Adaptive space-time BEM for the heat equation

Gantner, G..; van Venetië, R.

DOI
10.1016/j.camwa.2021.12.022
Publication date
2022
Document Version
Final published version
Published in
Computers and Mathematics with Applications

Computers and Mathematics with Applications 107 (2022) 117–131

Contents lists available at ScienceDirect

Computers and Mathematics with Applications

www.elsevier.com/locate/camwa

Adaptive space-time BEM for the heat equation

Gregor Gantner *, Raymond van Venetië

Korteweg-de Vries Institute for Mathematics, University of Amsterdam, P.O. Box 94248, 1090 GE Amsterdam, Netherlands

ARTICLE INFO

Keywords:
Space-time boundary element method
Heat equation
A posteriori error estimation
Adaptive mesh-refinement
Computation of singular integrals

ABSTRACT

We consider the space-time boundary element method (BEM) for the heat equation with prescribed initial and Dirichlet data. We propose a residual-type *a posteriori* error estimator that is a lower bound and, up to weighted L_2-norms of the residual, also an upper bound for the unknown BEM error. The possibly locally refined meshes are assumed to be prismatic, i.e., their elements are tensor-products $J \times K$ of elements in time J and space K. While the results do not depend on the local aspect ratio between time and space, assuming the scaling $|J| \approx \operatorname{diam}(K)^2$ for all elements and using Galerkin BEM, the estimator is shown to be efficient and reliable without the additional L_2-terms. In the considered numerical experiments on two-dimensional domains in space, the estimator seems to be equivalent to the error, independently of these assumptions. In particular for adaptive anisotropic refinement, both converge with the best possible convergence rate.

1. Introduction

In the last years, there has been a growing interest in simultaneous space-time boundary element methods (BEM) for the heat equation [6,18,19,17,5,7,9,23,25]. In contrast to the differential operator based variational formulation on the space-time cylinder, the variational formulation corresponding to space-time BEM is coercive [3,4] so that the discretized version always has a unique solution regardless of the chosen trial space which is even quasi-optimal in the natural energy norm. Moreover, it is naturally applicable on unbounded domains and only requires a mesh of the lateral boundary of the space-time cylinder (as well as a mesh of the spatial domain in case of nonhomogeneous initial data) resulting in a dimension reduction. The potential disadvantage that discretizations lead to dense matrices due to the nonlocality of the boundary integral operators has been tackled, e.g., in [18,19,17,24] via the fast multipole method and $\mathcal{H}$-matrices.

Two often mentioned advantages of simultaneous space-time methods are their potential for massive parallelization as well as their potential for fully adaptive refinement to resolve singularities local in both space and time. While the first advantage has been investigated in, e.g., [9,25], the latter requires suitable *a posteriori* computable error estimators, which have not been developed yet for the heat equation. Indeed, concerning *a posteriori* error estimation as well as adaptive refinement for BEM for time-dependent problems, we are only aware of the works [13,14] for the wave equation in two and three space dimensions, respectively.

In the present manuscript, we generalize the results [10,11] from Faermann for stationary PDEs to the heat equation: Let $\Omega \subset \mathbb{R}^n$, $n = 2, 3$, be a Lipschitz domain with boundary $\Gamma := \partial\Omega$ and $T > 0$ a given end time point with corresponding time interval $I := (0, T)$. We abbreviate the space-time cylinder $Q := I \times \Omega$ with lateral boundary $\Sigma := I \times \Gamma$ and corresponding outer normal vector $\mathbf{n} \in \mathbb{R}^n$. With the heat kernel

$$G(t, \mathbf{x}) := \begin{cases} \frac{1}{(4\pi t)^{n/2}} e^{-\frac{|\mathbf{x}|^2}{4t}} & \text{for } (t, \mathbf{x}) \in (0, \infty) \times \mathbb{R}^n, \\ 0 & \text{else,} \end{cases}$$

and a given function $f : \Sigma \to \mathbb{R}$, we consider the boundary integral equation

$$(\mathcal{V}\phi)(t, \mathbf{x}) := \int_\Sigma G(t - s, \mathbf{x} - \mathbf{y})\phi(s, \mathbf{y}) \, \mathrm{d}\mathbf{y} \, \mathrm{d}s = f(t, \mathbf{x}) \quad \text{for a.e. } (t, \mathbf{x}) \in \Sigma. \tag{1.1}$$

* Corresponding author.
E-mail addresses: G.Gantner@uva.nl (G. Gantner), R.vanVenetie@uva.nl (R. van Venetië).

https://doi.org/10.1016/j.camwa.2021.12.022
Received 6 August 2021; Received in revised form 7 December 2021; Accepted 28 December 2021
Available online 10 January 2022

Here, $\mathscr{V}$ is the single-layer operator. For given initial condition $u_0 : \Omega \to \mathbb{R}$ and Dirichlet data $u_D : \Sigma \to \mathbb{R}$, such equations arise from the heat equation

$$\begin{aligned} \partial_t u - \Delta u &= 0 \quad \text{on } Q, \\ u &= u_D \quad \text{on } \Sigma, \\ u(0,\cdot) &= u_0 \quad \text{on } \Omega. \end{aligned} \tag{1.2}$$

Let $\mathcal{P}$ be a mesh of the space-time boundary Σ consisting of prismatic elements $J \times K$ with $J \subseteq \overline{I}$ and $K \subseteq \Gamma$, and let Φ be an associated approximation of ϕ. Typically, Φ is a piecewise polynomial with respect to $\mathcal{P}$. As $\mathscr{V}$ is an isomorphism from the dual space $H^{-1/2,-1/4}(\Sigma) := H^{1/2,1/4}(\Sigma)'$ to the anisotropic Sobolev space $H^{1/2,1/4}(\Sigma)$, the discretization error $\|\phi - \Phi\|_{H^{-1/2,-1/4}(\Sigma)}$ is equivalent to the norm of the residual $\|f - \mathscr{V}\Phi\|_{H^{1/2,1/4}(\Sigma)}$. We show that the residual norm can be localized up to weighted L_2-terms, i.e.,

$$\sum_{J\times K\in\mathcal{P}} \eta_{\mathcal{P}}(\Phi, J\times K)^2 \lesssim \|f - \mathscr{V}\Phi\|^2_{H^{1/2,1/4}(\Sigma)} \lesssim \sum_{J\times K\in\mathcal{P}} \eta_{\mathcal{P}}(\Phi, J\times K)^2 + \zeta_{\mathcal{P}}(\Phi, J\times K)^2,$$

where $\eta_{\mathcal{P}}(\Phi, J\times K)^2$ measures the $H^{1/2,1/4}$-seminorm of the residual in a neighborhood of $J\times K$ and $\zeta_{\mathcal{P}}(\Phi) := (\operatorname{diam}(K)^{-1} + |J|^{-1/2})\|f - \mathscr{V}\Phi\|^2_{L_2(J\times K)}$. The hidden constants depend only on the regularity of the meshes found by fixing either the temporal or the spatial coordinate in $\mathcal{P}$. In particular, we do not require any assumption on the relation between the spatial and temporal size of the mesh elements, making anisotropically refined meshes possible.

If the elements satisfy the scaling $|J| \eqsim \operatorname{diam}(K)^2$ and if Φ is the Galerkin approximation of ϕ in a discrete space $\mathcal{X}$ that contains at least all $\mathcal{P}$-piecewise constant functions, then we can additionally prove that

$$\zeta_{\mathcal{P}}(\Phi, J\times K) \lesssim \eta_{\mathcal{P}}(\Phi, J\times K).$$

Indeed, numerical experiments (with $n = 2$) suggest that this is not the case in general: If the scaling condition is not enforced, we observe situations where the weighted L_2-terms ζ do not decay under mesh-refinement.

That being said, the estimator η does not only behave efficiently but also reliably in all considered examples. Moreover, anisotropic refinement steered by the space- and time-components of the estimator always yields the optimal algebraic convergence rate of both the estimator and the error. The source code that we used to generate the numerical results is available at [16].

Outline

The remainder of this work is organized as follows: Section 2 summarizes the general principles of the space-time boundary element method for the heat equation. Section 3 recalls the localization argument of [10,11] and applies it to anisotropic Sobolev spaces (Theorem 3.3). This result is then invoked in Corollary 3.5 for the residual, resulting in efficient and reliable *a posteriori* computable error bounds. In particular, a Poincaré-type inequality (Lemma 3.4) allows to estimate the weighted L_2-terms that are still present in the upper bound from Theorem 3.3. Finally, Section 4 introduces an adaptive algorithm for $n = 2$ which is based on the derived error estimator. Different marking and refinement strategies are presented. The adaptive algorithm is subsequently applied to several concrete examples with typical singularities in space and time. The stable implementation is discussed in Appendix A.

2. Preliminaries

2.1. General notation

Throughout and without any ambiguity, $|\cdot|$ denotes the absolute value of scalars, the Euclidean norm of vectors in $\mathbb{R}^m$, or the measure of a set in $\mathbb{R}^m$, e.g., the length of an interval or the area of a surface in $\mathbb{R}^3$. We write $A \lesssim B$ to abbreviate $A \le C\,B$ with some generic constant $C > 0$, which is clear from the context. Moreover, $A \eqsim B$ abbreviates $A \lesssim B \lesssim A$.

2.2. Anisotropic Sobolev spaces

For n-dimensional $\omega \subseteq \Omega$ or $(n-1)$-dimensional $\omega \subseteq \Gamma$, and $\mu \in (0,1]$, we first recall the Sobolev space

$$H^\mu(\omega) := \big\{ v \in L_2(\omega) \,:\, \|v\|_{H^\mu(\omega)} < \infty \big\}$$

associated with the Sobolev–Slobodeckij norm

$$\|v\|^2_{H^\mu(\omega)} := \|v\|^2_{L_2(\omega)} + |v|^2_{H^\mu(\omega)}, \quad |v|^2_{H^\mu(\omega)} := \begin{cases} \int_\omega \int_\omega \frac{|v(\mathbf{x}) - v(\mathbf{y})|^2}{|\mathbf{x}-\mathbf{y}|^{\dim(\omega)+2\mu}} \,\mathrm{d}\mathbf{y}\,\mathrm{d}\mathbf{x} & \text{if } \mu \in (0,1), \\ \|\nabla_\omega v\|^2_{L_2(\omega)} & \text{if } \mu = 1, \end{cases}$$

where $\dim(\omega)$ denotes the dimension of ω, i.e., n or $n-1$, and ∇_ω denotes the (weak) gradient on ω, i.e., the standard gradient or the surface gradient.

Moreover, we define for any subinterval $J \subseteq \overline{I}$, $\nu \in (0,1]$, and any Banach space X,

$$H^\nu(J;X) := \big\{ v \in L_2(J;X) \,:\, \|v\|_{H^\nu(J;X)} < \infty \big\}$$

associated with the norm

$$\|v\|^2_{H^\nu(J;X)} := \|v\|^2_{L_2(J;X)} + |v|^2_{H^\nu(J;X)}, \quad |v|^2_{H^\nu(J;X)} := \begin{cases} \int_J \int_J \frac{\|v(t) - v(s)\|^2_X}{|t-s|^{1+2\nu}} \,\mathrm{d}s\,\mathrm{d}t & \text{if } \nu \in (0,1), \\ \|\partial_t v\|^2_{L_2(J;X)} & \text{if } \nu = 1, \end{cases}$$

where ∂_t denotes the (weak) time derivative. If $X = \mathbb{R}$, we simply write $H^\nu(J)$, $\|v\|_{H^\nu(J)}$, and $|v|_{H^\nu(J)}$. Finally, we recall the anisotropic Sobolev space

$$H^{\mu,\nu}(J\times\omega) := L_2(J;H^\mu(\omega))\cap H^\nu(J;L_2(\omega))$$

with corresponding norm

$$\|v\|^2_{H^{\mu,\nu}(J\times\omega)} := \|v\|^2_{L_2(J;H^\mu(\omega))} + \|v\|^2_{H^\nu(J;L_2(\omega))} \quad \text{for all } v\in H^{\mu,\nu}(J\times\omega).$$

We will sometimes use the abbreviation

$$|v|^2_{L_2(J;H^\mu(\omega))} := \int_J |v(t,\cdot)|^2_{H^\mu(\omega)}\,\mathrm{d}t \quad \text{for all } v\in L_2(J;H^\mu(\omega)).$$

For $\omega\in\{\Omega,\Gamma\}$, we denote by $H^{-\mu,-\nu}(I\times\omega)$ the dual space of $H^{\mu,\nu}(I\times\omega)$ with duality pairing $\langle\cdot,\cdot\rangle_{I\times\omega}$. We interpret $L_2(I\times\omega)$ as subspace of $H^{-\mu,-\nu}(I\times\omega)$ via

$$\langle v,\psi\rangle_{I\times\omega} := \int_I\int_\omega v(t,\mathbf{x})\psi(t,\mathbf{x})\,\mathrm{d}\mathbf{x}\,\mathrm{d}t \quad \text{for all } v\in H^{\mu,\nu}(I\times\omega) \text{ and } \psi\in L_2(I\times\omega).$$

2.3. Boundary integral equations

It is well-known that for $u_0\in L^2(\Omega)$ and $u_D\in H^{1/2,1/4}(\Sigma)$, the heat equation (1.2) admits a unique solution $u\in H^{1,1/2}(Q)$. With the normal derivative $\phi_N := \partial_{\mathbf{n}}u\in H^{-1/2,-1/4}(\Sigma)$, u satisfies the representation formula

$$u = \widetilde{\mathscr{M}}_0 u_0 + \widetilde{\mathscr{V}}\phi_N - \widetilde{\mathscr{K}}u_D, \tag{2.1}$$

where

$$(\widetilde{\mathscr{M}}_0 u_0)(t,\mathbf{x}) := \int_\Omega G(t,\mathbf{x}-\mathbf{y})u_0(\mathbf{y})\,\mathrm{d}\mathbf{y} \quad \text{for all } (t,\mathbf{x})\in Q \tag{2.2}$$

denotes the initial potential,

$$(\widetilde{\mathscr{V}}\phi_N)(t,\mathbf{x}) := \int_\Sigma G(t-s,\mathbf{x}-\mathbf{y})\phi_N(s,\mathbf{y})\,\mathrm{d}\mathbf{y}\,\mathrm{d}s \quad \text{for all } (t,\mathbf{x})\in Q \tag{2.3}$$

denotes the single-layer potential, and

$$(\widetilde{\mathscr{K}}u_D)(t,\mathbf{x}) := \int_\Sigma \partial_{\mathbf{n}(\mathbf{y})}G(t-s,\mathbf{x}-\mathbf{y})u_D(s,\mathbf{y})\,\mathrm{d}\mathbf{y}\,\mathrm{d}s \quad \text{for all } (t,\mathbf{x})\in Q \tag{2.4}$$

denotes the double-layer potential. These linear operators satisfy the mapping properties $\widetilde{\mathscr{M}}_0 : L^2(\Omega)\to H^{1,1/2}(Q)$, $\widetilde{\mathscr{V}}_0 : H^{-1/2,-1/4}(\Sigma)\to H^{1,1/2}(Q)$, and $\widetilde{\mathscr{K}}_0 : H^{1/2,1/4}(\Sigma)\to H^{1,1/2}(Q)$. The lateral trace $(\cdot)|_\Sigma : H^{1,1/2}(Q)\to H^{1/2,1/4}(\Sigma)$ of these potentials is given by

$$(\widetilde{\mathscr{M}}_0 u_0)|_\Sigma = \mathscr{M}_0 u_0, \quad (\widetilde{\mathscr{V}}\phi_N)|_\Sigma = \mathscr{V}\phi_N, \quad (\widetilde{\mathscr{K}}u_D)|_\Sigma = (\mathscr{K}-1/2)u_D,$$

where the initial operator $\mathscr{M}_0$, the single-layer operator $\mathscr{V}$, and the double-layer operator $\mathscr{K}$ are defined as in (2.2)–(2.4) for $(t,\mathbf{x})\in\Sigma$. Applying the lateral trace to (2.1) thus results in

$$\mathscr{V}\phi_N = (\mathscr{K}+1/2)u_D - \mathscr{M}_0 u_0, \tag{2.5}$$

i.e., (1.1) with $f := (\mathscr{K}+1/2)u_D - \mathscr{M}_0 u_0$. As the single-layer operator $\mathscr{V}$ is also coercive, i.e.,

$$\langle\mathscr{V}\psi,\psi\rangle_\Sigma \geq c_{\mathrm{coe}}\|\psi\|^2_{H^{-1/2,-1/4}(\Sigma)} \quad \text{for all } \psi\in H^{-1/2,-1/4}(\Sigma) \tag{2.6}$$

with some constant $c_{\mathrm{coe}}>0$, (2.5) is uniquely solvable and the solution ϕ_N is just the missing normal derivative $\partial_{\mathbf{n}}u$ to compute u via the representation formula (2.1).

Alternatively, one can make the ansatz $u = \widetilde{\mathscr{M}}_0 u_0 + \widetilde{\mathscr{V}}\phi$. Indeed, both $\widetilde{\mathscr{M}}_0 u_0$ and $\widetilde{\mathscr{V}}\phi$ satisfy the heat equation, where $\widetilde{\mathscr{M}}_0 u_0$ restricted to $\{0\}\times\Omega$ coincides with u_0 and $\widetilde{\mathscr{V}}\phi$ vanishes there. To satisfy the Dirichlet boundary conditions, one has to solve

$$\mathscr{V}\phi = u_D - \mathscr{M}_0 u_0, \tag{2.7}$$

i.e., (1.1) with $f := u_D - \mathscr{M}_0 u_0$. While (2.5) is called direct method as it directly provides the physically relevant quantity $\phi_N = \partial_{\mathbf{n}}u$, (2.7) is called indirect method.

For more details and proofs, we refer to the seminal works [3,20,4], which considered $u_0 = 0$, and to [7,8] for the general case.

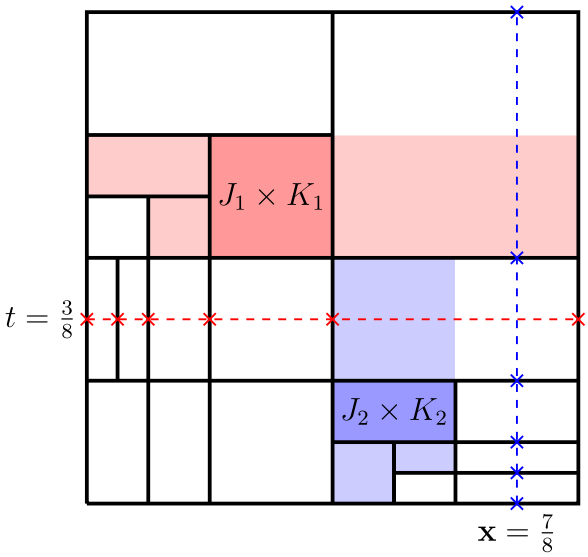

Fig. 2.1. Prismatic mesh $\mathcal{P}$ for $\Gamma = [0,1]$ and $T = 1$. The dashed blue and red lines indicate the meshes $\mathcal{P}|_t$ for $t = \frac{3}{8}$ and $\mathcal{P}|_{\mathbf{x}}$ for $\mathbf{x} = \frac{7}{8}$, respectively, where the corresponding elements are limited by crosses. For the elements $J_1 \times K_1 = [\frac{1}{2}, \frac{3}{4}] \times [\frac{1}{8}, \frac{1}{4}]$ and $J_2 \times K_2 = [\frac{1}{4}, \frac{1}{2}] \times [\frac{1}{2}, \frac{3}{4}]$, the corresponding integration domains $\bigcup_{\substack{\tilde{J}\times\tilde{K}\in\mathcal{P} \\ |J_1\cap\tilde{J}|>0 \\ K_1\cap\tilde{K}\neq\emptyset}} (J_1 \cap \tilde{J}) \times (K_1 \cup \tilde{K})$ and $\bigcup_{\substack{\tilde{J}\times\tilde{K}\in\mathcal{P} \\ J_2\cap\tilde{J}\neq\emptyset \\ |K_2\cap\tilde{K}|>0}} (J_2 \cup \tilde{J}) \times (K_2 \cap \tilde{K})$ from (3.5) are highlighted in (light) red and (light) blue, respectively.

2.4. Boundary meshes

Throughout this work, we consider prismatic meshes $\mathcal{P}$ of Σ:

- $\mathcal{P}$ is a finite set of prisms of the form $P = J \times K$, where $J \subseteq \overline{I} = [0,T]$ is some non-empty compact interval and $K \subseteq \Gamma$ is the image of some compact Lipschitz domain[1] $\hat{K} \subset \mathbb{R}^{n-1}$ under some bi-Lipschitz mapping;
- for all $P, \tilde{P} \in \mathcal{P}$ with $P \neq \tilde{P}$, the intersection has measure zero on Σ;
- $\mathcal{P}$ is a partition of Σ, i.e., $\Sigma = \bigcup_{P\in\mathcal{P}} P$.

For arbitrary $t \in \overline{I}$ and $\mathbf{x} \in \Gamma$, we abbreviate the induced sets

$$\mathcal{P}|_t := \big\{ K \subseteq \Gamma \,:\, (\{t\} \times \Gamma) \cap (J \times K) \neq \emptyset \text{ for some } J \times K \in \mathcal{P} \big\}$$

and

$$\mathcal{P}|_{\mathbf{x}} := \big\{ J \subseteq \overline{I} \,:\, (\overline{I} \times \{\mathbf{x}\}) \cap (J \times K) \neq \emptyset \text{ for some } J \times K \in \mathcal{P} \big\};$$

see Fig. 2.1 for a visualization. For almost all $t \in \overline{I}$, $\mathcal{P}|_t$ is a mesh of Γ, i.e., a partition of Γ into finitely many compact Lipschitz domains such that the intersection of two distinct elements has measure zero on Γ. Similarly, for almost all $\mathbf{x} \in \Gamma$, $\mathcal{P}|_{\mathbf{x}}$ is a mesh of $\overline{I}$, i.e., a partition of $\overline{I}$ into finitely many non-empty compact intervals such that the intersection of two different intervals is at most a point.

Note that for one fixed prismatic mesh $\mathcal{P}$ there exist constants $C_{\mathrm{nei}} \geq 1$, $C_{\mathrm{dist}} \geq 1$, $C_{\mathrm{shape}} \geq 1$, and $C_{\mathrm{lqu}} \geq 1$ such that:

- for almost all $t \in \overline{I}$, the number of neighbors of an element in $\mathcal{P}|_t$ is bounded, i.e.,

$$\#\big\{ \tilde{K} \in \mathcal{P}|_t \,:\, K \cap \tilde{K} \neq \emptyset \big\} \leq C_{\mathrm{nei}} \quad \text{for all } K \in \mathcal{P}|_t. \tag{2.8}$$

- for almost all $t \in \overline{I}$, the elements of $\mathcal{P}|_t$ are uniformly away from non-neighboring elements, i.e.,

$$\operatorname{diam}(K) \leq C_{\mathrm{dist}} \operatorname{dist}(K, \tilde{K}) \quad \text{for all } K, \tilde{K} \in \mathcal{P}|_t \text{ with } K \cap \tilde{K} = \emptyset; \tag{2.9}$$

- for almost all $t \in \overline{I}$, the elements of $\mathcal{P}|_t$ are shape-regular, i.e.,

$$C_{\mathrm{shape}}^{-1} |K| \leq \operatorname{diam}(K)^{n-1} \leq C_{\mathrm{shape}} |K| \quad \text{for all } K \in \mathcal{P}|_t; \tag{2.10}$$

- for almost all $\mathbf{x} \in \Gamma$, $\mathcal{P}|_{\mathbf{x}}$ is locally quasi-uniform, i.e.,

$$|J| \leq C_{\mathrm{lqu}} |\tilde{J}| \quad \text{for all } J, \tilde{J} \in \mathcal{P}|_{\mathbf{x}} \text{ with } J \cap \tilde{J} \neq \emptyset. \tag{2.11}$$

In the remainder of this work, we will always indicate the dependence of estimates on these particular constants.

[1] A compact Lipschitz domain is the closure of a bounded Lipschitz domain. For $n = 2$, it is a compact interval with non-empty interior.

Remark 2.1. *If, for $n=2$, the meshes $\mathcal{P}|_t$ are found by iteratively bisecting some initial mesh and the level difference of neighboring elements is bounded by 1, then the constants from (2.8)–(2.10) depend only on the initial mesh; cf. [1]. For $n=3$, the same holds true if the initial mesh is for instance a conforming (curvilinear) triangulation of Γ and one iteratively applies newest vertex bisection. The arguments for (2.8)–(2.9) are found in [2, Section 2.3 and 4.1].*

2.5. Boundary element method

Given a prismatic boundary mesh $\mathcal{P}$ and an associated finite-dimensional trial space $\mathcal{X} \subset H^{-1/2,-1/4}(\Sigma)$, e.g., the space of all $\mathcal{P}$-piecewise polynomials of some fixed degree in space and time, let $\Phi \in \mathcal{X}$ denote the Galerkin discretization of the solution ϕ of the boundary integral equation (1.1), i.e.,

$$\langle \mathscr{V}\Phi, \Psi\rangle_\Sigma = \langle f, \Psi\rangle_\Sigma \quad \text{for all } \Psi \in \mathcal{X}, \tag{2.12}$$

which is equivalent to the Galerkin orthogonality

$$\langle \mathscr{V}(\phi - \Phi), \Psi\rangle_\Sigma = 0 \quad \text{for all } \Psi \in \mathcal{X}. \tag{2.13}$$

Note that coercivity (2.6) guarantees unique solvability of the latter equations, and the Céa lemma applies

$$\|\phi - \Phi\|_{H^{-1/2,-1/4}(\Sigma)} \le \frac{C_{\mathrm{cont}}}{c_{\mathrm{coe}}} \min_{\Psi\in\mathcal{X}} \|\phi - \Psi\|_{H^{-1/2,-1/4}(\Sigma)}, \tag{2.14}$$

where C_{cont} is the operator norm of $\mathscr{V} : H^{-1/2,-1/4}(\Sigma) \to H^{1/2,1/4}(\Sigma)$. Suppose $\mathcal{P} = \{J \times K : J \in \mathcal{P}_{\overline{I}}, K \in \mathcal{P}_\Gamma\}$ is a full tensor-mesh corresponding to a mesh $\mathcal{P}_\Gamma$ of Γ with uniform mesh-size $h_{\mathbf{x}} \eqsim \operatorname{diam}(K)$ for all $K \in \mathcal{P}_\Gamma$ and a mesh $\mathcal{P}_{\overline{I}}$ of $\overline{I}$ with uniform step-size $h_t \eqsim h_{\mathbf{x}}^\sigma$ for some $\sigma > 0$. Using $\mathcal{P}$-piecewise polynomials of some degree $p_{\mathbf{x}} \in \mathbb{N}_0$ in space- and some degree $p_t \in \mathbb{N}_0$ in time-direction as trial space $\mathcal{X}$, then gives the error decay rate

$$\min_{\Psi\in\mathcal{X}} \|\phi - \Psi\|_{H^{-1/2,-1/4}(\Sigma)} \lesssim N^{-\frac{\min\{p_{\mathbf{x}}+3/2,(p_t+5/4)\sigma\}}{n-1+\sigma}} \quad \text{for all smooth } \phi; \tag{2.15}$$

see [5, Theorem 3.3]. Here, $N \eqsim h_{\mathbf{x}}^{-(n-1)} h_t^{-1} = h_{\mathbf{x}}^{n-1+\sigma}$ denotes the number of degrees of freedom. The optimal grading parameter is thus given by $\sigma = (p_{\mathbf{x}} + \frac{3}{2})/(p_t + \frac{5}{4})$ with resulting rate $\mathcal{O}\big(N^{-\frac{p_{\mathbf{x}}+3/2}{n-1+\sigma}}\big)$.

3. A posteriori error estimation

As $\mathscr{V}$ is an isomorphism, it holds that

$$\|\phi - \Phi\|_{H^{-1/2,-1/4}(\Sigma)} \eqsim \|f - \mathscr{V}\Phi\|_{H^{1/2,1/4}(\Sigma)}. \tag{3.1}$$

Here, $\Phi \in H^{-1/2,-1/4}(\Sigma)$ can be an arbitrary approximation of the solution ϕ of (1.1). While the right-hand side is in principle *a posteriori* computable, the computation of the Sobolev–Slobodeckij norm over the full space-time boundary Σ is expensive, and it does not provide any information on where to locally refine the given mesh to increase the accuracy of the approximation. According to (3.1), it is sufficient to derive suitable estimate for the residual $f - \mathscr{V}\Phi$ in the $H^{1/2,1/4}(\Sigma)$-norm. Recall that this term is $L_2(\Sigma)$-orthogonal to all functions $\Psi \in \mathcal{X}$ provided that Φ is the Galerkin approximation of ϕ in $\mathcal{X}$; see (2.13).

3.1. Localization of the anisotropic Sobolev–Slobodeckij norm

The following proposition provides the key argument for our *a posteriori* error estimation. While the first inequality is trivial, the original version of the second one already goes back to [10,11]. We make use of the slightly generalized version from [15, Lemma 4.5]; see [12, Lemma 5.3.2] for a detailed proof.

Proposition 3.1. *Let $\mu \in (0,1)$ and $\mathcal{P}_\Gamma$ be a mesh of Γ. Then, there exist constants $C_1, C_2 > 0$ such that for all $v \in H^\mu(\Gamma)$, there holds that*

$$C_1^{-1} \sum_{K\in\mathcal{P}_\Gamma} \sum_{\substack{\tilde K\in\mathcal{P}_\Gamma\\ K\cap\tilde K\neq\emptyset}} |v|^2_{H^\mu(K\cup\tilde K)} \le \|v\|^2_{H^\mu(\Gamma)} \le \sum_{K\in\mathcal{P}_\Gamma} \sum_{\substack{\tilde K\in\mathcal{P}_\Gamma\\ K\cap\tilde K\neq\emptyset}} |v|^2_{H^\mu(K\cup\tilde K)} + C_2 \sum_{K\in\mathcal{P}_\Gamma} \operatorname{diam}(K)^{-2\mu}\|v\|^2_{L_2(K)}. \tag{3.2}$$

The constant C_1 is given as $C_1 = 2(C_{\mathrm{nei}}+1)^2$ with C_{nei} from (2.8) (with $\mathcal{P}|_t$ replaced by $\mathcal{P}_\Gamma$), and C_2 depends only on the dimension n, μ, Γ, and the constant C_{dist} from (2.9) (with $\mathcal{P}|_t$ replaced by $\mathcal{P}_\Gamma$). □

Note that local quasi-uniformity (2.11) (with $\mathcal{P}|_{\mathbf{x}}$ replaced by $\mathcal{P}_{\overline{I}}$) of a time mesh $\mathcal{P}_{\overline{I}}$ is actually equivalent to

$$\operatorname{diam}(J) = |J| \le C_{\mathrm{lqu}} \operatorname{dist}(J,\tilde J) \quad \text{for all } J,\tilde J \in \mathcal{P}_{\overline{I}} \text{ with } J\cap\tilde J = \emptyset. \tag{3.3}$$

Moreover, for any element $J \in \mathcal{P}_{\overline{I}}$, there are at most three $\tilde J \in \mathcal{P}_{\overline{I}}$ with $J \cap \tilde J \neq \emptyset$. In particular, the same reference as before applies and we also obtain the following proposition.

Proposition 3.2. *Let $\nu \in (0,1)$ and $\mathcal{P}_{\overline{I}}$ be a mesh of $\overline{I}$. Then, there exist constants $C_1, C_2 > 0$ such that for all $v \in H^\nu(I)$, there holds that*

$$C_1^{-1} \sum_{J\in\mathcal{P}_{\overline{I}}} \sum_{\substack{\tilde J\in\mathcal{P}_{\overline{I}}\\ J\cap\tilde J\neq\emptyset}} |v|^2_{H^\nu(J\cup\tilde J)} \le \|v\|^2_{H^\nu(I)} \le \sum_{J\in\mathcal{P}_{\overline{I}}} \sum_{\substack{\tilde J\in\mathcal{P}_{\overline{I}}\\ J\cap\tilde J\neq\emptyset}} |v|^2_{H^\nu(J\cup\tilde J)} + C_2 \sum_{J\in\mathcal{P}_{\overline{I}}} |J|^{-2\nu}\|v\|^2_{L_2(J)}. \tag{3.4}$$

The constant C_1 is given as $C_1 = 32$, and C_2 depends only on ν, $|I|$, and the constant C_{lqu} from (2.11) (with $\mathcal{P}|_{\mathbf{x}}$ replaced by $\mathcal{P}_{\overline{I}}$). □

The latter two propositions allow to derive the following *a posteriori error estimation*, which can be employed for arbitrary approximations Φ.

Theorem 3.3. *Let $\mu, \nu \in (0,1)$ and $\mathcal{P}$ be a prismatic mesh of Σ. Then, there exist constants $C'_{\mathrm{eff}}, C''_{\mathrm{rel}} > 0$ such that for all $v \in H^{\mu,\nu}(\Sigma)$, there holds that*

$$\sum_{J\times K\in\mathcal{P}} \Bigg(\sum_{\substack{\tilde{J}\times\tilde{K}\in\mathcal{P}\\ |J\cap\tilde{J}|>0\\ K\cap\tilde{K}\neq\emptyset}} |v|^2_{L_2(J\cap\tilde{J};H^\mu(K\cup\tilde{K}))} + \sum_{\substack{\tilde{J}\times\tilde{K}\in\mathcal{P}\\ J\cap\tilde{J}\neq\emptyset\\ |K\cap\tilde{K}|>0}} |v|^2_{H^\nu(J\cup\tilde{J};L_2(K\cap\tilde{K}))} \Bigg) \leq (C'_{\mathrm{eff}})^2 \|v\|^2_{H^{\mu,\nu}(\Sigma)} \tag{3.5}$$

as well as

$$(C'_{\mathrm{rel}})^{-2}\|v\|^2_{H^{\mu,\nu}(\Sigma)} \leq \sum_{J\times K\in\mathcal{P}} \Bigg(\sum_{\substack{\tilde{J}\times\tilde{K}\in\mathcal{P}\\ |J\cap\tilde{J}|>0\\ K\cap\tilde{K}\neq\emptyset}} |v|^2_{L_2(J\cap\tilde{J};H^\mu(K\cup\tilde{K}))} + \sum_{\substack{\tilde{J}\times\tilde{K}\in\mathcal{P}\\ J\cap\tilde{J}\neq\emptyset\\ |K\cap\tilde{K}|>0}} |v|^2_{H^\nu(J\cup\tilde{J};L_2(K\cap\tilde{K}))} \Bigg) + \sum_{J\times K\in\mathcal{P}} \big(\operatorname{diam}(K)^{-2\mu} + |J|^{-2\nu}\big)\|v\|^2_{L_2(J\times K)}; \tag{3.6}$$

see Fig. 2.1 for a visualization of the involved integration domains. The constant C'_{eff} is given as $C'_{\mathrm{eff}} = \max(2(C_{\mathrm{nei}}+1)^2, 32)$ with C_{nei} from (2.8), and C'_{rel} depends only on n, μ, ν, Γ, $|I|$ and the constants C_{dist} from (2.9) as well as C_{lqu} from (2.11).

Proof. We split the proof into four steps.

Step 1: In this step, we bound $\|v\|_{L_2(I;H^\mu(\Gamma))}$ from below. Proposition 3.1 gives that

$$\|v\|^2_{L_2(I;H^\mu(\Gamma))} = \int_I \|v(t,\cdot)\|^2_{H^\mu(\Gamma)}\,\mathrm{d}t \gtrsim \int_I \sum_{K\in\mathcal{P}|_t} \sum_{\substack{\tilde{K}\in\mathcal{P}|_t\\ K\cap\tilde{K}\neq\emptyset}} |v(t,\cdot)|^2_{H^\mu(K\cup\tilde{K})}\,\mathrm{d}t.$$

Note that $K\in\mathcal{P}|_t$ is equivalent to $J\times K\in\mathcal{P}$ for some J with $t\in J$. With the indicator function $\mathbb{1}_S$ of a set S, the last term thus is equal to

$$\begin{aligned}\int_I \sum_{K\in\mathcal{P}|_t} \sum_{\substack{\tilde{K}\in\mathcal{P}|_t\\ K\cap\tilde{K}\neq\emptyset}} |v(t,\cdot)|^2_{H^\mu(K\cup\tilde{K})}\,\mathrm{d}t &= \int_I \sum_{J\times K\in\mathcal{P}} \mathbb{1}_J(t) \sum_{\substack{\tilde{J}\times\tilde{K}\in\mathcal{P}\\ K\cap\tilde{K}\neq\emptyset}} \mathbb{1}_{\tilde{J}}(t)|v(t,\cdot)|^2_{H^\mu(K\cup\tilde{K})}\,\mathrm{d}t\\ &= \sum_{J\times K\in\mathcal{P}} \sum_{\substack{\tilde{J}\times\tilde{K}\in\mathcal{P}\\ |J\cap\tilde{J}|>0\\ K\cap\tilde{K}\neq\emptyset}} |v|^2_{L_2(J\cap\tilde{J};H^\mu(K\cup\tilde{K}))}.\end{aligned}$$

Step 2: In this step, we bound $\|v\|_{L_2(I;H^\mu(\Gamma))}$ from above. Proposition 3.1 gives that

$$\begin{aligned}\|v\|^2_{L_2(I;H^\mu(\Gamma))} &= \int_I \|v(t,\cdot)\|^2_{H^\mu(\Gamma)}\,\mathrm{d}t\\ &\lesssim \int_I \sum_{K\in\mathcal{P}|_t} \sum_{\substack{\tilde{K}\in\mathcal{P}|_t\\ K\cap\tilde{K}\neq\emptyset}} |v(t,\cdot)|^2_{H^\mu(K\cup\tilde{K})} + \sum_{K\in\mathcal{P}|_t} \operatorname{diam}(K)^{-2\mu}\|v(t,\cdot)\|^2_{L_2(K)}\,\mathrm{d}t.\end{aligned} \tag{3.7}$$

The first term in (3.7) has already been treated in Step 1. As $K\in\mathcal{P}|_t$ is equivalent to $J\times K\in\mathcal{P}$ for some J with $t\in J$, the second term reads

$$\begin{aligned}\int_I \sum_{K\in\mathcal{P}|_t} \operatorname{diam}(K)^{-2\mu}\|v(t,\cdot)\|^2_{L_2(K)}\,\mathrm{d}t &= \int_I \sum_{J\times K\in\mathcal{P}} \mathbb{1}_J(t)\operatorname{diam}(K)^{-2\mu}\|v(t,\cdot)\|^2_{L_2(K)}\,\mathrm{d}t\\ &= \sum_{J\times K\in\mathcal{P}} \operatorname{diam}(K)^{-2\mu}\|v\|^2_{L_2(J\times K)}.\end{aligned}$$

Step 3: In this step, we bound $\|v\|_{H^\nu(I;L_2(\Gamma))}$ from below. The Fubini theorem, Proposition 3.2, and the same argument as in Step 1 give that

$$\begin{aligned}\|v\|^2_{H^\nu(I;L_2(\Gamma))} &= \int_\Gamma \|v(\cdot,\mathbf{x})\|^2_{H^\nu(I)}\,\mathrm{d}\mathbf{x} \gtrsim \int_\Gamma \sum_{J\in\mathcal{P}|_{\mathbf{x}}} \sum_{\substack{\tilde{J}\in\mathcal{P}|_{\mathbf{x}}\\ J\cap\tilde{J}\neq\emptyset}} |v(\cdot,\mathbf{x})|^2_{H^\nu(J\cup\tilde{J})}\,\mathrm{d}\mathbf{x}\\ &= \int_\Gamma \sum_{J\times K\in\mathcal{P}} \mathbb{1}_K(\mathbf{x}) \sum_{\substack{\tilde{J}\times\tilde{K}\in\mathcal{P}\\ J\cap\tilde{J}\neq\emptyset}} \mathbb{1}_{\tilde{K}}(\mathbf{x})|v(\cdot,\mathbf{x})|^2_{H^\nu(J\cup\tilde{J})}\,\mathrm{d}\mathbf{x}\\ &= \sum_{J\times K\in\mathcal{P}} \sum_{\substack{\tilde{J}\times\tilde{K}\in\mathcal{P}\\ J\cap\tilde{J}\neq\emptyset\\ |K\cap\tilde{K}|>0}} |v|^2_{H^\nu(J\cup\tilde{J};L_2(K\cap\tilde{K}))}.\end{aligned}$$

Step 4: In this step, we bound $\|v\|_{H^\nu(I;L_2(\Gamma))}$ from above. The Fubini theorem and Proposition 3.2 give that

$$\begin{aligned}\|v\|^2_{H^\nu(I;L_2(\Gamma))} &= \int_\Gamma \|v(\cdot,\mathbf{x})\|^2_{H^\nu(I)}\,\mathrm{d}\mathbf{x}\\ &\lesssim \int_\Gamma \sum_{J\in\mathcal{P}|_{\mathbf{x}}} \sum_{\substack{\tilde{J}\in\mathcal{P}|_{\mathbf{x}}\\ J\cap\tilde{J}\neq\emptyset}} |v(\cdot,\mathbf{x})|^2_{H^\nu(J\cup\tilde{J})} + \sum_{J\in\mathcal{P}|_{\mathbf{x}}} |J|^{-2\nu}\|v(\cdot,\mathbf{x})\|^2_{L_2(J)}\,\mathrm{d}\mathbf{x}.\end{aligned} \tag{3.8}$$

The first term in (3.8) has already been treated in Step 3. The second term reads

$$\int_\Gamma \sum_{J\in\mathcal{P}|_{\mathbf{x}}} |J|^{-2\nu}\|v(\cdot,\mathbf{x})\|^2_{L_2(J)}\,\mathrm{d}\mathbf{x} = \int_\Gamma \sum_{J\times K\in\mathcal{P}} \mathbb{1}_K(\mathbf{x})\,|J|^{-2\nu}\|v(\cdot,\mathbf{x})\|^2_{L_2(J)}\,\mathrm{d}\mathbf{x}$$
$$= \sum_{J\times K\in\mathcal{P}} |J|^{-2\nu}\|v\|^2_{L_2(J\times K)}.$$

This concludes the proof. □

3.2. Poincaré-type inequality

Assuming the grading $|J| \approx \operatorname{diam}(K)^{\mu/\nu}$ as well as $L_2(\Sigma)$-orthogonality of v to piecewise constants, the following local Poincaré-type inequality allows to get rid of the weighted L_2-terms in (3.6); see Corollary 3.5. The proof works essentially as in [4, Proposition 5.3], where a global version on uniform meshes is considered.

Lemma 3.4. *Let $\mu,\nu\in(0,1)$ and $\mathcal{P}$ be a prismatic mesh of Σ. Then, there holds for all $v\in H^{\mu,\nu}(\Sigma)$ and all $J\times K\in\mathcal{P}$ with $\langle v,1\rangle_{L_2(J\times K)}=0$ that*

$$\|v\|^2_{L_2(J\times K)} \le C_{\mathrm{shape}}\big(\operatorname{diam}(K)^{2\mu}|v|^2_{L_2(J;H^\mu(K))} + |J|^{2\nu}|v|^2_{H^\nu(J;L_2(K))}\big). \tag{3.9}$$

Here, $C_{\mathrm{shape}}\ge 1$ is the constant from (2.10).

Proof. Let Π_J, Π_K, and $\Pi_{J\times K}$ denote the L_2-orthogonal projection onto the space of constants on J, K, and $J\times K$, respectively. Note that $\Pi_{J\times K}=\Pi_J\otimes\Pi_K$ and thus

$$\begin{aligned}\|v\|_{L_2(J\times K)} &= \|(1-\Pi_{J\times K})v\|_{L_2(J\times K)}\\ &\le \|(1-\Pi_J\otimes\mathrm{Id})v\|_{L_2(J\times K)} + \|(\Pi_J\otimes\mathrm{Id}-\Pi_J\otimes\Pi_K)v\|_{L_2(J\times K)}.\end{aligned} \tag{3.10}$$

As Π_J has operator norm 1, a standard Poincaré-type inequality, see, e.g., [11, Lemma 3.4] for the elementary proof, shows for the second term in (3.10) that

$$\begin{aligned}\|(\Pi_J\otimes\mathrm{Id}-\Pi_J\otimes\Pi_K)v\|^2_{L_2(J\times K)} &\le \|(1-\mathrm{Id}\otimes\Pi_K)v\|^2_{L_2(J\times K)}\\ &= \int_J \|(1-\Pi_K)v(t,\cdot)\|^2_{L_2(K)}\,\mathrm{d}t\\ &\le \frac{\operatorname{diam}(K)^{n-1+2\mu}}{2|K|}\int_J |v(t,\cdot)|^2_{H^\mu(K)}\,\mathrm{d}t\\ &\le \frac{C_{\mathrm{shape}}}{2}\operatorname{diam}(K)^{2\mu}|v|^2_{L_2(J;H^\mu(K))}.\end{aligned}$$

The first term in (3.10) can be estimated similarly

$$\|(1-\Pi_J\otimes\mathrm{Id})v\|^2_{L_2(J\times K)} \le \frac{1}{2}|J|^{2\nu}|v|^2_{H^\nu(J;L_2(K))},$$

which concludes the proof. □

3.3. A posteriori error estimators

For arbitrary prismatic meshes $\mathcal{P}$ of Σ with some associated trial space $\mathcal{X}\subset H^{-1/2,-1/4}(\Sigma)$ and $\Phi\in\mathcal{X}$, we define the following error indicators for all $J\times K\in\mathcal{P}$,

$$\eta^{\mathbf{x}}_{\mathcal{P}}(\Phi,J\times K)^2 := \sum_{\substack{\tilde J\times\tilde K\in\mathcal{P}\\ |J\cap\tilde J|>0\\ K\cap\tilde K\neq\emptyset}} |f-\mathscr{V}\Phi|^2_{L_2(J\cap\tilde J;H^{1/2}(K\cup\tilde K))},$$
$$\eta^{t}_{\mathcal{P}}(\Phi,J\times K)^2 := \sum_{\substack{J\times\tilde K\in\mathcal{P}\\ J\cap\tilde J\neq\emptyset\\ |K\cap\tilde K|>0}} |f-\mathscr{V}\Phi|^2_{H^{1/4}(J\cup\tilde J;L_2(K\cap\tilde K))},$$
$$\zeta^{\mathbf{x}}_{\mathcal{P}}(\Phi,J\times K)^2 := \operatorname{diam}(K)^{-1}\|f-\mathscr{V}\Phi\|^2_{L_2(J\times K)},$$
$$\zeta^{t}_{\mathcal{P}}(\Phi,J\times K)^2 := |J|^{-1/2}\|f-\mathscr{V}\Phi\|^2_{L_2(J\times K)}.$$

The corresponding error estimators read as

$$\eta_{\mathcal{P}}(\Phi)^2 := \sum_{J\times K\in\mathcal{P}} \eta_{\mathcal{P}}(\Phi,J\times K)^2 \quad\text{with } \eta_{\mathcal{P}}(\Phi,J\times K)^2 := \eta^{\mathbf{x}}_{\mathcal{P}}(\Phi,J\times K)^2 + \eta^{t}_{\mathcal{P}}(\Phi,J\times K)^2,$$
$$\zeta_{\mathcal{P}}(\Phi)^2 := \sum_{J\times K\in\mathcal{P}} \zeta_{\mathcal{P}}(\Phi,J\times K)^2 \quad\text{with } \zeta_{\mathcal{P}}(\Phi,J\times K)^2 := \zeta^{\mathbf{x}}_{\mathcal{P}}(\Phi,J\times K)^2 + \zeta^{t}_{\mathcal{P}}(\Phi,J\times K)^2.$$

With (3.1), we overall obtain the following *a posteriori* estimates.

Corollary 3.5. *Let ϕ be the solution of (1.1) and $\mathcal{P}$ be a prismatic mesh of Σ with some associated discrete trial space $\mathcal{X} \subset H^{-1/2,-1/4}(\Sigma)$. Then, there exist constants $C_{\mathrm{eff}}, \tilde{C}_{\mathrm{rel}} > 0$ such that for arbitrary $\Phi \in \mathcal{X}$, there holds that*

$$C_{\mathrm{eff}}^{-1}\eta_{\mathcal{P}}(\Phi) \le \|\phi - \Phi\|_{H^{-1/2,-1/4}(\Sigma)} \le \tilde{C}_{\mathrm{rel}}\big(\eta_{\mathcal{P}}(\Phi)^2 + \zeta_{\mathcal{P}}(\Phi)^2\big)^{1/2}. \tag{3.11}$$

If the space $\mathcal{X}$ contains all $\mathcal{P}$-piecewise constant functions and $\Phi \in \mathcal{X}$ is the Galerkin approximation of ϕ, there further holds that

$$\zeta_{\mathcal{P}}(\Phi, J\times K)^2 \le C_{\mathrm{shape}}\big(\mathrm{diam}(K)^{-1} + |J|^{-1/2}\big)\big(\mathrm{diam}(K) + |J|^{1/2}\big)\eta_{\mathcal{P}}(\Phi, J\times K)^2 \tag{3.12}$$

for all $J \times K \in \mathcal{P}$. If $C_{\mathrm{grad}}^{-1}\mathrm{diam}(K) \le |J|^{1/2} \le C_{\mathrm{grad}}\mathrm{diam}(K)$ is satisfied for all $J \times K \in \mathcal{P}$ and a uniform constant $C_{\mathrm{grad}} \ge 1$, this implies the existence of a constant $C_{\mathrm{rel}} > 0$ such that

$$\|\phi - \Phi\|_{H^{-1/2,-1/4}(\Sigma)} \le C_{\mathrm{rel}}\eta_{\mathcal{P}}(\Phi). \tag{3.13}$$

The constants C_{eff} and $\tilde{C}_{\mathrm{rel}}$ are given as $C_{\mathrm{eff}} = C'_{\mathrm{eff}}C_{\mathrm{cont}}$ and $C_{\mathrm{rel}} = C'_{\mathrm{rel}}/c_{\mathrm{coe}}$ with C'_{eff} and C'_{rel} from Theorem 3.3, the operator norm C_{cont} of $\mathcal{V}$, and c_{coe} from (2.6). The constant C_{rel} is given as $C_{\mathrm{rel}} = \tilde{C}_{\mathrm{rel}}\sqrt{2C_{\mathrm{shape}}(1 + C_{\mathrm{grad}})}$. □

Remark 3.6. *According to (2.15), the required scaling $|J| \approx \mathrm{diam}(K)^2$, i.e., $\sigma = 2$, is the optimal scaling for approximating smooth solutions ϕ if the polynomial degrees of $\mathcal{X}$ satisfy $p_{\mathbf{x}} = 2p_t + 1$.*

4. Numerical experiments

In this section, we employ the error estimator η within an adaptive algorithm using different refinement strategies, and investigate the resulting convergence rates. We restrict ourselves to the case $n = 2$, with $\Gamma = \partial\Omega$ being the boundary of a polygonal domain $\Omega \subset \mathbb{R}^2$, and set the time domain to be $I = (0, 1)$. In all considered examples, we employ the indirect approach (2.7) (coinciding with the direct approach (2.5) in Section 4.3 and 4.6 due to homogeneous Dirichlet data).

For $\mathcal{P}$ a prismatic mesh of the space-time boundary, i.e., a quadrilateral mesh as $n = 2$, we consider the trial space $\mathcal{X}$ of piecewise constants with respect to $\mathcal{P}$. In particular, this allows us to perform integration in time analytically for all integrals that are involved in the computation of the Galerkin matrix and the evaluation of the single-layer operator $\mathcal{V}$; see, e.g., [4]. The remaining integrals over Γ have a logarithmic singularity, for which we use the quadrature rules from [22]. For the computation of the Sobolev–Slobodeckij seminorm in the Faermann estimator $\eta_{\mathcal{P}}(\Phi)$, we use Duffy transformations and Gauss quadrature for the regularized integrands. Further details on the numerical computation of the involved singular integrals are found in Appendix A. The source code that we used to generate the numerical results is available at [16].

4.1. Adaptive algorithm

In our numerical experiments below, we employ the following adaptive algorithm with $\theta = 0.9$.

Algorithm 4.1. *Let $0 < \theta \le 1$ be a marking parameter and $\mathcal{P} = \{J \times K : J \in \mathcal{P}_{\overline{I}}, K \in \mathcal{P}_\Gamma\}$ be an initial tensor-mesh corresponding to a mesh $\mathcal{P}_\Gamma$ of Γ and a mesh $\mathcal{P}_{\overline{I}}$ of $\overline{I} = [0, T]$. For each $\ell = 0, 1, 2, \dots$, iterate the following steps:*

(i) *Compute Galerkin approximation Φ_ℓ of ϕ in the space $\mathcal{X}_\ell$ of all $\mathcal{P}_\ell$-piecewise constant functions on Σ.*
(ii) *Compute indicators $\eta^{\mathbf{x}}_{\mathcal{P}_\ell}(\Phi_\ell, J\times K)$ and $\eta^{t}_{\mathcal{P}_\ell}(\Phi_\ell, J\times K)$ for all elements $J \times K \in \mathcal{P}_\ell$.*
(iii) *Determine two minimal sets of marked elements $\mathcal{M}^{\mathbf{x}}_\ell, \mathcal{M}^{t}_\ell \subseteq \mathcal{P}_\ell$ such that*

$$\theta^2\eta_{\mathcal{P}_\ell}(\Phi_\ell)^2 \le \sum_{J\times K\in\mathcal{M}^{\mathbf{x}}_\ell}\eta^{\mathbf{x}}_{\mathcal{P}_\ell}(\Phi_\ell, J\times K)^2 + \sum_{J\times K\in\mathcal{M}^{t}_\ell}\eta^{t}_{\mathcal{P}_\ell}(\Phi_\ell, J\times K)^2. \tag{4.1}$$

(iv) *Refine at least all marked elements to obtain a new mesh $\mathcal{P}_{\ell+1}$.*

We will focus on *isotropic* and *anisotropic* adaptive strategies:

- In isotropic refinement, we require $\mathcal{M}^{\mathbf{x}}_\ell = \mathcal{M}^{t}_\ell$ in the marking step (iii), so that (4.1) simplifies to $\theta^2\eta_{\mathcal{P}_\ell}(\Phi_\ell)^2 \le \sum_{J\times K\in\mathcal{M}_\ell}\eta_{\mathcal{P}_\ell}(\Phi_\ell, J\times K)^2$. In the refinement step (iv), we iteratively mark additional elements to ensure that, after subdividing all marked elements into four congruent rectangles, the new mesh $\mathcal{P}_{\ell+1}$ has only one hanging node per edge.
- In anisotropic refinement, we bisect the elements $\mathcal{M}^{\mathbf{x}}_\ell \setminus \mathcal{M}^{t}_\ell$ in space, bisect the elements $\mathcal{M}^{t}_\ell \setminus \mathcal{M}^{\mathbf{x}}_\ell$ in time, and subdivide all elements $\mathcal{M}^{\mathbf{x}}_\ell \cap \mathcal{M}^{t}_\ell$ into four congruent rectangles. Then, we iteratively bisect additional elements in space and/or time to ensure that the level difference in space and in time between elements sharing an edge in the new mesh $\mathcal{P}_{\ell+1}$ is bounded by 1. Here, the level in space and the level in time of elements are defined as the number of bisections in space and time, respectively, to obtain the element from the initial mesh $\mathcal{P}_0$.

For comparison, we also include uniform refinement, where $\mathcal{P}_{\ell+1}$ is obtained from $\mathcal{P}_\ell$ by subdividing each element into four congruent rectangles. For all considered refinement strategies, it is easy to see that the mesh constants from (2.8)–(2.11) corresponding to $(\mathcal{P}_\ell)_{\ell\in\mathbb{N}_0}$ depend only on the initial mesh $\mathcal{P}_0$.

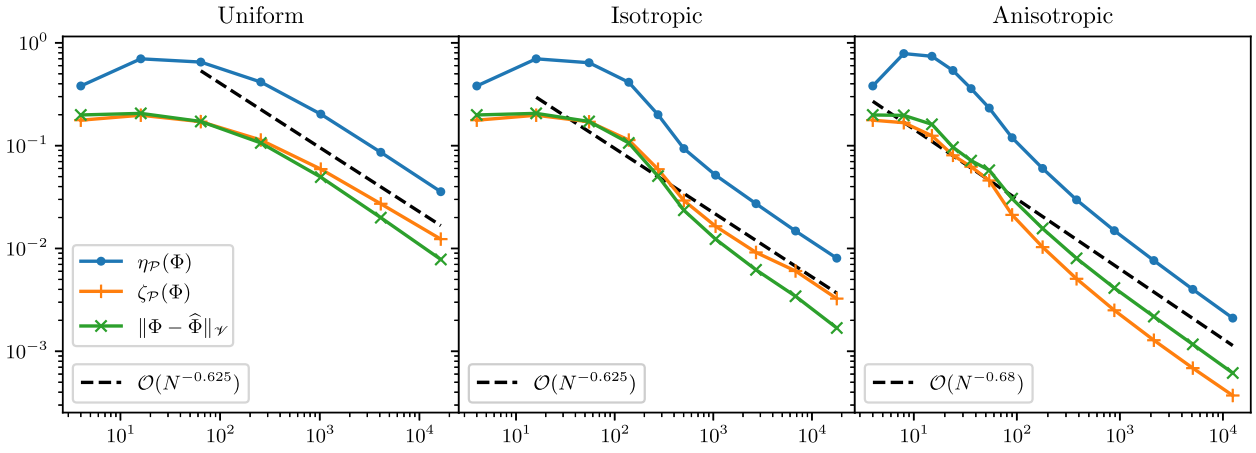

Fig. 4.1. Error estimators for the smooth problem of Section 4.3 plotted double-logarithmically over the degrees of freedom $N = \#\mathcal{P}$: uniform refinement (left), isotropic refinement (middle), and anisotropic refinement (right).

4.2. Reference for exact error

As the exact error $\|\phi - \Phi\|_{H^{-1/2,-1/4}(\Sigma)}$ cannot be readily computed in the examples below, we compare the error estimator $\eta_{\mathcal{P}}$ and the weighted L_2-terms $\zeta_{\mathcal{P}}$ from Section 3.3 with the following $(h - h/2)$-estimator: For a mesh $\mathcal{P}$, define the uniformly refined mesh as $\widehat{\mathcal{P}}$. With the Galerkin approximation $\widehat{\Phi}$ from the refined trial space, we define the $(h - h/2)$-estimator as

$$\|\Phi - \widehat{\Phi}\|_{\mathscr{Y}}^2 := \langle \mathscr{V}(\Phi - \widehat{\Phi}), \Phi - \widehat{\Phi} \rangle_\Sigma.$$

Under the saturation assumption $\|\phi - \widehat{\Phi}\|_{\mathscr{Y}} \le q_{\mathrm{sat}} \|\phi - \Phi\|_{\mathscr{Y}}$, the triangle inequality shows that this estimator is equivalent to $\|\phi - \Phi\|_{\mathscr{Y}}$, and therefore to the error $\|\phi - \Phi\|_{H^{-1/2,-1/4}(\Sigma)}$ by coercivity of $\mathscr{V}$. Note that the saturation assumption is indeed satisfied under the realistic (asymptotic) assumption that $\|\phi - \Phi\|_{\mathscr{Y}} = \mathcal{O}((\#\mathcal{P})^{-s})$ for some arbitrary rate $s > 0$.

4.3. Smooth problem

Let $\Omega = (0,1)^2$ and consider the (smooth) solution $u(t, x_1, x_2) := \exp(-2\pi^2 t)\sin(\pi x_1)\sin(\pi x_2)$ with initial condition $u_0(x_1, x_2) := \sin(\pi x_1)\sin(\pi x_2)$ and Dirichlet data $u_D \equiv 0$. We choose $\mathcal{P}_0 := \{[0,1] \times K : K \in \mathcal{P}_\Gamma\}$ with the uniform mesh $\mathcal{P}_\Gamma$ of Γ being aligned with the corners and consisting of four elements, as initial mesh of the space-time boundary Σ.

Fig. 4.1 displays the results in double-logarithmic plots so that the slopes of the lines indicate the corresponding convergence rates. With the number of degrees of freedom $N = \#\mathcal{P}$, we see the expected rate $\mathcal{O}(N^{-5/8}) = \mathcal{O}(N^{-0.625})$ from (2.15) for both uniform refinement and isotropic refinement (with still slightly worse rate for the weighted L_2-terms $\zeta_{\mathcal{P}}(\Phi)$ for uniform refinement), albeit adaptive isotropic refinement offers quantitively better results. For anisotropic refinement, the rate is improved to $\mathcal{O}(N^{-15/22}) \approx \mathcal{O}(N^{-0.68})$. According to (2.15), this coincides with the best possible rate that can be achieved with uniform tensor-meshes, where the optimal scaling parameter in $h_t \approx h_{\mathbf{x}}^\sigma$ is given by $\sigma = 6/5$. Note that we do not require setting an explicit scaling in our anisotropic adaptive algorithm, it recovers the optimal rate automatically.

4.4. Mildly singular problem

Let $\Omega = (0,1)^2$, with initial condition $u_0 \equiv 0$ and Dirichlet data $u_D(t, x_1, x_2) := t^2$. We expect the solution here to be only singular in the four corners of the unit square as the initial condition is compatible with the Dirichlet data. The initial mesh $\mathcal{P}_0$ is chosen as in Section 4.3. Fig. 4.2 displays the results. The asymptotic decay rate for all estimators under uniform refinement seems to be $\mathcal{O}(N^{-1/3})$, which is improved to $\mathcal{O}(N^{-1/2})$ for isotropic refinement, and finally, under anisotropic refinement this becomes the optimal rate $\mathcal{O}(N^{-15/22})$.

4.5. Singular problem

Let $\Omega = (0,1)^2$ with initial condition $u_0 \equiv 0$ and Dirichlet data $u_D \equiv 1$. The solution to this problem is known to have a strong singularity for $t = 0$ due to the incompatibility of initial and boundary conditions, in addition to singularities in the four corners of the unit square. The initial mesh $\mathcal{P}_0$ is chosen as in Section 4.3.

Fig. 4.3 displays the results. The Faermann estimator $\eta_{\mathcal{P}}(\Phi)$ and the $(h - h/2)$-estimator $\|\Phi - \widehat{\Phi}\|_{\mathscr{Y}}$ show both the same sensible convergence behavior for this problem. For uniform refinement, they display the rate $\mathcal{O}(N^{-1/8})$, which is then improved by isotropic refinement to $\mathcal{O}(N^{-1/4})$. Finally, for anisotropic refinement, they achieve the best possible rate $\mathcal{O}(N^{-15/22})$, recovering the rate for a smooth problem. Looking at Fig. 4.4, we see strong anisotropic refinement towards $t = 0$ with elements of size $h_{\mathbf{x}} = 1, h_t = 2^{-18}$, and some mild refinement towards the corners of the unit square.

On the other hand, the weighted L_2-terms $\zeta_{\mathcal{P}}(\Phi)$ do not seem to decay for uniform or isotropic refinement, and seem to degenerate for anisotropic refinement. This is problematic for the reliability bound in Corollary 3.5. Further inspection suggests that this is a theoretical shortcoming rather than a practical one. This is hinted by the $(h - h/2)$-estimator, which one generally assumes to be reliable. Note that this does not contradict the theoretical results from Corollary 3.5, which states $\zeta_{\mathcal{P}}(\Phi) \lesssim \eta_{\mathcal{P}}(\Phi) \lesssim \|\phi - \Phi\|_{H^{-1/2,-1/4}(\Sigma)}$ only under the additional parabolic scaling assumption $h_t \approx h_{\mathbf{x}}^2$ for all space-time elements.

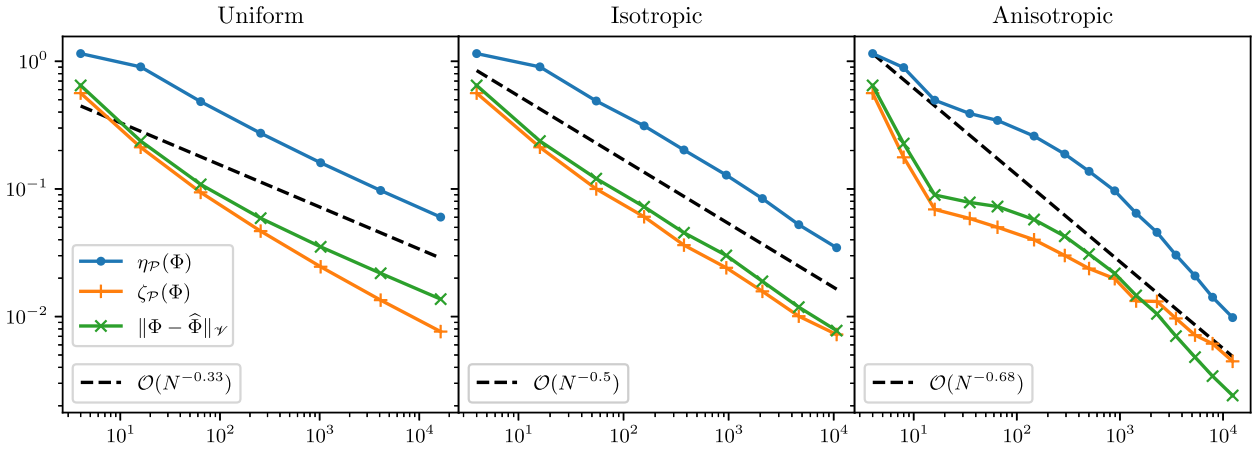

Fig. 4.2. Error estimators for the mildly singular problem of Section 4.4 plotted double-logarithmically over the degrees of freedom $N = \#\mathcal{P}$: uniform refinement (left), isotropic refinement (middle), and anisotropic refinement (right).

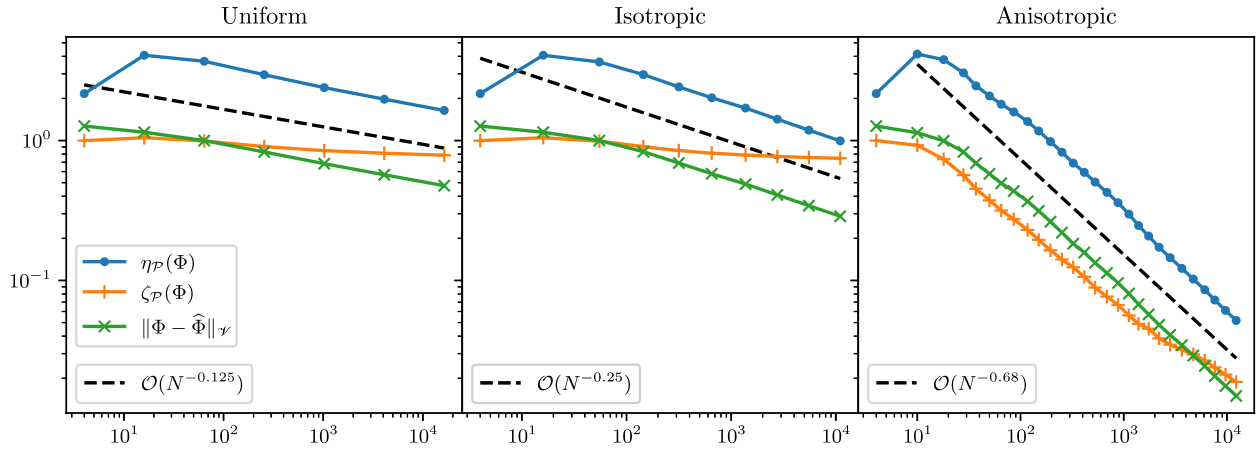

Fig. 4.3. Error estimators for the singular problem of Section 4.5 plotted double-logarithmically over the degrees of freedom $N = \#\mathcal{P}$: uniform refinement (left), isotropic refinement (middle), and anisotropic refinement (right).

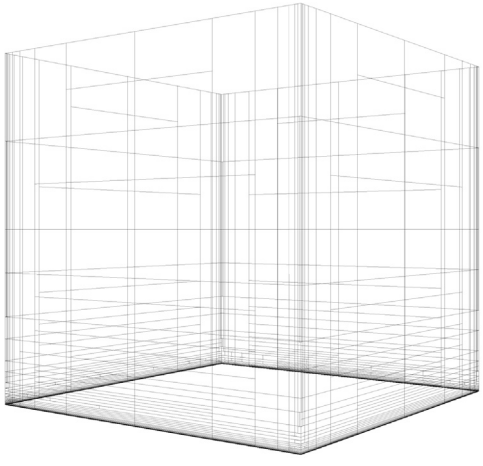

Fig. 4.4. Mesh with $N = 1391$ elements, generated by anisotropic refinement for the singular problem of Section 4.5.

Under this parabolic scaling assumption, the optimal error decay rate for smooth problems becomes $\mathcal{O}(N^{-1/2})$; see (2.15). Fig. 4.5 displays the results of uniform and adaptive refinement, with meshes that satisfy this scaling constraint,[2] providing convergence rates $\mathcal{O}(N^{-0.18})$ and $\mathcal{O}(N^{-0.4})$, respectively, for *all* considered estimators.

[2] For uniform refinement, all elements are split into eight new elements by bisecting once in space-direction and three times in time-direction. For adaptive refinement, we assume $\mathcal{M}_\ell^{\mathbf{x}} = \mathcal{M}_\ell^t$. All these marked elements are subdivided into four congruent rectangles, where we use additional bisections in space and/or time to guarantee that the level differences between elements sharing an edge are bounded by 1 and that $\frac{1}{2}h_{\mathbf{x}}^2 \le h_t \le 2h_{\mathbf{x}}^2$.

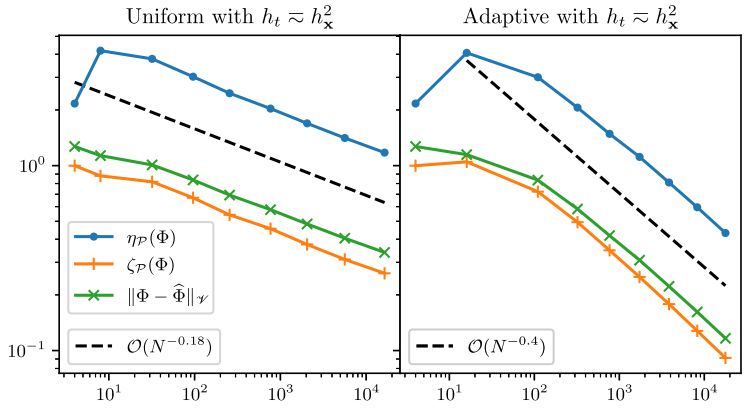

Fig. 4.5. Error estimators for the singular problem of Section 4.5 plotted double-logarithmically over the degrees of freedom $N = \#\mathcal{P}$: uniform refinement (left) and adaptive refinement (right) with parabolic scaling $h_t \eqsim h_{\mathbf{x}}^2$.

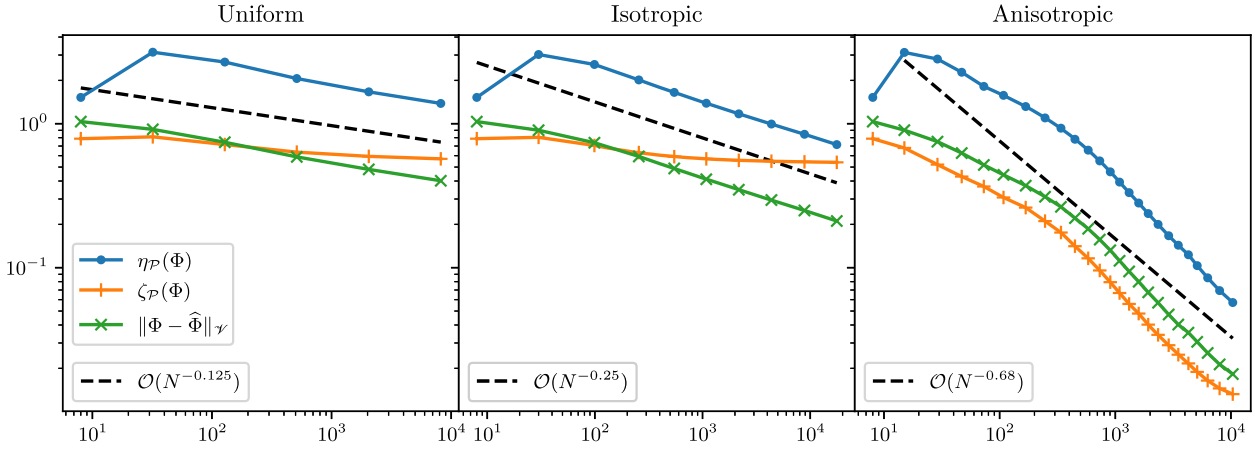

Fig. 4.6. Error estimators for the singular L-shape problem of Section 4.6 plotted double-logarithmically over the degrees of freedom $N = \#\mathcal{P}$: uniform refinement (left), isotropic refinement (middle), and anisotropic refinement (right).

4.6. Singular L-shape problem

We consider the L-shaped domain $\Omega := (-1,1)^2 \setminus [-1,0]^2$ with data $u_0 \equiv 1$ and $u_D \equiv 0$. The solution has a strong singularity for $t = 0$, in addition to a singularity at the re-entrant corner $(0,0)$. We choose $\mathcal{P}_0 := \{[0,1] \times K : K \in \mathcal{P}_\Gamma\}$, with the uniform mesh $\mathcal{P}_\Gamma$ of Γ being aligned with the corners and consisting of eight elements, as initial mesh of the space-time boundary Σ. Fig. 4.6 displays the results, which are similar to those of Section 4.5 with a better behavior of the weighted L_2-terms $\zeta_{\mathcal{P}}(\Phi)$ for anisotropic refinement.

Enforcing the parabolic scaling $h_t \eqsim h_{\mathbf{x}}^2$ as in Section 4.5, *all* estimators converge again with the same rates, being $\mathcal{O}(N^{-0.18})$ for uniform refinement and $\mathcal{O}(N^{-0.45})$ for adaptive refinement (not displayed).

Acknowledgement

The first author has been supported by the Austrian Science Fund (FWF) under grant J4379-N. The second author has been supported by the Netherlands Organisation for Scientific Research (NWO) under contract no. 613.001.652.

Appendix A. Numerical computation

Let $\Omega \subset \mathbb{R}^2$ be a simply connected Lipschitz domain and $\gamma : [0, L] \to \Gamma$ be a parametrization of its boundary $\Gamma = \partial\Omega$. For a given prismatic mesh $\mathcal{P}$, i.e., a quadrilateral mesh, of the space-time boundary Σ, we briefly explain how to numerically compute the corresponding Galerkin solution Φ of (2.12) in the trial space of piecewise constants with respect to $\mathcal{P}$ as well as the corresponding error estimator $\eta_{\mathcal{P}}(\Phi)$ and the weighted L_2-terms $\zeta_{\mathcal{P}}(\Phi)$. For all $t \in \overline{I}$, we assume that γ is piecewise smooth with respect to the corresponding spatial mesh $\mathcal{P}|_t$, where for simplicity $|\gamma'| = 1$.

We start with the following analytic observations which will be used to integrate the involved integrals in time: With the exponential integral $\mathrm{Ei}(x) := -\int_{-x}^{\infty} y^{-1} e^{-y}\,\mathrm{d}y$ and

$$\mathfrak{g}_t(\mathbf{x}) := \begin{cases} \frac{1}{4\pi}\,\mathrm{Ei}\left(-\frac{|\mathbf{x}|^2}{4t}\right) & \text{for } (t,\mathbf{x}) \in (0,\infty) \times \mathbb{R}^2, \\ 0 & \text{else,} \end{cases}$$

it holds that

$$\int_{\tilde a}^{\tilde b} G(t-s,\mathbf{x})\,\mathrm{d}s = \mathfrak{g}_{t-\tilde b}(\mathbf{x}) - \mathfrak{g}_{t-\tilde a}(\mathbf{x}) \quad \text{for all } 0 \le \tilde a < \tilde b \text{ and } (t,\mathbf{x}) \in [0,\infty)\times\mathbb{R}^2. \tag{A.1}$$

With

$$\mathfrak{G}_t(\mathbf{x}) := \begin{cases} \frac{1}{4\pi}\Big(t e^{-\frac{|\mathbf{x}|^2}{4t}} + t(1+\frac{|\mathbf{x}|^2}{4t})\,\mathrm{Ei}\big(-\frac{|\mathbf{x}|^2}{4t}\big)\Big) & \text{for } (t,\mathbf{x}) \in (0,\infty)\times\mathbb{R}^2, \\ 0 & \text{else,} \end{cases}$$

it further holds that

$$\int_a^b \int_{\tilde a}^{\tilde b} G(t-s,\mathbf{x})\,\mathrm{d}s\,\mathrm{d}t = \mathfrak{G}_{b-\tilde b}(\mathbf{x}) - \mathfrak{G}_{b-\tilde a}(\mathbf{x}) + \mathfrak{G}_{a-\tilde a}(\mathbf{x}) - \mathfrak{G}_{a-\tilde b}(\mathbf{x}); \tag{A.2}$$

for all $0 \le a < b$, $0 \le \tilde a < \tilde b$, and $\mathbf{x} \in \mathbb{R}^2$; see, e.g., [4] or [21] for more details. As $\mathrm{Ei}(\cdot) - \log|\cdot|$ is smooth, $\mathfrak{g}_t$ as well as $\mathfrak{G}_t$ have a logarithmic singularity for $\mathbf{x} \to 0$ (provided they are not identically 0).

A.1. Galerkin discretization

To compute the Galerkin discretization (2.12) for the trial space $\mathcal{X}$ of piecewise constants with respect to $\mathcal{P}$, we have to compute

$$\langle \mathcal{V}\mathbb{1}_{\tilde J\times\tilde K}, \mathbb{1}_{J\times K}\rangle_\Sigma \quad \text{and} \quad \langle f, \mathbb{1}_{J\times K}\rangle_\Sigma \quad \text{for all } J\times K, \tilde J\times\tilde K \in \mathcal{P}.$$

Let $J=[a,b]$, $\tilde J=[\tilde a,\tilde b]$, $K=\gamma([c,d])$, and $\tilde K=\gamma([\tilde c,\tilde d])$. We abbreviate $\gamma_K(\hat x) := \gamma(c+\hat x(d-c))$ and $\gamma_{\tilde K} := \gamma(\tilde c + \hat y(\tilde d - \tilde c))$.

A.1.1. Galerkin matrix

The Fubini theorem and (A.2) show that

$$\langle \mathcal{V}\mathbb{1}_{\tilde J\times\tilde K}, \mathbb{1}_{J\times K}\rangle_\Sigma = \int_K \int_{\tilde K} \mathfrak{G}_{b-\tilde b}(\mathbf{x}-\mathbf{y}) - \mathfrak{G}_{b-\tilde a}(\mathbf{x}-\mathbf{y}) + \mathfrak{G}_{a-\tilde a}(\mathbf{x}-\mathbf{y}) - \mathfrak{G}_{a-\tilde b}(\mathbf{x}-\mathbf{y})\,\mathrm{d}\mathbf{y}\,\mathrm{d}\mathbf{x}.$$

To compute terms of the form $\int_K \int_{\tilde K} \mathfrak{G}_t(\mathbf{x}-\mathbf{y})\,\mathrm{d}\mathbf{y}\,\mathrm{d}\mathbf{x}$ for $t>0$, we first use the transformation formula

$$\int_K \int_{\tilde K} \mathfrak{G}_t(\mathbf{x}-\mathbf{y})\,\mathrm{d}\mathbf{y}\,\mathrm{d}\mathbf{x} = (d-c)(\tilde d-\tilde c)\int_0^1\int_0^1 \mathfrak{G}_t\big(\gamma_K(\hat x)-\gamma_{\tilde K}(\hat y)\big)\,\mathrm{d}\hat y\,\mathrm{d}\hat x.$$

If $K\cap\tilde K = \emptyset$, the integrand $F(\hat x,\hat y) := \mathfrak{G}_t(\gamma_K(\hat x)-\gamma_{\tilde K}(\hat y))$ is smooth and we can use standard Gauss quadrature in both directions.

If $K\cap\tilde K \neq \emptyset$, we assume without loss of generality that $K=\tilde K$ or that K and $\tilde K$ intersect only in one point, i.e., $\#(K\cap\tilde K)=1$, otherwise we can just split K and $\tilde K$.

If $K=\tilde K$, the integrand has a logarithmic singularity along the diagonal $\hat x=\hat y$. More precisely, it has the form $F(\hat x,\hat y) = f_1(\hat x,\hat y) + f_2(\hat x,\hat y)\log(|\hat x-\hat y|^2)$ with smooth functions f_1 and f_2. We employ the Duffy transformations $\tau_1(\hat x,\hat y) := (\hat x,\hat x\hat y)$ and $\tau_2(\hat x,\hat y) := (\hat x\hat y,\hat x)$, which both map the (open) unit square bijectively onto some (open) triangle, where $[0,1]^2 = \bigcup_{i=1}^2 \tau_i([0,1]^2)$ with intersection of measure zero between the sets. As $|\det(D\tau_i(\hat x,\hat y))| = \hat x$ for $i=1,2$, the integral can be written as

$$\int_0^1\int_0^1 F(\hat x,\hat y,\hat z)\,\mathrm{d}\hat y\,\mathrm{d}\hat x = \int_0^1\int_0^1 \hat x \sum_{i=1}^2 F\big(\tau_i(\hat x,\hat y)\big)\,\mathrm{d}\hat y\,\mathrm{d}\hat x.$$

The final integrand of the form $\tilde f_1(\hat x,\hat y) + \tilde f_2(\hat x,\hat y)\log(\hat x) + \tilde f_3(\hat x,\hat y)\log(\hat y)$ with smooth functions $\tilde f_1$, $\tilde f_2$, and $\tilde f_3$, and we can use the quadrature from [22] in both $\hat x$- and $\hat y$-direction.

Finally, if $\#(K\cap\tilde K)=1$, the integrand has a logarithmic singularity at $(\hat x,\hat y)=(0,1)$ or $(\hat x,\hat y)=(1,0)$. Without loss of generality, we suppose that the singularity is at $(\hat x,\hat y)=(0,1)$ so that the integrand has the form $F(\hat x,\hat y) = f_1(\hat x,\hat y) + f_2(\hat x,\hat y)\log(|1+\hat x-\hat y|^2)$ with smooth functions f_1 and f_2. We rotate the integration domain by $\pi/2$, i.e., $(\hat x,\hat y)\mapsto(\hat y,1-\hat x)$, which transforms the singularity to $(\hat x,\hat y)=(0,0)$, and then employ the same transformations as for the case $K=\tilde K$. Similar as before, the final integrand is of the form $\tilde f_1(\hat x,\hat y) + \tilde f_2(\hat x,\hat y)\log(\hat x)$ with smooth functions $\tilde f_1$ and $\tilde f_2$, and we can use the quadrature from [22] in $\hat x$-direction and standard Gauss quadrature in $\hat y$-direction.

We remark that $\int_K\int_{\tilde K}\mathfrak{G}_t(\mathbf{x}-\mathbf{y})\,\mathrm{d}\mathbf{y}\,\mathrm{d}\mathbf{x}$ can even be computed exactly if K and $\tilde K$ lie both on *one* straight line.

A.1.2. Right-hand side vector

We consider the indirect boundary element method from Section 2.3 so that $f = u_D - \mathcal{M}_0 u_0$. Provided that u_D is $\mathcal{P}$-piecewise smooth, the term $\langle u_D, \mathbb{1}_{J\times K}\rangle_\Sigma$ can be easily computed by first transforming $J\times K$ onto $[0,1]^2$ and subsequently applying Gauss quadrature in both directions. For $\mathcal{M}_0 u_0$ we employ the Fubini theorem and (A.1),

$$\langle \mathcal{M}_0 u_0, \mathbb{1}_{J\times K}\rangle = \int_K\int_\Omega \big(\mathfrak{g}_a(\mathbf{x}-\mathbf{y}) - \mathfrak{g}_b(\mathbf{x}-\mathbf{y})\big)u_0(\mathbf{y})\,\mathrm{d}\mathbf{y}\,\mathrm{d}\mathbf{x}.$$

The integrand has a logarithmic singularity for $\mathbf{y}\in\partial\Omega$.

Let $\mathcal{T}_K$ be a partition of Ω into curvilinear triangles of the form $T=\gamma_T(\hat T)$ with the reference triangle $\hat T = \{(\hat y,\hat z)\in[0,1]^2 : \hat z \le 1-\hat y\}$ and some smooth diffeomorphism $\gamma_T:\hat T\to T$ such that for all $T,\tilde T\in\mathcal{T}_K$ with $T\neq\tilde T\in\mathcal{T}_K$, the intersection has measure zero. Moreover, we suppose that there exists a unique element $T\in\mathcal{T}_K$ with $K\cap T = K$. With $t\in\{a,b\}$ and the abbreviation $\tilde u_{0,T} := (u_0\circ\gamma_T)|\det(D\gamma_T)|$, we have that

$$\int_K \int_\Omega \mathfrak{g}_t(\mathbf{x}-\mathbf{y})u_0(\mathbf{y})\,\mathrm{d}\mathbf{y}\,\mathrm{d}\mathbf{x} = (d-c)\sum_{T\in\mathcal{T}_K}\int_0^1\int_{\hat T}\mathfrak{g}_t\big(\gamma_K(\hat x)-\gamma_T(\hat y,\hat z)\big)\,\tilde u_{0,T}(\hat y,\hat z)\,\mathrm{d}\hat y\,\mathrm{d}\hat z\,\mathrm{d}\hat x.$$

Remark A.1. *To construct $\mathcal{T}_K$ in our examples from Section 4, we start from some initial mesh of Ω consisting of one square for $\Omega=(0,1)^2$ and three squares for $\Omega := (-1,1)^2\setminus[-1,0]^2$, and proceed as follows: First, we dyadically refine the element that contains K until K becomes the edge of one of the resulting squares. We use further dyadic refinements to ensure that there is at most one hanging node per edge. To obtain a triangular mesh $\mathcal{T}_K$, we finally bisect the elements along one diagonal. Note that the resulting $\mathcal{T}_K$ is* not *conforming. The number of elements in $\mathcal{T}_K$ is proportional to the level of K.*

If $K\cap T=\emptyset$, the integrand $F(\hat x,\hat y,\hat z) := \mathfrak{g}_t(\gamma_K(\hat x)-\gamma_T(\hat y,\hat z))\,\tilde u_{0,T}(\hat y,\hat z)$ is smooth and we can use standard Gauss quadrature in all three directions.
If $K\cap T=K$, we suppose that $\gamma_K=\gamma_T(\cdot,0)$ so that the integrand has a logarithmic singularity for $(\hat x,\hat z)=(\hat y,0)$. More precisely, it has the form

$$F(\hat x,\hat y,\hat z) = f_1(\hat x,\hat y,\hat z) + f_2(\hat x,\hat y,\hat z)\log(|\mathbf{F}(\hat x,\hat y,\hat z)(\hat x-\hat y,\hat z)^\top|^2)$$

for some smooth functions f_1, f_2 with values in $\mathbb{R}$, and $\mathbf{F}$ with values in $\mathbb{R}^{2\times2}$ and $\det\mathbf{F}\neq0$. We employ the following transformations

$$\tau_1(\hat x,\hat y,\hat z) := \big(\hat x,\hat x(1-\hat y),\hat x\hat y\hat z\big),$$
$$\tau_2(\hat x,\hat y,\hat z) := \big(\hat x(1-\hat y),\hat x(1-\hat y\hat z),\hat x\hat y\hat z\big),$$
$$\tau_3(\hat x,\hat y,\hat z) := \big(\hat x(1-\hat y+\hat y\hat z),\hat x(1-\hat y),\hat x\hat y\big),$$

which all map the (open) unit cube bijectively onto some (open) tetrahedron, where $[0,1]\times\hat T=\bigcup_{i=1}^3\tau_i([0,1]^3)$ with intersection of measure zero between the sets. As $|\det(D\tau_i(\hat x,\hat y,\hat z))|=\hat x^2\hat y$ for $i=1,2,3$, the integral can be written as

$$\int_0^1\int_{\hat T}F(\hat x,\hat y,\hat z)\,\mathrm{d}\hat y\,\mathrm{d}\hat z\,\mathrm{d}\hat x = \int_0^1\int_0^1\int_0^1\hat x^2\hat y\sum_{i=1}^3 F\big(\tau_i(\hat x,\hat y,\hat z)\big)\,\mathrm{d}\hat y\,\mathrm{d}\hat z\,\mathrm{d}\hat x.$$

Note that the vector $(\hat x-\hat y,\hat z)^\top$ from the definition of F is transformed under τ_1, τ_2, and τ_3 to $\hat x\hat y(1,z)^\top$, $\hat x\hat y(z-1,z)^\top$, and $\hat x\hat y(z,1)^\top$, respectively. We infer that we can use the quadrature from [22] in $\hat x$- and $\hat y$-direction, and standard Gauss quadrature in $\hat z$-direction.

If $\#(K\cap T)=1$, we suppose that $\gamma_K(0)=\gamma_T(0,0)$. We further suppose that K and T can be parametrized via *one* smooth diffeomorphism $\gamma_{K\cup T}$: The parameter domain of $\gamma_{K\cup T}$ should contain at least the line $\{\hat x(\hat v,\hat w):\hat x\in[0,1]\}$ for some $(\hat v,\hat w)\in\mathbb{R}^2\setminus[0,\infty)^2$ and $\hat T$. Moreover, $\gamma_K(\hat x)=\gamma_{K\cup T}(\hat x(\hat v,\hat w))$ for $\hat x\in[0,1]$ and $\gamma_T(\hat y,\hat z)=\gamma_{K\cup T}(\hat y,\hat z)$ for $(\hat y,\hat z)\in\hat T$. Then the integrand has a logarithmic singularity for $(\hat x,\hat y,\hat z)=(0,0,0)$. More precisely, it has the form

$$F(\hat x,\hat y,\hat z) := f_1(\hat x,\hat y,\hat z) + f_2(\hat x,\hat y,\hat z)\log(|\mathbf{F}(\hat x,\hat y,\hat z)(\hat x\hat v-\hat y,\hat x\hat w-\hat z)^\top|^2)$$

for some smooth functions f_1, f_2 with values in $\mathbb{R}$, and $\mathbf{F}$ with values in $\mathbb{R}^{2\times2}$ and $\det\mathbf{F}\neq0$. Applying the transformations τ_i from before, the vector $(\hat x\hat v-\hat y,\hat x\hat w-\hat z)^\top$ from the definition of F is transformed to $\hat x((\hat v,\hat w)^\top-(1-\hat y,\hat y\hat z)^\top)$, $\hat x((1-\hat y)(v,w)^\top-(1-yz,yz)^\top)$, and $\hat x((1-\hat y+\hat y\hat z)(v,w)^\top-(1-y,y)^\top)$, respectively. Due to our assumption on (v,w), up to the factor $\hat x$, each of these terms is uniformly away from $(0,0)^\top$ for all $(\hat y,\hat z)\in\hat T$. Thus, we can again use the quadrature from [22] in $\hat x$-direction, and standard Gauss quadrature in $\hat y$ and $\hat z$-direction.

A.2. Evaluation of residual

To compute the error estimator $\eta_{\mathcal{P}}(\Phi)$ as well as the weighted L_2-terms $\zeta_{\mathcal{P}}(\Phi)$, we have to evaluate the residual $f-\mathscr{V}\Phi$.

A.2.1. Single-layer operator

To evaluate the single-layer operator for piecewise constants with respect to $\mathcal{P}$, we have to compute

$$(\mathscr{V}\mathbb{1}_{\tilde J\times\tilde K})(t,\mathbf{x}) = \int_{\tilde J}\int_{\tilde K}G(t-s,\mathbf{x}-\mathbf{y})\,\mathrm{d}\mathbf{y}\,\mathrm{d}s \quad \text{for all } \tilde J\times\tilde K\in\mathcal{P} \text{ and all } (t,\mathbf{x})\in\Sigma.$$

Let $\tilde J=[\tilde a,\tilde b]$ and $\tilde K=\gamma([\tilde c,\tilde d])$, and abbreviate again $\gamma_{\tilde K}:=\gamma(\tilde c+\hat y(\tilde d-\tilde c))$. The Fubini theorem and (A.1) show that

$$(\mathscr{V}\mathbb{1}_{\tilde J\times\tilde K})(t,\mathbf{x}) = \int_{\tilde K}\mathfrak{g}_{t-\tilde b}(\mathbf{x}-\mathbf{y})-\mathfrak{g}_{t-\tilde a}(\mathbf{x}-\mathbf{y})\,\mathrm{d}\mathbf{y}.$$

To compute terms of the form $\int_{\tilde K}\mathfrak{g}_s(\mathbf{x}-\mathbf{y})\,\mathrm{d}\mathbf{y}$ for $s>0$, we first use the transformation formula

$$\int_{\tilde K}\mathfrak{g}_s(\mathbf{x}-\mathbf{y})\,\mathrm{d}\mathbf{y} = \int_0^1\mathfrak{g}_s\big(\mathbf{x}-\gamma_{\tilde K}(\hat y)\big)\,\mathrm{d}\hat y.$$

If $\mathbf{x}\notin\tilde K$, the integrand is smooth and we can use standard Gauss quadrature.
If $\mathbf{x}\in\tilde K$, we assume without loss of generality that $\mathbf{x}=\gamma(\tilde c)$ or $\mathbf{x}=\gamma(\tilde d)$, otherwise we can just split $\tilde K$. The integrand is of the form $f_1(\hat y)+f_2(\hat y)\log(\hat y)$ with smooth functions f_1 and f_2, and we can use the quadrature from [22].
We remark that $\int_{\tilde K}\mathfrak{g}_s(\mathbf{x}-\mathbf{y})\,\mathrm{d}\mathbf{y}$ can even be computed exactly if $\mathbf{x}$ and $\tilde K$ lie both on *one* straight line.

A.2.2. Initial operator

We consider the indirect boundary element method from Section 2.3 so that $f = u_D - \mathscr{M}_0 u_0$. To evaluate f at $(t,\mathbf{x}) \in J \times K$ with $t > 0$ and $J \times K \in \mathcal{P}$, let $\mathcal{T}_K$ be again a curvilinear triangulation of Ω as in Appendix A.1.2. With the abbreviation $\tilde{u}_{0,T} := (u_0 \circ \gamma_T)|\det(D\gamma_T)|$, we have that

$$(\mathscr{M}_0 u_0)(t,\mathbf{x}) = \sum_{T \in \mathcal{T}_K} \int_0^1 \int_0^1 G\big(t, \mathbf{x} - \gamma_T(\hat{y},\hat{z})\big)\, \tilde{u}_{0,T}(\hat{y},\hat{z})\, \mathrm{d}\hat{y}\, \mathrm{d}\hat{z}.$$

As $\mathbf{x} \neq \gamma_T(\hat{y},\hat{z})$, the integrand is smooth and we can use standard Gauss quadrature in both directions.

A.3. Error estimator and L_2-terms

Now that we can evaluate the residual $r := f - \mathscr{V}\Phi$, we explain how to compute the estimator $\eta_{\mathcal{P}}(\Phi)$ as well as the weighted L_2-terms $\zeta_{\mathcal{P}}(\Phi)$. We assume that r is, at least $\mathcal{P}$-piecewise, sufficiently smooth. In particular, $\zeta_{\mathcal{P}}(\Phi)$ can be easily computed via the transformation formula and standard Gauss quadrature in both directions.

For $\eta_{\mathcal{P}}(\Phi)$, we need to compute terms of the form $|r|_{L_2(J,H^{1/2}(K\cup\tilde{K}))}$ and $|r|_{H^{1/4}(J\cup\tilde{J},L_2(K))}$ with $K \cap \tilde{K} \neq \emptyset$ and $J \cap \tilde{J} \neq \emptyset$.

The first term reads as

$$|r|^2_{L_2(J,H^{1/2}(K\cup\tilde{K}))} = \int_J |r(t,\cdot)|^2_{H^{1/2}(K)} + 2\int_K \int_{\tilde{K}} \frac{|r(t,\mathbf{x}) - r(t,\mathbf{y})|^2}{|\mathbf{x}-\mathbf{y}|^2}\, \mathrm{d}\mathbf{y}\, \mathrm{d}\mathbf{x} + |r(t,\cdot)|^2_{H^{1/2}(\tilde{K})}\, \mathrm{d}t.$$

We consider the integrand for fixed t. The first and last term can be transformed as in the case $K = \tilde{K}$ of Section A.1.1 and subsequently be computed by standard Gauss quadrature in both directions. Similarly, the middle term can be transformed as in the case $\#(K \cap \tilde{K}) = 1$ of Section A.1.1 and subsequently be computed by standard Gauss quadrature in both directions. Finally, we use standard Gauss quadrature in t-direction for all three terms.

Now, we consider

$$|r|^2_{H^{1/4}(J\cup\tilde{J},L_2(K))} = \int_K |r(\cdot,\mathbf{x})|^2_{H^{1/4}(J)} + 2\int_J \int_{\tilde{J}} \frac{|r(t,\mathbf{x}) - r(s,\mathbf{x})|^2}{|t-s|^{3/2}}\, \mathrm{d}s\, \mathrm{d}t + |r(\cdot,\mathbf{x})|^2_{H^{1/4}(\tilde{J})}\, \mathrm{d}\mathbf{x}.$$

The first and last term can be transformed as in the case $K = \tilde{K}$ of Section A.1.1. With γ_J defined analogously as γ_K, this shows for the first one that

$$\begin{aligned} |r(\cdot,\mathbf{x})|^2_{H^{1/4}(J)} &= 2|J|^2 \int_0^1 \int_0^1 \frac{|r(\gamma_J(\hat{t}),\mathbf{x}) - r(\gamma_J(\hat{t}\hat{s}),\mathbf{x})|^2}{|\gamma_J(\hat{t}) - \gamma_J(\hat{t}\hat{s})|^{3/2}}\, \hat{t}\, \mathrm{d}\hat{s}\, \mathrm{d}\hat{t} \\ &= 2|J|^{1/2} \int_0^1 \int_0^1 \frac{|r(\gamma_J(\hat{t}),\mathbf{x}) - r(\gamma_J(\hat{t}(1-\hat{s})),\mathbf{x})|^2}{\hat{s}}\, \hat{s}^{-1/2}\, \hat{t}^{-1/2}\, \mathrm{d}\hat{s}\, \mathrm{d}\hat{t}. \end{aligned}$$

As r is piecewise smooth, the quotient is smooth and we can use Gauss quadrature with weight $\hat{t}^{-1/2}$ in $\hat{t}$-direction and with weight $\hat{s}^{-1/2}$ in $\hat{s}$-direction. Similarly, the second term can be transformed as in the case $\#(K \cap \tilde{K}) = 1$ of Section A.1.1 so that for $\tilde{J}$ left from J and $\gamma_J, \gamma_{\tilde{J}}$ defined analogously as $\gamma_K, \gamma_{\tilde{K}}$, we get that

$$\int_J \int_{\tilde{J}} \frac{|r(t,\mathbf{x}) - r(s,\mathbf{x})|^2}{|t-s|^{3/2}}\, \mathrm{d}s\, \mathrm{d}t = |J|\,|\tilde{J}| \int_0^1 \int_0^1 \frac{|r(\gamma_J(\hat{s}\hat{t}),\mathbf{x}) - r(\gamma_{\tilde{J}}(1-\hat{t}),\mathbf{x})|^2}{|\gamma_J(\hat{s}\hat{t}) - \gamma_{\tilde{J}}(1-\hat{t})|^{3/2}}\, \hat{t}\, \mathrm{d}\hat{s}\, \mathrm{d}\hat{t} + |J|\,|\tilde{J}| \int_0^1 \int_0^1 \frac{|r(\gamma_J(\hat{t}),\mathbf{x}) - r(\gamma_{\tilde{J}}(1-\hat{s}\hat{t}),\mathbf{x})|^2}{|\gamma_J(\hat{t}) - \gamma_{\tilde{J}}(1-\hat{s}\hat{t})|^{3/2}}\, \hat{t}\, \mathrm{d}\hat{s}\, \mathrm{d}\hat{t}.$$

We can apply Gauss quadrature with weight $\hat{t}^{-1/2}$ in $\hat{t}$-direction and with weight 1 in $\hat{s}$-direction. Finally, we use the transformation formula and standard Gauss quadrature in $\hat{x}$-direction for all three terms.

References

[1] Markus Aurada, Michael Feischl, Thomas Führer, Michael Karkulik, Dirk Praetorius, Efficiency and optimality of some weighted-residual error estimator for adaptive 2D boundary element methods, Comput. Methods Appl. Math. 13 (3) (2013) 305–332.

[2] Markus Aurada, Michael Feischl, Thomas Führer, Michael Karkulik, J. Markus Melenk, Dirk Praetorius, Local inverse estimates for non-local boundary integral operators, Math. Comput. 86 (308) (2017) 2651–2686.

[3] Douglas N. Arnold, Patrick J. Noon, Boundary integral equations of the first kind for the heat equation, in: Boundary Elements IX, vol. 3, Springer, 1987, pp. 213–229.

[4] Martin Costabel, Boundary integral operators for the heat equation, Integral Equ. Oper. Theory 13 (4) (1990) 498–552.

[5] Alexey Chernov, Anne Reinarz, Sparse grid approximation spaces for space–time boundary integral formulations of the heat equation, Comput. Math. Appl. 78 (11) (2019) 3605–3619.

[6] Alexey Chernov, Christoph Schwab, Sparse space-time Galerkin BEM for the nonstationary heat equation, Z. Angew. Math. Mech. 93 (6–7) (2013) 403–413.

[7] Stefan Dohr, Kazuki Niino, Olaf Steinbach, Space-time boundary element methods for the heat equation, in: Space-Time Methods: Applications to Partial Differential Equations, De Gruyter, 2019, pp. 1–60.

[8] Stefan Dohr, Distributed and Preconditioned Space–Time Boundary Element Methods for the Heat Equation, PhD thesis, TU Graz, 2019.

[9] Stefan Dohr, Jan Zapletal, Günther Of, Michal Merta, Michal Kravčenko, A parallel space–time boundary element method for the heat equation, Comput. Math. Appl. 78 (9) (2019) 2852–2866.

[10] Birgit Faermann, Localization of the Aronszajn-Slobodeckij norm and application to adaptive boundary elements methods. Part I. The two-dimensional case, IMA J. Numer. Anal. 20 (2) (2000) 203–234.

[11] Birgit Faermann, Localization of the Aronszajn-Slobodeckij norm and application to adaptive boundary element methods. Part II. The three-dimensional case, Numer. Math. 92 (3) (2002) 467–499.

[12] Gregor Gantner, Optimal adaptivity for splines in finite and boundary element methods, PhD thesis, TU Wien, 2017.

[13] Matthias Gläfke, Adaptive methods for time domain boundary integral equations, PhD thesis, Brunel University London, 2012.

[14] Heiko Gimperlein, Ceyhun Özdemir, David Stark, Ernst P. Stephan, A residual a posteriori error estimate for the time–domain boundary element method, Numer. Math. 146 (2) (2020) 239–280.
[15] Gregor Gantner, Dirk Praetorius, Adaptive BEM for elliptic PDE systems, part I: abstract framework, for weakly-singular integral equations, Appl. Anal. (2020) 1–34, https://doi.org/10.1080/00036811.2020.1800651.
[16] Gregor Gantner, Raymond van Venetië, Implementation of: adaptive space-time BEM for the heat equation, Software, zenodo:5165043, 2021.
[17] Helmut Harbrecht, Johannes Tausch, A fast sparse grid based space–time boundary element method for the nonstationary heat equation, Numer. Math. 140 (1) (2018) 1–26.
[18] Michael Messner, Martin Schanz, Johannes Tausch, A fast Galerkin method for parabolic space–time boundary integral equations, J. Comput. Phys. 258 (2014) 15–30.
[19] Michael Messner, Martin Schanz, Johannes Tausch, An efficient Galerkin boundary element method for the transient heat equation, SIAM J. Sci. Comput. 37 (3) (2015) A1554–A1576.
[20] Patrick James Noon, The single layer heat potential and Galerkin boundary element methods for the heat equation, PhD thesis, University of Maryland, 1988.
[21] Anne Reinarz, Sparse space-time boundary element methods for the heat equation, PhD thesis, University of Reading, 2015.
[22] R.N.L. Smith, Direct Gauss quadrature formulae for logarithmic singularities on isoparametric elements, Eng. Anal. Bound. Elem. 24 (2) (2000) 161–167.
[23] Johannes Tausch, Nyström method for BEM of the heat equation with moving boundaries, Adv. Comput. Math. 45 (5) (2019) 2953–2968.
[24] Raphael Watschinger, Michal Merta, Günther Of, Jan Zapletal, A parallel fast multipole method for a space-time boundary element method for the heat equation, Preprint, arXiv:2106.15911, 2021.
[25] Jan Zapletal, Raphael Watschinger, Günther Of, Michal Merta, Semi-analytic integration for a parallel space-time boundary element method modeling the heat equation, Preprint, arXiv:2102.09811, 2021.