Sparse space-time Galerkin BEM for the nonstationary heat equation

Alexey Chernov[1,*] **and Christoph Schwab**[2,**]

[1] Hausdorff Center for Mathematics and Institute for Numerical Simulation, University of Bonn
Endenicher Allee 60, 53115 Bonn, Germany
[2] Seminar for Applied Mathematics, ETH Zentrum, HG G57.1, ETH Zürich, Rämistrasse 101, 8092 Zurich, Switzerland

Received XXXX, revised XXXX, accepted XXXX
Published online XXXX

Key words Sparse tensor product, Space-time discretization, Boundary Integral Equations, Boundary Element Methods

MSC (2000) Primary: 65N30 Secondary: 41A25, 65R20, 65N12, 65N38

Dedicated to W. L. Wendland on the occasion of this 75th anniversary

We construct and analyze sparse tensorized space-time Galerkin discretizations for boundary integral equations resulting from the boundary reduction of nonstationary diffusion equations with either Dirichlet or Neumann boundary conditions. The approach is based on biorthogonal multilevel subspace decompositions and a weighted sparse tensor product construction. We compare the convergence behavior of the proposed method to the standard full tensor product discretizations. In particular, we show for the problem of nonstationary heat conduction in a bounded two- or three-dimensional spatial domain that low order sparse space-time Galerkin schemes are competitive with high order full tensor product discretizations in terms of the asymptotic convergence rate of the Galerkin error in the energy norms, under lower regularity requirements on the solution.

1 Introduction

Numerical solution of parabolic evolution problems is required in numerous applications. Traditional numerical methods are based on Finite Element or Finite Difference discretization in physical space combined with a suitable time stepping approach in the time variable. Implicit time stepping methods (e.g. backward Euler or Crank-Nicholson) require solution of a d-dimensional elliptic problem in every time step, whereas the explicit forward Euler scheme results in the strong stability condition on the time steps $h_t \lesssim h_x^2$, where h_t is the size of the time interval and h_x is the characteristic mesh size in the spatial finite element mesh.

We refer to [13] and to the references therein for a survey of single-step, multi-step and Discontinuous Galerkin time stepping schemes for abstract, parabolic equations. All timestepping schemes described in [13] are *implicit* and of *finite order* in the timestep h_t. These timestepping schemes reduce the parabolic problem to a sequence of elliptic problems to be solved in the bounded spatial domain $D \subset \mathbb{R}^d$. If these problems have been solved by Finite Element Methods (FEM) on a quasiuniform meshes of meshwidth h_x, and if their solution is performed in linear complexity, this entails, per timestep, work of order $O(h_x^{-d})$, i.e. the total work for solution is of asymptotic order $O(h_t^{-1} h_x^{-d})$. Using the time-analyticity of the semigroup generated by the parabolic evolution problem, in [9] an hp-discretization with respect to the time variable with *exponential convergence* has been proposed; this approach, when combined with an optimal order multilevel solver in the spatial domain, will reduce the total work to order $O(h_x^{-d} |\log h_x|^2)$. Recently, in [10], a completely different approach based on a *space-time compressive, adaptive Galerkin discretization* of parabolic evolution problems on sparse tensor products of multilevel spaces in the spatial domain has been proposed and analyzed. It has been shown in [10] that this approach allows to reduce the total work to order $O(h_x^{-d})$ while retaining the convergence rates of the previous discretizations. In the present paper, we combine the idea of [10] with a *boundary reduction* of nonstationary heat equations with homogeneous volume source which consists of the prior reduction to the "mantle" Σ of the space-time cylinder as has been proposed earlier by [1, 3, 7] and recently, in the context of shape-sensitivity calculus, in [2]. Similarly to the elliptic case [8], this approach leads to first and second kind boundary integral equations involving integral operators on Σ. Moreover, from [3, 7] it is known, that the heat single layer operator and the hypersingular first kind boundary integral

* Corresponding author, e-mail: chernov@hcm.uni-bonn.de, Phone: +49 228 73 62246, Fax: +49 228 73 62251, Supported by the Hausdorff Center for Mathematics, Bonn
** Supported by the European Research Council under FP7 grant AdG247277.

operator for the heat equation are elliptic and continuous in appropriate anisotropic Sobolev spaces on Σ. Remarkably, this is *not* true for the domain heat operator. The coercivity of the boundary integral operators provides a basis for Galerkin discretizations of the first kind boundary integral equations by means of space-time Galerkin Boundary Element Methods (BEM), whose stability is ensured by the classical Lemma of Ceá.

Due to $\Sigma = (0,T) \times \partial D$, it is natural to use tensor product bases on Σ in Galerkin discretizations; full tensor product discretization schemes for the weakly singular integral equations were analyzed in [1, 3, 7]. Positivity for the hypersingular parabolic boundary integral operator of the first kind and the full Calderón projector for the heat equation was shown in [3, 7]. In both cases, for Finite Element spaces of fixed polynomial degree, the Galerkin error in the energy norm decays as $O\big((\#\mathrm{dof})^{-b/(d+1)}\big)$ where the constant $b > 0$ depends only on the polynomial degree and is, in particular, independent of the mesh size and of the spatial dimension d of the physical domain. For the physically relevant cases $d = 2$ and $d = 3$, this results in low convergence rates, when expressed in terms of the number of degrees of freedom. In this paper we introduce and analyze a new weighted sparse space-time Galerkin discretization spaces with the aim to improve this unfavorable convergence behavior. Sparse-tensor space-time Galerkin discretization of heat potentials has been introduced first in [2] for the case of the equal weight, and applied for numerical solution of the heat equation in a randomly perturbed domain. In the present paper we show, in particular, that anisotropic sparse tensor product bases can yield better convergence rates in terms of the number of degrees of freedom in the case $d = 3$. As a main result, we prove that the Galerkin error of the sparse space-time discretization in the energy norm decays asymptotically at least as $O\big((\#\mathrm{dof})^{-\tilde{b}/(d-1)}\big)$ up to logarithmic terms, where $\tilde{b} > 0$ is again a constant which is independent on the mesh size and of the space dimension $d \geq 2$ which is a significant improvement compared to the convergence of straightforward full tensor Galerkin discretizations of heat operator as were proposed in [3]. Using *sparse tensor space-time multilevel bases* for the Galerkin discretization of first kind boundary integral operators for the heat equation allows therefore to realize complexity gains of both, *the boundary reduction by integral equation formulations* and of *the sparse tensor, compressive space-time discretization* proposed and analyzed recently in [10] for space-time discretizations without boundary reduction. In particular, we show for the problem of nonstationary heat conduction in two- and three-dimensional spatial domain that low order sparse space-time Galerkin schemes are comparable with high order full tensor product discretizations in terms of the asymptotic convergence rate of the Galerkin error in the energy norm. We point out at this stage that the Galerkin stiffness matrix of the space-time boundary integral operators is densely populated in general, due to the nonlocal nature of these integral operators. Exponentially convergent matrix compressions based on Fast Multipole type methods for heat potentials, such as a (suitably modified) *fast Gauss transform* (see, e.g. [4, 11, 12]), have to be employed in addition to the Galerkin discretizations investigated here in order to obtain algorithms of complexity $O(h_x^{-(d-1)}|\log h_x|^a)$. The analysis of compression of the stiffness matrix is beyond the scope of the present paper, but can be accounted for straightforwardly by a Strang-type perturbation argument, using the coercivity of the boundary integral operators.

The paper is organized as follows. In Sect. 2 we introduce the nonstationary heat equation with Dirichlet and Neumann boundary conditions and set up the functional framework. In Sect. 3 we perform the boundary reduction of the volume formulations and, using the direct and indirect approach, we derive the associated first kind boundary integral equations whose solution is equivalent to the solution of the original nonstationary heat equations. In Sect. 4 we introduce sparse space-time Galerkin discretization schemes for these BIEs and obtain the *a priori* error bounds. In Sect. 5 we compare the complexity of the full and the presently proposed, new sparse space-time discretization schemes for the practically relevant cases $d = 2$ and $d = 3$. In the Appendix we collect a technical Lemma on inclusion of different scales of anisotropic Sobolev spaces.

Throughout, we use the notation $f \lesssim g$ if there exists a constant $C > 0$, independent on the parameters which f and g might depend on, such that $f \leq Cg$. The relation $f \sim g$ is equivalent to $f \lesssim g$ and $g \lesssim f$.

2 Weak solution of the nonstationary heat equation

Let $D \subset \mathbb{R}^d$, $d \geq 2$ denote a bounded domain with boundary $\Gamma := \partial D$ and exterior unit normal vector field n which we assume in this paper for simplicity to be smooth; most of our results hold for $\Gamma \in C^k$ for k being finite, but sufficiently large. With $T > 0$ we denote a fixed finite time horizon, $I = (0,T)$ the time interval of interest, by $Q = I \times D$ the space-time cylinder and by $\Sigma = I \times \partial D$ its "mantle". For some $t \in [0,T]$ we write $D_t := D \times \{t\}$ and observe $\partial Q = \overline{\Sigma} \cup D_0 \cup D_T$. In Q we consider a linear nonstationary heat equation with either Dirichlet or Neumann boundary conditions. The Neumann problem reads: given $f : Q \to \mathbb{R}$ and $h : \Sigma \to \mathbb{R}$, find $u : Q \to \mathbb{R}$ satisfying

$$\begin{aligned} (\partial_t - \Delta)u &= f \quad \text{in } Q, \\ \gamma_1 u &= h \quad \text{on } \Sigma, \\ u &= 0 \quad \text{in } D_0. \end{aligned} \tag{1}$$

Here γ_1 is the conormal derivative on Σ. For classical solutions $u \in C^2(\overline{Q})$, we have $\gamma_1 u = \partial_n u|_\Sigma$. With these notations, we can state the Dirichlet problem of the heat equation: given $f : Q \to \mathbb{R}$ and $g : \Sigma \to \mathbb{R}$, find $w : Q \to \mathbb{R}$ satisfying

$$\begin{aligned} (\partial_t - \Delta)w &= f \quad \text{in } Q, \\ \gamma_0 w &= g \quad \text{on } \Sigma, \\ w &= 0 \quad \text{in } D_0, \end{aligned} \tag{2}$$

where γ_0 is the trace operator, i.e. $\gamma_0 w = w|_\Sigma$. The question of well-posedness of (1), (2) has been addressed in the fundamental article by Costabel [3]. In order to state these results we introduce first the functional analytic framework. For $r, s \geq 0$ it is appropriate to work with the anisotropic Sobolev spaces

$$H^{r,s}(Q) := L^2(I; H^r(D)) \cap H^s(I; L^2(D)), \qquad H^{r,s}(\Sigma) := L^2(I; H^r(\Gamma)) \cap H^s(I; L^2(\Gamma)) \tag{3}$$

equipped with the graph norm, and their duals $H^{-r,-s} := (H^{r,s})'$. Note that they are well-defined for $(r, s) \in [-1, 1] \times \mathbb{R}$ if $\Gamma \in C^{0,1}$ and for all $r, s \in \mathbb{R}$ if $\Gamma \in C^\infty$. The variational treatment of essential boundary and initial conditions requires subspaces of $H^{r,s}(Q)$ with homogeneous boundary conditions

$$\begin{aligned} &\tilde{H}^{r,s}(Q) := \{u|_Q : u \in H^{r,s}(\mathbb{R} \times D), u(t, x) = 0 \text{ for } t < 0\}, \quad H_0^{r,s}(Q) := L^2(I; H_0^r(D)) \cap H^s(I; L^2(D)), \\ &\hat{H}_0^{r,s}(Q) := \{u|_Q : u \in H_0^{r,s}(\mathbb{R} \times D), u(t, x) = 0 \text{ for } t > T\}, \quad \tilde{H}^{-r,-s}(Q) := (\hat{H}_0^{r,s}(Q))', \quad r - \frac{1}{2} \notin \mathbb{Z} \end{aligned}$$

and analogously defined subspaces $\tilde{H}_0^{r,s}(Q)$, $\tilde{H}^{r,s}(\Sigma)$, etc. Note that the inclusions of the subspaces are strict for $r, s > 1/2$, but become an identity if $r, s < 1/2$.

Theorem 2.1 *[3, Lemma 2.21, see also (2.3)] For every $f \in L^2(Q)$ and $h \in L^2(I; H^{-\frac{1}{2}}(\Gamma))$ there exists a unique $u \in \tilde{H}^{1,\frac{1}{2}}(Q)$ satisfying* (1).

Theorem 2.2 *[3, Theorem 2.9] For every $f \in \tilde{H}^{-1,-\frac{1}{2}}(Q)$ and $g \in H^{\frac{1}{2},\frac{1}{4}}(\Sigma)$ there exists a unique $w \in \tilde{H}^{1,\frac{1}{2}}(Q)$ satisfying* (2).

3 Boundary reduction

In this section we reduce the Neumann and Dirichlet problems (1) and (2) to the "mantle" Σ of the space-time cylinder Q. In particular, we observe that the boundary integral operators for the heat equation can be bounded and positive in one and the same norm. Note that this property is not true for the heat operator itself. In this section we consider formulations (1), (2) with homogeneous volume source term $f = 0$. In our exposition, we follow [3], see also [1, 2, 7]. Let

$$G(t, x) := (4\pi|x|)^{-\frac{d}{2}} \exp\left(-\frac{|x|^2}{4t}\right) \vartheta(t), \qquad \vartheta(t) := \frac{1}{2}\big(1 + \operatorname{sign}(t)\big) \tag{4}$$

be the fundamental solution of the heat equation. The following theorem provides the representation formula for the solution of a homogeneous heat equation.

Theorem 3.1 *[3, Theorem 2.20] Suppose $u \in \tilde{H}^{1,\frac{1}{2}}(Q)$ satisfying $(\partial_t - \Delta)u = 0$ in Q. Then there holds the representation formula*

$$u = K_0(\gamma_1 u) - K_1(\gamma_0 u) \quad \text{in } Q, \tag{5}$$

where for $(t_0, x_0) \in Q$ the single layer heat potential K_0 *and the* double layer heat potential K_1 *are defined by*

$$\begin{aligned} K_0(\varphi)(t_0, x_0) &:= \int_\Sigma \varphi(t, x) G(t_0 - t, x_0 - x)\, d\sigma_x\, dt, \\ K_1(\psi)(t_0, x_0) &:= \int_\Sigma \psi(t, x) \gamma_{1,x} G(t_0 - t, x_0 - x)\, d\sigma_x\, dt. \end{aligned} \tag{6}$$

where $\gamma_{1,x}$ denotes the conormal derivative applied in the point x.

Similarly to the elliptic case, single and double layer potentials satisfy the jump relations on the "mantle" Σ of the time-space cylinder.

Theorem 3.2 *[3, Theorem 3.4] For all $\psi \in H^{-\frac{1}{2},-\frac{1}{4}}(\Sigma)$ and all $\varphi \in H^{\frac{1}{2},\frac{1}{4}}(\Sigma)$ there hold the jump relations*

$$[\gamma_0 K_0 \psi] = 0, \qquad [\gamma_1 K_0 \psi] = -\psi, \qquad [\gamma_0 K_1 \varphi] = \varphi, \qquad [\gamma_1 K_1 \varphi] = 0.$$

Theorem 3.2 implies that for every $\psi \in H^{-\frac{1}{2},-\frac{1}{4}}(\Sigma)$ and for every $\varphi \in H^{\frac{1}{2},\frac{1}{4}}(\Sigma)$, the single layer operator V, the hypersingular integral operator W, the double layer operator K and the related operator N

$$V\psi := \gamma_0 K_0 \psi, \qquad W\varphi := -\gamma_1 K_1 \varphi, \qquad K\varphi := \gamma_0 (K_1 \varphi)|_Q + \frac{1}{2}\varphi, \qquad N\psi := \gamma_1 (K_0 \psi)|_Q - \frac{1}{2}\psi$$

are well-defined maps. Furthermore, V and W are positive and define isomorphisms in anisotropic trace spaces:

Theorem 3.3 *[3, Corollary 3.13] The single layer operator $V : H^{-\frac{1}{2},-\frac{1}{4}}(\Sigma) \to H^{\frac{1}{2},\frac{1}{4}}(\Sigma)$ is an isomorphism, and*

$$\exists c_V > 0 : \qquad \langle \psi, V\psi \rangle \geq c_V \|\psi\|^2_{H^{-\frac{1}{2},-\frac{1}{4}}(\Sigma)} \qquad \forall \psi \in H^{-\frac{1}{2},-\frac{1}{4}}(\Sigma)\ . \tag{7}$$

The hypersingular integral operator $W : H^{\frac{1}{2},\frac{1}{4}}(\Sigma) \to H^{-\frac{1}{2},-\frac{1}{4}}(\Sigma)$ is an isomorphism, and

$$\exists c_W > 0 : \qquad \langle \varphi, W\varphi \rangle \geq c_W \|\varphi\|^2_{H^{\frac{1}{2},\frac{1}{4}}(\Sigma)} \qquad \forall \varphi \in H^{\frac{1}{2},\frac{1}{4}}(\Sigma). \tag{8}$$

The analysis of the corresponding *Calderón projector* implies in particular well-posedness of the direct (9), (11) and of the indirect (10), (12) boundary integral formulations for the heat equation.

Theorem 3.4 *The unique solution $w \in \tilde{H}^{1,\frac{1}{2}}(Q)$ of the Dirichlet problem* (2) *with $f = 0$ can be represented*

(a) *as $w = K_0\psi - K_1 g$, where $\psi \in H^{-\frac{1}{2},-\frac{1}{4}}(\Sigma)$ is the unique solution of the first kind integral equation*

$$V\psi = (\frac{1}{2}I + K)g. \tag{9}$$

Then $\psi = \gamma_1 w$ on Σ.

(b) *as $w = K_0\psi$, where $\psi \in H^{-\frac{1}{2},-\frac{1}{4}}(\Sigma)$ is the unique solution of the first kind integral equation*

$$V\psi = g. \tag{10}$$

Theorem 3.5 *The unique solution $u \in \tilde{H}^{1,\frac{1}{2}}(Q)$ of the Neumann problem* (1) *with $f = 0$ can be represented*

(a) *as $u = K_0 h - K_1\varphi$, where $\varphi \in H^{\frac{1}{2},\frac{1}{4}}(\Sigma)$ is the unique solution of the first kind integral equation*

$$W\varphi = (\frac{1}{2}I - N)h\ . \tag{11}$$

Then $\varphi = \gamma_0 u$ on Σ.

(b) *as $u = K_1\varphi$, where $\varphi \in H^{\frac{1}{2},\frac{1}{4}}(\Sigma)$ is the unique solution of the first kind integral equation*

$$W\varphi = -h\ . \tag{12}$$

Moreover, if Γ is sufficiently smooth, the integral operators in (9) – (12) are one-to-one mappings in the scale of spaces.

Proposition 3.6 *([3, Proposition 4.3]) Assume that $\Gamma \in C^\infty$. Then for any $s \geq 0$ the mappings*

$$\begin{array}{rccc} V: & \tilde{H}^{-\frac{1}{2}+s,(-\frac{1}{2}+s)/2}(\Sigma) & \to & \tilde{H}^{\frac{1}{2}+s,(\frac{1}{2}+s)/2}(\Sigma), \\ (\frac{1}{2}I + K), (\frac{1}{2}I - N): & \tilde{H}^{\frac{1}{2}+s,(\frac{1}{2}+s)/2}(\Sigma) & \to & \tilde{H}^{\frac{1}{2}+s,(\frac{1}{2}+s)/2}(\Sigma), \\ W: & \tilde{H}^{\frac{1}{2}+s,(\frac{1}{2}+s)/2}(\Sigma) & \to & \tilde{H}^{-\frac{1}{2}+s,(-\frac{1}{2}+s)/2}(\Sigma) \end{array} \tag{13}$$

are isomorphisms.

The positivity (7) and (8) of V and W provides the basis for analysis of Galerkin discretizations of (9) – (12).

4 Space-time Galerkin discretizations of the first kind integral equations

The coercivity of (7) and (8) of the first kind boundary integral operators implies stability and quasioptimality of conforming Galerkin approximations of (9) – (12) by the classical Ceá lemma. For any closed subspaces $\mathcal{X}_L \subset \mathcal{X} := H^{-\frac{1}{2},-\frac{1}{4}}(\Sigma)$ and $\mathcal{V}_L \subset \mathcal{V} := H^{\frac{1}{2},\frac{1}{4}}(\Sigma)$ the Galerkin equations: to find $\psi_L \in \mathcal{X}_L$ resp. $\varphi_L \in \mathcal{V}_L$ such that

$$\langle \eta_L, V\psi_L\rangle = \langle \eta_L, g\rangle \qquad \forall \eta_L \in \mathcal{X}_L, \qquad\qquad \langle \eta_L, V\psi_L\rangle = \langle \eta_L, (\frac{1}{2}I + K)g\rangle \qquad \forall \eta_L \in \mathcal{X}_L, \tag{14}$$

$$\langle \zeta_L, W\varphi_L\rangle = \langle \zeta_L, -h\rangle \qquad \forall \zeta_L \in \mathcal{V}_L, \qquad\qquad \langle \zeta_L, W\varphi_L\rangle = \langle \zeta_L, (\frac{1}{2}I - N)h\rangle \qquad \forall \zeta_L \in \mathcal{V}_L, \tag{15}$$

are uniquely solvable and the Galerkin solutions $\psi_L \in \mathcal{X}_L$, resp. $\varphi_L \in \mathcal{V}_L$ are quasioptimal:

$$\begin{aligned} \|\psi - \psi_L\|_{H^{-\frac{1}{2},-\frac{1}{4}}(\Sigma)} &\le \frac{\|V\|}{c_V} \inf_{\eta_L \in \mathcal{X}_L} \|\psi - \eta_L\|_{H^{-\frac{1}{2},-\frac{1}{4}}(\Sigma)}, \\ \|\varphi - \varphi_L\|_{H^{\frac{1}{2},\frac{1}{4}}(\Sigma)} &\le \frac{\|W\|}{c_W} \inf_{\zeta_L \in \mathcal{V}_L} \|\varphi - \zeta_L\|_{H^{\frac{1}{2},\frac{1}{4}}(\Sigma)}. \end{aligned} \tag{16}$$

The key ingredient in efficient Galerkin approximation is therefore the proper choice of the discrete spaces $\mathcal{V}_L$, $\mathcal{X}_L$. The Cartesian product structure of the domain $\Sigma = I \times \Gamma$ allows for a natural *tensor product* discretization of $\mathcal{V}, \mathcal{X}$ with piecewise polynomials. In particular, suppose

$$\begin{aligned} &\mathcal{X}_0^x \subset \mathcal{X}_1^x \subset \cdots \subset \mathcal{X}_{\ell_x}^x \subset \cdots \subset H^{-\frac{1}{2}}(\Gamma), \quad \mathcal{X}_0^t \subset \mathcal{X}_1^t \subset \cdots \subset \mathcal{X}_{\ell_t}^t \subset \cdots \subset H^{-\frac{1}{4}}(I), \\ &\mathcal{V}_0^x \subset \mathcal{V}_1^x \subset \cdots \subset \mathcal{V}_{\ell_x}^x \subset \cdots \subset H^{\frac{1}{2}}(\Gamma), \qquad \mathcal{V}_0^t \subset \mathcal{V}_1^t \subset \cdots \subset \mathcal{V}_{\ell_t}^t \subset \cdots \subset H^{\frac{1}{4}}(I) \end{aligned} \tag{17}$$

are the nested finite element spaces of piecewise polynomials of degree p_x, p_t, q_x, q_t, respectively, which are associated to a finite element mesh of refinement level ℓ_x, ℓ_t with the mesh width $h_x \sim 2^{-\ell_x}$, $h_t \sim 2^{-\ell_t}$. We remark that elements of $\{\mathcal{V}_{\ell_x}^x\}_{\ell_x \ge 0}$ are globally continuous piecewise polynomial functions, whereas the other discrete families may contain discontinuous functions. Galerkin discretizations based on *full tensor product* spaces have been considered in [1,3,7]:

$$\mathcal{V}_L := \mathcal{V}_{\ell_x}^x \otimes \mathcal{V}_{\ell_t}^t, \qquad \mathcal{X}_L := \mathcal{X}_{\ell_x}^x \otimes \mathcal{X}_{\ell_t}^t \tag{18}$$

for some fixed relation of the mesh widths (the optimal choice is $\ell_t = 2\ell_x$, i.e. $h_t \sim h_x^2$) and fixed, usually low polynomial degrees. Along the lines of [7, Theorem 7.5] and [3, Proposition 5.3, Corollary 5.5] we obtain

Theorem 4.1 *Suppose $\mathcal{X}_L = \mathcal{X}_{\ell_x}^x \otimes \mathcal{X}_{\ell_t}^t$ for some polynomial degrees $p_x, p_t \ge 0$. Then for the solution $\psi \in \tilde{H}^{p_x+1,p_t+1}(\Sigma)$ of the weakly singular integral equation* (9) *or* (10) *and the Galerkin solution $\psi_L \in \mathcal{X}_L$ of* (14) *holds*

$$\|\psi - \psi_L\|_{H^{-\frac{1}{2},-\frac{1}{4}}(\Sigma)} \lesssim (h_x^{\frac{1}{2}} + h_t^{\frac{1}{4}})(h_x^{p_x+1} + h_t^{p_t+1})\|\psi\|_{H^{p_x+1,p_t+1}(\Sigma)}, \tag{19}$$

where the total number of unknowns is $N_L := \dim(\mathcal{X}_L) \sim h_x^{-(d-1)} h_t^{-1}$.

Corollary 4.2 *Suppose in addition to the assumptions of Theorem 4.1 that $g \in \tilde{H}^{2p_t+3,p_t+\frac{3}{2}}(\Sigma)$,*

$$h_t \sim h_x^2 \quad \text{and} \quad p_x \le 2p_t + 1. \tag{20}$$

Then $N_L \sim h_x^{-(d+1)}$ and there holds the bound

$$\begin{aligned} \|\psi - \psi_L\|_{H^{-\frac{1}{2},-\frac{1}{4}}(\Sigma)} &\lesssim h_x^{p_x+\frac{3}{2}} \|\psi\|_{H^{p_x+1,p_t+1}(\Sigma)} \\ &\lesssim N_L^{-\frac{p_x+3/2}{d+1}} \|\psi\|_{H^{2(p_t+1),(p_t+1)}(\Sigma)}, \\ &\sim N_L^{-\frac{p_x+3/2}{d+1}} \|g\|_{H^{2p_t+3,p_t+\frac{3}{2}}(\Sigma)}\,. \end{aligned} \tag{21}$$

P r o o f. The first line of (21) follows directly from (19) and from the relations (20) yielding in particular $h_x^{p_x+1} \gtrsim h_t^{p_t+1}$. The second line of (21) follows from the inclusion $H^{2(p_t+1),(p_t+1)}(\Sigma) \subset H^{p_x+1,p_t+1}(\Sigma)$ and the third line from the mapping properties of V and $(\frac{1}{2}I + K)$ from Proposition 3.6 with $s = 2p_t + \frac{5}{2}$. □

Theorem 4.3 *Suppose $\mathcal{V}_L = \mathcal{V}^x_{\ell_x} \otimes \mathcal{V}^t_{\ell_t}$ for some fixed polynomial degrees $q_x \geq 1$, $q_t \geq 0$. Then for the solution $\varphi \in \tilde{H}^{q_x+1,q_t+1}(\Sigma)$ of the hypersingular integral equation* (11) *or* (12) *and the Galerkin solution $\varphi_L \in \mathcal{V}_L$ of* (15) *holds*

$$\|\varphi - \varphi_L\|_{H^{\frac{1}{2},\frac{1}{4}}(\Sigma)} \lesssim (h_x^\alpha + h_t^\beta)\|\varphi\|_{H^{q_x+1,q_t+1}(\Sigma)} \tag{22}$$

where $\alpha = \min\{q_x + \frac{1}{2}, q_x + 1 - \frac{\mu}{4}\}$, $\beta = \min\{q_t + \frac{3}{4}, q_t + 1 - \frac{1}{2\mu}\}$, $\mu = \frac{q_x+1}{q_t+1}$ and $N_L := \dim(\mathcal{V}_L) \sim h_x^{-(d-1)} h_t^{-1}$.

Corollary 4.4 *Suppose in addition to the assumptions of Theorem 4.3 that $h \in \tilde{H}^{2q_t+1,q_t+\frac{1}{2}}(\Sigma)$,*

$$h_t \sim h_x^2 \quad \text{and} \quad q_x \leq 2q_t + 1. \tag{23}$$

Then $N_L \sim h_x^{-(d+1)}$ and we have the bound

$$\begin{aligned} \|\varphi - \varphi_L\|_{H^{\frac{1}{2},\frac{1}{4}}(\Sigma)} &\lesssim h_x^{q_x+\frac{1}{2}} \|\varphi\|_{H^{q_x+1,q_t+1}(\Sigma)} \\ &\lesssim N_L^{-\frac{q_x+1/2}{d+1}} \|\varphi\|_{H^{2(q_t+1),(q_t+1)}(\Sigma)} \\ &\sim N_L^{-\frac{q_x+1/2}{d+1}} \|h\|_{H^{2q_t+1,q_t+\frac{1}{2}}(\Sigma)}. \end{aligned} \tag{24}$$

Proof. By assumption, $\mu \leq 2$. Hence

$$2\beta \geq \mu\beta = \mu(q_t + 1) - \frac{1}{2} = q_x + \frac{1}{2} = \alpha$$

and the first line of (24) follows. The second line follows from the inclusion $H^{2(q_t+1),(q_t+1)}(\Sigma) \subset H^{q_x+1,q_t+1}(\Sigma)$ and the third line from the mapping properties of W and $(\frac{1}{2}I - N)$ from Proposition 3.6 with $s = 2p_t + \frac{3}{2}$ and $s = 2p_t + \frac{1}{2}$. $\square$

From (21) and (24) it is clear that the convergence rate of the Galerkin error when measured in terms of the total number of degrees of freedom, $\#\mathrm{dof}$, scales unfavorably with the space dimension d: the Galerkin error decays as $O\big((\#\mathrm{dof})^{-b/(d+1)}\big)$ where the constant b depends on the (fixed) polynomial degree and is independent of h_x and d. In what follows, instead of the *standard full tensor product Galerkin discretization*, we introduce a *sparse space-time tensor Galerkin discretization*, and prove that it yields a milder dependence on d, namely $O\big((\#\mathrm{dof})^{-\tilde{b}/(d-1)}\big)$ again up to logarithmic terms, under the provision of sufficient smoothness of the solution.

The sparse space-time tensor Galerkin discretization is based on the multilevel decompositions

$$\mathcal{V}^x_L = \mathcal{W}^x_0 \oplus \cdots \oplus \mathcal{W}^x_L, \quad \mathcal{V}^t_L = \mathcal{W}^t_0 \oplus \cdots \oplus \mathcal{W}^t_L, \quad \mathcal{X}^x_L = \mathcal{Y}^x_0 \oplus \cdots \oplus \mathcal{Y}^x_L, \quad \mathcal{X}^t_L = \mathcal{Y}^t_0 \oplus \cdots \oplus \mathcal{Y}^t_L. \tag{25}$$

Explicit biorthogonal spline-wavelet bases of the detail spaces $\mathcal{W}^{(\cdot)}_\ell$, $\mathcal{Y}^{(\cdot)}_\ell$ are required for implementation, but are available, see e.g. [6]. Utilizing the weighted sparse tensor product construction [5], we build the families $\{\hat{\mathcal{V}}^\sigma_L\}_{L\geq 0}$, $\{\hat{\mathcal{X}}^\sigma_L\}_{L\geq 0}$ of sparse space-time tensor subspaces from (25) by

$$\hat{\mathcal{V}}^\sigma_L := \bigotimes_{\ell_t/\sigma + \ell_x\sigma \leq L} \mathcal{W}^t_{\ell_t} \otimes \mathcal{W}^x_{\ell_x} \subset \mathcal{V} = H^{\frac{1}{2},\frac{1}{4}}(\Sigma), \quad \hat{\mathcal{X}}^\sigma_L := \bigotimes_{\ell_t/\sigma + \ell_x\sigma \leq L} \mathcal{Y}^t_{\ell_t} \otimes \mathcal{Y}^x_{\ell_x} \subset \mathcal{X} = H^{-\frac{1}{2},-\frac{1}{4}}(\Sigma) \tag{26}$$

for suitable fixed parameters $0 < \sigma < \infty$. The corresponding to $\hat{\mathcal{X}}^\sigma_L$ and $\hat{\mathcal{V}}^\sigma_L$ Galerkin solutions of (14), (15) will be denoted by $\hat{\psi}_L$ and $\hat{\varphi}_L$ respectively. From (16), the rate of convergence of sparse Galerkin solutions will be determined by the consistency order of the sparse subspaces $\hat{\mathcal{X}}^\sigma_L$ and $\hat{\mathcal{V}}^\sigma_L$ for solutions ψ and φ being sufficiently smooth.

In [2], sparse space-time tensor Galerkin discretizations with equal weight $\sigma = 1$ have been first introduced and analyzed. As we will see below, the results in [2] can be somewhat improved for the case $d = 3$, where $\sigma = \sqrt{2}$ yields better convergence rates.

We concentrate first on the hypersingular integral equations (11) and (12) and their sparse space-time Galerkin approximations (15) based on the sparse tensor product spaces $\hat{\mathcal{V}}^\sigma_L$. We recall [5, Theorem 7.1] ensuring existence of $\hat{\zeta}_L \in \hat{\mathcal{V}}^\sigma_L$ satisfying

$$\|\varphi - \hat{\zeta}_L\|_{H^{\frac{1}{2},\frac{1}{4}}_{\mathrm{mix}}(\Sigma)} \lesssim \hat{N}_L^{-\alpha} (\log \hat{N}_L)^\beta \|\varphi\|_{H^{s_x,s_t}_{\mathrm{mix}}(\Sigma)} \tag{27}$$

for every $1/2 < s_x \leq q_x + 1$, $1/4 < s_t \leq q_t + 1$ and $\hat{N}_L = \dim(\hat{\mathcal{V}}^\sigma_L)$. Here $H^{r,s}_{\mathrm{mix}}(I \times \Gamma) := H^r(\Gamma) \otimes H^s(I)$ can be understood as a space of "square integrable mixed highest derivatives", see e.g. [5] and references therein. In what follows we also need the space $\tilde{H}^{r,s}_{\mathrm{mix}}(I \times \Gamma) := H^r(\Gamma) \otimes \tilde{H}^s(I)$ whose elements satisfy the homogeneous initial conditions. Note that $H^{r,s}_{\mathrm{mix}}(\Sigma) = \tilde{H}^{r,s}_{\mathrm{mix}}(\Sigma)$ if $0 \le s < \frac{1}{2}$. The exponent α in the leading order term in (27) is given by

$$\alpha = \frac{\min\{(s_x - \frac{1}{2}), (s_t - \frac{1}{4})\sigma^2\}}{\max\{d-1, \sigma^2\}} \tag{28}$$

where $(d-1)$ is the dimension of the domain boundary $\Gamma = \partial D$. Simple calculations show that α attains its maximum if σ^2 is between $(d-1)$ and $(s_x - \frac{1}{2})/(s_t - \frac{1}{4})$ and is equal to

$$\alpha_{\max} = \min\left\{(s_x - \tfrac{1}{2})/(d-1), (s_t - \tfrac{1}{4})\right\}. \tag{29}$$

We fix the value $\nu := s_t - \frac{1}{4}$ and seek for the smallest s_x (and accordingly for the largest space $H^{s_x,s_t}_{\mathrm{mix}}(\Sigma)$) so that (27) holds with highest rate $\alpha = \alpha_{\max} = \nu$. Obviously, $s_x = (d-1)\nu + \frac{1}{2}$, yielding $\alpha = \nu$ and $\sigma^2 = d-1$. In this case $\beta = \nu + \frac{1}{2}$. There holds

Theorem 4.5 *Suppose Γ is sufficiently smooth and $\varphi \in \tilde{H}^{(d-1)\nu+\frac{1}{2},\nu+\frac{1}{4}}_{\mathrm{mix}}(\Sigma)$ for*

$$\nu = \frac{q_x + 1/2}{d-1} \quad \text{and} \quad q_t \ge \nu - \frac{3}{4}. \tag{30}$$

The error of the sparse Galerkin solution $\hat{\varphi}_L \in \hat{\mathcal{V}}_L^{\sqrt{d-1}}$ admits the bound

$$\|\varphi - \hat{\varphi}_L\|_{H^{\frac{1}{2},\frac{1}{4}}(\Sigma)} \lesssim \hat{N}_L^{-\nu}(\log \hat{N}_L)^{\nu+\frac{1}{2}} \|\varphi\|_{H^{(d-1)\nu+\frac{1}{2},\nu+\frac{1}{4}}_{\mathrm{mix}}(\Sigma)}. \tag{31}$$

P r o o f. For the anisotropic spaces (3) being intersection spaces, we obtain

$$\begin{aligned}
\|\varphi - \hat{\zeta}_L\|_{H^{\frac{1}{2},\frac{1}{4}}(\Sigma)} &\simeq \|\varphi - \hat{\zeta}_L\|_{L^2(I;H^{\frac{1}{2}}(\Gamma))} + \|\varphi - \hat{\zeta}_L\|_{H^{\frac{1}{4}}(I;L^2(\Gamma))} \\
&\simeq \|\varphi - \hat{\zeta}_L\|_{H^{\frac{1}{2},0}_{\mathrm{mix}}(\Sigma)} + \|\varphi - \hat{\zeta}_L\|_{H^{0,\frac{1}{4}}_{\mathrm{mix}}(\Sigma)} \\
&\lesssim \|\varphi - \hat{\zeta}_L\|_{H^{\frac{1}{2},\frac{1}{4}}_{\mathrm{mix}}(\Sigma)}.
\end{aligned} \tag{32}$$

We choose $s_x := q_x + 1$, yielding the first relation in (30), and estimate the right-hand of (32) side by (27). □

Corollary 4.6 *Suppose Γ is sufficiently smooth, that $h \in \tilde{H}^{(d+1)\nu,\frac{(d+1)\nu}{2}}(\Sigma)$ and ν, q_x, q_t satisfy* (30). *Then*

$$\|\varphi - \hat{\varphi}_L\|_{H^{\frac{1}{2},\frac{1}{4}}(\Sigma)} \lesssim \hat{N}_L^{-\nu}(\log \hat{N}_L)^{\nu+\frac{1}{2}} \|h\|_{H^{(d+1)\nu,\frac{(d+1)\nu}{2}}(\Sigma)}. \tag{33}$$

P r o o f. There holds $H^{k,\frac{k}{2}}(\Sigma) \subset H^{a,b}_{\mathrm{mix}}(\Sigma)$ when $k \ge a + 2b$, see Lemma 5.2 below. Hence for $k = (d+1)\nu + 1$

$$\|\varphi - \hat{\varphi}_L\|_{H^{\frac{1}{2},\frac{1}{4}}(\Sigma)} \lesssim \hat{N}_L^{-\nu}(\log \hat{N}_L)^{\nu+\frac{1}{2}} \|\varphi\|_{H^{k,\frac{k}{2}}(\Sigma)}. \tag{34}$$

Then the assertion follows by Proposition 3.6, since the mappings

$$W : \tilde{H}^{k,\frac{k}{2}}(\Sigma) \to \tilde{H}^{r,\frac{r}{2}}(\Sigma), \qquad (\frac{1}{2}I - N) : \tilde{H}^{r,\frac{r}{2}}(\Sigma) \to \tilde{H}^{r,\frac{r}{2}}(\Sigma), \qquad r = k-1 \tag{35}$$

are isomorphisms for $r \ge \frac{1}{2}$ yielding $\|\varphi\|_{H^{k,\frac{k}{2}}(\Sigma)} \lesssim \|h\|_{H^{r,\frac{r}{2}}(\Sigma)}$ for $r = (d+1)\nu$. □

Construction of a suitable projector $\hat{\eta}_L$ for the solution ψ of the single layer equations also relies on an Aubin-Nitsche duality argument: we *assume* $\psi \in L^2(\Sigma)$ and denote by $\hat{\Pi}^\sigma_L : L^2(\Sigma) \to \hat{\mathcal{X}}^\sigma_L$ the $L^2(\Sigma)$-orthogonal projection. Then

$$\begin{aligned}
\|\psi - \hat{\Pi}^\sigma_L \psi\|_{H^{-\frac{1}{2},-\frac{1}{4}}(\Sigma)} &= \sup_{\xi \in H^{\frac{1}{2},\frac{1}{4}}(\Sigma)} \frac{\langle \psi - \hat{\Pi}^\sigma_L \psi, \xi \rangle_{L^2(\Sigma)}}{\|\xi\|_{H^{\frac{1}{2},\frac{1}{4}}(\Sigma)}} = \sup_{\xi \in H^{\frac{1}{2},\frac{1}{4}}(\Sigma)} \frac{\langle \psi - \hat{\Pi}^\sigma_L \psi, \xi - \hat{\Pi}^\sigma_L \xi \rangle_{L^2(\Sigma)}}{\|\xi\|_{H^{\frac{1}{2},\frac{1}{4}}(\Sigma)}} \\
&\le \|\psi - \hat{\Pi}^\sigma_L \psi\|_{L^2(\Sigma)} \sup_{\xi \in H^{\frac{1}{2},\frac{1}{4}}(\Sigma)} \frac{\|\xi - \hat{\Pi}^\sigma_L \xi\|_{L^2(\Sigma)}}{\|\xi\|_{H^{\frac{1}{2},\frac{1}{4}}(\Sigma)}}.
\end{aligned} \tag{36}$$

By [5, Theorem 4.3], for $0 < s_x \le p_x + 1$ and $0 < s_t \le p_t + 1$

$$\|\xi - \hat{\Pi}^\sigma_L \xi\|_{L^2(Q)} = \inf_{\hat{\xi}_L \in \hat{\mathcal{X}}^\sigma_L} \|\xi - \hat{\xi}_L\|_{L^2(Q)} \lesssim \hat{N}_L^{-\alpha} (\log \hat{N}_L)^\beta \|\xi\|_{H^{s_x,s_t}_{\mathrm{mix}}(\Sigma)} \text{ for } \alpha = \frac{\min\{s_x, s_t\sigma^2\}}{\max\{d-1, \sigma^2\}} \tag{37}$$

and some $\beta \ge 0$. The exponent α attains its maximum for σ^2 between $(d-1)$ and s_x/s_t and equals

$$\alpha_{\max} = \min\big\{s_x/(d-1), s_t\big\}. \tag{38}$$

For fixed $\mu := s_t$ and $s_x = (d-1)\mu$ we have $\alpha = \alpha_{\max} = \mu$ when $\sigma^2 = d-1$. In this case $\beta = \mu + \frac{1}{2}$. Moreover, $H^{k,\frac{k}{2}}(\Sigma) \subset H^{(d-1)\mu,\mu}_{\mathrm{mix}}(\Sigma)$ for $k \ge (d+1)\mu$, see Lemma 5.2 below. This yields for the particular values $k := \frac{1}{2}$ and $\mu := 1/(2(d+1))$ the estimate

$$\sup_{\xi \in H^{\frac{1}{2},\frac{1}{4}}(\Sigma)} \frac{\|\xi - \hat{\Pi}^\sigma_L \xi\|_{L^2(\Sigma)}}{\|\xi\|_{H^{\frac{1}{2},\frac{1}{4}}(\Sigma)}} \lesssim \hat{N}_L^{-\frac{1}{2(d+1)}} (\log \hat{N}_L)^{\frac{1}{2(d+1)} + \frac{1}{2}} . \tag{39}$$

Combining this with (36) and (37) we obtain the following convergence estimate.

Theorem 4.7 *Suppose* $\psi \in \tilde{H}^{(d-1)\mu,\mu}_{\mathrm{mix}}(\Sigma)$ *for* μ, p_x, p_t *satisfying*

$$\mu = \frac{p_x + 1}{d-1}, \quad \textit{and} \quad p_t + 1 \ge \mu. \tag{40}$$

The error of the sparse tensor Galerkin approximation $\hat{\psi}_L \in \hat{\mathcal{X}}^{\sqrt{d-1}}_L$ *admits the bound*

$$\|\psi - \hat{\psi}_L\|_{H^{-\frac{1}{2},-\frac{1}{4}}(\Sigma)} \lesssim \hat{N}_L^{-\lambda} (\log \hat{N}_L)^{\lambda+1} \|\psi\|_{H^{(d-1)\mu,\mu}_{\mathrm{mix}}(\Sigma)} \tag{41}$$

where

$$\lambda = \mu + \frac{1}{2(d+1)} = \begin{cases} p_x + \frac{7}{6}, & d = 2, \\ \frac{p_x}{2} + \frac{5}{8}, & d = 3 . \end{cases} \tag{42}$$

P r o o f. The first term in the right-hand side of (36) admits the bound

$$\|\psi - \hat{\Pi}^\sigma_L \psi\|_{L^2(\Sigma)} \lesssim \hat{N}_L^{-\mu} \log(\hat{N}_L)^{\mu + \frac{1}{2}} \|\psi\|_{H^{(d-1)\mu,\mu}_{\mathrm{mix}}(\Sigma)}, \tag{43}$$

the second in bounded by (39). Combination of the two estimates yields the assertion with the exponent λ from (42). □

Corollary 4.8 *Suppose* $g \in \tilde{H}^{r,\frac{r}{2}}(\Sigma)$ *and* $r, \mu, \lambda, p_x, p_t$ *satisfy* (40), (42) *and*

$$r = \frac{d+1}{d-1}(p_x + 1) + 1. \tag{44}$$

Then the error of the sparse Galerkin solution $\hat{\psi}_L \in \hat{\mathcal{X}}^{(d-1)}_L$ *admits the bound*

$$\|\psi - \hat{\psi}_L\|_{H^{-\frac{1}{2},-\frac{1}{4}}(\Sigma)} \lesssim \hat{N}_L^{-\lambda} (\log \hat{N}_L)^{\lambda+1} \|g\|_{H^{r,\frac{r}{2}}(\Sigma)}. \tag{45}$$

P r o o f. We utilize $H^{k,\frac{k}{2}}(\Sigma) \subset H^{(d-1)\mu,\mu}_{\mathrm{mix}}(\Sigma)$ for $k = (d+1)\mu$, see Lemma 5.2 below, and $\mu = \frac{p_x+1}{d-1}$ yielding

$$\|\psi\|_{H^{(d-1)\mu,\mu}_{\mathrm{mix}}(\Sigma)} \lesssim \|\psi\|_{H^{k,\frac{k}{2}}(\Sigma)}, \qquad k = \frac{d+1}{d-1}(p_x + 1) . \tag{46}$$

The assertion follows from Proposition 3.6 yielding that the following mappings are isomorphisms for $r \ge \frac{1}{2}$:

$$V : \tilde{H}^{k,\frac{k}{2}}(\Sigma) \to \tilde{H}^{r,\frac{r}{2}}(\Sigma), \qquad (\frac{1}{2}I + K) : \tilde{H}^{r,\frac{r}{2}}(\Sigma) \to \tilde{H}^{r,\frac{r}{2}}(\Sigma), \qquad r = k + 1 . \tag{47}$$

□

Table 1 Convergence parameters of the full and sparse tensor product discretizations for the weakly singular integral equations (9), (10).

V, Full tensor product, $d=2$

(p_x,p_t)	$\gamma=\frac{p_x+3/2}{3}$	$r=2p_t+3$	$p_x\le 2p_t+1$
$(0,0)$	$\frac{1}{2}$	3	true
$(1,0)$	$\frac{5}{6}$	3	true
$(1,1)$	$\frac{5}{6}$	5	true
$(3,1)$	$\frac{3}{2}$	5	true

V, Sparse tensor product, $d=2$

(p_x,p_t)	$\gamma=p_x+\frac{7}{6}$	$r=3p_x+4$	$p_x\le p_t$
$(0,0)$	$\frac{7}{6}$	4	true
$(1,0)$	–	–	false
$(1,1)$	$\frac{13}{6}$	7	true
$(3,1)$	–	–	false

V, Full tensor product, $d=3$

(p_x,p_t)	$\gamma=\frac{p_x+3/2}{4}$	$r=2p_t+3$	$p_x\le 2p_t+1$
$(0,0)$	$\frac{3}{8}$	3	true
$(1,0)$	$\frac{5}{8}$	3	true
$(1,1)$	$\frac{5}{8}$	5	true
$(3,1)$	$\frac{9}{8}$	5	true

V, Sparse tensor product, $d=3$

(p_x,p_t)	$\gamma=\frac{p_x}{2}+\frac{5}{8}$	$r=2p_x+3$	$p_x\le 2p_t+1$
$(0,0)$	$\frac{5}{8}$	3	true
$(1,0)$	$\frac{9}{8}$	5	true
$(1,1)$	$\frac{9}{8}$	5	true
$(3,1)$	$\frac{17}{8}$	9	true

5 Discussion: Sparse space-time tensor product vs. full tensor product BEM

In Table 1 we summarize the results of Corollary 4.2 and Corollary 4.8 on convergence of the full and sparse space-time Galerkin solutions for the weakly singular integral equations (9) and (10). The convergence estimates (21) and (45) have the form

$$\|\psi-\tilde{\psi}_L\|_{H^{-\frac{1}{2},-\frac{1}{4}}(\Sigma)}\lesssim \tilde{N}_L^{-\gamma}(\log\tilde{N}_L)^{\tilde{\gamma}}\|g\|_{H^{r,\frac{r}{2}}(\Sigma)}$$

where $\tilde{\psi}_L\in\{\psi_L,\hat{\psi}_L\}$, $\tilde{N}_L\in\{N_L,\hat{N}_L\}$ and $\tilde{\gamma}\in\{0,\gamma+1\}$. In Table 1 we list the values γ, r and the restriction on polynomial degrees for practically relevant cases $d=2$ and $d=3$.

We observe that the sparse tensor product approximation achieves significantly better convergence rates than the full tensor product approximation for the same choice of polynomial degrees (p_x,p_t). But this comes at the price of stronger regularity assumptions on the data $g\in H^{r,\frac{r}{2}}(\Sigma)$: e.g. $r=4$ instead of $r=3$ for $(p_x,p_t)=(0,0)$ and $r=7$ instead of $r=5$ for $(p_x,p_t)=(1,1)$. The convergence parameters for $(p_x,p_t)=(1,0)$ and $(3,1)$ are not shown, because these combinations violate the restriction $p_x\le p_t$.

Analyzing the diagonal entries of Table 1 we find e.g. that in the case $d=2$, the simplest sparse tensor approximation with $(p_x,p_t)=(0,0)$ achieves an asymptotic convergence rate $\gamma=\frac{7}{6}$ which is higher than the rate obtained by the full tensor product approximation with $(p_x,p_t)=(0,0)$ and $(1,0)$, yielding the rates $\gamma=\frac{1}{2}$ and $\gamma=\frac{5}{6}$ respectively; it is outperformed only by the full tensor product approximation with $(p_x,p_t)=(3,1)$, however, at the price of a higher regularity assumption on the data ($r=5$ instead of $r=4$). In the case $d=3$, the lowest order sparse approximation achieves essentially (i.e., up to logarithmic terms) the same convergence rate as the full tensor product approximation with $(p_x,p_t)=(1,0)$ ($\gamma=\frac{5}{8}$); the sparse tensor discretization with $(p_x,p_t)=(1,0)$ yields the same rate as the full tensor product approximation with $(p_x,p_t)=(3,1)$ ($\gamma=\frac{9}{8}$).

In Table 2 we summarize the results of Corollary 4.4 and Corollary 4.6 on convergence of the full and sparse space-time Galerkin solutions for the hypersingular integral equations (11) and (12). The convergence estimates (24) and (33) have the form

$$\|\varphi-\tilde{\varphi}_L\|_{H^{-\frac{1}{2},-\frac{1}{4}}(\Sigma)}\lesssim \tilde{N}_L^{-\gamma}(\log\tilde{N}_L)^{\tilde{\gamma}}\|h\|_{H^{r,\frac{r}{2}}(\Sigma)}$$

where $\tilde{\varphi}_L\in\{\varphi_L,\hat{\varphi}_L\}$, $\tilde{N}_L\in\{N_L,\hat{N}_L\}$ and $\tilde{\gamma}\in\{0,\gamma+\frac{1}{2}\}$. The relation of the convergence parameters is similar to the case of the weakly singular equations discussed above. We sum up our findings.

Remark 5.1 For diffusion in a bounded, two and three-dimensional domain, the lowest order sparse tensor product approximation outperforms or matches (up to logarithmic terms) the convergence rate of the lowest and the first order approximations by the full tensor products. It does *not* require stronger regularity assumptions on the data (the case

Table 2 Convergence parameters of the full and sparse tensor product discretizations for the hypersingular integral equations (11), (12).

W, Full tensor product, $d=2$			
(q_x,q_t)	$\gamma=\frac{q_x+1/2}{3}$	$r=2q_t+1$	$q_x\le 2q_t+1$
$(1,0)$	$\frac{1}{2}$	1	true
$(1,1)$	$\frac{1}{2}$	3	true
$(3,1)$	$\frac{7}{6}$	3	true

W, Sparse tensor product, $d=2$			
(q_x,q_t)	$\gamma=q_x+\frac{1}{2}$	$r=3(q_x+\frac{1}{2})$	$q_x\le q_t$
$(1,0)$	–	–	false
$(1,1)$	$\frac{3}{2}$	$\frac{9}{2}$	true
$(3,1)$	–	–	false

W, Full tensor product, $d=3$			
(q_x,q_t)	$\gamma=\frac{q_x+1/2}{4}$	$r=2q_t+1$	$q_x\le 2q_t+1$
$(1,0)$	$\frac{3}{8}$	1	true
$(1,1)$	$\frac{3}{8}$	3	true
$(3,1)$	$\frac{7}{8}$	3	true

W, Sparse tensor product, $d=3$			
(q_x,q_t)	$\gamma=\frac{q_x+1/2}{2}$	$r=2q_x+1$	$q_x\le 2q_t+1$
$(1,0)$	$\frac{3}{4}$	3	true
$(1,1)$	$\frac{3}{4}$	3	true
$(3,1)$	$\frac{7}{4}$	7	true

$d=2$, $(q_x,q_t)=(1,1)$ is an exception). Consequently, for sufficiently smooth data, either the lowest/first order sparse discretization or higher order full tensor product should be chosen to ensure a favourable convergence rate of the Galerkin scheme. The final choice of the discretization depends on the implementation complexity of the low order sparse- or the high order full tensor product method, respectively.

Appendix

We prove the following technical lemma.

Lemma 5.2 *For $a,b,k\ge 0$ there holds $H^{k,\frac{k}{2}}(\Omega_1\times\Omega_2)\subset H^{a,b}_{\mathrm{mix}}(\Omega_1\times\Omega_2)$ if $k\ge a+2b$.*

Proof. Suppose $\Omega_1=\mathbb{R}^m$, $\Omega_2=\mathbb{R}^n$. Then

$$\|u\|_{H^{a,b}_{\mathrm{mix}}(\mathbb{R}^m\times\mathbb{R}^n)}=\int_{\mathbb{R}^m}\int_{\mathbb{R}^n}(1+|\xi|^2)^a(1+|\tau|^2)^b|\mathcal{F}u|^2\,d\tau d\xi$$

$$\|u\|_{H^{k,\frac{k}{2}}(\mathbb{R}^m\times\mathbb{R}^n)}=\int_{\mathbb{R}^m}\int_{\mathbb{R}^n}\left((1+|\xi|^2)^k+(1+|\tau|^2)^{\frac{k}{2}}\right)|\mathcal{F}u|^2\,d\tau d\xi$$

where $\mathcal{F}u$ is the Fourier transform of u with respect to all variables. By Young's inequality

$$(1+|\xi|^2)^a(1+|\tau|^2)^b\le\frac{(1+|\xi|^2)^{ap}}{p}+\frac{(1+|\tau|^2)^{bq}}{q},\qquad \frac{1}{p}+\frac{1}{q}=1\,.$$

The assertion follows for $k\ge k_0$ where $k_0=ap$ and $k_0=2bq$. Solving this system we obtain $k_0=a+2b$. The assertion for bounded domains with a smooth boundary follows from this result by suitable prolongation and restriction operators. □

Acknowledgements This work was performed while the second author was visiting the Hausdorff Institute for Mathematics (HIM) at the University of Bonn (Germany) during the HIM Trimester on "High Dimensional Approximation" from May - August 2011.

References

[1] C. Bourgeois and S. Nicaise, Prewavelet analysis of the heat equation, Numer. Math. **87**(3), 407–434 (2001).
[2] A. Chernov and Ch. Schwab, First order k-th moment finite element analysis of nonlinear operator equations with stochastic data, to appear in Math. Comput., also available as Tech. Rep. 2011b06, HIM, University of Bonn, August 2011, URL: `http://131.220.77.52/files/preprints/AnalysisandNumerics/ChernovSchwab_FOkM.pdf`
[3] M. Costabel, Boundary integral operators for the heat equation, Integral Equations Operator Theory **13**(4), 498–552 (1990).
[4] L. Greengard and J. Strain, A fast algorithm for the evaluation of heat potentials, Comm. Pure Appl. Math., **43**(8) 949–963 (1990).
[5] M. Griebel and H. Harbrecht, On the construction of sparse tensor product spaces, Tech. rep., 2011, (to appear in Math. Comp.).
[6] B. Han and Z. Shen Wavelets with short support. SIAM J. Math. Anal. **38**, 530556 (2006).

[7] P. J. Noon, The single layer heat potential and Galerkin Boundary Element Methods for the heat equation, PhD thesis, University of Maryland, 1988.
[8] S. Sauter and Ch. Schwab, Boundary Element Methods, Springer Verlag, Berlin, Heidelberg, New-York (2011).
[9] D. Schötzau and Ch. Schwab, Time discretization of parabolic problems by the hp-version of the discontinuous Galerkin finite element method, SIAM J. Numer. Anal. **38**(3), 837–875 (electronic) (2000).
[10] Ch. Schwab and R. Stevenson, Space–time adaptive wavelet methods for parabolic evolution problems, Math. Comput. **78**, 1293–1318 (2009).
[11] M. Spivak and S.K. Veerapaneni and L. Greengard, The fast generalized Gauss transform, SIAM J. Sci. Comput. **32**(5),3092–3107 (2010).
[12] J. Tausch and A. Weckiewicz, Multidimensional Fast Gauss Transforms by Chebyshev Expansions, SIAM J. Scientific Computing, **31** (5) 3547-3565 (2009).
[13] V. Thomée, Galerkin Finite Element Methods for Parabolic Problems, second edition (Springer, 2006).