda|r}$ for the quadrature error in approximating an integral $\int g\partial^\beta \psi_\lambda$ where $g \in C^\infty$. This is what we are going to do in the following.

The usual way wavelets are constructed on the n-cube is by taking *tensor products of univariate wavelets and scaling functions*. In order to obtain the best possible estimates we will exploit this fact. We start with considering the *univariate case* $n=1$.

So we consider ψ_λ to be a univariate spline wavelet with $\tilde{d}$ *vanishing moments* and of *order* $d \geqslant t+1$ with respect to a subdivision of $[0,1]$ into subintervals, that in this section are assumed to have *equal length* $h = h_{|\lambda|} \eqsim 2^{-|\lambda|}$. Since the wavelets satisfy homogeneous Dirichlet boundary conditions of order $t-1$, for $\beta \leqslant t$ integration by parts shows that the β*-th derivative* $\psi_\lambda^{(\beta)}$ *has* $\tilde{d}+\beta$ *vanishing moments*.

From now on, we will only consider those ψ_λ that satisfy *homogeneous Dirichlet boundary conditions* of the *maximal possible order* $d-2$, which in case $d > t+1$ is not satisfied by each wavelet whose support has non-empty intersection with $\partial\Omega$. For such wavelets, there exist scalars $a_\ell = a_{\ell,\lambda,\beta}$ such that

$$\psi_\lambda^{(\beta)} = \sum_{\ell\in\mathbb{Z}} a_\ell \zeta(\cdot+\ell h),$$

where ζ is the *cardinal B-spline of order* $d-\beta$ with knot distance h. The number of non-zero a_ℓ is bounded, uniformly in λ.

LEMMA 4.1 For any $q \in P_{d-\beta-1}$, $\sum_{i\in\mathbb{Z}} q(\cdot+ih)\zeta(\cdot+ih) \in P_0$, and $\sum_{i\in\mathbb{Z}} \zeta(\cdot+ih) = 1$.

REMARK 4.1 For comparison, using (Cohen; 2000, Theorem 2.8.1) one infers that a general compactly supported $\zeta \in L_1$ with $\int_{\mathbb{R}} \zeta \neq 0$ satisfies the Strang-Fix conditions of order $d-\beta-1$ (implying that the shift invariant space generated by ζ contains $P_{d-\beta-1}$) if and only if $\sum_{i\in\mathbb{Z}}(\cdot+ih)^k\zeta(\cdot+ih) \in P_{k-1}$ for $k=1,\ldots,d-\beta-1$, and $0 \neq \sum_{i\in\mathbb{Z}} \zeta(\cdot+ih) \in P_0$.

Proof. It is sufficient to give the proof for $h=1$. With ζ_k denoting the B-spline of order k, ζ_1 is the characteristic function of $[0,1]$, and for $k>1$,

$$\zeta_k(x) = \tfrac{x}{k-1}\zeta_{k-1}(x) + (1-\tfrac{x-1}{k-1})\zeta_{k-1}(x-1). \tag{4.1}$$

For $k=d-\beta=1$, both statements of the lemma are obviously true. Assume that both statements are true for some $k-1=d-\beta \geqslant 1$. Using (4.1), we find that

$$\sum_{i\in\mathbb{Z}} q(x+i)\zeta_k(x+i) = \sum_{i\in\mathbb{Z}} \left[q(x+i)\tfrac{x+i}{k-1} + q(x+i+1)(1-\tfrac{x+i}{k-1})\right]\zeta_{k-1}(x+i),$$

and so in particular $\sum_{i\in\mathbb{Z}} \zeta_k(x+i) = \sum_{i\in\mathbb{Z}} \zeta_{k-1}(x+i) = 1$. Substituting $q(x)=x^r$ for $r\in\mathbb{N}$, we have

$$\begin{aligned}(x+i)^r\tfrac{x+i}{k-1} + (x+i+1)^r(1-\tfrac{x+i}{k-1}) &= \tfrac{(x+i)^{r+1}}{k-1} + \sum_{\ell=0}^{r}\binom{r}{\ell}(x+i)^\ell - \tfrac{1}{k-1}\sum_{\ell=0}^{r}\binom{r}{\ell}(x+i)^{\ell+1}\\ &= \sum_{\ell=0}^{r}\binom{r}{\ell}(x+i)^\ell - \tfrac{1}{k-1}\sum_{\ell=0}^{r-1}\binom{r}{\ell}(x+i)^{\ell+1} \in P_r,\end{aligned}$$

where in particular for $r=k-1$,

$$\begin{aligned}&(x+i)^{k-1}\tfrac{x+i}{k-1} + (x+i+1)^{k-1}(1-\tfrac{x+i}{k-1})\\ &\qquad= \sum_{\ell=0}^{k-1}\binom{k-1}{\ell}(x+i)^\ell - \tfrac{1}{k-1}\sum_{\ell=0}^{k-2}\binom{k-1}{\ell}(x+i)^{\ell+1}\\ &\qquad= \sum_{\ell=0}^{k-2}\binom{k-1}{\ell}(x+i)^\ell - \tfrac{1}{k-1}\sum_{\ell=0}^{k-3}\binom{r}{\ell}(x+i)^{\ell+1} \in P_{k-2}.\end{aligned}$$

We conclude that for any $q\in P_{k-1}$, $\sum_{i\in\mathbb{Z}} q(\cdot+i)\zeta_k(\cdot+i) = \sum_{i\in\mathbb{Z}} \tilde{q}(\cdot+i)\zeta_{k-1}(\cdot+i)$ for some $\tilde{q}\in P_{k-2}$, which completes the proof by the induction hypothesis. $\square$

Thanks to the fact that $\psi_\lambda^{(\beta)}$ has $\tilde{d}+\beta$ vanishing moments we have the following result.

LEMMA 4.2 $\sum_{\ell\in\mathbb{Z}} a_\ell q(\ell) = 0$ for any $q\in P_{\tilde{d}+\beta-1}$.

Proof. Without loss of generality we take $h=1$, and drop λ from the notations. We have

$$\sum_{i\in\mathbb{Z}} \psi^{(\beta)}(x+i) = \sum_{i\in\mathbb{Z}}\sum_{\ell\in\mathbb{Z}} a_\ell\zeta(x+i+\ell) = \sum_{\ell\in\mathbb{Z}} a_\ell \sum_{i\in\mathbb{Z}}\zeta(x+i+\ell) = \sum_{\ell\in\mathbb{Z}} a_\ell\sum_{i\in\mathbb{Z}}\zeta(x+i),$$

so that $0=\int_{\mathbb{R}}\psi^{(\beta)}(x)dx = \int_0^1\sum_{i\in\mathbb{Z}}\psi^{(\beta)}(x+i)dx$ together with $\sum_{i\in\mathbb{Z}}\zeta(\cdot+ih)=1\neq 0$ implies $\sum_{\ell\in\mathbb{Z}} a_\ell = 0$.

Now suppose that the statement of the lemma is valid for any $q\in P_{\tilde{d}+\beta-2}$. Then writing $(x+$

$i)^{\tilde{d}+\beta-1} = \sum_{k=0}^{\tilde{d}+\beta-1} \binom{\tilde{d}+\beta-1}{k} (-\ell)^k (x+i+\ell)^{\tilde{d}+\beta-1-k}$, we have

$$\begin{aligned}\sum_{i\in\mathbb{Z}} (x+i)^{\tilde{d}+\beta-1} \psi^{(\beta)}(x+i) &= \sum_{k=0}^{\tilde{d}+\beta-1} \binom{\tilde{d}+\beta-1}{k} \sum_{\ell\in\mathbb{Z}} a_\ell (-\ell)^k \sum_{i\in\mathbb{Z}} (x+i+\ell)^{\tilde{d}+\beta-1-k} \zeta(x+i+\ell) \\ &= \sum_{k=0}^{\tilde{d}+\beta-1} \binom{\tilde{d}+\beta-1}{k} \sum_{\ell\in\mathbb{Z}} a_\ell (-\ell)^k \sum_{i\in\mathbb{Z}} (x+i)^{\tilde{d}+\beta-1-k} \zeta(x+i) \qquad (4.2) \\ &= \sum_{\ell\in\mathbb{Z}} a_\ell (-\ell)^{\tilde{d}+\beta-1} \sum_{i\in\mathbb{Z}} \zeta(x+i),\end{aligned}$$

so that $0 = \int_0^1 \sum_{i\in\mathbb{Z}} (x+i)^{\tilde{d}+\beta-1} \psi^{(\beta)}(x+i)dx$ together with $\sum_{i\in\mathbb{Z}} \zeta(\cdot + ih) = 1 \neq 0$ implies $\sum_{\ell\in\mathbb{Z}} a_\ell \ell^{\tilde{d}+\beta-1} = 0$, with that completing the proof. $\square$

REMARK 4.2 Above proof confirms the known fact that $\sum_{\ell\in\mathbb{Z}} a_\ell q(\ell) = 0$ for any $q \in P_{\tilde{d}+\beta-1}$ is also a sufficient condition for $\psi_\lambda^{(\beta)}$ to have $\tilde{d}+\beta$ vanishing moments.

To approximately compute $\int_{\mathbb{R}} g\psi_\lambda^{(\beta)}$, it is written as $\sum_{i\in\mathbb{Z}} \int_{ih}^{(i+1)h} g\psi_\lambda^{(\beta)}$, and the individual integrals over the intervals $[ih,(i+1)h]$ are approximated by *composite quadrature rules* of *order* $p > 0$ and *rank* N, which in this section are assumed to be *shift invariant*, i.e., of type $\sum_{j=1}^N w_j g(x_j+ih)\psi_\lambda^{(\beta)}(x_j+ih)$ with w_j and x_j independent of i. For the resulting approximation for $\int_{\mathbb{R}} g\psi_\lambda^{(\beta)}$ we have the following result.

PROPOSITION 4.1 For any polynomial q of degree less or equal to $\tilde{d} + \max(d-1, p-d+2\beta)$,

$$\sum_{i\in\mathbb{Z}} \sum_{j=1}^N w_j q(x_j+ih)\psi_\lambda^{(\beta)}(x_j+ih) = \int_{\mathbb{R}} q\psi_\lambda^{(\beta)}.$$

Proof. Again, without loss of generality we take $h = 1$, and drop λ from the notations. We write $\int_{\mathbb{R}} q\psi^{(\beta)}$ and the quadrature approximation as $\int_0^1 \sum_{i\in\mathbb{Z}} q(x+i)\psi^{(\beta)}(x+i)dx$ and $\sum_{j=1}^N w_j \sum_{i\in\mathbb{Z}} q(x_j+i)\psi^{(\beta)}(x_j+i)$, respectively. For $r \in \mathbb{N}$, as in (4.2) we have

$$\begin{aligned}\sum_{i\in\mathbb{Z}} (x+i)^r \psi^{(\beta)}(x+i) &= \sum_{k=0}^{r} \binom{r}{k} \sum_{\ell\in\mathbb{Z}} a_\ell (-\ell)^k \sum_{i\in\mathbb{Z}} (x+i)^{r-k} \zeta(x+i) \\ &= \sum_{k=\tilde{d}+\beta}^{r} \binom{r}{k} \sum_{\ell\in\mathbb{Z}} a_\ell (-\ell)^k \sum_{i\in\mathbb{Z}} (x+i)^{r-k} \zeta(x+i) \\ &= \sum_{m=0}^{r-\tilde{d}-\beta} \binom{r}{m} \sum_{\ell\in\mathbb{Z}} a_\ell (-\ell)^{r-m} \sum_{i\in\mathbb{Z}} (x+i)^m \zeta(x+i),\end{aligned}$$

where the second equality follows from Lemma 4.2. For $r \leqslant \tilde{d} + \max(d-1, p-d+2\beta)$, i.e., $r - \tilde{d} - \beta \leqslant \max(d-\beta-1, p-d+\beta)$, the last expression is either a constant thanks to Lemma 4.1, or, since $\zeta \in P_{d-\beta-1}$, it is a polynomial of degree less or equal to $p-1$. Since in any case $p > 0$, in both cases the quadrature approximation is exact. $\square$

So the order of exactness of the quadrature rule is an increasing function of $\tilde{d}$, and a non-decreasing, and eventually increasing function of both d and p. It is remarkable that even for $p = 1$, the rule is already exact for any polynomial of degree $\tilde{d}+d-1$. On the other hand it is fair to say that the required number of function evaluations grows with p, d and $\tilde{d}$. E.g. thinking of $N=1$, a $\frac{p}{2}$-point Gauss rule being of order p, and ψ_λ being a biorthogonal spline wavelet of order d with $\tilde{d}$ vanishing moments as introduced in Cohen et al. (1992) whose support extends to $2(d+\tilde{d}-1)$ intervals of type $[ih,(i+1)h]$, this number of function evaluations is $p(d+\tilde{d}-1)$.

An alternative way to approximate $\int_{\mathbb{R}} g\psi_\lambda^{(\beta)}$ for smooth g has been analyzed in Barinka et al. (2002). After splitting the integral into $\int_{\mathbb{R}} g(\psi_\lambda^{(\beta)}+C)-C\int_{\mathbb{R}} g\psi_\lambda^{(\beta)}$ for some constant C with $\psi_\lambda^{(\beta)}+C\geqslant 0$, both integrals were approximated by Gauss quadrature, the first using $\psi_\lambda^{(\beta)}+C$ as weight function. In this way exactness of order p is obtained at the expense of only $2*p/2=p$ function evaluations. Of course additional work is required in setting up this Gauss rule.

PROPOSITION 4.2 $|\int_{\mathbb{R}} g\psi_\lambda^{(\beta)}|\lesssim 2^{-|\lambda|(\frac{1}{2}+t+\tilde{d})}|g|_{W_\infty^{\tilde{d}+\beta}}$ and

$$|\int_{\mathbb{R}} g\psi_\lambda^{(\beta)}-\sum_{i\in\mathbb{Z}}\sum_{j=1}^{N} w_j g(x_j+ih)\psi_\lambda^{(\beta)}(x_j+ih)| \lesssim N^{-p}2^{-|\lambda|(\frac{1}{2}+t+\tilde{d}+\max(d-\beta,p+\beta+1-d))}\|g\|_{W_\infty^{\max(p,\tilde{d}+\max(d-1,p-d+2\beta)+1)}}.$$

Proof. Since $\psi_\lambda^{(\beta)}$ has $\tilde{d}+\beta$ vanishing moments, using (2.1) we have

$$\begin{aligned}|\int_{\mathbb{R}} g\psi_\lambda^{(\beta)}| &\lesssim \operatorname{vol}(\operatorname{supp}\psi_\lambda)\|\psi_\lambda^{(\beta)}\|_{L_\infty}\operatorname{diam}(\operatorname{supp}\psi_\lambda)^{\tilde{d}+\beta}|g|_{W_\infty^{\tilde{d}+\beta}}\\ &\lesssim 2^{-|\lambda|}2^{|\lambda|(\frac{1}{2}+\beta-t)}2^{-|\lambda|(\tilde{d}+\beta)}|g|_{W_\infty^{\tilde{d}+\beta}}.\end{aligned}$$

Similarly to the proof of Proposition 3.1, but now using the fact that by Proposition 4.1 we may subtract from g say its Taylor polynomial q of order $m:=\tilde{d}+\max(d-1,p-d+2\beta)+1$ around some point in $\operatorname{supp}\psi_\lambda$, the quadrature error can be bounded on some multiple of

$$\begin{aligned}&N2^{-|\lambda|}N^{-1}(2^{-|\lambda|}N^{-1})^p|(g-q)\psi_\lambda^{(\beta)}|_{W_\infty^p}\\ &\lesssim 2^{-|\lambda|(p+1)}N^{-p}\max_{0\leqslant\ell\leqslant d-1-\beta}\|\psi_\lambda^{(\beta+\ell)}\|_{L_\infty}|g-q|_{W_\infty^{p-\ell}}\\ &\lesssim 2^{-|\lambda|(p+1)}N^{-p}\max_{0\leqslant\ell\leqslant d-1-\beta}2^{|\lambda|(\frac{1}{2}+\beta+\ell-t)}2^{-|\lambda|\max(0,m-p+\ell)}|g|_{W_\infty^{\max(0,m-p+\ell)+p-\ell}}\\ &\lesssim 2^{-|\lambda|(p+1)}N^{-p}2^{|\lambda|(d-\frac{1}{2}-t)}2^{-|\lambda|(m-p+d-1-\beta)}\|g\|_{W_\infty^{\max(p,m)}},\end{aligned}$$

where we used $\psi_\lambda\in P_{d-1}$, (2.1), and that $m-p+d-1-\beta\geqslant 0$ by $\tilde{d}+\beta\geqslant 0$. By substituting the expression for m, the proof is completed. $\square$

Returning to the *multivariate setting*, we consider wavelets ψ_λ of the following type

$$\psi_\lambda=\xi_{\lambda_1}\otimes\cdots\otimes\xi_{\lambda_n},$$

where $|\lambda_1|=\cdots=|\lambda_n|=|\lambda|$, and for $1\leqslant\ell\leqslant n$, ξ_{λ_ℓ} is *either* a univariate wavelet ψ_{λ_ℓ} of the type we studied in this section, except that for $\gamma\leqslant d-1$ it satisfies $\|\psi_{\lambda_\ell}\|_{W_\infty^\gamma}\lesssim 2^{|\lambda_\ell|(\frac{1}{2}+\gamma-\frac{t}{n})}$ instead of $\|\psi_{\lambda_\ell}\|_{W_\infty^\gamma}\lesssim 2^{|\lambda_\ell|(\frac{1}{2}+\gamma-t)}$, *or* such a function without vanishing moments, i.e., a univariate scaling function, where at least one of the factors $\xi_{\lambda_1},\ldots,\xi_{\lambda_n}$ is a wavelet. Note that the scaling of an L_2-orthonormalized wavelet ψ_λ with a factor $2^{-|\lambda|t}$ to give it an "energy-norm" of order 1 is independent of the space dimension n. Here, we distributed this scaling evenly over the factors. Furthermore, note that above assumption on the multivariate wavelet means that it satisfies homogeneous Dirichlet boundary conditions of the maximal order $d-2$, meaning that for $d-1>t$ we exclude some wavelets that are mapped onto the physical boundary $\partial\Omega$.

To approximate $\int_{\mathbb{R}^n} g\partial^\beta\psi_\lambda$ for $|\beta|\leqslant t$, we apply the product of the quadrature rules for the univariate integrals in the coordinate directions, i.e., sums of shift invariant composite rules of order p over the subintervals on which the univariate wavelet or scaling function is polynomial. So denoting $I_{\lambda_\ell,\beta_\ell}(g)=\int_{\mathbb{R}} g\xi_{\lambda_\ell}^{(\beta_\ell)}$ and $Q_{\lambda_\ell,\beta_\ell,M}(g)$ its quadrature approximation using rank M, given $N\in\mathbb{N}^n$ we approximate $\int_{\mathbb{R}^n} g\partial^\beta\psi_\lambda=(I_{\lambda_1,\beta_1}\otimes\cdots\otimes I_{\lambda_n,\beta_n})(g)$ by $(Q_{\lambda_1,\beta_1,N^{1/n}}\otimes\cdots\otimes Q_{\lambda_n,\beta_n,N^{1/n}})(g)$. Note that the total number of abscissae is $\eqsim N$.

PROPOSITION 4.3 We have

$$|(I_{\lambda_1,\beta_1}\otimes\cdots\otimes I_{\lambda_n,\beta_n}-Q_{\lambda_1,\beta_1,N^{1/n}}\otimes\cdots\otimes Q_{\lambda_n,\beta_n,N^{1/n}})(g)|\lesssim$$
$$N^{-p/n}2^{-|\lambda|(\frac{n}{2}+t+\tilde{d}+\min_{1\leqslant\ell\leqslant n}\max(d-\beta_\ell,p+\beta_\ell+1-d))}\max_{\{\gamma:\gamma_\ell\leqslant\max(p,\tilde{d}+\max(d-1,p-d+2\beta_\ell)+1)\}}\|\partial^\gamma g\|_{L_\infty(\mathbb{R}^n)}.$$

Proof. It is sufficient to consider the case that ξ_{λ_1} is a wavelet and, since the other cases give never worse bounds, the other factors are scaling functions. With obvious simplifications of the notations, Proposition 4.2 (with N, t reading as $N^{1/n}$, t/n) shows that for $k=1$,

$$|(I_1\otimes\cdots\otimes I_k-Q_1\otimes\cdots\otimes Q_k)(g)|\lesssim$$
$$N^{-p/n}2^{-|\lambda|(k(\frac{1}{2}+\frac{t}{n})+\tilde{d}+\max(d-\beta_1,p+\beta_1+1-d))}\max_{\{\gamma:\gamma_\ell\leqslant\max(p,\tilde{d}+\max(d-1,p-d+2\beta_\ell)+1),\,1\leqslant\ell\leqslant k\}}\|\partial^\gamma g\|_{L_\infty(\mathbb{R}^k)}.$$

For $k>1$, we write

$$\begin{aligned}&I_1\otimes\cdots\otimes I_k-Q_1\otimes\cdots\otimes Q_k=\\&(I_1\otimes\cdots\otimes I_{k-1}-Q_1\otimes\cdots\otimes Q_{k-1})\otimes Q_k+I_1\otimes\cdots\otimes I_{k-1}\otimes(I_k-Q_k).\end{aligned}$$

From

$$\begin{aligned}|(I_k-(I_k-Q_k))(g)|&\lesssim 2^{-|\lambda|(\frac{1}{2}+\frac{t}{n})}\|g\|_{W_\infty^{\max(p,\tilde{d}+\max(d-1,p-d+2\beta_k)+1)}(\mathbb{R})},\\|(I_k-Q_k)(g)|&\lesssim N^{-p/n}2^{-|\lambda|(\frac{1}{2}+t+\max(d-\beta_k,p+\beta_k+1-d))}\|g\|_{W_\infty^{\max(p,\max(d-1,p-d+2\beta_k)+1)}(\mathbb{R})},\\|(I_1\otimes\cdots\otimes I_{k-1})(g)|&\lesssim 2^{-|\lambda|(\frac{1}{2}+\frac{t}{n}+\tilde{d}+(k-2)(\frac{1}{2}+\frac{t}{n}))}\|\partial_1^{\tilde{d}+\beta_1}\partial_2^{\beta_2}\cdots\partial_{k-1}^{\beta_{k-1}}g\|_{L_\infty(\mathbb{R}^{k-1})},\end{aligned}$$

by Proposition 4.2 for both wavelets and scaling functions ($\tilde{d}=0$), and, for the last estimate, additionally a tensor product argument (cf. Light and Cheney (1985)), by again applying this tensor product argument and the induction hypothesis we arrive at the statement of the proposition. □

Recalling that $\mathbf{M}_{\lambda,\lambda'}=\sum_{|\alpha|,|\beta|\leqslant t}\int_\Omega a_{\alpha,\beta}\partial^\alpha\psi_{\lambda'}\partial^\beta\psi_\lambda$, and using that

$$\min_{|\beta|\leqslant t}\min_{1\leqslant\ell\leqslant n}\max(d-\beta_\ell,p+\beta_\ell+1-d)=\max(d-t,p+1-d,\lceil\tfrac{p+1}{2}\rceil),$$

we arrive at the conclusion that for the basically shift invariant tensor product setting discussed in this section, the following estimate is valid.

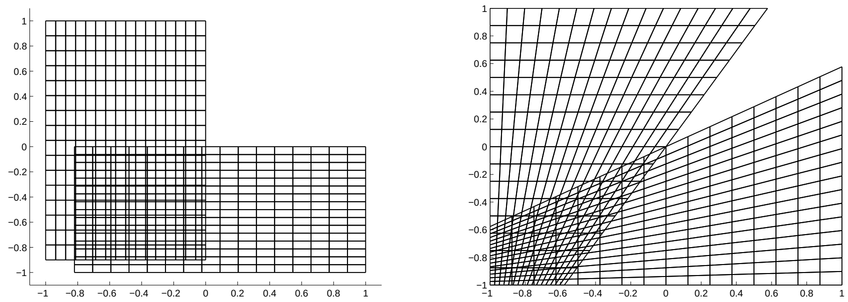

FIG. 5. Two simple domains made up of two overlapping patches with non-matching dyadic grids.

COROLLARY 4.1 Let $|\lambda| \geqslant |\lambda'|$ with supp $\psi_\lambda \subset \overline{\Xi_{\lambda',i'}}$ for some $1 \leqslant i' \leqslant m$, and such that ψ_λ satisfies homogeneous Dirichlet boundary conditions of order $d-2$. Then

$$|\mathbf{M}^{(\mathrm{r})}_{\lambda,\lambda'} - \mathbf{M}^{(\mathrm{r}),*}_{\lambda,\lambda'}| \lesssim N^{-p/n} 2^{-(|\lambda|-|\lambda'|)(\frac{n}{2}+t+\tilde{d}+\max(d-t,p+1-d,\lceil\frac{p+1}{2}\rceil))}.$$

To compare with the upper bound from Proposition 3.1, note that $\frac{n}{2}+t+\tilde{d}+\max(d-t,p+1-d,\lceil\frac{p+1}{2}\rceil)-(\frac{n}{2}+p-d+1) = \max(2d+\tilde{d}-p-1,t+d,t+d+\tilde{d}-\lfloor\frac{(p+1)}{2}\rfloor)$.

5. Numerical Tests

The numerical experiments in this section intend to confirm the sharpness of the different estimates given in Lemma 2.1, Propositions 3.1, 3.2 and 3.3, and Corollary 4.1. On domains $\Omega \subset \mathbb{R}^2$, we consider an operator $L: H^1_0(\Omega) \to H^{-1}(\Omega)$ of order $2t = 2$ defined by

$$(Lw)(v) := \sum_{k=1}^{2} \int_\Omega \partial_k w \partial_k v, \tag{5.1}$$

which results from the variational formulation of Poisson's problem with homogeneous Dirichlet boundary conditions. We are concerned with the size and the approximation of the entries in the stiffness matrix representing L with respect to aggregated wavelet frames.

As reference systems $\Psi^\square_i = \{\psi^\square_{i,\mu} \ : \ \mu \in \Lambda^\square_i\} \subset H^1_0(0,1)^2$, we use biorthogonal spline wavelet bases of order $d=2$ or $d=3$, having $\tilde{d}=2$ or $\tilde{d}=3$ vanishing moments, respectively. The dual wavelets are chosen not to satisfy any boundary conditions, so that indeed all primal wavelets have the aforementioned number of vanishing moments (Dahmen and Schneider (1998)). The wavelets are constructed as tensor products of univariate wavelets and scaling functions with respect to uniform partitions of the unit interval as discussed in Sect. 4.

The elements of the aggregated frame are obtained by lifting according to

$$\psi_{i,\mu}(x) = \frac{\psi^\square_{i,\mu}(\kappa_i^{-1}(x))}{|\det D\kappa_i(\kappa_i^{-1}(x))|^{1/2}}, \quad \text{for } x \in \mathrm{Im}\,\kappa_i, \tag{5.2}$$

and zero elsewhere on Ω, where $\kappa_i : (0,1)^2 \to \Omega_i \subset \Omega$ represents a smooth parametrization of the $i-$th patch of the open, overlapping covering $\Omega = \bigcup_{i=1}^m \Omega_i$. Although not required for the current application, in our software we included the additional scaling given by the denominator in (5.2) in order that $\|\psi_{i,\mu}\|_{L_2(\Omega)} = \|\psi^{\square}_{i,\mu}\|_{L_2(0,1)^2}$, and that the lifted primal wavelets are biorthogonal to similarly lifted dual wavelets. Except for affine κ_i, effectively it yields smooth, non-polynomial coefficients in the differential operator that we therefore have omitted in (5.1). For $d > 2$, the reference system $\Psi_i^{\square}$ depends (weakly) on i, in the sense that on those edges that are mapped into the interior of Ω, homogeneous Dirichlet boundary conditions of order $d-2 > t-1 = 0$ are prescribed, which guarantees that all $\psi_{i,\mu} \in C^{d-2}(\Omega)$. The aggregated wavelet frame on Ω is now defined as

$$\Psi := \{\psi_\lambda \;:\; \lambda = (i,\mu) \in \bigcup_{i=1}^m \{i\} \times \Lambda_i^{\square}\}. \tag{5.3}$$

We consider parametrizations of type

$$\kappa_i(r,s) = (1-r)(1-s)\mathbf{b}^{(0,0)} + (1-r)s\mathbf{b}^{(0,1)} + r(1-s)\mathbf{b}^{(1,0)} + rs\mathbf{b}^{(1,1)},$$

where $\mathbf{b}^{(k,\ell)} \in \mathbb{R}^2$, $(k,\ell) \in \{0,1\}^2$. Thus, provided that the vertices $\mathbf{b}^{(k,\ell)}$ are ordered appropriately, κ_i maps the unit square to an arbitrary quadrangle in $\mathbb{R}^2$. In case the vertices describe a parallelogram, κ_i is affine and so the denominator in (5.2) is a constant.

We consider two different types of overlapping decompositions of Ω. The first type refers to the situation of overlapping rectangular patches with non-matching dyadic grids both being aligned with the Cartesian coordinates, as shown in Figure 5 (*left*).

First of all, we address the decay estimates in Lemma 2.1, which are the essential ingredients for the proof of Theorem 2.1, stating a sufficient compressibility of $\mathbf{M}$. For the grids in Figure 5 (*left*), we could compute the entries of $\mathbf{M}$ exactly, whereas for the grids in Figure 5 (*right*) we applied our composite quadrature scheme with, for this goal, $N \gg 1$ and a high order p such that the quadrature error is neglectable.

For fixed columns of $\mathbf{M}^{(\mathrm{r})}$ and $\mathbf{M}^{(\mathrm{s})}$, we have computed the largest entry in modulus as function of level difference of row and column indices. The decay of the modulus of this entry is illustrated by the results given in Figure 6. Lemma 2.1 predicts the exponential decay rate $n/2+\tilde{d}+t$ or $n/2+d-1-t$ in base 2 for $\mathbf{M}^{(\mathrm{r})}$ or $\mathbf{M}^{(\mathrm{s})}$, respectively. For $\mathbf{M}^{(\mathrm{r})}$, we observed the rate 4 or 5 for $d=\tilde{d}=2$ or $d=\tilde{d}=3$, and for $\mathbf{M}^{(\mathrm{s})}$, we observed the rate 1 or 2 for $d=\tilde{d}=2$ or $d=\tilde{d}=3$, all in accordance with the predicted rates.

For investigating the *quadrature errors* of our composite schemes we used product Gaussian quadrature formulas of fixed order p as building block. Figure 7 addresses the rate of convergence of the composite quadrature scheme for a fixed entry from $\mathbf{M}^{(\mathrm{r})}$ as function of the granularity or rank N. We have used a quadrature rule of order $p=4$ for the case $d=\tilde{d}=2$ and $p=2$ for $d=\tilde{d}=3$. We observe the polynomial rates $2=4/2=p/n$ and $1=2/2$, respectively, as predicted by both Proposition 3.1 and Corollary 4.1.

For fixed N and $p=d=\tilde{d}=2$, the decay of the quadrature error in a fixed column of $\mathbf{M}^{(\mathrm{r})}$ as function of the level difference $||\lambda|-|\lambda'||$ of the involved wavelets is examined in Figure 8. We observe the exponential rate $n/2+t+\tilde{d}+\max(d-t,p+1-d,\lceil(p+1)/2\rceil)=6$ in base 2 as predicted by Corollary 4.1, which is much better than the rate $n/2+p-d+1=2$ predicted by Proposition 3.1.

Unlike that from Corollary 4.1, as stated in Remark 3.3 the bound from Proposition 3.1 also applies to entries from $\mathbf{M}^{(\mathrm{s})}$ when the singular supports of the corresponding wavelets are nested as function of

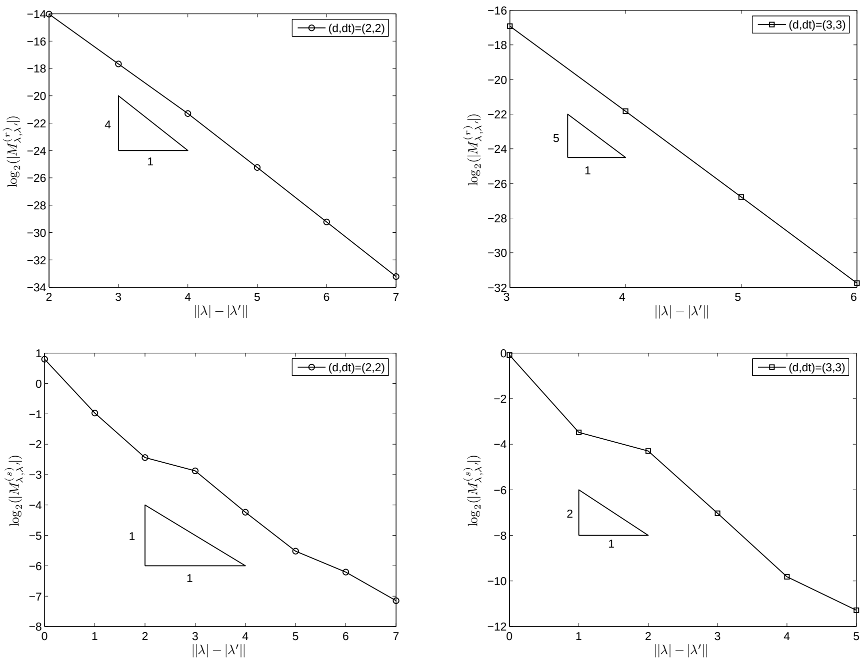

FIG. 6. Decay of the entries in a column of $\mathbf{M}^{(\mathrm{r})}$ (*upper part*) and $\mathbf{M}^{(\mathrm{s})}$ (*lower part*) for $d = \tilde{d} = 2$ and $d = \tilde{d} = 3$.

the level, cf. Figure 3. The results given in Figure 9 for $p = d = \tilde{d} = 2$ indicate that for those entries the bound from this proposition as function of the level difference is sharp.

Figure 10 addresses the rate of convergence of the composite quadrature scheme for entries in $\mathbf{M}^{(\mathrm{s})}$ as function of the rank N. We have used a quadrature rule of order $p = 2$ for the case $d = \tilde{d} = 2$ and $p = 4$ for $d = \tilde{d} = 3$. We observe the polynomial rates $(d-t)/n = 1/2$ and 1, respectively, in accordance with the second term from the bound of Proposition 3.2. Since in these cases $p \geqslant \max\{d-t, 2d-2-t\}$, as stated in the proof of Corollary 3.2, the second term in this bound is always dominating. Recall that for the entries of $\mathbf{M}^{(\mathrm{s})}$ the integrand of any integral in (3.3) may be discontinuous. Consequently, if N is successively increased, the ratio of the number of quadrature knots on either side of the singularity may differ, even for uniform dyadic N-refinement, as we have used it in our experiments. This causes the oscillatory behaviour of the error that can be observed in Figure 10.

For fixed N, the decay of the quadrature error in a fixed column of $\mathbf{M}^{(\mathrm{s})}$ as function of the level difference $||\lambda| - |\lambda'||$ of the involved wavelets is examined in Figure 11. The results confirm the exponential rate $n/2 + d - 1 - t$ in base 2 given by the second term from the bound of Proposition 3.2.

Since for $d = 3$, $d - 1 > t = 1$, for this case alternatively we can apply the composite quadrature to the right hand side in (3.8), i.e., *after integration by parts.* The results shown in the lower error diagram of Figure 10, obtained with $p = 4$, confirm the improved polynomial rate $3/2 = (d - t + \min\{t, d-1-t)\})/n$ predicted by Proposition 3.3, and illustrate an improved quantitative performance. Indeed, in the lower error diagram of Figure 10 the initial error for $N = 1$ is more than ten times smaller than without the integration by parts trick.

To conclude we can say that in our tests all estimates have shown to be sharp. Moreover, also the

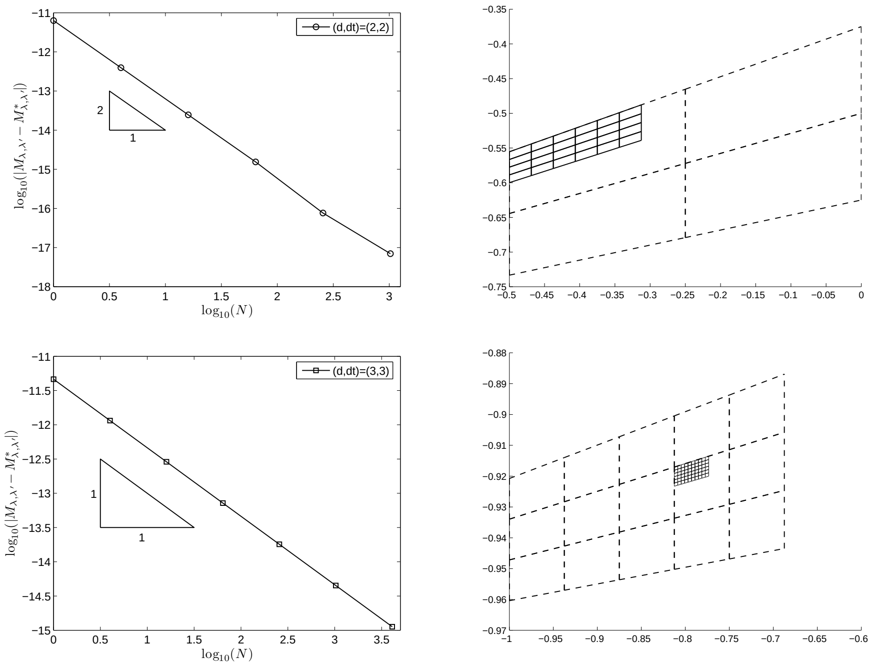

FIG. 7. Quadrature errors for single entries of $\mathbf{M}^{(\mathrm{r})}$ for different granularities N and $p = 4$ (and $d = \tilde{d} = 2$) and $p = 2$ ($d = \tilde{d} = 3$), respectively. The right pictures show the the singular supports of the wavelets involved.

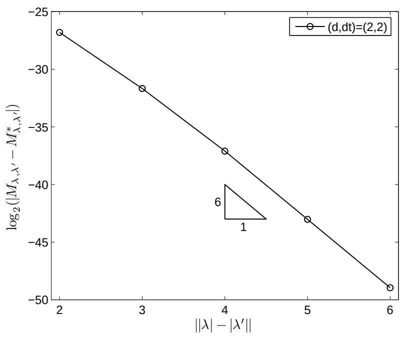

FIG. 8. Decay of the quadrature error in a column of $\mathbf{M}^{(\mathrm{r})}$ as function of $||\lambda| - |\lambda'||$ for fixed N, and $p = d = \tilde{d} = 2$.

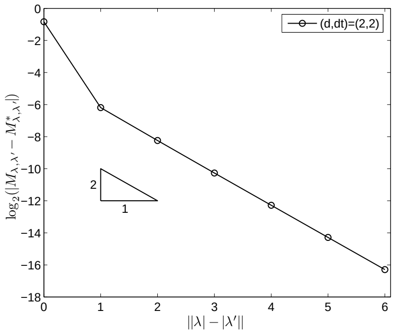

FIG. 9. Decay of the quadrature error for entries $\mathbf{M}^{(s)}_{\lambda,\lambda'}$ as function of $||\lambda| - |\lambda'||$ for fixed λ' and N, $p = d = \tilde{d} = 2$, where the singular supports of ψ_λ and $\psi_{\lambda'}$ are nested like in Figure 3, because $\lambda = (i,\mu)$ and $\lambda' = (i,\mu')$, i.e., they are lifted by the same κ_i.

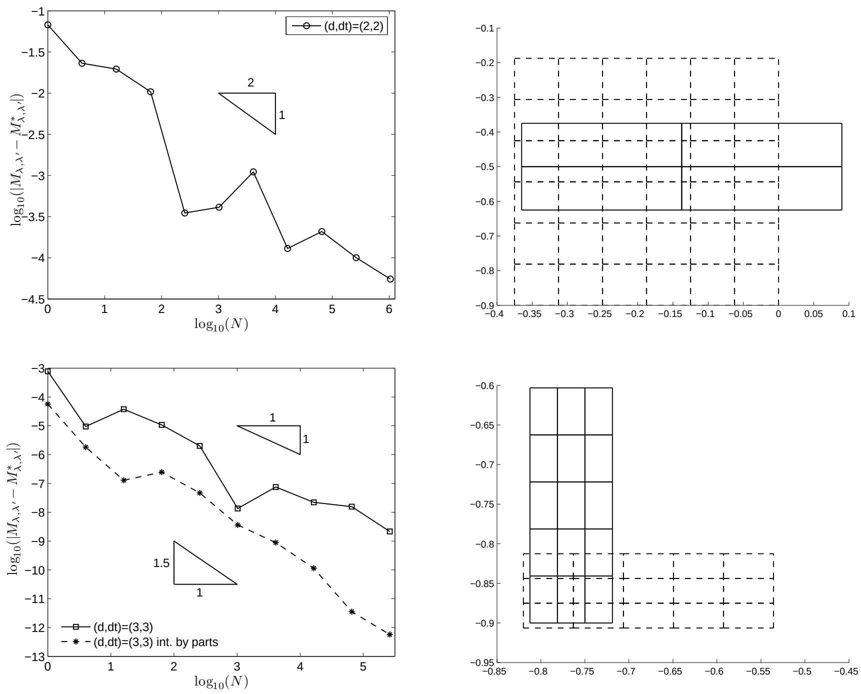

FIG. 10. Quadrature errors for single entries of $\mathbf{M}^{(s)}$ for different granularities N with $p = d = \tilde{d} = 2$*(top left)* and $p = 4$, $d = \tilde{d} = 3$ *(bottom left)*, respectively. The right pictures show the singular supports of the wavelets involved. The dotted line in the lower left picture refers to the case where the composite rule was applied to the right hand side of (3.8).

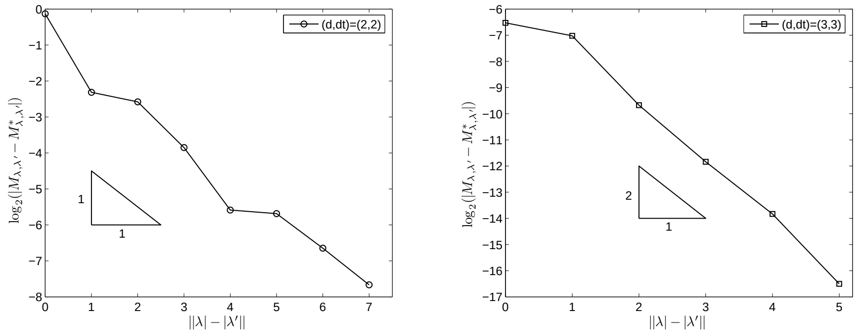

FIG. 11. Decay of the quadrature error in a column of $\mathbf{M}^{(\mathrm{s})}$ as function of $||\lambda| - |\lambda'||$ for fixed N and $p = 2$, $d = \tilde{d} = 2$, and $p = 4$, $d = \tilde{d} = 3$.

quantitative performances of our quadrature scheme for the approximation of the matrices $\mathbf{M}^{(\mathrm{r})}$ have turned out to be quite promising, cf. Figure 7 and 8. Naturally, the computation of entries in $\mathbf{M}^{(\mathrm{s})}$ is much harder. Nevertheless, applying an additional integration by parts as suggested in (3.8), both higher convergence rates as function of the rank N and an improvement of constants can be achieved. Therefore, it can be expected that the application of this trick, possibly in combination with an adaptive quadrature scheme as mentioned at the end of Section 5, will allow for an efficient computation of stiffness matrices also for domains with more complex geometries.

REFERENCES

Barinka, A., Barsch, T., Dahlke, S., Mommer, M. and Konik, M. (2002). Quadrature formulas for refinable functions and wavelets. II. Error analysis, *J. Comput. Anal. Appl.* **4(4)**: 339–361.

Canuto, C., Tabacco, A. and Urban, K. (1999). The wavelet element method part I: Construction and analysis, *Appl. Comput. Harmon. Anal.* **6**: 1–52.

Cohen, A. (2000). *Wavelet methods in numerical analysis*, In P.G. Ciarlet and J. L. Lions, editors, *Handbook of Numerical Analysis. Vol. VII.*, pages 417–711. North-Holland, Amsterdam.

Cohen, A., Dahmen, W. and DeVore, R. (2001). Adaptive wavelet methods for elliptic operator equations – Convergence rates, *Math. Comp.* **70**: 27–75.

Cohen, A., Dahmen, W. and DeVore, R. (2002). Adaptive wavelet methods II - Beyond the elliptic case, *Found. Comput. Math.* **2(3)**: 203–245.

Cohen, A., Daubechies, I. and Feauveau, J. (1992). Biorthogonal bases of compactly supported wavelets, *Comm. Pur. Appl. Math.* **45**: 485–560.

Cohen, A. and Masson, R. (2000). Wavelet adaptive method for second order elliptic problems: Boundary conditions and domain decomposition, *Numer. Math.* **86**: 193–238.

Dahlke, S., Fornasier, M. and Raasch, T. (2004). Adaptive frame methods for elliptic operator equations, *Adv. Comput. Math., doi:10.1007/s10444-005-7501-6* .

Dahlke, S., Fornasier, M., Raasch, T., Stevenson, R. P. and Werner, M. (2006). Adaptive frame methods for elliptic operator equations: The steepest descent approach, *IMA J. Numer. Math., doi:10.1093/imanum/drl035* .

Dahmen, W. and Schneider, R. (1998). Wavelets with complementary boundary conditions—function spaces on the cube, *Results Math.* **34(3-4)**: 255–293.

Dahmen, W. and Schneider, R. (1999a). Composite wavelet bases for operator equations, *Math. Comp.* **68**: 1533–1567.

Dahmen, W. and Schneider, R. (1999b). Wavelets on manifolds I: Construction and domain decomposition, *SIAM*

J. Math. Anal. **31**: 184–230.

Dahmen, W., Vorloeper, J. and Urban, K. (2002). Adaptive wavelet methods — basic concepts and applications to the Stokes problem, *in* D.-X. Zhou (ed.), *Proceedings of the International Conference of Computational Harmonic Analysis*, World Scientific, pp. 39–80.

Gantumur, T. (2006). An optimal adaptive wavelet method for nonsymmetric and indefinite elliptic problems, *J. Comp. Appl. Math.* . Available online 29 December 2006. doi:10.1016/j.cam.2006.11.013.

Gantumur, T., Harbrecht, H. and Stevenson, R. (2007). An optimal adaptive wavelet method without coarsening of the iterands, *Math. Comp.* **77**: 615–629.

Gantumur, T. and Stevenson, R. P. (2006a). Computation of differential operators in wavelet coordinates, *Math. Comp.* **75**: 697–709.

Gantumur, T. and Stevenson, R. P. (2006b). Computation of singular integral operators in wavelet coordinates, *Computing* **76**: 77–107.

Harbrecht, H. and Stevenson, R. (2006). Wavelets with patchwise cancellation properties, *Math. Comp.* **75(256)**: 1871–1889.

Kunoth, A. and Sahner, J. (2006). Wavelets on manifolds: An optimized construction, *Math. Comp.* **75**: 1319–1349.

Light, W. A. and Cheney, E. W. (1985). *Approximation theory in tensor product spaces, volume 1169 of* Lecture Notes in Mathematics, Springer-Verlag, Berlin.

Stevenson, R. (2007). Composite wavelet bases with extended stability and cancellation properties, *SIAM J. Numer. Anal.* **45**(1): 133–162.

Stevenson, R. P. (2003). Adaptive solution of operator equations using wavelet frames, *SIAM J. Numer. Anal.* **41(3)**: 1074–1100.

Stevenson, R. P. (2004). On the compressibility of operators in wavelet coordinates, *SIAM J. Math. Anal.* **35(5)**: 1110–1132.