Eidgenössische Technische Hochschule Zürich

Ecole polytechnique fédérale de Zurich
Politecnico federale di Zurigo
Swiss Federal Institute of Technology Zurich

Two-scale Finite Element Discretizations for Infinitesimal Generators of Jump Processes in Finance

Fang Liu[†], Nils Reich, Aihui Zhou[‡]

Research Report No. 2008-23
June 2008

Seminar für Angewandte Mathematik
Eidgenössische Technische Hochschule
CH-8092 Zürich
Switzerland

[†]School of Applied Mathematics, Central University of Finance and Economics, Beijing 100081, China (fliu@lsec.cc.ac.cn).

[‡]LSEC, ICMSEC, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190, China (azhou@lsec.cc.ac.cn).

Two-scale Finite Element Discretizations for Infinitesimal Generators of Jump Processes in Finance

Fang Liu[†], Nils Reich, Aihui Zhou[‡]

Seminar für Angewandte Mathematik
Eidgenössische Technische Hochschule
CH-8092 Zürich
Switzerland

Research Report No. 2008-23 June 2008

Abstract

We propose and analyze a number of two-scale discretization schemes for infinitesimal generators of jump processes arising in Finance. It is shown that the number of degrees of freedom of the the two-scale discretization is significantly smaller than that of the standard one-scale finite element approach while at the same time preserving the accuracy of the one-scale discretization. Based on the two-scale discretizations, we present some local and parallel algorithms. The main idea of these algorithms is to use a coarse grid to approximate the low frequencies and then to use a fine grid to correct the relatively high frequencies by some local/parallel procedures. In particular, we present two so-called combination based two-scale finite element discretization schemes on tensor product domains. They can be carried out in parallel. As a result, both the computational time and the storage can be reduced considerably. Finally, we illustrate how a combination of wavelet and Lagrangian finite element basis functions can be applied to further reduce the complexity arising from the non-locality of the integro-differential operators under consideration.

Keywords: Finite element, combination, infinitesimal generator, sparse grids, two-scale discretization, wavelets.

2000 Mathematics Subject Classification. 45K05, 65N15, 65N30, 65N55.

[†]School of Applied Mathematics, Central University of Finance and Economics, Beijing 100081, China (fliu@lsec.cc.ac.cn).

[‡]LSEC, ICMSEC, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190, China (azhou@lsec.cc.ac.cn).

1 Introduction

Consider arbitrage-free values $u(x,T)$ of contingent claims on baskets of $d \in \mathbb{N}$ assets. The log-returns of the underlying assets are modeled by a Lévy or, more generally, a Feller process X with state space $\mathbb{R}^d$ and $X_0 = x$.

By the fundamental theorem of asset pricing (see [19]), the arbitrage free price u of a European contingent claim with payoff function $g(\cdot)$ and maturity $T > 0$ is given by the conditional expectation

$$u(x,t) = \mathbb{E}\left(g(X_T) \mid X_t = x\right),$$

under an a-priori chosen (risk-neutral) martingale measure equivalent to the historical measure (see e.g. [18, 20] for measure selection criteria).

Deterministic methods to compute $u(x,T)$ are based on the solution of the corresponding backward Kolmogorov equation

$$u_t + \mathcal{A}u = 0, \qquad u|_{t=T} = g. \tag{1.1}$$

Here $\mathcal{A}$ denotes the (integro-differential) infinitesimal generator of X with domain $\mathcal{D}(\mathcal{A})$. For the Galerkin based finite element implementation, equation (1.1) is converted into variational form. Formally, the resulting problem reads: Find u such that

$$\langle \frac{d}{dt}u,\, v\rangle + \mathcal{E}(u,\, v) = 0 \quad \forall v \in \mathcal{D}(\mathcal{E}), \tag{1.2}$$

where $\mathcal{E}(u,v) = \langle \mathcal{A}u, v\rangle$ denotes the bilinear form associated to $\mathcal{A}$ and X.

In the classical setting of Black-Scholes, X is a geometric Brownian Motion and $\mathcal{A}$ is a diffusion operator so that a closed form solution of (1.1) and (1.2) for plain vanilla contracts is possible in certain cases. For more general Lévy or Feller price processes X, $\mathcal{A}$ is in general a pseudodifferential operator with symbol ψ, i.e.

$$(\mathcal{A}u)(x) = (\psi(x,D)u)(x) = -\int_{\mathbb{R}^d} e^{i\langle \xi, x\rangle}\psi(x,\xi)\widehat{u}(\xi)d\xi. \tag{1.3}$$

In one dimension, the pricing equations (1.1) and (1.2) have been studied by several authors, e.g. [3, 11, 29, 38, 39]. In [22, 45] this setting was extended to $d \geq 2$ dimensions using Lévy processes with 1-homogeneous Lévy copulas. There, the domain of the infinitesimal generator $\mathcal{A}$ was characterized explicitly as an anisotropic Sobolev space. Furthermore, it was shown that the corresponding variational problem is well-posed.

In order to solve (1.2) using finite elements, there are two main challenges that need to be addressed:

- The "curse of dimension": The number of degrees of freedom on a tensor product finite element mesh of width h in dimension d grows like $\mathcal{O}(h^{-d})$ as $h \to 0$.
- The non-locality of the underlying operator $\mathcal{A}$: The finite element stiffness matrix is dense and thus consists of $\mathcal{O}(h^{-2d})$ non-zero entries.

To overcome these two challenges, a wavelet based finite element discretizations of (1.2) was introduced in [22, 44, 52] based on a sparse tensor product approach in combination with wavelet

compression techniques. Under certain conditions on the wavelets' number of vanishing moments, this approach yields asymptotically optimal, essentially dimension-independent complexity $\mathcal{O}(h^{-1}|\log h|^{2(d-1)})$, see [44]. For the analysis of general wavelet based finite element methods we refer to [16, 50, 51, 52] and the references therein. Construction techniques for arbitrary order spline wavelets on the interval with prescribed number of vanishing moments can be found in [15, 41].

Even though a wavelet discretization of (1.2) is proved to be *asymptotically* optimal of complexity $\mathcal{O}(h^{-1}|\log h|^{2(d-1)})$ for a very general class of *non-local* operators, the constants involved in the complexity estimates can be significant. More precisely, on rather low but practically important levels of refinement ($h = 2^{-j}$ with $j \leq 8$, say) there often occurs a computational overhead that, especially for *local* operators, can be reduced considerably by using classical (Lagrangian) finite element basis functions. Following this observation, in this paper we shall split the non-local operator $\mathcal{A}$ and its corresponding bilinear form into a local and a non-local part. Then, to reduce the computational complexity for solving the stationary part of the Kolmogorov equation (1.1), we shall introduce a two-scale approach, including a *basic* two-scale discretization scheme and a *combination based* two-scale discretization scheme. Moreover, based on the two-scale discretizations, we shall construct certain local and parallel algorithms to reduce the computational cost even further. These algorithms are motivated by the observation that, for a solution to some elliptic problems, low frequency components can be approximated well by a relatively coarse grid and high frequency components can be computed on a fine grid using local and parallel procedures. Note that the same observation is also crucial in the theory of multigrid and wavelet methods (see, e.g., [1, 10, 28, 58]). As we will prove below, from a Financial Mathematics point of view low frequency components correspond to the (non-local) jump part of the underlying stochastic process, whereas high frequencies are arising from its (local) diffusion part.

The two-scale approach is an iterative method, which is, in a way, related to that in Lin [33]. The two-scale approach has been used for solving a variety of partial differential equations and integral equations with different types of discretization methods, see e.g. [17, 25, 32, 35, 36, 37, 40, 53, 54, 55, 56]. We should mention that the combination based two-scale method is closely related to the sparse grid method that was developed in [4, 26, 59]. The sparse grid method is a powerful tool in the numerical solution of classical partial differential equations (c.f. [5, 6] and references cited therein) as well as high-dimensional equations arising in Finance (c.f. [22, 52]). The so-called (multiscale) combination technique [21, 27], an extrapolation-type sparse grid variant, has been investigated in a number of papers (see e.g. [5, 24, 27, 43, 57]). In Finance it has been applied in the Black-Scholes setting by e.g. [46]. Instead of the multi-level basis approach [1, 58], a two-level basis approach in the two-scale finite element discretization was used in [35, 36, 37], which is known to be more flexible than the multi-level basis approach [34, 35].

Let us give a little more detailed illustration of the two-scale combination method on tensor product domains, for instance. The main idea of the two-scale finite element combination method is to use a coarse grid to approximate low frequencies and to combine univariate fine and coarse grids to handle high frequencies by parallel procedures. For instance, in three dimensional cases, with the same approximation accuracy, the degrees of freedom for getting the two-scale finite element combination approximation $u^h_{H,H,H}$ is only of complexity $O(h^{-2})$ when $H = O(h^{1/2})$ is chosen for the corresponding univariate fine and coarse grids. The complexity of the standard finite element solution $u_{h,h,h}$ is $O(h^{-3})$. This approach turns out to be advantageous in two respects. First, the possibility of using existing codes allows the straightforward application of two-scale combination discretization to large scale problems. Second, since the different subproblems can be solved fully in parallel, there is a very elegant and efficient inherent coarse-grain parallelism that makes the two-scale combination discretization perfectly suitable for modern high-performance computers.

The remainder of this paper is organized as follows. In Section 2, some basic assumptions and descriptions for the Lévy copula processes are introduced. Several fundamental auxiliary results are provided. A basic time stepping scheme is illustrated so that the time-dependence of (1.2) can be handled. In Sections 3 and 4 the main results of this work are presented. In Section 3, at first we present and analyze a basic two-scale discretization approach, including the associated local and parallel algorithms. In Section 4, to reduce the computational complexity further, we develop a two-scale finite element combination discretization for elliptic (log price) problems. In Section 5, we briefly illustrate how the advantages of wavelet-based sparse grid methods can be combined with the algorithms constructed in Sections 3 & 4.

2 The variational formulation for Lévy copula processes

To obtain an explicit realization of the abstract variational problem (1.2), in this section we briefly recall the basic properties and definitions of Lévy copula processes. Most results presented in this section are based on [22, 30, 45].

Recall, that a càdlàg stochastic process $X = \{X_t : t > 0\}$ with state space $\mathbb{R}^d$ such that $X_0 = 0$ a.s. is called a Lévy process if it has independent and stationary increments and is stochastically continuous.

For its characteristic function we have the Lévy-Khinchin representation (c.f. [47]),

$$\mathbb{E}\left(e^{i\langle \xi, X_t\rangle}\right) = e^{-t\psi(\xi)}, \quad \xi \in \mathbb{R}^d\,, \tag{2.1}$$

with *characteristic exponent* $\psi : \mathbb{R}^d \to \mathbb{C}$ given by

$$\psi(\xi) = \psi_D(\xi) + \psi_J(\xi), \tag{2.2}$$

where

$$\psi_D(\xi) := i\langle \gamma, \xi\rangle - \frac{1}{2}\langle \xi, Q\xi\rangle\,, \tag{2.3}$$

$$\psi_J(\xi) := \int_{\mathbb{R}^d} \left(1 - e^{i\langle \xi, z\rangle} + i\langle \xi, z\rangle 1_{\{|z|\le 1\}}\right) \nu(\mathrm{d}z)\,. \tag{2.4}$$

Here $Q \in \mathbb{R}^{d\times d}$ denotes a covariance matrix, $\gamma \in \mathbb{R}^d$ is a drift vector and ν a Lévy measure which satisfies

$$\int_{\mathbb{R}^d} 1 \wedge |z|^2 \, \nu(\mathrm{d}z) < \infty. \tag{2.5}$$

The triplet (Q, ν, γ) is called characteristic triplet of the process X. Throughout this work, we shall assume

Assumption 2.1. *The diffusion part of X does not vanish, i.e.*

$$Q > 0. \tag{2.6}$$

Since the law of a Lévy process X is described in a time-independent fashion in terms of its characteristic triplet (Q, ν, γ), it seems natural to use the triplet for a characterization of the dependence among the one-dimensional margins of X. Since the drift vector γ corresponds to

the deterministic drift part of X, it has no effect on the dependence structure between margins. The dependence of the Brownian motion part of X is characterized entirely by its covariance matrix Q. Since the continuous part and the jump part of a Lévy process are independent, it remains to describe the dependence structure of the jump part of X. To this end, we apply Lévy copulas as originally defined in [30, 49].

Throughout this work, we shall consider a Lévy copula process X as defined in [22, 45]. For sake of brevity, we refer to these sources for the detailed definitions. In particular, the process X with characteristic triplet (Q, ν, γ) admits a Lévy density $k(z)dz = \nu(dz)$. Furthermore, we assume the following:

Assumption 2.2. *The marginal Lévy measures ν_i, $i = 1, \dots, d$, of X are absolutely continuous with densities k_i that satisfy:*

- Semiheavy tails: *There are constants $G_i > 0$, $M_i > 0$, $i = 1, \dots, d$, such that*

$$k_i(z) \lesssim \begin{cases} e^{G_i z}, & z < -1 \\ e^{-M_i z}, & z > 1 \end{cases} . \tag{2.7}$$

- Quasi-stable margins: *There are constants $0 < Y_i \leq 1$ and $c_i^+, c_i^- \geq 0$, $c_i^+ + c_i^- > 0$, $i = 1, \dots, d$, such that*

$$\begin{aligned} k_i(z) &\gtrsim c_i^- \frac{1}{|z|^{1+Y_i}} 1_{\{z<0\}}(z) + c_i^+ \frac{1}{|z|^{1+Y_i}} 1_{\{0\leq z\}}(z) \quad 0 < |z| \leq 1 \,, \\ k_i(z) &\lesssim c_i^- \frac{1}{|z|^{1+Y_i}} 1_{\{z<0\}}(z) + c_i^+ \frac{1}{|z|^{1+Y_i}} 1_{\{0\leq z\}}(z) \quad 0 < |z| \leq 1 \,. \end{aligned} \tag{2.8}$$

Remark 2.3. *The assumption of semiheavy tails* (2.7) *is required to obtain a suitable explicit form of the infinitesimal generator $\mathcal{A}$ of X as in [45, Corollary 4.3]. Furthermore, it is necessary for the efficient localization of the corresponding variational problem* (1.2) *on $\mathbb{R}^d$ to the bounded domain $\Omega := [-R, R]^d$, $R > 0$, as described in [45, Section 4.3].*

Remark 2.4. *Assumption* (2.8) *with $Y \leq 1$ on the intensity of the margins' singularities at the origin is required to prove optimal convergence of our numerical schemes below.*

Remark 2.5. *Assumption 2.2 is satisfied by many Lévy processes used in financial modeling, for example Kou's model [31], Normal Inverse Gaussian processes [2], Meixner processes [48], and tempered stable or CGMY processes [8] with $Y \leq 1$.*

With (2.6) and Assumption 2.2, from [22, Theorem 3.6] and [45, Theorem 4.6] we infer the following concrete realization and well-posedness of the abstract problem (1.2):

$$\begin{aligned} &\text{Find } u \in L^2((0,T); H^1(\mathbb{R}^d)) \cap H^1((0,T); H^{-1}(\mathbb{R}^d)) \text{ such that} \\ &\langle \frac{\partial u}{\partial \tau}, v \rangle_{H^{-1}(\mathbb{R}^d), H^1(\mathbb{R}^d)} + \mathcal{E}(u, v) = 0 \,, \ \tau \in (0,T), \, \forall v \in H^1(\mathbb{R}^d) \,, \\ &u(0) = u_0(x) := g(e^{x_1}, \dots, e^{x_d}) \,. \end{aligned} \tag{2.9}$$

After so-called "removal of drift", i.e. a suitable transformation to remove the linear terms in (2.3) under the assumption of semiheavy tails (2.7) as described in e.g. [45, Corollary 4.3], the bilinear form $\mathcal{E}(\cdot, \cdot)$ in (2.9) is given by

$$\mathcal{E}(u,v) \;=\; D(u,v) \;+\; J(u,v), \quad u,v \in H^1(\mathbb{R}^d), \tag{2.10}$$

with diffusion part

$$D(u,v) \;=\; \frac{1}{2}\sum_{i,j=1}^{d} Q_{ij} \int_{\mathbb{R}^d} \frac{\partial u}{\partial x_i}(x)\frac{\partial v}{\partial x_j}(x)dx\,, \tag{2.11}$$

and jump part

$$\begin{aligned} J(u,v) &= -\int_{\mathbb{R}^d} \psi_J(\xi)\widehat{u}(\xi)\overline{\widehat{v}(\xi)}d\xi \\ &= -\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} \left(u(x+z) - u(x) - \sum_{i=1}^{d} z_i \frac{\partial u}{\partial x_i}(x) \right) v(x)dx\,\nu(dz)\,. \end{aligned} \tag{2.12}$$

Here, the symbol ψ_J is given by (2.4). The well-posedness of (2.9) is ensured by

Theorem 2.6. *The bilinear form $\mathcal{E}(\cdot,\cdot)$ satisfies a Gårding inequality, i.e. there exist constants $\gamma > 0$ and $c \geq 0$ such that*

$$\mathcal{E}(u,u) \geq \gamma \|u\|^2_{H^1(\mathbb{R}^d)} - c\|u\|^2_{L^2(\mathbb{R}^d)} \quad \forall u \in H^1(\mathbb{R}^d), \tag{2.13}$$

and is continuous, i.e.

$$\mathcal{E}(u,v) \lesssim \|u\|_{H^1(\mathbb{R}^d)}\|v\|_{H^1(\mathbb{R}^d)} \;\forall u,v \in H^1(\mathbb{R}^d).$$

Proof. See [45, Theorem 4.6]. □

Using the exponential shift in time $u \mapsto e^{-ct}u$ with $c \geq 0$ as in (2.13) to erase the L^2-term in (2.13), one may furthermore assume without loss of generality that $\mathcal{E}(\cdot,\cdot)$ is coercive on $H^1(\mathbb{R}^d) \times H^1(\mathbb{R}^d)$, i.e. there holds

$$\mathcal{E}(u,u) \gtrsim \|u\|^2_{H^1(\mathbb{R}^d)} \;\forall u \in H^1(\mathbb{R}^d). \tag{2.14}$$

For our analysis, we also need to recall the following well-known result:

Lemma 2.7. *There holds*

$$\|u\|^2_{H^1(\mathbb{R}^d)} \lesssim D(u,u) \;\forall u \in H^1(\mathbb{R}^d)\,, \tag{2.15}$$

where $D(\cdot,\cdot)$ denotes the Laplace-type bilinear form given by (2.11).

Additionally, from (2.8) and the Fourier representation in (2.12) we obtain the following crucial lemma.

Lemma 2.8. *For any u, $v \in H^1(\mathbb{R}^d)$ there holds*

$$|J(u,v)| \;\lesssim\; \|u\|_{L^2(\mathbb{R}^d)}\|v\|_{H^1(\mathbb{R}^d)}. \tag{2.16}$$

Proof. By [22, Theorem 3.4] and [45, Proposition 3.5], assumption (2.8) implies that the symbol ψ_J in (2.4) is equivalent to an anisotropic distance function that satisfies $\psi_J(\xi) \geq 0$ for all $\xi \in \mathbb{R}^d$ and furthermore

$$\psi_J(\xi) \lesssim |\xi_1|^{Y_1} + \ldots + |\xi_d|^{Y_d} + 1$$

with $Y_1, \dots, Y_d$ given by (2.8). Denoting $\overline{Y} = \max\{Y_1, \dots, Y_d\}$, we thus obtain

$$\psi_J(\xi) \lesssim (1+|\xi|^2)^{\overline{Y}/2} \lesssim (1+|\xi|^2)^{1/2} \ \forall \xi \in \mathbb{R}^d,$$

where the last inequality follows from (2.8). Hence, using the Cauchy-Schwarz inequality and Plancherel's Theorem with the representation (2.12), we obtain

$$\begin{aligned} |J(u,v)|^2 &\lesssim \int_{\mathbb{R}^d} |\widehat{u}(\xi)|^2 d\xi \cdot \int_{\mathbb{R}^d} |\psi_J(\xi)\overline{\widehat{v}(\xi)}|^2 d\xi \\ &\lesssim \|\widehat{u}\|^2_{L^2(\mathbb{R}^d)} \cdot \int_{\mathbb{R}^d} (1+|\xi|^2)|\widehat{v}(\xi)|^2 d\xi \\ &\lesssim \|u\|^2_{L^2(\mathbb{R}^d)} \cdot \|v\|^2_{H^1(\mathbb{R}^d)}. \end{aligned}$$

□

Remark 2.9. *In case the constraint $Y_i \leq 1$, $i = 1, \dots, d$, in (2.8) is not satisfied, for any sufficiently smooth function v the arguments of Lemma 2.8 directly imply*

$$|J(u,v)| \lesssim \|u\|_{L^2(\mathbb{R}^d)} \|v\|_{H^{\overline{Y}}(\mathbb{R}^d)} \tag{2.17}$$

with $\overline{Y} = \max\{Y_1, \dots, Y_d\}$.

Furthermore, if $\Omega \subset \mathbb{R}^d$ is a bounded convex domain, then it is seen from a standard argument that $\forall f \in L^2(\Omega)$, there exists a unique $u \equiv \mathcal{A}^{-1} f \in H^1_0(\Omega) \cap H^2(\Omega)$ such that

$$\|u\|_{2,\Omega} \lesssim \|f\|_{0,\Omega}. \tag{2.18}$$

Even though the main goal of this work is to obtain an efficient space discretization scheme for (2.9), for sake of completeness we also briefly illustrate how the time dependence of (2.9) can be discretized. To this end, we describe the θ-scheme here. In practice, very often more sophisticated tools, e.g. *hp*-DG time-stepping (see [29]), can be applied to obtain exponential convergence in time.

For the moment, assume that we have fixed a finite dimensional space $V^h \subset H^1(\Omega)$ corresponding to a given meshwidth $h > 0$. After removal of drift, the Kolmogorov equation (1.1) is parabolic, and the θ-scheme can be described as follows:

Let $0 \leq \theta \leq 1$. For $T < \infty$ and $M \in \mathbb{N}$, define the time step $k = T/M$, and $t^m = mk$, $m = 0, \dots, M$. The fully discrete θ-scheme reads:

$$\begin{aligned} &\text{First find } u_h^0 \in V^h \text{ satisfying } u_h^0 = u_{0,J}. \\ &\text{Then for } m = 0, 1, \dots, M-1, \text{ find } u_h^{m+1} \in V^h \text{ such that} \\ &\langle \frac{u_h^{m+1} - u_h^m}{k}, v_J \rangle + \mathcal{E}(u_h^{m+\theta}, v_J) = 0 \ \ \forall v_J \in V^h \,. \end{aligned} \tag{2.19}$$

Here $u_h^{m+\theta} := \theta u_h^{m+1} + (1-\theta) u_h^m$. The approximation of the initial data could be chosen as a finite element projection $u_{0,J} = P_h(u_0|_\Omega)$ or as an interpolant of $u_0|_{\Omega_R}$. The efficient numerical solution of the stationary elliptic problems in (2.19) is studied in the following sections. If we denote the resulting coefficient vector of the solution by $\underline{u}^m$ then the sequence of problems (2.19) is equivalent to the matrix equations

$$(k^{-1}\boldsymbol{M} + \theta \boldsymbol{A})\underline{u}^{m+1} = k^{-1}\boldsymbol{M}\underline{u}^m - (1-\theta)\boldsymbol{A}\underline{u}^m, \quad m = 0, 1, \dots, M-1, \tag{2.20}$$

where $\boldsymbol{M}$ denotes the mass-matrix and $\boldsymbol{A}$ the stiffness matrix of $\mathcal{E}(\cdot,\cdot)$ with respect to the basis of V^h.

Note that for $\theta = 1/2$, the scheme in (2.19) coincides with the Crank-Nicholson scheme whereas for $\theta = 0$ one obtains the explicit and for $\theta = 1$ the implicit Euler scheme. For stability and convergence considerations, see e.g. [39, 51].

3 Basic two-scale discretizations

In this section we define certain finite element discretization schemes to solve the elliptic (log price) problem. More generally, we solve: Find $u \in H_0^1(\Omega)$ such that

$$\mathcal{E}(u,v) = \langle f, v\rangle \quad \forall v \in H_0^1(\Omega)\,, \tag{3.1}$$

where $\mathcal{E}(\cdot,\cdot)$ is given by (2.10) and Ω is a bounded domain in $\mathbb{R}^d$.

In order to discretize the variational equation (3.1), we employ a sequence of piecewise linear finite element spaces associated with a shape-regular finite element mesh. Assume that $T^h(\Omega) = \{\tau\}$ is a mesh of Ω with mesh-size function $h(x)$ whose value is the diameter h_τ of the element τ containing x. One basic assumption on the mesh is that it is not exceedingly over-refined locally, namely,

Assumption 3.1. *There exists $\nu \geq 1$ such that*

$$h_\Omega^\nu \lesssim h(x), \quad x \in \Omega, \tag{3.2}$$

where $h_\Omega = \max_{x\in\Omega} h(x)$ is the (largest) mesh size of $T^h(\Omega)$.

This is obviously a very mild assumption from theoretical point of view. Usually, we will drop the subscript and simply write h instead of h_Ω for the mesh size on a domain that is clear from the context.

Let $T^h(\Omega)$ consist of shape-regular simplices and define $S^{h,r}(\Omega)$ to be a space of continuous functions on Ω such that for $v \in S^{h,r}(\Omega), v$ restricted to each τ is a polynomial of total degree $\leq r$, namely

$$S^{h,r}(\Omega) = \{v \in C(\bar{\Omega}) : v\,|_\tau \in P_\tau^r \quad \forall \tau \in T^h(\Omega)\}, \tag{3.3}$$

where P_τ^r is the space of polynomials of degree not greater than a positive integer r. Set $S_0^{h,r}(\Omega) = S^{h,r}(\Omega) \cap H_0^1(\Omega)$. These are the Lagrange finite element spaces and we refer to [9, 54] for their basic properties that will be used in our analysis. For simplicity, in this paper we shall focus our study only on the piecewise linear Lagrange finite element approximation. Let $S^h(\Omega) = S^{h,1}(\Omega)$ and $S_0^h(\Omega) = S_0^{h,1}(\Omega)$.

Note that, the analysis of this work does not depend on the particular choice of piecewise linear basis functions for $S_0^h(\Omega)$. For instance, one may choose a classical Lagrangian ("nodal") finite element bases (see e.g. [9, Section II.7]) or piecewise linear wavelet basis functions (see e.g. [10, Chapter 1]). In particular, if Ω is a tensor product domain choosing a wavelet basis for $S_0^h(\Omega)$ allows for the very efficient discretization of the non-local part $J(\cdot,\cdot)$ of $\mathcal{E}(\cdot,\cdot)$ as described in [22, 44, 51, 52].

The (standard) one-scale finite element discretization for (3.1) reads: Find $u_h \in S_0^h(\Omega)$ such that

$$\mathcal{E}(u_h, v) = \langle f, v\rangle \quad \forall v \in S_0^h(\Omega). \tag{3.4}$$

We also require the Galerkin-projection $P_h : H_0^1(\Omega) \mapsto S_0^h(\Omega)$, defined by

$$\mathcal{E}(u - P_h u, v) = 0 \quad \forall v \in S_0^h(\Omega). \tag{3.5}$$

Using the coercivity (2.14) of $\mathcal{E}(\cdot,\cdot)$ and Theorem 2.6, one obtains the following well-known error bounds (see e.g. [9, Chapter III]):

Theorem 3.2. *If $u_h \in S_0^h(\Omega)$ is the solution of (3.4), then*

$$\|u - u_h\|_{1,\Omega} \lesssim h|u|_{2,\Omega}, \tag{3.6}$$

$$\|u - u_h\|_{0,\Omega} \lesssim h^2|u|_{2,\Omega}. \tag{3.7}$$

Remark 3.3. *As shown in [45, Section 4.5], in the theory of asset pricing the bounded domain Ω arises from a localization procedure. In fact, Ω can be interpreted as the active area of a knock-out barrier option with the boundary of Ω being the barrier. In the notation of Section 1, the price of this barrier contract is given by*

$$u_\Omega(x,t) = \mathbb{E}\left(g(X_T)1_{\{T<\tau_\Omega\}}|X_t = x\right), \tag{3.8}$$

where $\tau_\Omega = \inf\{s \geq 0 | X_s \notin \Omega\}$ denotes the first exit time of Ω by X.

3.1 A basic two-scale discretization

Due to the non-locality of $\mathcal{E}(\cdot,\cdot)$ and the high-dimensionality of Ω, the straightforward finite element discretization (3.4) yields a dense matrix of substantial size, which ususally is not practicable to implement. In order to reduce the computational cost of solving the original problem (3.1), we may introduce a so-called basic two-scale method.

The main idea of the basic two-scale method is to use a coarse mesh of size H, to approximate the low frequencies and to use a fine mesh of size h, $h \ll H$, to handle the high frequencies (c.f. [53, 54]). Based on (2.10), Theorem 2.6, Lemma 2.7 and [22, Theorem 3.5], we may indeed treat $J(\cdot,\cdot)$ as a low frequency perturbation of the high frequency part $D(\cdot,\cdot)$ of $\mathcal{E}(\cdot,\cdot)$. More precisely, for $h \ll H$ our basic two-scale algorithm is defined as follows:

Algorithm 3.4.

1. *Solve (3.1) on a coarse gird: Find $u_H \in S_0^H(\Omega)$ such that*

$$\mathcal{E}(u_H, v) = \langle f, v\rangle \quad \forall v \in S_0^H(\Omega).$$

2. *Compute a linear boundary value problem on a fine gird: Find $u^h \in S_0^h(\Omega)$ such that*

$$D(u^h, v) = \langle f, v\rangle - J(u_H, v) \quad \forall v \in S_0^h(\Omega).$$

3. *Find a further coarse grid correction $e_H \in S_0^H(\Omega)$ such that*

$$\mathcal{E}(e_H, v) = \langle f, v\rangle - \mathcal{E}(u^h, v) \quad \forall v \in S_0^H(\Omega)$$

and set $\tilde{u}^h = u^h + e_H$ in Ω.

Theorem 3.5. *Assume that u^h and $\tilde{u}^h$ are obtained by Algorithm 3.4, then*

$$\|u^h - u_h\|_{1,\Omega} \lesssim H^2|u|_{2,\Omega}, \tag{3.9}$$
$$\|\tilde{u}^h - u_h\|_{0,\Omega} \lesssim H^3|u|_{2,\Omega}, \tag{3.10}$$

where u_h denotes the solution of the one-scale discretization (3.4)*. Consequently,*

$$\|u - u^h\|_{1,\Omega} \lesssim (h + H^2)|u|_{2,\Omega}, \tag{3.11}$$
$$\|u - \tilde{u}^h\|_{0,\Omega} \lesssim (h^2 + H^3)|u|_{2,\Omega}, \tag{3.12}$$

where u denotes the solution of the original variational equation (3.1)*.*

Proof. From (3.4), we have

$$D(u_h, v) + J(u_h, v) = \langle f, v\rangle \quad \forall v \in S_0^h(\Omega),$$

which together with Lemma 2.8 implies

$$D(u^h - u_h, v) = -J(u_H - u_h, v) \lesssim \|u_H - u_h\|_{0,\Omega}\|v\|_{1,\Omega} \quad \forall v \in S_0^h(\Omega).$$

Using Theorem 3.2, we then obtain

$$\begin{aligned}\|u^h - u_h\|_{1,\Omega}^2 &\lesssim D(u^h - u_h, u^h - u_h)\\ &\lesssim \|u_H - u_h\|_{0,\Omega}\|u^h - u_h\|_{1,\Omega} \lesssim H^2|u|_{2,\Omega}\|u^h - u_h\|_{1,\Omega}.\end{aligned}$$

Hence, we get (3.9). Note that there holds

$$\|\tilde{u}^h - u_h\|_{0,\Omega} = \|(I - P_H)(u^h - u_h)\|_{0,\Omega} \lesssim H\|u^h - u_h\|_{1,\Omega} \lesssim H^3|u|_{2,\Omega}.$$

Here I denotes the identity operator. By Theorem 3.2 and the triangle inequality, we obtain (3.11) and (3.12). This completes the proof. □

Remark 3.6. *If instead of the finite element space $S_0^h(\Omega)$ with piecewise linear basis functions one chooses a higher order basis, i.e. $S_0^{h,r}(\Omega)(r \geq 1)$ with the mesh of size h, then it can be seen from the above arguments that the solutions u^h and $\tilde{u}^h$ of the corresponding basic two-scale Algorithm 3.4 satisfy*

$$\|u - u^h\|_{1,\Omega} \lesssim (h^r + H^{r+1})|u|_{r+1,\Omega},$$
$$\|u - \tilde{u}^h\|_{0,\Omega} \lesssim (h^{r+1} + H^{r+2})|u|_{r+1,\Omega}.$$

Naturally, in this case one may choose Lagrangian finite element basis functions of order r or piecewise polynomial wavelets of degree r to generate $S_0^{h,r}(\Omega)$.

3.2 Local and parallel algorithms

In this subsection, following [54] we propose a local and a parallel algorithm. We shall first discuss the local algorithm. The generalization of the local to the parallel algorithm is straightforward.

The main idea of the local algorithm is that the global components of an approximation may be obtained by a relatively coarse grid and the rest of the computation can then be localized.

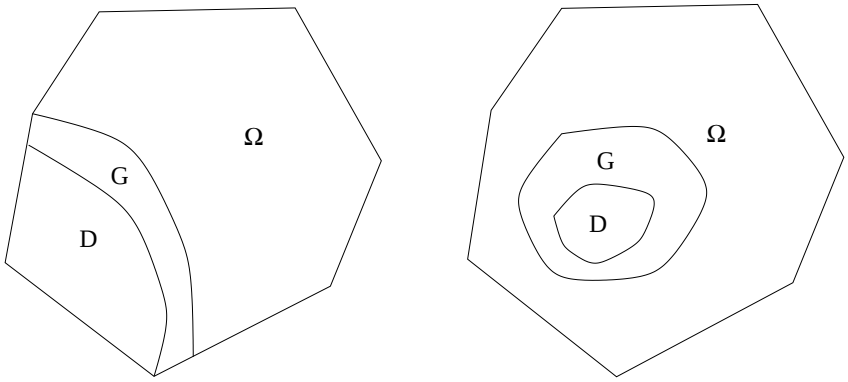

Figure 1: Subdomains

For $D \subset G \subset \Omega$, we use the notation $D \subset\subset G$ to mean that $\text{dist}(\partial D \setminus \partial\Omega, \partial G \setminus \partial\Omega) > 0$ (see Figure 1). For simplicity, we assume that Ω and its any subdomain involved in this subsection are polytopic. Note that any $w \in H_0^1(\Omega_0)$ can be naturally extended to be a function in $H_0^1(\Omega)$ with zero outside of Ω_0 when $\Omega_0 \subset \Omega$, thus we shall state this fact by the slightly abused notation $H_0^1(\Omega_0) \subset H_0^1(\Omega)$.

Given $G \subset \Omega$, we define $S^h(G)$ and $T^h(G)$ to be the restriction of $S^h(\Omega)$ and $T^h(\Omega)$ to G, respectively, and

$$S_0^h(G) = \{v \in S^h(\Omega) : \text{supp } v \subset\subset G\}.$$

For any $G \subset \Omega$ mentioned in this paper, we assume that it aligns with $T^h(\Omega)$ whenever necessary.

Let Ω_0 be a subdomain of Ω containing another slightly smaller subdomain $D \subset \Omega$ (namely $D \subset\subset \Omega_0$). As above, suppose $h \ll H$. We let $T_H^h(\Omega)$ denote a locally refined shape-regular mesh that may be viewed as being obtained by refining $T^H(\Omega)$ locally around the subdomain D in such a way that $T_H^h(\Omega_0) = T^h(\Omega_0)$. We are interested in obtaining the approximation solution in the given subdomain D with an accuracy comparable to that from $T^h(\Omega)$.

The local algorithm reads as follows.

Algorithm 3.7. *1. Find a global coarse grid solution $u_H \in S_0^H(\Omega)$:*

$$\mathcal{E}(u_H, v) = \langle f, v\rangle \quad \forall v \in S_0^H(\Omega).$$

2. Find a local fine grid correction $e_h \in S_0^h(\Omega_0)$:

$$D(e_h, v) = \langle f, v\rangle - \mathcal{E}(u_H, v) \quad \forall v \in S_0^h(\Omega_0).$$

3. Update: $u^h = u_H + e_h$, in Ω_0.

To analyze Algorithm 3.7, we need the following result that was shown in [54].

Lemma 3.8. *[54] Suppose that $f \in H^{-1}(\Omega)$ and $D \subset\subset \Omega_0$. If $w \in S^h(\Omega_0)$ satisfies*

$$D(w, v) = f(v) \quad \forall\ v \in S_0^h(\Omega_0), \tag{3.13}$$

then

$$\|w\|_{1,D} \lesssim \|w\|_{0,\Omega_0} + \|f\|_{-1,\Omega_0}, \tag{3.14}$$

where

$$\|f\|_{-1,\Omega_0} = \sup_{\phi\in H_0^1(\Omega_0),\|\phi\|_{1,\Omega_0}=1} f(\phi).$$

Theorem 3.9. *Assume that $u^h \in S_0^h(\Omega_0)$ is obtained by Algorithm 3.7. Then*

$$\|u_h - u^h\|_{1,D} \lesssim \|u_h - u_H\|_{0,\Omega} + H\|u_h - u_H\|_{1,\Omega} \lesssim H^2|u|_{2,\Omega}.$$

Proof. From the construction of Algorithm 3.7, we have

$$D(u^h - u_h, v) = J(u_h - u_H, v) \quad \forall v \in S_0^h(\Omega_0), \tag{3.15}$$

which together with Lemma 3.8 yields

$$\|u^h - u_h\|_{1,D} \lesssim \|u^h - u_h\|_{0,\Omega_0} + \|u_h - u_H\|_{0,\Omega_0} \lesssim \|u_h - u_H\|_{0,\Omega_0} + \|e_h\|_{0,\Omega_0}.$$

It remains to estimate $\|e_h\|_{0,\Omega_0}$, for which we use the Aubin-Nitsche duality argument. Given any $\phi \in L^2(\Omega_0)$, there exists $w \in H_0^1(\Omega_0)$ such that

$$D(v, w) = \langle \phi, v\rangle \quad \forall v \in H_0^1(\Omega_0).$$

Let $w_h^0 \in S_0^h(\Omega_0)$ and $w_H^0 \in S_0^H(\Omega_0)$ satisfy

$$D(v_h, w_h^0) = D(v_h, w) \quad \forall v_h \in S_0^h(\Omega_0), \quad D(v_H, w_H^0) = D(v_H, w) \quad \forall v_H \in S_0^H(\Omega_0).$$

Then

$$\|w - w_h^0\|_{1,\Omega_0} \lesssim h\|\phi\|_{0,\Omega_0}, \quad \|w - w_H^0\|_{1,\Omega_0} \lesssim H\|\phi\|_{0,\Omega_0}.$$

From (3.15) and

$$\langle e_h, \phi\rangle = D(e_h, w) = D(e_h, w_h^0) = D(u^h - u_h, w_h^0) + D(u_h - u_H, w_h^0),$$

it follows that

$$\begin{aligned}
\langle e_h, \phi\rangle &= J(u_h - u_H, w_h^0) + D(u_h - u_H, w_h^0) = \mathcal{E}(u_h - u_H, w_h^0) \\
&= \mathcal{E}(u_h - u_H, w_h^0 - w) + \mathcal{E}(u_h - u_H, w) \\
&= \mathcal{E}(u_h - u_H, w_h^0 - w) + \mathcal{E}(u_h - u_H, w - w_H^0) \\
&\lesssim H\|u_h - u_H\|_{1,\Omega}\|\phi\|_{0,\Omega_0},
\end{aligned}$$

which implies

$$\|e_h\|_{0,\Omega_0} \lesssim H\|u_h - u_H\|_{1,\Omega}.$$

The desired result then follows. □

Next, we shall propose a parallel algorithm that achieves the following goals:

1. Smooth u^h to obtain a global $H^1(\Omega)$ approximation;
2. Improve the L^2 error.

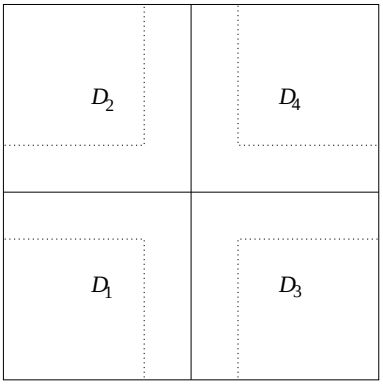

Figure 2: Domain decomposition: $D_j \subset\subset \Omega_j$ and D_{m+1}

We assume that $\{\Omega_1, \Omega_2, \ldots, \Omega_m\}$ is a partition of Ω (namely, $\Omega_i \cap \Omega_j = \emptyset$, $i \neq j$, and $\cup_{j=1}^m \bar{\Omega}_j = \bar{\Omega}$) and set $D_{m+1} = \Omega \setminus (\cup_{j=1}^m \bar{D}_j)$, where $D_j \subset\subset \Omega_j$, $j = 1, 2, \ldots, m$, (see Figure 2 for an illustration in two dimensions).

Algorithm 3.10. *1. Find a global coarse grid solution $u_H \in S_0^H(\Omega)$:*

$$\mathcal{E}(u_H, v) = \langle f, v \rangle \quad \forall v \in S_0^H(\Omega).$$

2. Find local fine grid corrections $e_h^j \in S_0^h(\Omega_j)(j = 1, 2, \ldots, m)$ in parallel:

$$D(e_h^j, v) = \langle f, v \rangle - \mathcal{E}(u_H, v) \quad \forall v \in S_0^h(\Omega_j).$$

3. Set $u^h = u_H + e_h^j$, in $D_j (j = 1, 2, \ldots, m)$, and u^h on $\bar{D}_{m+1}$ is defined by: $u^h \mid_{\partial D_j \cap \partial D_{m+1}} = u_H + e_h^j (j = 1, 2, \ldots, m)$ and satisfying

$$D(u^h, v) = \langle f, v \rangle - J(u_H, v) \quad \forall v \in S_0^h(D_{m+1}).$$

4. Find a further coarse grid correction $e_H \in S_0^H(\Omega)$:

$$\mathcal{E}(e_H, v) = \langle f, v \rangle - \mathcal{E}(u^h, v) \quad \forall v \in S_0^H(\Omega).$$

5. Update: $\tilde{u}^h = u^h + e_H$ in Ω.

In the above algorithm, Step 3 is for obtaining a global H^1 solution and Step 4 is for improving the L^2 error.

Theorem 3.11. *Assume that u^h and $\tilde{u}^h$ are the solutions obtained by Algorithm 3.10. Then*

$$\|u_h - u^h\|_{1,\Omega} \lesssim H \|u_h - u_H\|_{1,\Omega} \lesssim H^2 |u|_{2,\Omega}$$

and

$$\|u_h - \tilde{u}^h\|_{0,\Omega} \lesssim H^2 \|u_h - u_H\|_{1,\Omega} \lesssim H^3 |u|_{2,\Omega}.$$

Proof. Let

$$D_{D_{m+1}}(u,v)=\frac{1}{2}\sum_{i,j=1}^{d}Q_{ij}\int_{D_{m+1}}\frac{\partial u}{\partial x_i}(x)\frac{\partial v}{\partial x_j}(x)dx,$$

then there holds

$$\|\nabla(u_h-u^h)\|^2_{0,D_{m+1}}\lesssim D_{D_{m+1}}(u_h-u^h,u_h-u^h).$$

Note that from the construction of Algorithm 3.10, we have

$$D(u_h-u^h,v)=J(u_H-u_h,v)\quad\forall v\in S_0^h(D_{m+1}). \tag{3.16}$$

Hence, for any $v\in S_0^h(D_{m+1})$, by Lemma 2.8 we obtain

$$\|\nabla(u_h-u^h)\|^2_{0,D_{m+1}}\lesssim D_{D_{m+1}}(u_h-u^h,u_h-u^h-v)+\|u_h-u_H\|_{0,\Omega}\|v\|_{1,\Omega}$$

and

$$\begin{aligned}\|\nabla(u_h-u^h)\|^2_{0,D_{m+1}}\lesssim{}&\|\nabla(u_h-u^h)\|_{0,D_{m+1}}\inf_{\chi\in S_0^h(D_{m+1})}\|u_h-u^h-\chi\|_{1,D_{m+1}}\\&+\|u_h-u_H\|_{0,\Omega}\inf_{\chi\in S_0^h(D_{m+1})}\left(\|u_h-u^h-\chi\|_{1,D_{m+1}}+\|\chi\|_{1,\Omega}\right).\end{aligned}$$

Since

$$\inf_{\chi\in S_0^h(G)}\|v-\chi\|_{1,G}\lesssim\|v\|_{1/2,\partial G}\quad\forall v\in S^h(G) \tag{3.17}$$

is valid for any polytopic domain $G\subset\Omega$ [54], we get

$$\begin{aligned}\|\nabla(u_h-u^h)\|^2_{0,D_{m+1}}\lesssim{}&\|\nabla(u_h-u^h)\|_{0,D_{m+1}}\|u_h-u^h\|_{1/2,\partial D_{m+1}}\\&+\|u_h-u_H\|_{0,\Omega}(\|u_h-u^h\|_{1/2,\partial D_{m+1}}+\|u_h-u^h\|_{1,D_{m+1}}).\end{aligned}$$

Using the estimation

$$\|u_h-u^h\|_{1/2,\partial D_{m+1}}\lesssim\left(\sum_{j=1}^{m}\|u_h-u^h\|^2_{1/2,\partial D_j}\right)^{1/2}\lesssim\left(\sum_{j=1}^{m}\|u_h-u^h\|^2_{1,D_j}\right)^{1/2},$$

or

$$\|u_h-u^h\|_{1/2,\partial D_{m+1}}\lesssim|\!|\!|u_h-u^h|\!|\!|_{1,\Omega},$$

we conclude

$$\begin{aligned}\|\nabla(u_h-u^h)\|^2_{0,D_{m+1}}&\lesssim|\!|\!|u_h-u^h|\!|\!|^2_{1,\Omega}\\&+(|\!|\!|u_h-u^h|\!|\!|_{1,\Omega}+\|u_h-u^h\|_{1,D_{m+1}})\|u_h-u_H\|_{0,\Omega},\end{aligned}$$

where

$$|\!|\!|u_h-u^h|\!|\!|_{1,\Omega}=\left(\sum_{j=1}^{m}\|u_h-u^h\|^2_{1,D_j}\right)^{1/2}.$$

Thus we may estimate as follows

$$\|u_h - u^h\|_{1,\Omega}^2 \lesssim \|\nabla(u_h - u^h)\|_{0,\Omega}^2 = \|\nabla(u_h - u^h)\|_{0,D_{m+1}}^2 + \|\nabla(u_h - u^h)\|_{0,\Omega\setminus D_{m+1}}^2$$

$$\lesssim \quad |\!|\!| u_h - u^h |\!|\!|_{1,\Omega}^2 + \|u_h - u_H\|_{0,\Omega} |\!|\!| u_h - u^h |\!|\!|_{1,\Omega} + \|u_h - u^h\|_{1,\Omega} \|u_h - u_H\|_{0,\Omega},$$

namely, we arrive at

$$\|u_h - u^h\|_{1,\Omega} \lesssim |\!|\!| u_h - u^h |\!|\!|_{1,\Omega} + \|u_h - u_H\|_{0,\Omega}. \tag{3.18}$$

Note that

$$\|u_h - u_H\|_{0,\Omega} \lesssim H \|u_h - u_H\|_{1,\Omega}.$$

By Theorem 3.9 and

$$u_h - \tilde{u}^h = (I - P_H)(u_h - u^h), \tag{3.19}$$

we complete the proof. □

4 A combination based two-scale discretization

In this section, we shall discuss a combination based two-scale finite element discretization to reduce the computational complexity further over tensor product domains. For sake of brevity, we only give a detailed description of three-dimensional problems over the domain $\Omega = [0,1]^3$ here. The extension to $\Omega = [0,1]^d$, $d \geq 3$, is addressed at the end of this section. The results in this section can easily be generalized to the domain $[-R,R]^d$, $R > 0$. Note that such tensor product domains arise in most of the classical asset pricing problems (c.f. e.g. [11, 22]).

In our discussion, we require a so-called mixed Sobolev space (see e.g. [35, 43]):

$$W_2^{G,3}(\Omega) := \{w \in H^2(\Omega) |\ \partial_{x_i}\partial_{x_j}\partial_{x_k}(w) \in L^2(\Omega), i,j,k = 1,\ldots,d, x_i \neq x_j \text{ or } x_i \neq x_k\}$$

with its natural norm $\|\cdot\|_{W_2^{G,3}(\Omega)}$. In this notation similar spaces have already been introduced in [42]. The superscript G stems from the German "gemischt" meaning "mixed".

Furthermore, we shall introduce two-scale interpolations for the three-dimensional case. Assume that $T^{h_{x_i}}([0,1])$ is a uniform mesh with mesh size h_{x_i} on $[0,1]$ and $S^{h_{x_i}}([0,1]) \subset H^1([0,1])$ is the associated piecewise linear finite element space, where $i = 1,2,3$. Set $S_0^{h_{x_i}}([0,1]) = S^{h_{x_i}}([0,1]) \cap H_0^1([0,1])$ and $T^{h_{x_1},h_{x_2},h_{x_3}}(\Omega) = T^{h_{x_1}}([0,1]) \times T^{h_{x_2}}([0,1]) \times T^{h_{x_3}}([0,1])$ is the tensor product mesh. The operator $I_h : C([0,1]) \to S^h([0,1])$ is the standard Lagrangian interpolation operator defined on $T^h([0,1])$.

Let $I_{h_{x_1},h_{x_2},h_{x_3}}$ be the usual trilinear interpolation operator on the partition $T^{h_{x_1},h_{x_2},h_{x_3}}(\Omega)$. One sees that $I_{h_{x_1},0,0}$ is the interpolation operator which interpolates only in x_1-direction on lines of mesh size h_{x_1}, etc. Obviously,

$$I_{h_{x_1},h_{x_2},h_{x_3}} = I_{h_{x_1},0,0} \cdot I_{0,h_{x_2},0} \cdot I_{0,0,h_{x_3}}.$$

It is shown in the following lemma that a one-scale interpolation on a fine grid can be obtained by a combination of two-scale interpolations asymptotically. Let $H \gg h$, and define the two-scale interpolation by

$$I_{H,H,H}^h u = I_{h,H,H} u + I_{H,h,H} u + I_{H,H,h} u - 2 I_{H,H,H} u.$$

From the identity

$$\begin{aligned} I =& I_{h,h,h} + (I - I_{h,0,0}) + (I - I_{0,h,0}) + (I - I_{0,0,h}) - (I - I_{h,0,0})(I - I_{0,h,0}) \\ & - (I - I_{h,0,0})(I - I_{0,0,h}) - (I - I_{0,h,0})(I - I_{0,0,h}) + (I - I_{h,0,0})(I - I_{0,h,0})(I - I_{0,0,h}), \end{aligned}$$

we obtain the following two results (see [35] for details).

Lemma 4.1. *[35] If $u \in W_2^{G,3}(\Omega)$, then*

$$H\|I^h_{H,H,H}u - I_{h,h,h}u\|_{1,\Omega} + \|I^h_{H,H,H}u - I_{h,h,h}u\|_{0,\Omega} \lesssim H^3\|u\|_{W_2^{G,3}(\Omega)}. \tag{4.1}$$

Lemma 4.2. *[35] If $u \in H_0^1(\Omega) \cap W_2^{G,3}(\Omega)$, then*

$$D((I - I_{h_{x_1},h_{x_2},h_{x_3}})u, v) \lesssim (\max\{h_{x_1}, h_{x_2}, h_{x_3}\})^2\|u\|_{W_2^{G,3}(\Omega)}\|v\|_{1,\Omega} \ \forall v \in S_0^{h_{x_1},h_{x_2},h_{x_3}}(\Omega). \tag{4.2}$$

Recall that the standard trilinear finite element scheme on Ω is: Find $u_{h_{x_1},h_{x_2},h_{x_3}} \in S_0^{h_{x_1},h_{x_2},h_{x_3}}(\Omega)$ such that

$$\mathcal{E}(u_{h_{x_1},h_{x_2},h_{x_3}}, v) = \langle f, v\rangle \ \forall v \in S_0^{h_{x_1},h_{x_2},h_{x_3}}(\Omega). \tag{4.3}$$

Following [35, 36], we may define a two-scale finite element combination approximation $u^h_{H,H,H}$ by

$$u^h_{H,H,H} = u_{h,H,H} + u_{H,h,H} + u_{H,H,h} - 2u_{H,H,H}. \tag{4.4}$$

Theorem 4.3. *If $u \in H_0^1(\Omega) \cap W_2^{G,3}(\Omega)$, then*

$$\|u_{h,h,h} - u^h_{H,H,H}\|_{1,\Omega} \lesssim H^2\|u\|_{W_2^{G,3}(\Omega)}, \tag{4.5}$$

$$\|u_{h,h,h} - u^h_{H,H,H}\|_{0,\Omega} \lesssim H^3\|u\|_{W_2^{G,3}(\Omega)}. \tag{4.6}$$

Consequently,

$$\|u - u^h_{H,H,H}\|_{1,\Omega} \lesssim (h + H^2)\|u\|_{W_2^{G,3}(\Omega)}, \tag{4.7}$$

$$\|u - u^h_{H,H,H}\|_{0,\Omega} \lesssim (h^2 + H^3)\|u\|_{W_2^{G,3}(\Omega)}. \tag{4.8}$$

Proof. By the triangle inequality, we have

$$\begin{aligned} & \|u_{h,h,h} - u^h_{H,H,H}\|_{1,\Omega} \\ \lesssim\ & \|u_{h,H,H} - I_{h,H,H}u\|_{1,\Omega} + \|u_{H,h,H} - I_{H,h,H}u\|_{1,\Omega} \\ & + \|u_{H,H,h} - I_{H,H,h}u\|_{1,\Omega} + 2\|u_{H,H,H} - I_{H,H,H}u\|_{1,\Omega} + \|u_{h,h,h} - I_{h,h,h}u\|_{1,\Omega} \\ & + \|I_{h,H,H}u + I_{H,h,H}u + I_{H,H,h}u - 2I_{H,H,H}u - I_{h,h,h}u\|_{1,\Omega}. \end{aligned} \tag{4.9}$$

Next, we want to estimate $\|u_{h,H,H} - I_{h,H,H}u\|_{1,\Omega}$. For all $v \in S_0^{h,H,H}(\Omega)$, there holds

$$\mathcal{E}(u_{h,H,H} - I_{h,H,H}u, v) = \mathcal{E}(u - I_{h,H,H}u, v) = D(u - I_{h,H,H}u, v) + J(u - I_{h,H,H}u, v).$$

Using Lemmas 4.2 and 2.8, we then obtain

$$\begin{aligned} \mathcal{E}(u_{h,H,H} - I_{h,H,H}u, v) &\lesssim H^2\|u\|_{W_2^{G,3}(\Omega)}\|v\|_{1,\Omega} + \|u - I_{h,H,H}u\|_{0,\Omega}\|v\|_{1,\Omega} \\ &\lesssim H^2\|u\|_{W_2^{G,3}(\Omega)}\|v\|_{1,\Omega}. \end{aligned}$$

Choose $v := u_{h,H,H} - I_{h,H,H}u$ and apply the coercivity (2.14), we get

$$\|u_{h,H,H} - I_{h,H,H}u\|_{1,\Omega} \lesssim H^2\|u\|_{W_2^{G,3}(\Omega)}, \tag{4.10}$$

which together with Lemma 4.1 and similar estimations of the other terms in (4.9) yields (4.5).

For the L^2-norm error estimate, we use a duality argument. Let $w \in H_0^1(\Omega)$ such that

$$\mathcal{E}(w,\phi) = \langle u_{h,h,h} - u^h_{H,H,H}, \phi\rangle \quad \forall \phi \in H_0^1(\Omega).$$

Let $\phi := u_{h,h,h} - u^h_{H,H,H} \in H_0^1(\Omega)$, then we have

$$\begin{aligned}
&\langle u_{h,h,h} - u^h_{H,H,H}, u_{h,h,h} - u^h_{H,H,H}\rangle \\
&= \mathcal{E}(w, u_{h,h,h} - u^h_{H,H,H}) = \mathcal{E}(u_{h,h,h} - u^h_{H,H,H}, w - I_{H,H,H}w) \\
&\lesssim \|u_{h,h,h} - u^h_{H,H,H}\|_{1,\Omega}\|w - I_{H,H,H}w\|_{1,\Omega} \\
&\lesssim H\|u_{h,h,h} - u^h_{H,H,H}\|_{1,\Omega}|w|_{2,\Omega}.
\end{aligned}$$

By (2.18), there holds

$$|w|_{2,\Omega} \lesssim \|u_{h,h,h} - u^h_{H,H,H}\|_{0,\Omega},$$

so we arrive at

$$\|u_{h,h,h} - u^h_{H,H,H}\|_{0,\Omega} \lesssim H\|u_{h,h,h} - u^h_{H,H,H}\|_{1,\Omega} \lesssim H^3\|u\|_{W_2^{G,3}(\Omega)}.$$

By the triangle inequality, we complete the proof. □

It is concluded from Theorem 4.3 that the two-scale finite element combination approximation $u^h_{H,H,H}$ is a much more efficient approximate solution in terms of computational cost as compared to $u_{h,h,h}$. In fact, with the same approximate accuracy, the degrees of freedom for getting $u^h_{H,H,H}$ is only of $O(h^{-2})$ when $H = O(h^{1/2})$ is chosen while that for the standard finite element solution $u_{h,h,h}$ is of $O(h^{-3})$. In addition, it may be very important that the two-scale finite element combination approximation $u^h_{H,H,H}$ can be carried out in parallel. As a result, both the computational time and the storage can be reduced.

It is seen from Theorem 2.6, Lemma 2.7 and [22, Theorem 3.5] that $D(\cdot,\cdot)$ can be interpreted as the high frequency part while $J(\cdot,\cdot)$ is the low frequency part of $\mathcal{E}(\cdot,\cdot)$. We therefore propose a refined two-scale finite element combination algorithm as follows.

Algorithm 4.4.

1. *Solve* (3.1) *on a coarse grid: Find* $u_{H,H,H} \in S_0^{H,H,H}(\Omega)$ *such that*

$$\mathcal{E}(u_{H,H,H}, v) = \langle f, v\rangle \quad \forall v \in S_0^{H,H,H}(\Omega).$$

2. *Compute linear boundary value problems on partially fine grids in parallel:*
   Find $e_{h,H,H} \in S_0^{h,H,H}(\Omega)$ *such that*

$$D(e_{h,H,H}, v) = \langle f, v\rangle - \mathcal{E}(u_{H,H,H}, v) \quad \forall v \in S_0^{h,H,H}(\Omega);$$

   Find $e_{H,h,H} \in S_0^{H,h,H}(\Omega)$ *such that*

$$D(e_{H,h,H}, v) = \langle f, v\rangle - \mathcal{E}(u_{H,H,H}, v) \quad \forall v \in S_0^{H,h,H}(\Omega);$$

   Find $e_{H,H,h} \in S_0^{H,H,h}(\Omega)$ *such that*

$$D(e_{H,H,h}, v) = \langle f, v\rangle - \mathcal{E}(u_{H,H,H}, v) \quad \forall v \in S_0^{H,H,h}(\Omega).$$

3. Set

$$\tilde{u}^h_{H,H,H} = u_{H,H,H} + e_{h,H,H} + e_{H,h,H} + e_{H,H,h} \quad \text{in } \Omega.$$

4. Find a further coarse grid correction $\hat{e}_{H,H,H} \in S_0^{H,H,H}(\Omega)$ *such that*

$$\mathcal{E}(\hat{e}_{H,H,H}, v) = \langle f, v\rangle - \mathcal{E}(\tilde{u}^h_{H,H,H}, v) \quad \forall v \in S_0^{H,H,H}(\Omega).$$

5. Set $\hat{u}^h_{H,H,H} = \tilde{u}^h_{H,H,H} + \hat{e}_{H,H,H}$ *in* Ω.

Theorem 4.5. *Assume that* $\tilde{u}^h_{H,H,H}$ *and* $\hat{u}^h_{H,H,H}$ *are obtained by Algorithm 4.4, then*

$$\|u_{h,h,h} - \tilde{u}^h_{H,H,H}\|_{1,\Omega} \lesssim H^2 \|u\|_{W_2^{G,3}(\Omega)}, \tag{4.11}$$

$$\|u_{h,h,h} - \hat{u}^h_{H,H,H}\|_{0,\Omega} \lesssim H^3 \|u\|_{W_2^{G,3}(\Omega)}. \tag{4.12}$$

Consequently,

$$\|u - \tilde{u}^h_{H,H,H}\|_{1,\Omega} \lesssim (h + H^2)\|u\|_{W_2^{G,3}(\Omega)}, \tag{4.13}$$

$$\|u - \hat{u}^h_{H,H,H}\|_{0,\Omega} \lesssim (h^2 + H^3)\|u\|_{W_2^{G,3}(\Omega)}. \tag{4.14}$$

Proof. Set $u^{h,H,H} = u_{H,H,H} + e_{h,H,H}$, $u^{H,h,H} = u_{H,H,H} + e_{H,h,H}$ and $u^{H,H,h} = u_{H,H,H} + e_{H,H,h}$, then from the definition, we obtain $\tilde{u}^h_{H,H,H} = u^{h,H,H} + u^{H,h,H} + u^{H,H,h} - 2u_{H,H,H}$. Hence,

$$\begin{aligned}
&\|u_{h,h,h} - \tilde{u}^h_{H,H,H}\|_{1,\Omega} \\
&\lesssim \|u^{h,H,H} + u^{H,h,H} + u^{H,H,h} - u_{h,H,H} - u_{H,h,H} - u_{H,H,h}\|_{1,\Omega} \\
&\quad + \|u_{h,H,H} + u_{H,h,H} + u_{H,H,h} - 2u_{H,H,H} - u_{h,h,h}\|_{1,\Omega} \\
&\lesssim \|u^{h,H,H} - u_{h,H,H}\|_{1,\Omega} + \|u^{H,h,H} - u_{H,h,H}\|_{1,\Omega} + \|u^{H,H,h} - u_{H,H,h}\|_{1,\Omega} \\
&\quad + \|u_{h,H,H} + u_{H,h,H} + u_{H,H,h} - 2u_{H,H,H} - u_{h,h,h}\|_{1,\Omega}.
\end{aligned}$$

Note that Theorem 3.5 implies

$$\|u^{h,H,H} - u_{h,H,H}\|_{1,\Omega} + \|u^{H,h,H} - u_{H,h,H}\|_{1,\Omega} + \|u^{H,H,h} - u_{H,H,h}\|_{1,\Omega} \lesssim H^2\|u\|_{2,\Omega}. \tag{4.15}$$

Therefore, combining (4.15) and Theorem 4.3, we get (4.11) and then (4.13). There holds

$$\|u_{h,h,h} - \hat{u}^h_{H,H,H}\|_{0,\Omega} = \|(I - P_{H,H,H})(u_{h,h,h} - \tilde{u}^h_{H,H,H})\|_{0,\Omega} \lesssim H\|u_{h,h,h} - \tilde{u}^h_{H,H,H}\|_{1,\Omega}.$$

This completes the proof. □

Remark 4.6. *We may also develop some local and parallel algorithms for the combination based two-scale finite element method (c.f. [36]).*

Remark 4.7. *The combination based two-scale discretization approach can be generalized to any dimensions. For* $\Omega = [0,1]^d$, $d \geq 3$, *recall that the standard Galerkin projection* $P_{\mathbf{h}} : H_0^1(\Omega) \mapsto S_0^{\mathbf{h}}(\Omega)$ *is defined by*

$$\mathcal{E}(u - P_{\mathbf{h}}u, v) = 0 \ \forall v \in S_0^{\mathbf{h}}(\Omega) \tag{4.16}$$

for $\mathbf{h} = (h_1, h_2, \ldots, h_d)$. *Then we can construct the two-scale finite element Galerkin projection as follows:*

$$B^h_{H\mathbf{e}} P_{h\mathbf{e}} u = \sum_{i=1}^{d} P_{H\hat{\mathbf{e}}_i + h\mathbf{e}_i} u - (d-1) P_{H\mathbf{e}} u,$$

where $\mathbf{e} = (1, \ldots, 1) \in \mathbb{R}^d, \hat{\mathbf{e}}_i = \mathbf{e} - \mathbf{e}_i$, $\mathbf{e}_i = (0, \ldots, 0, 1, 0, \ldots, 0) \in \mathbb{R}^d$ *whose i-th component is one and zero otherwise, and* $\mathbf{h}\alpha = (h_1\alpha_1, \ldots, h_d\alpha_d)$ *for* $\alpha_i \in \{0, 1\}$, $i \in \{1, 2, \ldots, d\}$. *For instance,* $B^h_{H,H,H} P_{h,h,h} u = P_{h,H,H} u + P_{H,h,H} u + P_{H,H,h} u - 2P_{H,H,H} u$. *Following [23], we can expect similar results for d−dimensions. For instance, if* $u \in H^1_0(\Omega) \cap W^{G,3}_2(\Omega)$, *then*

$$\|B^h_{H\mathbf{e}} P_{h\mathbf{e}} u - P_{h\mathbf{e}} u\|_{0,\Omega} + H \|B^h_{H\mathbf{e}} P_{h\mathbf{e}} u - P_{h\mathbf{e}} u\|_{1,\Omega} \lesssim H^3 \|u\|_{W^{G,3}_2(\Omega)}. \tag{4.17}$$

5 Wavelet-Lagrangian finite element implementation

Recall that the algorithms of Sections 3 & 4 do not depend on any particular choice of basis in $S^h_0(\Omega)$ or $S^{\mathbf{h}}_0(\Omega)$, respectively. Since Lagrangian finite element functions as well as certain piecewise polynomial wavelet basis functions generate $S^h_0(\Omega)$, in this section we briefly illustrate how one may use a wavelet basis to discretize the non-local bilinear form $\mathcal{E}(\cdot,\cdot)$ on the coarse mesh and employ classical Lagrangian functions for the parallel discretization of $D(\cdot,\cdot)$ on the fine meshes in Algorithms 3.4, 3.10 and 4.4. As already indicated in the introduction, the basic idea behind this approach is the following: First, wavelet discretization yields an almost sparse representation of the non-local form $J(\cdot,\cdot)$ defined by (2.12), see [14, 51, 44]. Second, as illustrated in the above Sections, the local form $D(\cdot,\cdot)$ can be discretized very efficiently with significantly less computational overhead using plain Lagrangian basis functions.

We need to introduce some notation: Assume that any meshwidth under consideration can be represented by a negative power of two and we therefore can associate a *level index* $j > 0$ to each meshwidth $h = 2^{-j}$. Denoting $h_0 = 2^0$, $h_1 = 2^{-1}$, $h_2 = 2^{-2}$, ..., one obtains that the spaces

$$S^{h_0}_0(\Omega) \subset S^{h_1}_0(\Omega) \subset S^{h_2}_0(\Omega) \subset \ldots L^2(\Omega),$$

define a multiresolution in the sense of [10, 13]. The spaces $S^{h_j}_0(\Omega) = \operatorname{span}(\Phi_j)$, $j = 0, 1, 2, \ldots$, are spanned by *single scale bases* $\Phi_j = \{\phi_{j,k} : k \in \Delta_j\}$ consisting of the Lagrangian finite element (or "nodal") basis functions $\phi_{j,k}$. Here Δ_j denotes a suitable index set of cardinality $\dim(S^{h_j}_0(\Omega))$.

Using the methodology of e.g. [15, 41], to the collections Φ_j one can associate a set of *dual* bases $\widetilde{\Phi}_j = \{\widetilde{\phi}_{j,k} : k \in \Delta_j\}$, i.e. one has $\langle \phi_{j,k}, \widetilde{\phi}_{j,k'} \rangle = \delta_{k,k'}$, $k, k' \in \Delta_j$. With $\nabla_j = \Delta_{j+1} \backslash \Delta_j$, for these single-scale bases one can then construct biorthogonal *complement* or *wavelet bases* $\Psi_j = \{\psi_{j,k} : k \in \nabla_j\}$, $\widetilde{\Psi}_j = \{\widetilde{\psi}_{j,k} : k \in \nabla_j\}$, i.e. $\langle \psi_{j,k}, \widetilde{\psi}_{j',k'} \rangle = \delta_{(j,k),(j',k')}$.

Remark 5.1. *As illustrated in [10, Section 2.11] and [15, 41], one can construct several different wavelet bases for* $S^{h_j}_0(\Omega)$, $j \geq 0$, *depending on which properties are desired. For example, one may obtain wavelets with an arbitrarily large number of vanishing moments, which is very desireable for the compression of non-local operators (cf. e.g. [14, 44]). Increasing the number of vanishing moments however expands the wavelets' supports. Here, in dimension* $d = 1$, *we give an explicit example of a piecewise linear wavelet basis with two vanishing moments which has turned out to be very useful in practice (multivariate wavelets on* $[0, 1]^d$ *can be obtained as suitable tensor products of these):*

The wavelets comprise of piecewise linear continuous functions on $[0,1]$ *vanishing at the endpoints. The mesh for level* $j \geq 0$ *is defined by the nodes* $x_{j,k} := k2^{-(j+1)}$ *with* $k \in \Delta_j := \{0, \ldots, 2^{j+1}\}$. *There holds* $N_j := \dim S_0^{h_j} = 2^{j+1} - 1$ *and therefore* $M_j := \dim \Psi_j = \dim S_0^{h_j} - \dim S_0^{h_{j-1}} = 2^j$.

On level $j = 0$ *we have* $N_0 = M_0 = 1$ *and* $\psi_{0,1}$ *is defined as the piecewise linear function with value* $c_0 := \sqrt{3}/2 > 0$ *at* $x_{0,1} = \frac{1}{2}$ *and* 0 *at the endpoints* 0, 1. *This choice of* c_0 *ensures the* L^2 *normalization of the wavelets.*

For $j > 0$ *we firstly define* $c_j := c_0 2^{j/2}$. *Then the boundary wavelet* $\psi_{j,0}$ *is defined as the piecewise linear function such that* $\psi_{j,0}(x_{j,1}) = 2c_j$, $\psi_{j,0}(x_{j,2}) = -c_j$ *and* $\psi_{j,0}(x_{j,s}) = 0$ *for all other* $s \neq 1, 2$. *Similarly, the boundary wavelet* ψ_{j,M_j-1} *takes the values* $\psi_{j,M_j-1}(x_{j,N_j}) = 2c_j$, $\psi_{j,M_j-1}(x_{j,N_j-1}) = -c_j$ *and zero at all other nodes. For the remaining location indices* $0 < k < M_j - 1$ *the wavelet* $\psi_{j,k}$ *is defined by* $\psi_{j,k}(x_{j,2k}) = -c_j$, $\psi_{j,k}(x_{j,2k+1}) = 2c_j$, $\psi_{j,k}(x_{j,2k+2}) = -c_j$ *and* $\psi_{j,k}(x_{j,s}) = 0$ *for all other* $s \neq 2k$, $2k+1$, $2k+2$.

Since the corresponding dual wavelet bases $\widetilde{\Psi}_j$, $j \geq 0$, *are solely of analytic importance and do not have to be computed in practice, for sake of brevity we refer to [10] for their illustration.*

Denoting by W^j the span of Ψ_j there holds

$$S_0^{h_{j+1}}(\Omega) = W^j \oplus S_0^{h_j}(\Omega), \quad j > 0. \tag{5.1}$$

Thus, for any $j > 0$, the finite element space $S_0^{h_j}(\Omega)$ can be written as a direct sum of the wavelet spaces $W^{j'}$, $j' < j$, (using the convention $W^0 := S_0^{h_0}(\Omega)$). Figure 3 shows the decomposition of the finite element space $S_0^{h_j}$, $j = 4$, spanned by continuous, piecewise linear ("nodal") Lagrangian basis functions $\phi_{j,k}$ into its increment spaces $W^{j'}$, $j' = 0, \ldots, 3$, spanned by the wavelets defined in Remark 5.1.

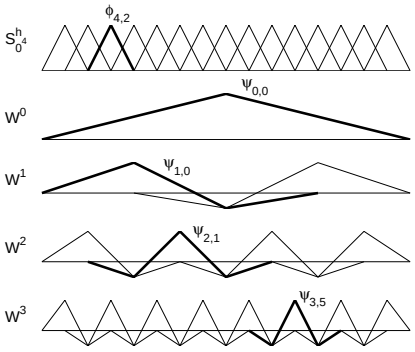

Figure 3: Schematic of single-scale space $S_0^{h_j}$ and its decomposition into multiscale wavelet spaces $W^{j'}$

By (5.1), for any $u_j \in S_0^{h_j}(\Omega)$ one has two equivalent representations

$$u_j = \sum_{j'=0}^{j-1} \sum_{k' \in \nabla_{j'}} d_{j',k'} \psi_{j',k'} = \sum_{k \in \Delta_j} c_k \phi_{j,k}. \tag{5.2}$$

The corresponding arrays of single-scale, respectively wavelet coefficients $\mathbf{c}$, $\mathbf{d}$ are interrelated by the explicitly known *multiscale transformation* $\mathbf{T}_j : \mathbf{d} \mapsto \mathbf{c}$, c.f. e.g. [13]. Based on the constructions in [15, 41], one readily infers that $\bigcup_{j'<j} \Psi_{j'}$, $j > 0$, forms a Riesz-basis in $L^2(\Omega)$, i.e. there holds

$$\|u_j\|^2 \sim \sum_{j'=0}^{j-1} \sum_{k' \in \nabla_{j'}} d_{j',k'}^2 \quad \forall u_j \in S_0^{h_j}(\Omega).$$

Thus, by [7, 12], there holds

$$\|\mathbf{T}_j\|, \|\mathbf{T}_j^{-1}\| = \mathcal{O}(1),$$

which is crucial for the realizations of Algorithms 3.4, 3.10 and 4.4 that we describe in the following:

In order to exploit the advantages of the wavelet basis when dealing with the non-local bilinear form $\mathcal{E}(\cdot,\cdot)$ as well as the efficient parallelization of the discretization of the local form $D(\cdot,\cdot)$ using classical Lagrangian basis functions, Algorithms 3.4, 3.10 and 4.4 can be realized by

1. For the coarse grid discretization of the form

   $$\text{Find } u_H \in S_0^H(\Omega) \text{ such that } \mathcal{E}(u_H, v) = \langle f, v\rangle \quad \forall v \in S_0^H(\Omega),$$

   employ the sparse tensor product wavelet methods of [22] with additional wavelet compression as in [39, 44] to efficiently obtain the wavelet representation

   $$u_H = \sum_{j'=0}^{J} \sum_{k' \in \nabla_{j'}} d_{j',k'}^H \psi_{j',k'},$$

   where $H = 2^{-J}$.

2. Employ the multiscale transformation $\mathbf{T}_j \mathbf{d}^H$ to obtain the corresponding single-scale representation of u_H in terms of Lagrangian basis functions with coefficient vector $\mathbf{c}^H$ as in (5.2).

3. Proceed from Step *2* in Algorithms 3.4, 3.10 and 4.4 using existing methodology of e.g. [23, 35, 36] to efficiently discretize the local form $D(\cdot,\cdot)$ and obtain the final numerical solution.

The main reason for combining wavelet methods with classical finite element methods in the above algorithms is the non-locality of the form $\mathcal{E}(\cdot,\cdot)$ due to the existence of jumps in the underlying stochastic process X. Even on the coarse grid, standard finite element schemes are of complexity $\mathcal{O}(H^{-2d})$ and therefore hard to apply even in moderate dimensions. It is known however that sparse tensor product wavelets yield a quasi-sparse representation even of non-local bilinear forms resulting in asymptotically optimal complexity $\mathcal{O}(H^{-1}|\log H|^{2(d-1)})$, see [22, 44]. For the fine-grid discretization of $D(\cdot,\cdot)$ one may still employ the same wavelet methods, but, since the form is local, for moderate dimensions one obtains an efficient discretization with significantly less computational overhead by using the algorithms of Sections 3 & 4 with Lagrangian finite element functions.

Remark 5.2. *Numerical results for the presented methods will appear in a separate paper.*

Acknowledgments. The authors wish to thank Professor C. Schwab of ETH Zürich for valuable comments and discussions.

References

[1] R.E. Bank. Hierarchical bases and the finite element method. In *Acta Numer.*, volume 5, pages 1–43. Cambridge University Press, Cambridge, 1996.

[2] O.E. Barndorff-Nielsen. Normal inverse Gaussian distributions and stochastic volatility modelling. *Scand. J. Statist.*, 24(1):1–13, 1997.

[3] M. Briani, R. Natalini, and G. Russo. Implicit-explicit numerical schemes for jump–diffusion processes. *Calcolo*, 44:33–57, 2007.

[4] H.-J. Bungartz and M. Griebel. A note on the complexity of solving Poisson's equation for spaces of bounded mixed derivative. *J. Complexity*, 15:167–199, 1999.

[5] H.-J. Bungartz and M. Griebel. Sparse grids. *Acta. Numer.*, 13:1–123, 2004.

[6] H.-J. Bungartz, M. Griebel, and U. Rüde. Extrapolation, combination, and sparse grid techniques for elliptic boundary value problems. *Comput. Methods Appl. Mech. Engrg.*, 116:243–252, 1994.

[7] J.M. Carnicer, W. Dahmen, and J.M. Peña. Local decomposition of refinable spaces and wavelets. *Appl. Comput. Harmon. Anal.*, 3(2):127–153, 1996.

[8] P. Carr, H. Geman, D.B. Madan, and M. Yor. The fine structure of assets returns: An empirical investigation. *J. Business*, 75(2):305–332, 2002.

[9] P.G. Ciarlet. Basic error estimates for elliptic problems. In *Handbook of numerical analysis, Vol. II*, pages 17–351. North-Holland, Amsterdam, 1991.

[10] A. Cohen. *Numerical Analysis of Wavelet Methods*, volume 32 of *Studies in Mathematics and its Applications*. North-Holland Publishing Co., Amsterdam, 2003.

[11] R. Cont and E. Voltchkova. A finite difference scheme for option pricing in jump diffusion and exponential Lévy models. *SIAM J. Numer. Anal.*, 43(4):1596–1626, 2005.

[12] W. Dahmen. Stability of multiscale transformations. *J. Fourier Anal. Appl.*, 2(4):341–361, 1996.

[13] W. Dahmen. Multiscale and wavelet methods for operator equations. In *Multiscale problems and methods in numerical simulations*, volume 1825 of *Lecture Notes in Math.*, pages 31–96. Springer, Berlin, 2003.

[14] W. Dahmen, H. Harbrecht, and R. Schneider. Compression techniques for boundary integral equations—asymptotically optimal complexity estimates. *SIAM J. Numer. Anal.*, 43(6):2251–2271 (electronic), 2006.

[15] W. Dahmen, A. Kunoth, and K. Urban. Biorthogonal spline wavelets on the interval - stability and moment conditions. *Appl. Comput. Harm. Anal.*, 6:259–302, 1999.

[16] W. Dahmen, S. Prößdorf, and R. Schneider. Wavelet approximation methods for periodic pseudodifferential equations. Part II - fast solution and matrix compression. *Adv. Comput. Math.*, 1:259–335, 1993.

[17] C.N. Dawson and M.F. Wheeler. Two-grid methods for mixed finite element approximations of nonlinear parabolic equations. *Contemp. Math.*, 180:191–203, 1994.

[18] F. Delbaen, P. Grandits, T. Rheinländer, D. Samperi, M. Schweizer, and C. Stricker. Exponential hedging and entropic penalties. *Math. Finance*, 12:99–123, 2002.

[19] F. Delbaen and W. Schachermayer. A general version of the fundamental theorem of asset pricing. *Math. Ann.*, 300:463–520, 1994.

[20] F. Delbaen and W. Schachermayer. The variance-optimal martingale measure for continuous processes. *Bernoulli*, 2:81–105, 1996.

[21] F.J. Delvos. d-Variate Boolean interpolation. *J. Approx. Theory*, 34:99–114, 1982.

[22] W. Farkas, N. Reich, and C. Schwab. Anisotropic stable Lévy copula processes - analytical and numerical aspects. *Math. Model. Methods Appl. Sci.*, 17:1405–1443, 2007.

[23] X. Gao, F. Liu, and A. Zhou. Three-scale finite element discretization schemes for eigenvalue problems. submitted.

[24] J. Garcke and M. Griebel. On the computation of the eigenproblems of hydrogen and helium in strong magnetic and electric fields with the sparse grid combination technique. *J. Comput. Phys.*, 165:694–716, 2000.

[25] V. Girault and J.L. Lions. Two-grid finite element schemes for the transient Navier-Stokes problem. *M2AN Math. Model. Numer. Anal.*, 35:945–980, 2001.

[26] M. Griebel, P. Oswald, and T. Schiekofer. Sparse grids for boundary integral equations. *Numer. Math.*, 83:279–312, 1999.

[27] M. Griebel, M. Schneider, and C. Zenger. A combination technique for the solution of sparse grid problem. In *Iterative Methods in Linear Algebra (P. de Groen and P. Beauwens, eds)*, pages 263–281. IMACS, Elsevier, North Holland, 1992.

[28] W. Hackbusch. *Multigrid Methods and Applications*, volume 4 of *Springer Series in Computational Mathematics*. Springer-Verlag, Berlin, 1985.

[29] N. Hilber, A.-M. Matache, and C. Schwab. Sparse wavelet methods for option pricing under stochastic volatility. *J. Comput. Finance*, 8:1–42, 2005.

[30] J. Kallsen and P. Tankov. Characterization of dependence of multidimensional Lévy processes using Lévy copulas. *J. Multivariate Anal.*, 97:1551–1572, 2006.

[31] G. Kou. A jump diffusion model for option pricing. *Mange. Sci.*, 48:1086–1101, 2002.

[32] W. Layton and L. Tobiska. A two-level method with backtracking for the Navier-Stokes equations. *SIAM J. Numer. Anal.*, 35:2035–2054, 1998.

[33] Q. Lin. Some problems concerning approximate solutions of operator equations. *Acta Math. Sinica*, 22:219–230, 1979.

[34] Q. Lin, N. Yan, and A. Zhou. A sparse finite element method with high accuracy. I. *Numer. Math.*, 88:731–742, 2001.

[35] F. Liu and A. Zhou. Two-scale finite element discretizations for partial differential equations. *J. Comput. Math.*, 24(3):373–392, 2006.

[36] F. Liu and A. Zhou. Localizations and parallelizations for two-scale finite element discretizations. *Commun. Pure Appl. Anal.*, 6(3):757–773, 2007.

[37] F. Liu and A. Zhou. Two-scale Boolean Galerkin discretizations for Fredholm integral equations of the second kind. *SIAM J. Numer. Anal.*, 45:296–312, 2007.

[38] A.-M. Matache, C. Schwab, and T.P. Wihler. Fast numerical solution of parabolic integrodifferential equations with applications in finance. *SIAM J. Sci. Comput.*, 27(2):369–393, 2005.

[39] A.-M. Matache, T. von Petersdorff, and C. Schwab. Fast deterministic pricing of options on Lévy driven assets. *M2AN Math. Model. Numer. Anal.*, 38(1):37–71, 2004.

[40] M. Mu and J. Xu. A two-grid method of a mixed Stokes-Darcy model for coupling fluid flow with porous media flow. *SIAM J. Numer. Anal.*, 45:1801–1813, 2007.

[41] H. Nguyen and R. Stevenson. Finite-element wavelets on manifolds. *IMA J. Numer. Math.*, 23:149–173, 2003.

[42] C. Pflaum. Convergence of the combination technique for second-order elliptic differential equations. *SIAM J. Numer. Anal.*, 34(6):2431–2455, 1997.

[43] C. Pflaum and A. Zhou. Error analysis of the combination technique. *Numer. Math.*, 84:327–350, 1999.

[44] N. Reich. *Wavelet Compression of Anisotropic Integrodifferential Operators on Sparse Tensor Product Spaces*. PhD Thesis 17661, ETH Zürich, 2008.

[45] N. Reich, C. Schwab, and C. Winter. *On Kolmogorov Equations for Anisotropic Multivariate Lévy Processes*. Submitted. Research report 2008-3, Seminar for Applied Mathematics, ETH Zürich, 2008.

[46] C. Reisinger and G. Wittum. Efficient hierarchical approximation of high-dimensional option pricing problems. *SIAM J. Sci. Comput.*, 29(1):440–458, 2007.

[47] K.-I. Sato. *Lévy Processes and Infinitely Divisible Distributions*, volume 68 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 1999.

[48] W. Schoutens. *Lévy Processes in Finance*. John Wiley & Sons, Chichester, 2003.

[49] P. Tankov. Dependence structure of Lévy processes with applications to risk management. Rapport Interne No. 502, CMAPX École Polytechnique, Mars 2003.

[50] T. von Petersdorff, R. Schneider, and C. Schwab. Multiwavelets for second kind integral equations. *SIAM J. Numer. Anal.*, 34:2212–2227, 1997.

[51] T. von Petersdorff and C. Schwab. Wavelet discretizations of parabolic integro-differential equations. *SIAM J. Numer. Anal.*, 41:159–180, 2003.

[52] T. von Petersdorff and C. Schwab. Numerical solution of parabolic equations in high dimensions. *M2AN Math. Model. Numer. Anal.*, 38(1):93–127, 2004.

[53] J. Xu. Two-grid discretization techniques for linear and nonlinear PDEs. *SIAM J. Numer. Anal.*, 33(5):1759–1777, 1996.

[54] J. Xu and A. Zhou. Local and parallel finite element algorithms based on two-grid discretizations. *Math. Comput.*, 69:881–909, 2000.

[55] J. Xu and A. Zhou. Local and parallel finite element algorithms based on two-grid discretizations for nonlinear problems. *Adv. Comput. Math.*, 14:293–327, 2001.

[56] J. Xu and A. Zhou. Local and parallel finite element algorithms for eigenvalue problems. *Acta Math. Appl. Sinica (English series)*, 18:185–200, 2002.

[57] Y. Xu and A. Zhou. Fast Boolean approximation methods for solving integral equations in high dimensions. *J. Integral Equations Appl.*, 16:83–110, 2004.

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

[59] C. Zenger. Sparse grids. In *Parallel algorithms for partial differential equations (Kiel, 1990)*, volume 31 of *Notes Numer. Fluid Mech.*, pages 241–251. Vieweg, Braunschweig, 1991.