ESAIM: M2AN 54 (2020) 2099–2123
https://doi.org/10.1051/m2an/2020033

**ESAIM: Mathematical Modelling and Numerical Analysis**
www.esaim-m2an.org

# GALERKIN–COLLOCATION APPROXIMATION IN TIME FOR THE WAVE EQUATION AND ITS POST-PROCESSING

Mathias Anselmann[1], Markus Bause[1,*], Simon Becher[2] and Gunar Matthies[2]

**Abstract.** We introduce and analyze families of Galerkin–collocation discretization schemes in time for the wave equation. Their conceptual basis is the establishment of a connection between the Galerkin method for the time discretization and the classical collocation methods, with the perspective of achieving the accuracy of the former with reduced computational costs provided by the latter in terms of less complex algebraic systems. Firstly, continuously differentiable in time discrete solutions are studied. Optimal order error estimates are proved. Then, the concept of Galerkin–collocation approximation is extended to twice continuously differentiable in time discrete solutions. A direct link between the two families by a computationally cheap post-processing is presented. A key ingredient of the proposed methods is the application of quadrature rules involving derivatives. The performance properties of the schemes are illustrated by numerical experiments.

**Mathematics Subject Classification.** 65M60, 65M12.

Received August 20, 2019. Accepted May 1, 2020.

## 1. Introduction

In this work we introduce and analyze a Galerkin–collocation (cGP-$\mathrm{C}^k$, $k \in \{1,2\}$) approach in time combined with a continuous Galerkin (cG) finite element method in space to approximate the solution to the second order hyperbolic wave problem

$$\begin{aligned} \partial_t^2 u - \Delta u &= f \quad \text{in} \quad \Omega \times (0,T], \\ u &= 0 \quad \text{on} \quad \partial\Omega \times (0,T], \\ u(\cdot,0) = u_0, \quad \partial_t u(\cdot,0) &= u_1 \quad \text{in} \quad \Omega, \end{aligned} \tag{1.1}$$

with $\mathrm{C}^k$ regular functions in time. In (1.1), $T > 0$ denotes some final time and $\Omega$ is a polygonal or polyhedral bounded domain in $\mathbb{R}^d$, with $d = 2$ or $d = 3$. The function $f : \Omega \times (0,T] \to \mathbb{R}$ and the initial values $u_0, u_1 : \Omega \to \mathbb{R}$ are given data. The system (1.1) is studied as a prototype model for more sophisticated wave phenomena of practical interest like, for instance, elastic wave propagation governed by the Lamé–Navier equations, the Maxwell system, or wave equations in coupled systems such as fluid–structure interaction and fully dynamic poroelasticity [42].

*Keywords and phrases.* Wave equation, variational time discretization, collocation methods, space-time finite element methods, error estimates, post-processing.

[1] Helmut Schmidt University, Faculty of Mechanical Engineering, Holstenhofweg 85, 22043 Hamburg, Germany.

[2] Technical University of Dresden, Institute of Numerical Mathematics, 01062 Dresden, Germany.

*Corresponding author: bause@hsu-hh.de

Our modification of the standard continuous Galerkin–Petrov method (cGP) for time discretization (*cf.*, *e.g.*, [8, 9, 16, 34]) and the innovation of this work comes through imposing collocation conditions involving the discrete solution's derivatives at the discrete time nodes while on the other hand downsizing the test space of the discrete variational problem compared with the standard cGP approach. This idea was recently introduced in [18] by two of the authors of this work for first-order systems of ordinary differential equations. We refer to our schemes as Galerkin–collocation methods. The collocation equations at the discrete time nodes then enable us to ensure regularity of higher order in time of the discrete solutions. A further key ingredient in the construction of the Galerkin–collocation approach comes through the application of a special quadrature formula, investigated in [33], and the definition of a related interpolation operator for the right-hand side term of the variational equation. Both of them use derivatives of the given function. The Galerkin–collocation schemes rely in an essential way on the perfectly matching set of polynomial spaces (trial and test space), quadrature formula, and interpolation operator. For the discretization of the spatial variables a continuous finite element approach is used here. This is done for the sake of brevity. Usually, discontinuous Galerkin methods are preferred; *cf.* [6, 20, 36]. Beyond the higher order regularity in the time, the Galerkin–collocation schemes offer appreciable advantages for the solution of the arising linear systems by a favorable impact on the matrix block structure; *cf.* [6] for details.

For the subclass of discrete solutions being once continuously differentiable in time an error analysis with optimal order error estimates in time and space and in various norms is given. We will stress the key ideas of our error analysis and present a fundamental concept for analyzing generalized Galerkin approximations to wave problems. One key point of our convergence proof for second-order hyperbolic problems is the weak stability result of Lemma 4.9. Compared with usual stability results for parabolic problems or for first-order hyperbolic problems (*cf.*, *e.g.*, [23], Lem. 4.2) a stability is obtained such that in the resulting error analysis some contributions can no longer be absorbed by terms on the left-hand side of the error inequality like it is typically done. Therefore, to prove error estimates of optimal order, the error in the time derivatives $\left(\partial_t u^0_{\tau,h}, \partial_t u^1_{\tau,h}\right)$ for the discrete approximation pair $\left(u^0_{\tau,h}, u^1_{\tau,h}\right)$ of $(u, \partial_t u)$ is bounded firstly. For this, a variational problem that is satisfied by $\left(\partial_t u^0_{\tau,h}, \partial_t u^1_{\tau,h}\right)$ is identified. Then, a minor extension of a result of [34] becomes applicable to the thus obtained problem. This yields an estimate for $\partial_t u - \partial_t u^0_{\tau,h}$ and $\partial_t^2 u - \partial_t u^1_{\tau,h}$. These auxiliary results then enable us to prove the desired optimal-order error estimates for $u - u^0_{\tau,h}$ and $\partial_t u - u^1_{\tau,h}$.

Space-time finite element methods with continuous and discontinuous discretizations of the time and space variables for parabolic and hyperbolic problems are well-known and have been studied carefully in the literature; *cf.*, *e.g.*, [1, 2, 8–10, 13–16, 19, 21, 22, 25–27, 30–32, 34, 36, 37, 44, 46] and the references therein. The space-time approaches of these works differ by the choices of the trial and, in particular, of the test spaces. Depending on the construction of the test basis functions, either time-marching schemes defined by local problems on the respective subintervals $\left(t_{n-1}, t_n\right]$ of $(0, T]$ (*cf.*, *e.g.*, [1, 2, 16, 22, 30, 31]) or schemes where all time steps are solved simultaneously (*cf.*, *e.g.*, [22, 27, 45]) are obtained. Here, by choosing basis test functions supported on a single subinterval $\left(t_{n-1}, t_n\right]$, we end up with a time-marching approach. Further, strong relations and equivalences between cGP schemes, collocation, and Runge–Kutta methods have been observed. In the literature, the relations are only exploited in the formulation and analysis of the schemes. For the latter we particularly refer to [4, 5]. In equation (2.2) of [5] nodal superconvergence properties of the cGP method are shown along such a line. Similarly, links between Runge–Kutta type and Galerkin schemes are also identified and applied in [28, 29, 47]. To the best of our knowledge, aside from (the works of two of the authors) [17, 18] the Galerkin and collocation method are usually not combined to conceptually new time discretization schemes as it is done here.

In a recent work [16], co-authored by one author of this work, a recursive post-processing of the original continuous in time cGP solution is presented and analyzed. The post-processed approximation is built on each time interval upon the Gauss–Lobatto quadrature points of the actual time interval, at which the classical cGP solution is superconvergent with one extra order of accuracy. On the one hand, the post-processing lifts the superconvergence of the original cGP solution at the Gauss–Lobatto quadrature points to all points of the time interval by adding a higher order correction term which vanishes at the Gauss–Lobatto quadrature points. On the other hand, the post-processing, which is done sequentially on the advancing time intervals and is of low computational costs, yields a numerical approximation that is globally $C^1$-regular in time. In Section 3.2 of [23] and p. 494 of [40], similar post-processing techniques and lifting operators were studied for discontinuous Galerkin approximations in time. The post-processing can nicely be exploited, for instance, for an *a-posteriori* error control in time and an adaptive choice of the time mesh. We point out that in contrast to [16], where the continuous differentiability is obtained by a post-processing of the continuous Galerkin–Petrov approximation, the higher order regularity in time that is built in this work is an inherent part of the construction of the discrete solution itself. This requires a different quadrature formula and interpolation operator for the right-hand side function.

In this work we focus on the convergence analysis and thus the theoretical validation of the proposed Galerkin–collocation methods. In Section 7.3, the resulting algebraic system is summarized for one member of the considered family of schemes, but not studied further. For this we refer to [6].

This work is organized as follows. In Section 2 we introduce our notation and summarize preliminaries. In particular, quadrature formulas and related interpolation operators are introduced. In Section 3 our class of Galerkin–collocation schemes is presented. Section 4 contains our error analysis for the family of once continuously differentiable in time Galerkin–collocation methods. In Section 5 *a priori* stability estimates for the discrete solution and the conservation of energy by the numerical schemes are studied. Our construction principle is extended in Section 6 to define a class of twice continuously differentiable in time Galerkin–collocation approximation schemes for the wave equation. A link to the first class of schemes by a post-processing procedure is presented. Finally, in Section 7 our error estimates are illustrated and confirmed by numerical experiments.

## 2. Notation and preliminaries

### 2.1. Function spaces and evolution form of continuous problem

We use standard notation. $H^m(\Omega)$ is the Sobolev space of $L^2(\Omega)$ functions with derivatives up to order $m$ in $L^2(\Omega)$ and $\langle\cdot,\cdot\rangle$ denotes the inner product in $L^2(\Omega)$. Further, $\langle\!\langle\cdot,\cdot\rangle\!\rangle$ defines the $L^2$ inner product on the product space $L^2(\Omega)\times L^2(\Omega)$. We let $H_0^1(\Omega) := \{u\in H^1(\Omega) \,:\, u=0 \text{ on } \partial\Omega\}$. For short, we put

$$H := L^2(\Omega) \quad\text{and}\quad V := H_0^1(\Omega).$$

We denote by $V'$ the dual space of $V$ and use the notation

$$\|\cdot\| := \|\cdot\|_{L^2(\Omega)}, \quad \|\cdot\|_m := \|\cdot\|_{H^m(\Omega)}, \quad m\in\mathbb{N},$$

for the norms of the Sobolev spaces where we do not differ between the scalar- and vector-valued cases. Throughout, the meaning will be obvious from the context. For a Banach space $B$, we let $L^2(0,T;B)$, $H^1(0,T;B)$, $C([0,T];B)$, and $C^m([0,T];B)$, $m\in\mathbb{N}$, be the Bochner spaces of $B$-valued functions, equipped with their natural norms. For a subinterval $J\subseteq[0,T]$, we use the notations $L^2(J;B)$, $H^1(J;B)$, $C^m(J;B)$, and $C^0(J;B) := C(J;B)$.

In what follows, for non-negative numbers $a$ and $b$, the expression $a\lesssim b$ stands for the inequality $a\le C\,b$ with a generic constant $C$ that is independent of the sizes of the spatial and temporal meshes. The value of $C$ can depend on the regularity of the space mesh, the polynomial degrees used for the space-time discretization, and the data (including $\Omega$).

For any given $u\in V$, let the operator $A: V\to V'$ be uniquely defined by

$$\langle Au, v\rangle := \langle\nabla u, \nabla v\rangle \quad \forall v\in V,$$

where $\langle\cdot,\cdot\rangle$ on the left-hand side is understood as duality pairing between $V'$ and $V$. Further, we denote by $\mathcal{A}: V\times H\to H\times V'$ the operator

$$\mathcal{A} = \begin{pmatrix} 0 & -I \\ A & 0 \end{pmatrix}$$

with the identity mapping $I : H \to H$. We let

$$X := L^2(0,T;V) \times L^2(0,T;H).$$

Introducing the unknowns $u^0 = u$ and $u^1 = \partial_t u$, problem (1.1) can be recovered in evolution form.

**Problem 2.1.** Let $f \in L^2(0,T;H)$ and $(u_0,u_1) \in V \times H$ be given and $F = (0,f)$. Find $U = \left(u^0,u^1\right) \in X$ such that

$$\partial_t U + \mathcal{A}U = F \quad \text{in} \quad (0,T), \quad U(0) = U_0 = (u_0,u_1)\,. \tag{2.1}$$

Problem (2.1) admits a unique solution $U \in X$ and the mapping $(f,u_0,u_1) \mapsto \left(u^0,u^1\right)$ is a linear continuous map from $L^2(0,T;H) \times V \times H$ to $X$; *cf.*, p. 273 and Theorem 1.1 of [38]. Further, $u^0 \in C([0,T];V)$ and $u^1 \in C([0,T];H)$ are satisfied; *cf.*, p. 275 and Theorem 8.2 of [39]. It follows from (2.1) that $\partial_t u^1 \in L^2(0,T;V')$.

**Assumption 2.2.** *Throughout, we tacitly assume that the solution $u$ of (1.1) satisfies all the additional regularity conditions that are required in our analysis. In addition, let $f \in C^s([0,T];H)$ for some sufficiently large parameter $s \in \mathbb{N}$ be satisfied.*

The first of the conditions in Assumption 2.2 implies further assumptions on the data $f,u_0,u_1$ and the boundary $\partial\Omega$ of $\Omega$. Improved regularity results for solutions to the wave problem (1.1) can be found in, *e.g.*, Section 7.2 of [24]. The second condition in Assumption 2.2 will allow us to apply an interpolation in time that is based on derivatives of the right-hand side function $f$.

## 2.2. Time and space discretization

For the time discretization, we decompose the time interval $I = (0,T]$ into $N$ subintervals $I_n = (t_{n-1},t_n]$, $n = 1,\dots,N$, where $0 = t_0 < t_1 < \dots < t_{N-1} < t_N = T$ such that $I = \bigcup_{n=1}^N I_n$. We put $\tau = \max_{n=1,\dots,N} \tau_n$ with $\tau_n = t_n - t_{n-1}$. Further, the set $\mathcal{M}_\tau := \{I_1,\dots,I_N\}$ of time intervals is called the time mesh. For a Banach space $B$ and any $k \in \mathbb{N}_0$, we let

$$\mathbb{P}_k(I_n;B) = \left\{ w_\tau : I_n \to B \ : \ w_\tau(t) = \sum_{j=0}^k W^j t^j \ \forall t \in I_n, W^j \in B \ \forall j \right\}.$$

For an integer $k \in \mathbb{N}$, we introduce the space

$$X_\tau^k(B) := \left\{ w_\tau \in C\left(\overline{I};B\right) \ : \ w_\tau|_{I_n} \in \mathbb{P}_k(I_n;B) \ \forall I_n \in \mathcal{M}_\tau \right\} \tag{2.2}$$

of globally continuous functions in time and for an integer $l \in \mathbb{N}_0$ the space

$$Y_\tau^l(B) := \left\{ w_\tau \in L^2(I;B) \ : \ w_\tau|_{I_n} \in \mathbb{P}_l(I_n;B) \ \forall I_n \in \mathcal{M}_\tau \right\}$$

of global $L^2$-functions in time.

For any non-negative integer $s$ and a function $w : I \to B$ that is piecewise sufficiently smooth with respect to the time mesh $\mathcal{M}_\tau$, we define by

$$\partial_t^s w\left(t_n^+\right) := \lim_{t \to t_n + 0} \partial_t^s w(t) \quad \text{and} \quad \partial_t^s w\left(t_n^-\right) := \lim_{t \to t_n - 0} \partial_t^s w(t)$$

the one-sided limits of the $s$th derivative of $w$.

For the space discretization, let $\mathcal{T}_h$ be a shape-regular mesh of $\Omega$ consisting of quadrilateral or hexahedral elements with mesh size $h > 0$. For some integer $r \in \mathbb{N}$, we let $V_h = V_h^{(r)}$ be given by

$$V_h = V_h^{(r)} = \left\{ v_h \in C\left(\overline{\Omega}\right) \ : \ v_h|_T \in \mathbb{Q}_r(K) \, \forall K \in \mathcal{T}_h \right\} \cap H_0^1(\Omega) \tag{2.3}$$

where $\mathbb{Q}_r(K)$ is the space defined by the multilinear reference mapping of polynomials on the reference element with maximum degree $r$ in each variable. Our restriction in this work to continuous finite elements in space is only done for simplicity and in order to reduce the technical methodology of analyzing our Galerkin–collocation discretization scheme to its key points. In the literature it has been mentioned that discontinuous finite element methods in space offer appreciable advantages over continuous ones for the discretization of wave equations; *cf.*, *e.g.*, [3, 20, 35, 36] and the references therein.

We denote by $P_h : H \to V_h$ the $L^2$-orthogonal projection onto $V_h$ such that for $w \in H$,

$$\langle P_h w, v_h\rangle = \langle w, v_h\rangle$$

for all $v_h \in V_h$. The operator $R_h : V \to V_h$ defines the elliptic projection onto $V_h$ such that for $w \in V$,

$$\langle \nabla R_h w, \nabla v_h\rangle = \langle \nabla w, \nabla v_h\rangle$$

for all $v_h \in V_h$. Finally, by $\mathcal{P}_h : H \times H \to V_h \times V_h$ we denote the $L^2$-projection onto the product space $V_h \times V_h$ and by $\mathcal{R}_h : V \times V \to V_h \times V_h$ the elliptic projection onto the product space $V_h \times V_h$. Let $A_h : V \to V_h$ be the operator that is defined by

$$\langle A_h w, v_h\rangle = \langle \nabla w, \nabla v_h\rangle$$

for all $v_h \in V_h$. Then, for $w \in V \cap H^2(\Omega)$ it holds that

$$\langle A_h w, v_h\rangle = \langle \nabla w, \nabla v_h\rangle = \langle A w, v_h\rangle \quad \forall v_h \in V_h.$$

Hence, we have that $A_h w = P_h A w$ for $w \in V \cap H^2(\Omega)$. Let $\mathcal{A}_h : V \times H \to V_h \times V_h$ be defined by

$$\mathcal{A}_h = \begin{pmatrix} 0 & -P_h \\ A_h & 0 \end{pmatrix}.$$

Hence, we have for $W = \left(w^0, w^1\right) \in \left(V \cap H^2(\Omega)\right) \times H$ that

$$\langle\!\langle \mathcal{A}_h W, \Phi_h\rangle\!\rangle = \left\langle -w^1, \phi_h^0\right\rangle + \left\langle \nabla w^0, \nabla \phi_h^1\right\rangle = \left\langle -w^1, \phi_h^0\right\rangle + \left\langle A w^0, \phi_h^1\right\rangle = \langle\!\langle \mathcal{A} W, \Phi_h\rangle\!\rangle$$

for all $\Phi_h = \left(\phi_h^0, \phi_h^1\right) \in V_h \times V_h$. This provides the consistency of $\mathcal{A}_h$ on $\left(V \cap H^2(\Omega)\right) \times H$, *i.e.*,

$$\mathcal{A}_h W = \mathcal{P}_h \mathcal{A} W. \tag{2.4}$$

Finally, let $U_{0,h} \in V_h^2$ denote a suitable approximation of the initial value $U_0 \in V \times H$ in (2.1) that will we used as the initial value $U_{\tau,h}(0)$ of the discrete solution. Further restrictions will be made below.

## 2.3. Quadrature formulas and their natural interpolation operators

Throughout this work, the polynomial degree $k \geq 3$ is assumed to be fixed. Let $\hat{t}_1^{\mathrm{H}} = -1$, $\hat{t}_{k-1}^{\mathrm{H}} = 1$, and $\hat{t}_s^{\mathrm{H}}$, $s = 2, \ldots, k-2$, be the roots of the Jacobi polynomial on $\widehat{I} := [-1, 1]$ with degree $k-3$ associated to the weighting function $\left(1-\hat{t}\right)^2\left(1+\hat{t}\right)^2$. Let $\widehat{I}^{\mathrm{H}} : C^1\left(\widehat{I}; B\right) \to \mathbb{P}_k\left(\widehat{I}; B\right)$ denote the Hermite interpolation operator with respect to point value and first derivative at both $-1$ and $1$ as well as the point values at $\hat{t}_s^{\mathrm{H}}$, $s = 2, \ldots, k-2$. By

$$\widehat{Q}^{\mathrm{H}}\left(\hat{g}\right) := \int_{-1}^{1} \widehat{I}^{\mathrm{H}}\left(\hat{g}\right)\left(\hat{t}\right)\, \mathrm{d}\hat{t}$$

we define an Hermite-type quadrature on $[-1, 1]$ which can be written as

$$\widehat{Q}^{\mathrm{H}}\left(\hat{g}\right) = \widehat{\omega}_L \hat{g}'(-1) + \sum_{s=1}^{k-1} \widehat{\omega}_s \hat{g}\left(\hat{t}_s^{\mathrm{H}}\right) + \widehat{\omega}_R \hat{g}'(1)$$

with non-vanishing weights. The affine mapping $T_n : \widehat{I} \to \overline{I}_n$ with $T_n(-1) = t_{n-1}$ and $T_n(1) = t_n$ yields

$$Q_n^{\mathrm{H}}(g) = \left(\frac{\tau_n}{2}\right)^2 \widehat{\omega}_L \partial_t g\left(t_{n-1}^+\right) + \frac{\tau_n}{2}\sum_{s=1}^{k-1} \widehat{\omega}_s g\left(t_{n,s}^{\mathrm{H}}\right) + \left(\frac{\tau_n}{2}\right)^2 \widehat{\omega}_R \partial_t g\left(t_n^-\right) \tag{2.5}$$

as Hermite-type quadrature formula on $I_n$, where $t_{n,s}^{\mathrm{H}} := T_n\left(\hat{t}_s^{\mathrm{H}}\right)$, $s = 1, \ldots, k-1$. We note that $Q_n^{\mathrm{H}}$ given in (2.5) integrates all polynomials up to degree $2k-3$ exactly, *cf.* [33]. Using $\widehat{I}^{\,\mathrm{H}}$ and $T_n$, the local Hermite interpolation on $I_n$ is given by

$$I_n^{\mathrm{H}} : C^1\left(\overline{I}_n; B\right) \to \mathbb{P}_k\left(\overline{I}_n; B\right), \quad v \mapsto \left(\widehat{I}^{\,\mathrm{H}}\left(v \circ T_n\right)\right) \circ T_n^{-1}.$$

Moreover, we define the global Hermite interpolation $I_\tau^{\mathrm{H}} : C^1\left(\overline{I}; B\right) \to X_\tau^k(B)$ by means of

$$\left(I_\tau^{\mathrm{H}} w\right)|_{I_n} := I_n^{\mathrm{H}}\left(w|_{I_n}\right) \quad \text{for all} \quad n = 1, \ldots, N.$$

In addition to Hermite-type interpolation and quadrature formula, Gauss and Gauss–Lobatto quadrature formulas will be used. To this end, we denote by $\hat{t}_s^{\mathrm{G}}$, $s = 1, \ldots, k-1$, the roots of the Legendre polynomial with degree $k-1$ and by $\hat{t}_s^{\mathrm{GL}}$, $s = 2, \ldots, k-1$, the roots of the Jacobi polynomial on $\widehat{I}$ with degree $k-2$ associated to the weighting function $\left(1-\hat{t}\right)\left(1+\hat{t}\right)$. Furthermore, we set $\hat{t}_1^{\mathrm{GL}} = -1$ and $\hat{t}_k^{\mathrm{GL}} = 1$. The operators $\widehat{I}^{\,\mathrm{G}} : C\left(\widehat{I}; B\right) \to P_{k-2}\left(\widehat{I}; B\right)$ and $\widehat{I}^{\,\mathrm{GL}} : C\left(\widehat{I}; B\right) \to P_{k-1}\left(\widehat{I}; B\right)$ are the Lagrange interpolation using the Gauss points $\hat{t}_s^{\mathrm{G}}$, $s = 1, \ldots, k-1$, and the Gauss–Lobatto points $\hat{t}_s^{\mathrm{GL}}$, $s = 1, \ldots, k$, respectively. We define by

$$\widehat{Q}^{\,\mathrm{G}}\left(\hat{g}\right) := \int_{-1}^{1} \widehat{I}^{\,\mathrm{G}}\left(\hat{g}\right)\left(\hat{t}\right)\,\mathrm{d}\hat{t} \quad \text{and} \quad \widehat{Q}^{\,\mathrm{GL}}\left(\hat{g}\right) := \int_{-1}^{1} \widehat{I}^{\,\mathrm{GL}}\left(\hat{g}\right)\left(\hat{t}\right)\,\mathrm{d}\hat{t}.$$

Gauss and Gauss–Lobatto quadrature formulas on $[-1, 1]$, that are transformed to

$$Q_n^{\mathrm{G}}(g) = \frac{\tau_n}{2}\sum_{s=1}^{k-1} \widehat{\omega}_s^{\mathrm{G}} g\left(t_{n,s}^{\mathrm{G}}\right) \quad \text{and} \quad Q_n^{\mathrm{GL}}(g) = \frac{\tau_n}{2}\sum_{s=1}^{k} \widehat{\omega}_s^{\mathrm{GL}} g\left(t_{n,s}^{\mathrm{GL}}\right) \tag{2.6}$$

on $I_n$ by the affine mapping $T_n$. The Gauss and Gauss–Lobatto formulas also integrate polynomials up to degree $2k-3$ exactly. Local Lagrange-type interpolation operators on $I_n$ are given by

$$\begin{aligned} I_n^{\mathrm{G}} &: C\left(\overline{I}_n; B\right) \to \mathbb{P}_{k-2}\left(\overline{I}_n; B\right), \quad v \mapsto \left(\widehat{I}^{\,\mathrm{G}}\left(v \circ T_n\right)\right) \circ T_n^{-1}, \\ I_n^{\mathrm{GL}} &: C\left(\overline{I}_n; B\right) \to \mathbb{P}_{k-1}\left(\overline{I}_n; B\right), \quad v \mapsto \left(\widehat{I}^{\,\mathrm{GL}}\left(v \circ T_n\right)\right) \circ T_n^{-1}. \end{aligned}$$

Furthermore, we define the global Lagrange interpolation operators $I_\tau^{\mathrm{G}} : C\left(\overline{I}; B\right) \to Y_\tau^{k-2}(B)$ and $I_\tau^{\mathrm{GL}} : C\left(\overline{I}; B\right) \to X_\tau^{k-1}(B)$ by

$$\left(I_\tau^{\mathrm{G}} w\right)|_{I_n} := I_n^{\mathrm{G}}(w|_{I_n}) \quad \text{and} \quad \left(I_\tau^{\mathrm{GL}} w\right)|_{I_n} := I_n^{\mathrm{GL}}(w|_{I_n}) \quad \text{for all} \quad n = 1, \ldots, N.$$

### 2.4. Further interpolation operators

In our error analysis we use some further interpolants in time introduced in [16, 23]. To keep this work self-contained, their definition is briefly summarized here. Remember that $k \geq 3$.

In the following, let $B$ be a Banach space satisfying $B \subset H$ and $\ell \in \mathbb{N}$. We define for $n = 1, \ldots, N$ the local $L^2$-projections $\Pi_n^\ell : L^2\left(I_n; B\right) \to \mathbb{P}_\ell\left(I_n; B\right)$ by

$$\int_{I_n} \left\langle \Pi_n^\ell w, q\right\rangle\,\mathrm{d}t = \int_{I_n} \left\langle w, q\right\rangle\,\mathrm{d}t \quad \forall q \in \mathbb{P}_\ell\left(I_n; B\right).$$

Next, a special interpolant in time is constructed. To this end, we define the Hermite interpolation operator $I_\tau^{k+1} : C^1\left(\overline{I}; B\right) \to C^1\left(\overline{I}; B\right) \cap X_\tau^{k+1}(B)$ by

$$I_\tau^{k+1} u\left(t_n\right) = u\left(t_n\right), \quad \partial_t I_\tau^{k+1} u\left(t_n\right) = \partial_t u\left(t_n\right), \quad n = 0, \ldots, N,$$

and

$$I_\tau^{k+1} u\left(t_{n,\mu}^{\mathrm{GL}}\right) = u\left(t_{n,\mu}^{\mathrm{GL}}\right), \quad n = 1, \ldots, N, \quad \mu = 2, \ldots, k-1.$$

For a function $u \in C^1\left(\overline{I}; B\right)$, we construct a local interpolant $R_n^k u \in \mathbb{P}_k\left(I_n; B\right)$ by

$$\begin{aligned} R_n^k u\left(t_{n-1}^+\right) &= I_\tau^{k+1} u\left(t_{n-1}^+\right), \\ \partial_t R_n^k u\left(t_{n,\mu}^{\mathrm{GL}}\right) &= \partial_t I_\tau^{k+1} u\left(t_{n,\mu}^{\mathrm{GL}}\right), \quad \mu = 1, \ldots, k, \end{aligned}$$

on each time subinterval $I_n$ and a global interpolant $R_\tau^k u \in Y_\tau^k(B)$ by

$$\left(R_\tau^k u\right)|_{I_n} := R_n^k\left(u|_{I_n}\right), \quad n = 1, \ldots, N.$$

Finally, we put $R_\tau^k u(0) := u(0)$. The thus defined $R_\tau^k u$ is continuously differentiable in time on $\overline{I}$ with $R_\tau^k u\left(t_n\right) = u\left(t_n\right)$ and $\partial_t R_\tau^k u\left(t_n\right) = \partial_t u\left(t_n\right)$ for all $n = 0, \ldots, N$; *cf.* [16].

## 3. Galerkin–collocation discretization and auxiliaries

In this section we introduce the approximation of the wave problem (2.1) by our Galerkin–collocation approach that combines collocation conditions at the endpoints $t_{n-1}$ and $t_n$ of the subintervals $I_n$ with variational equations for reduced test spaces compared with the standard continuous finite element approximation of the wave equation (*cf.* [16, 26, 34]). A family of discrete solutions that are once continuously differentiable in time is obtained. For this family an optimal order error analysis is provided in Section 4. For completeness and in order to show the impact of the collocation conditions, the standard continuous Galerkin approximation (*cf.* [26, 34]) of problem (2.1) is briefly recalled first in Section 3.1.

### 3.1. Space-time discretization with continuous Galerkin–Petrov method cGP($k$)

Here, we briefly present the standard continuous Galerkin–Petrov method of order $k \geq 1$ (in short, cGP($k$)) as time discretization applied to the evolution problem (2.1). For the space discretization, the continuous Galerkin approach cG($r$) in $V_h$, defined in (2.3), is used for the sake of simplicity. This yields the following fully discrete problem; *cf.*, *e.g.*, [16, 34] for details.

**Problem 3.1** (Global, fully discrete problem of cGP($k$)–cG($r$))**.** Find $U_{\tau,h} \in \left(X_\tau^k\left(V_h\right)\right)^2$ such that $U_{\tau,h}(0) = U_{0,h}$ and

$$\int_0^T \left(\langle\langle \partial_t U_{\tau,h}, V_{\tau,h} \rangle\rangle + \langle\langle \mathcal{A}_h U_{\tau,h}, V_{\tau,h} \rangle\rangle\right) \mathrm{d}t = \int_0^T \langle\langle F, V_{\tau,h} \rangle\rangle \, \mathrm{d}t$$

for all $V_{\tau,h} \in \left(Y_\tau^{k-1}\left(V_h\right)\right)^2$.

Both components of $U_{\tau,h} = \left(u_{\tau,h}^0, u_{\tau,h}^1\right)$ are computed in the same discrete space $X_\tau^k\left(V_h\right)$. By choosing test functions supported on a single subinterval $I_n$ and using the $(k+1)$-point Gauss–Lobatto quadrature formula, we recast Problem 3.1 as a sequence of local problems on $I_n$.

**Problem 3.2** (Local, numerically integrated, fully discrete problem of cGP($k$)–cG($r$) on $I_n$)**.** Find $U_{\tau,h}|_{I_n} \in \left(\mathbb{P}_k\left(I_n; V_h\right)\right)^2$ with $U_{\tau,h}\left(t_{n-1}^+\right) = U_{\tau,h}\left(t_{n-1}^-\right)$ for $n > 1$ and $U_{\tau,h}\left(t_0^+\right) = U_{0,h}$ such that

$$Q_{n,k+1}^{\mathrm{GL}}\left(\langle\langle \partial_t U_{\tau,h}, V_{\tau,h} \rangle\rangle + \langle\langle \mathcal{A}_h U_{\tau,h}, V_{\tau,h} \rangle\rangle\right) = Q_{n,k+1}^{\mathrm{GL}}\left(\langle\langle F, V_{\tau,h} \rangle\rangle\right)$$

for all $V_{\tau,h} \in \left(\mathbb{P}_{k-1}\left(I_n; V_h\right)\right)^2$.

In Problem 3.2 we use a $(k+1)$-point Gauss–Lobatto quadrature formula, which is in contrast to $Q_n^{\mathrm{GL}}$ in (2.6) that uses $k$ points. The quadrature formula on the left-hand side can be replaced by exact integration or by any quadrature formula which is exact for polynomials of degree up to order $2k-1$.

### 3.2. Space-time discretization with Galerkin–collocation method cGP-C$^1(k)$

From now on we suppose that $k \geq 3$ is a fixed integer without always mentioning this explicitly.

**Problem 3.3** (Local, numerically integrated, fully discrete problem of cGP-C$^1(k)$–cG$(r)$ on $I_n$)**.** Given $U_{\tau,h}\left(t_{n-1}^-\right)$ for $n>1$ and $U_{\tau,h}\left(t_0^-\right) = U_{0,h}$ for $n=1$, find $U_{\tau,h}|_{I_n} \in \left(\mathbb{P}_k\left(I_n;V_h\right)\right)^2$ such that

$$U_{\tau,h}\left(t_{n-1}^+\right) = U_{\tau,h}\left(t_{n-1}^-\right), \tag{3.1a}$$

$$\partial_t U_{\tau,h}\left(t_{n-1}^+\right) = -\mathcal{A}_h U_{\tau,h}\left(t_{n-1}^+\right) + \mathcal{P}_h F\left(t_{n-1}^+\right), \tag{3.1b}$$

$$\partial_t U_{\tau,h}\left(t_n^-\right) = -\mathcal{A}_h U_{\tau,h}\left(t_n^-\right) + \mathcal{P}_h F\left(t_n^-\right), \tag{3.1c}$$

and

$$Q_n^{\mathrm{H}}\left(\langle\!\langle \partial_t U_{\tau,h}, V_{\tau,h}\rangle\!\rangle + \langle\!\langle \mathcal{A}_h U_{\tau,h}, V_{\tau,h}\rangle\!\rangle\right) = Q_n^{\mathrm{H}}\left(\langle\!\langle F, V_{\tau,h}\rangle\!\rangle\right) \tag{3.1d}$$

for all $V_{\tau,h} \in \left(\mathbb{P}_{k-3}\left(I_n;V_h\right)\right)^2$.

For this scheme we make the following observations.

**Remark 3.4.** It directly follows from the definition of the scheme that $U_{\tau,h} \in \left(C^1\left(\overline{I};V_h\right)\right)^2$ is satisfied. Instead of the condition (3.1b) at $t_{n-1}^+$ we could also demand that

$$\partial_t U_{\tau,h}\left(t_{n-1}^+\right) = \partial_t U_{\tau,h}\left(t_{n-1}^-\right), \tag{3.2}$$

where we set $\partial_t U_{\tau,h}\left(t_0^-\right) = -\mathcal{A}_h U_{0,h} + \mathcal{P}_h F(0)$.

Since the time discretization is of Galerkin–Petrov type, we refer to it as a continuously differentiable Galerkin–Petrov approximation, for short cGP-C$^1(k)$.

Compared to Problem 3.2, the test space of the variational constraint (3.1d) reduces from $\left(\mathbb{P}_{k-1}\left(I_n;V_h\right)\right)^2$ to $\left(\mathbb{P}_{k-3}\left(I_n;V_h\right)\right)^2$. For $k=3$ the test space just becomes the set $\left(\mathbb{P}_0\left(I_n;V_h\right)\right)^2$ of piecewise constant functions in time. The collocation conditions (3.1b) and (3.1c) along with the reduced test space of the variational condition impact the block structure of the resulting linear algebraic system. By (3.2) a condensation of internal degrees of freedom becomes feasible which leads to smaller algebraic systems. Further, collocation conditions lead to sparser system matrices compared to variational conditions. This might simplify the future construction of efficient iterative solvers and preconditioners; *cf.* [6] for details.

The existence of a unique solution to Problem 3.3 can be proved along the lines of [15], page 812 and Theorem A.3 by using the equivalence of existence and uniqueness in the finite dimensional case.

For the scheme (3.1) we still state the following auxiliary result. We note that compared to Problem 3.3 the quadrature formula has been changed in Lemma 3.5. In addition, the test space has been increased from $\mathbb{P}_{k-3}$ to $\mathbb{P}_{k-2}$. Lemma 3.5 represents an important tool for the proofs below.

**Lemma 3.5.** *The solution $U_{\tau,h} \in \left(X_{\tau,h}^k\left(V_h\right)\right)^2$ of Problem 3.3 satisfies for $n=1,\ldots,N$ that*

$$Q_n^{\mathrm{GL}}\left(\langle\!\langle \partial_t U_{\tau,h}, V_{\tau,h}\rangle\!\rangle + \langle\!\langle \mathcal{A}_h U_{\tau,h}, V_{\tau,h}\rangle\!\rangle\right) = Q_n^{\mathrm{GL}}\left(\left\langle\!\left\langle I_\tau^{\mathrm{H}} F, V_{\tau,h}\right\rangle\!\right\rangle\right) \tag{3.3}$$

*for all $V_{\tau,h} \in \left(\mathbb{P}_{k-2}\left(I_n;V_h\right)\right)^2$.*

*Proof.* For arbitrarily chosen $V_{\tau,h} \in (\mathbb{P}_{k-2}(I_n; V_h))^2$, there exists some $d_n = d_n(V_{\tau,h}) \in V_h^2$ such that $V_{\tau,h}$ admits the representation

$$V_{\tau,h} = \widetilde{V}_{\tau,h} + d_n(V_{\tau,h})\,\psi_n$$

with

$$\widetilde{V}_{\tau,h} \in (\mathbb{P}_{k-3}(I_n; V_h))^2 \quad \text{and} \quad \psi_n(t) = \prod_{\mu=2}^{k-1} \left(t - t_{n,\mu}^{\mathrm{GL}}\right) \in \mathbb{P}_{k-2}(I_n)$$

where $t_{n,\mu}^{\mathrm{GL}}$, $\mu = 2, \ldots, k-1$, denote the inner Gauss–Lobatto quadrature points on $\overline{I}_n$. From (3.1d) along with the exactness of the Hermite-type quadrature formula (2.5) for all polynomials in $\mathbb{P}_{2k-3}(I_n)$ and of the Gauss–Lobatto quadrature formula (2.6) for all polynomials in $\mathbb{P}_{2k-3}(I_n)$, it follows that

$$\begin{aligned}
Q_n^{\mathrm{GL}}\left(\left\langle\!\left\langle \partial_t U_{\tau,h}, \widetilde{V}_{\tau,h} \right\rangle\!\right\rangle + \left\langle\!\left\langle \mathcal{A}_h U_{\tau,h}, \widetilde{V}_{\tau,h} \right\rangle\!\right\rangle\right) &= Q_n^{\mathrm{H}}\left(\left\langle\!\left\langle \partial_t U_{\tau,h}, \widetilde{V}_{\tau,h} \right\rangle\!\right\rangle + \left\langle\!\left\langle \mathcal{A}_h U_{\tau,h}, \widetilde{V}_{\tau,h} \right\rangle\!\right\rangle\right) \\
&= Q_n^{\mathrm{H}}\left(\left\langle\!\left\langle F, \widetilde{V}_{\tau,h} \right\rangle\!\right\rangle\right) = Q_n^{\mathrm{H}}\left(\left\langle\!\left\langle I_\tau^{\mathrm{H}} F, \widetilde{V}_{\tau,h} \right\rangle\!\right\rangle\right) \\
&= Q_n^{\mathrm{GL}}\left(\left\langle\!\left\langle I_\tau^{\mathrm{H}} F, \widetilde{V}_{\tau,h} \right\rangle\!\right\rangle\right).
\end{aligned} \tag{3.4}$$

Therefore, it remains to prove that

$$Q_n^{\mathrm{GL}}\left(\langle\!\langle \partial_t U_{\tau,h}, d_n\psi_n \rangle\!\rangle + \langle\!\langle \mathcal{A}_h U_{\tau,h}, d_n\psi_n \rangle\!\rangle\right) = Q_n^{\mathrm{GL}}\left(\left\langle\!\left\langle I_\tau^{\mathrm{H}} F, d_n\psi_n \right\rangle\!\right\rangle\right) \tag{3.5}$$

is satisfied. Since $\psi_n$ vanishes in the interior Gauss–Lobatto quadrature nodes $t_{n,\mu}^{\mathrm{GL}}$, $\mu = 2, \ldots, k-1$, and the quantities $\langle\!\langle \partial_t U_{\tau,h}, d_n\psi_n \rangle\!\rangle + \langle\!\langle \mathcal{A}_h U_{\tau,h}, d_n\psi_n \rangle\!\rangle$ and $\left\langle\!\left\langle I_\tau^{\mathrm{H}} F, d_n\psi_n \right\rangle\!\right\rangle$ coincide in the endpoints $t_{n-1}^+$ and $t_n^-$ by means of the conditions (3.1b) and (3.1c), the variational problem (3.5) is satisfied. Along with (3.4), this proves the assertion (3.3). $\square$

Moreover, it can be shown that the solution $U_{\tau,h}$ of Problem 3.3 fulfills the evolution problem

$$\partial_t U_{\tau,h} + I_\tau^{\mathrm{GL}} \mathcal{A}_h U_{\tau,h} = \mathcal{P}_h I_\tau^{\mathrm{GL}} I_\tau^{\mathrm{H}} F \tag{3.6}$$

on the whole time interval $\overline{I}$. Here, we refer for further details to the preprint, see [7], Lemma 3.6 and Remark 3.7.

## 4. Error estimates

The overall goal of this work is to prove estimates for the error

$$E(t) := U(t) - U_{\tau,h}(t),$$

where the Galerkin–collocation approximation $U_{\tau,h}$ is the solution of Problem 3.3. We will use in the sequel the componentwise representation $E(t) = \left(e^0(t), e^1(t)\right)$. We observe that $E$ is continuously differentiable in time on $\overline{I}$ if we assume for the exact solution that $U = \left(u^0, u^1\right) \in \left(C^1\left(\overline{I}; V\right)\right)^2$.

For each time interval $I_n$, $n = 1, \ldots, N$, we define the bilinear form

$$B_n^{\mathrm{GL}}(W, V) := Q_n^{\mathrm{GL}}\left(\langle\!\langle \partial_t W, V \rangle\!\rangle\right) + Q_n^{\mathrm{GL}}\left(\langle\!\langle \mathcal{A}_h W, V \rangle\!\rangle\right),$$

where $W$ and $V$ have to satisfy some smoothness conditions to ensure that $B_n^{\mathrm{GL}}$ is well-defined.

Our analysis will follow the main lines given in [16] since the solution $U_{\tau,h}$ in this paper is related to $L_\tau U_{\tau,h}$ there with the difference that our polynomial order $k$ is related to $k+1$ in [16]. This relation is motivated by the fact that the solution of the numerically integrated cGP-C$^1$($k$)–cG($r$), given in Problem 3.3 could also be interpreted as the post-processed solution of a numerically integrated cGP($k$-1)–cG($r$) scheme, given in Problem 3.2, with a modified right-hand side in that $F$ is replaced by $I_\tau^{\mathrm{H}} F$. For brevity, we will cite the results used from [16] and focus on new arguments in the analysis.

### 4.1. Error estimates for $\partial_t U_{\tau,h}$

We start with proving an $L^\infty(L^2)$-norm estimate for the time derivative $\partial_t E(t)$ of the error as an auxiliary result. This represents an essential argument in our proof and is specific to the hyperbolic character of (2.1). Using the $L^\infty(L^2)$-bound for $\partial_t E(t)$ an estimate for $E(t)$ will be proved in Section 4.2.

In order to bound $\partial_t E(t)$, we derive in the following theorem a variational problem satisfied by $\partial_t U_{\tau,h}$.

**Theorem 4.1.** *Let $U_{\tau,h} \in \left(X^k_{\tau,h}(V_h)\right)^2$ be the solution of Problem 3.3. Then, its time derivative $\partial_t U_{\tau,h} \in \left(X^{k-1}_{\tau,h}(V_h)\right)^2$ satisfies for all $n = 1,\dots,N$ the equation*

$$B_n^{\mathrm{GL}}\left(\partial_t U_{\tau,h}, V_{\tau,h}\right) = Q_n^{\mathrm{GL}}\left(\left\langle\!\left\langle \partial_t I_\tau^{\mathrm{H}} F, V_{\tau,h}\right\rangle\!\right\rangle\right) = \int_{I_n} \left\langle\!\left\langle \partial_t I_\tau^{\mathrm{H}} F, V_{\tau,h}\right\rangle\!\right\rangle \,\mathrm{d}t \tag{4.1}$$

*for all $V_{\tau,h} \in \left(\mathbb{P}_{k-2}(I_n;V_h)\right)^2$.*

*Proof.* Recalling that $\partial_t U_{\tau,h} \in \left(\mathbb{P}_{k-1}(I_n;V_h)\right)^2$, we get by the exactness of the Gauss–Lobatto formula (2.6) for all polynomials in $\mathbb{P}_{2k-3}(I_n;\mathbb{R})$ along with integration by parts that

$$\begin{aligned} B_n^{\mathrm{GL}}\left(\partial_t U_{\tau,h}, V_{\tau,h}\right) &= Q_n^{\mathrm{GL}}\left(\left\langle\!\!\left\langle \underbrace{\partial_t^2 U_{\tau,h} + \mathcal{A}_h \partial_t U_{\tau,h}}_{\in (P_{k-1}(I_n;V_h))^2}, V_{\tau,h}\right\rangle\!\!\right\rangle\right) = \int_{I_n} \left\langle\!\left\langle \partial_t\left(\partial_t U_{\tau,h} + \mathcal{A}_h U_{\tau,h}\right), V_{\tau,h}\right\rangle\!\right\rangle \,\mathrm{d}t \\ &= -\int_{I_n} \left\langle\!\left\langle \partial_t U_{\tau,h} + \mathcal{A}_h U_{\tau,h}, \partial_t V_{\tau,h}\right\rangle\!\right\rangle \,\mathrm{d}t + \left\langle\!\left\langle \partial_t U_{\tau,h} + \mathcal{A}_h U_{\tau,h}, V_{\tau,h}\right\rangle\!\right\rangle \Big|_{t_{n-1}^+}^{t_n^-} \end{aligned} \tag{4.2}$$

for $V_{\tau,h} \in \left(\mathbb{P}_{k-2}(I_n,V_h)\right)^2$. Using the exactness of the Hermite quadrature formula $Q_n^{\mathrm{H}}$ for polynomials in $\mathbb{P}_{2k-3}(I_n;\mathbb{R})$ and (3.1d), we conclude from (4.2) that

$$\begin{aligned} B_n^{\mathrm{GL}}\left(\partial_t U_{\tau,h}, V_{\tau,h}\right) &= -Q_n^{\mathrm{H}}\left(\left\langle\!\left\langle F, \partial_t V_{\tau,h}\right\rangle\!\right\rangle\right) + \left\langle\!\left\langle \partial_t U_{\tau,h} + \mathcal{A}_h U_{\tau,h}, V_{\tau,h}\right\rangle\!\right\rangle \Big|_{t_{n-1}^+}^{t_n^-} \\ &= -\int_{I_n} \left\langle\!\left\langle I_\tau^{\mathrm{H}} F, \partial_t V_{\tau,h}\right\rangle\!\right\rangle \,\mathrm{d}t + \left\langle\!\left\langle \partial_t U_{\tau,h} + \mathcal{A}_h U_{\tau,h}, V_{\tau,h}\right\rangle\!\right\rangle \Big|_{t_{n-1}^+}^{t_n^-} \\ &= \int_{I_n} \left\langle\!\left\langle \partial_t I_\tau^{\mathrm{H}} F, V_{\tau,h}\right\rangle\!\right\rangle \,\mathrm{d}t - \left\langle\!\left\langle I_\tau^{\mathrm{H}} F, V_{\tau,h}\right\rangle\!\right\rangle \Big|_{t_{n-1}^+}^{t_n^-} + \left\langle\!\left\langle \partial_t U_{\tau,h} + \mathcal{A}_h U_{\tau,h}, V_{\tau,h}\right\rangle\!\right\rangle \Big|_{t_{n-1}^+}^{t_n^-}. \end{aligned} \tag{4.3}$$

From (3.1b) and (3.1c) along with the interpolation properties of $I_\tau^{\mathrm{H}}$, it follows that

$$\partial_t U_{\tau,h}(t_*) + \mathcal{A}_h U_{\tau,h}(t_*) = \mathcal{P}_h I_\tau^{\mathrm{H}} F(t_*) \tag{4.4}$$

for $t_* \in \left\{t_{n-1}^+, t_n^-\right\}$. Combining (4.3) with (4.4) shows that

$$B_n^{\mathrm{GL}}\left(\partial_t U_{\tau,h}, V_{\tau,h}\right) = \int_{I_n} \left\langle\!\left\langle \partial_t I_\tau^{\mathrm{H}} F, V_{\tau,h}\right\rangle\!\right\rangle \,\mathrm{d}t$$

for all $V_{\tau,h} \in \left(\mathbb{P}_{k-2}(I_n,V_h)\right)^2$. Recalling that $\partial_t I_\tau^{\mathrm{H}} F|_{I_n} \in \left(\mathbb{P}_{k-1}(I_n;V_h)\right)^2$ and the exactness of the Gauss–Lobatto quadrature for functions of $\mathbb{P}_{2k-3}(I_n;\mathbb{R})$, this proves the assertion of the theorem. □

**Remark 4.2.** If the solution $u$ of (1.1) is sufficiently regular, the time derivative $\partial_t U = \left(\partial_t u, \partial_t^2 u\right)$ solves the evolution problem

$$\partial_t\left(\partial_t U\right) + \mathcal{A}\left(\partial_t U\right) = \partial_t F \quad \text{in} \quad (0,T), \quad \partial_t U(0) = -\mathcal{A}U(0) + F(0). \tag{4.5}$$

Assumptions on the data such that (4.5) is satisfied can be found in, *e.g.*, p. 410 and Theorem 5 of [24].

Rewriting (4.1) as

$$B_n^{\mathrm{GL}}\left(\partial_t U_{\tau,h}, V_{\tau,h}\right) = \int_{I_n} \langle\!\langle \partial_t F, V_{\tau,h}\rangle\!\rangle \,\mathrm{d}t + \int_{I_n} \left\langle\!\!\left\langle \partial_t I_\tau^{\mathrm{H}} F - \partial_t F, V_{\tau,h}\right\rangle\!\!\right\rangle \,\mathrm{d}t, \tag{4.6}$$

its discrete solution can now be regarded as the cGP($k$-1)–cG($r$) approximation of the evolution problem (4.5) up to the perturbation term $\int_{I_n} \left\langle\!\left\langle \partial_t I_\tau^{\mathrm{H}} F - \partial_t F, V_{\tau,h}\right\rangle\!\right\rangle \,\mathrm{d}t$ on the right-hand side. Further, the collocation condition (3.1b) for $n = 1$ along with the initial condition $U_{\tau,h}(0) = U_{0,h}$ shows that $\partial_t U_{\tau,h}(0) = -\mathcal{A}_h U_{0,h} + \mathcal{P}_h F(0)$ is satisfied.

We point out that there is a strong analogy between Remark 4.2 and Remark 5.3 of [16]. The main difference of the two statements comes through the different perturbation terms. Regarding the relation of the polynomial orders, both perturbation terms are of the same approximation order. Hence, we can directly follow the further arguments used in [16]. Especially, some assumptions about the discrete initial value $\partial_t U_{\tau,h}(0)$ with respect to the continuous initial value $\partial_t U(0)$ have to be fulfilled.

**Lemma 4.3.** *Let $U_{0,h} := (R_h u_0, R_h u_1)$. Then there holds that*

$$\partial_t U_{\tau,h}(0) = \begin{pmatrix} R_h & 0 \\ 0 & P_h \end{pmatrix} \partial_t U(0).$$

We refer to Lemma 5.4 of [16] for the proof of Lemma 4.3 taking into consideration that $U_{\tau,h}$ here is associated to $L_\tau U_{\tau,h}$ in [16]. We also note that the analogue of Assumption 3.6 from [16] is satisfied by $U_{\tau,h}$ due to (3.1a) and (3.1b) for $n = 1$. To keep this work self-contained, we need to cite explicitly Theorem 5.5 of [16] that generalizes a result of [34] for the cGP($k$)–cG($r$) approximation of the wave equation.

**Theorem 4.4.** *Let $\hat{u}$ denote the solution of (1.1) with data $\hat{f}$, $\hat{u}_0$, $\hat{u}_1$ instead of $f$, $u_0$, $u_1$. Suppose $\ell \in \mathbb{N}$ and let $\hat{f}_\tau$ be an approximation of $\hat{f}$ such that*

$$\|\hat{f} - \hat{f}_\tau\|_{C(\overline{I}_n;H)} \le C_{\hat{f}}\, \tau_n^{\ell+1}, \quad n = 1, \dots, N, \tag{4.7}$$

*where the constant $C_{\hat{f}}$ depends on $\hat{f}$ but is independent of $n$, $N$, and $\tau_n$. Furthermore, let $\widehat{U}_{\tau,h} = \left(\hat{u}^0_{\tau,h}, \hat{u}^1_{\tau,h}\right) \in \left(X_\tau^\ell\left(V_h\right)\right)^2$ be the solution of the local (on $I_n$) perturbed* cGP($\ell$)–cG($r$) *problem*

$$\int_{I_n} \left(\left\langle\!\!\left\langle \partial_t \widehat{U}_{\tau,h}, V_{\tau,h}\right\rangle\!\!\right\rangle + \left\langle\!\!\left\langle \mathcal{A}_h \widehat{U}_{\tau,h}, V_{\tau,h}\right\rangle\!\!\right\rangle\right) \mathrm{d}t = \int_{I_n} \left\langle\!\!\left\langle \widehat{F}_\tau, V_{\tau,h}\right\rangle\!\!\right\rangle \mathrm{d}t \tag{4.8}$$

*for all test functions $V_{\tau,h} = \left(v^0_{\tau,h}, v^1_{\tau,h}\right) \in \left(\mathbb{P}_{\ell-1}\left(I_n; V_h\right)\right)^2$ with $\widehat{F}_\tau := \left(0, \hat{f}_\tau\right)$ and the initial value $\widehat{U}_{\tau,h}\left(t^+_{n-1}\right) = \widehat{U}_{\tau,h}\left(t^-_{n-1}\right)$ for $n > 1$ and $\widehat{U}_{\tau,h}\left(t_0\right) = \widehat{U}_{0,h} := \left(R_h \hat{u}_0, P_h \hat{u}_1\right)$. For a sufficiently smooth exact solution $\hat{u}$, the estimates*

$$\|\hat{u}(t) - \hat{u}^0_{\tau,h}(t)\| + \|\partial_t \hat{u}(t) - \hat{u}^1_{\tau,h}(t)\| \lesssim \tau^{l+1}\, C_t\left(\hat{u}\right) + h^{r+1}\, C_x\left(\hat{u}\right), \tag{4.9}$$

$$\|\nabla\left(\hat{u}(t) - \hat{u}^0_{\tau,h}(t)\right)\| \lesssim \tau^{l+1}\, C_t\left(\hat{u}\right) + h^{r}\, C_x\left(\hat{u}\right) \tag{4.10}$$

*hold for all $t \in \overline{I}$, where $C_t\left(\hat{u}\right)$ and $C_x\left(\hat{u}\right)$ are quantities depending on various temporal and spatial derivatives of $\hat{u}$.*

We conclude from Theorem 4.4 the following error estimates.

**Theorem 4.5.** *Let $U_{0,h} := (R_h u_0, R_h u_1)$ and assume that the exact solution $U = \left(u^0, u^1\right) := (u, \partial_t u)$ is sufficiently smooth. Then the error estimates*

$$\|\partial_t U(t) - \partial_t U_{\tau,h}(t)\| \lesssim \tau^k\, C_t\,(\partial_t u) + h^{r+1}\, C_x\,(\partial_t u) \lesssim \tau^k + h^{r+1}, \tag{4.11}$$

$$\|\nabla\left(\partial_t u^0(t) - \partial_t u^0_{\tau,h}(t)\right)\| \lesssim \tau^k\, C_t\,(\partial_t u) + h^r\, C_x\,(\partial_t u) \lesssim \tau^k + h^r \tag{4.12}$$

*hold for all $t \in \overline{I}$, where $C_t\,(\partial_t u)$ and $C_x\,(\partial_t u)$ are quantities depending on various temporal and spatial derivatives of $\partial_t u$.*

*Proof.* To prove (4.11) and (4.12), we apply Theorem 4.4. Since the solution $u$ is sufficiently smooth, the function $\hat{u} := \partial_t u$ is the solution of the wave equation (1.1) with the right-hand side $\hat{f} := \partial_t f$ and the initial conditions $\hat{u}(0) = \hat{u}_0 := u_1$ and $\partial_t \hat{u}(0) = \hat{u}_1 := f(0) - A u_0$. Let us define the modified right-hand side $\hat{f}_\tau := \partial_t I^{\mathrm{H}}_\tau f$ and $\widehat{F}_\tau := \left(0, \hat{f}_\tau\right)$. Then, the discrete function $\widehat{U}_{\tau,h} := \partial_t U_{\tau,h} \in \left(X^{k-1}_\tau\,(V_h)\right)^2$ satisfies all the conditions required for the discrete solution $\widehat{U}_{\tau,h}$ in Theorem 4.4 with $\ell = k-1$. In fact, by the construction of the discrete solution $U_{\tau,h}$ in Problem 3.3, the continuity of $\partial_t U_{\tau,h}$ in the discrete points $t_n$, $n = 0, \ldots, N$, is ensured by the conditions (3.1a)–(3.1c). Therefore, it holds that $\widehat{U}_{\tau,h} \in (\mathbb{P}_{k-1}\,(I_n; V_h))^2$ and that $\widehat{U}_{\tau,h}\left(t^+_{n-1}\right) = \widehat{U}_{\tau,h}\left(t^-_{n-1}\right)$. Moreover, from $U_{0,h} := (R_h u_0, R_h u_1)$ and Lemma 4.3, we get that $\widehat{U}_{0,h} = \widehat{U}_{\tau,h}(0) = \partial_t U_{\tau,h}(0) = (R_h \hat{u}_0, P_h \hat{u}_1)$. Theorem 4.1 implies for all $n = 1, \ldots, N$ and all $V_{\tau,h} \in (\mathbb{P}_{k-2}\,(I_n; V_h))^2$ that

$$B^{\mathrm{GL}}_n\left(\widehat{U}_{\tau,h}, V_{\tau,h}\right) = Q^{\mathrm{GL}}_n\left(\left\langle\!\left\langle \partial_t \widehat{U}_{\tau,h}, V_{\tau,h}\right\rangle\!\right\rangle + \left\langle\!\left\langle \mathcal{A}_h \widehat{U}_{\tau,h}, V_{\tau,h}\right\rangle\!\right\rangle\right) = Q^{\mathrm{GL}}_n\left(\left\langle\!\left\langle \widehat{F}_\tau, V_{\tau,h}\right\rangle\!\right\rangle\right).$$

Each quadrature formula in the previous equation is exact since all integrands are polynomials in $t$ with degree not greater than $2k-3$ such that equation (4.8) of Theorem 4.4 is satisfied. Thus, we have shown that $\widehat{U}_{\tau,h}$ is the discrete solution of Theorem 4.4 for the above defined data. To verify the approximation property for $\hat{f}_\tau$, we use the definition of $\hat{f}$ and $\hat{f}_\tau$, apply standard error estimates of Hermite interpolation, and obtain (4.7) with a constant $C_{\hat{f}} = C\|\partial_t^{k+1} f\|_{C(\overline{I};H)}$. Then, we use Theorem 4.4 with $\ell = k-1$. Recalling the representation by components, $\partial_t U = \left(\partial_t u^0, \partial_t u^1\right) = (\hat{u}, \partial_t \hat{u})$ and $\widehat{U}_{\tau,h} = \left(\hat{u}^0_{\tau,h}, \hat{u}^1_{\tau,h}\right) = \left(\partial_t u^0_{\tau,h}, \partial_t u^1_{\tau,h}\right)$, we directly get assertion (4.11) from (4.9) and assertion (4.12) from (4.10). ☐

## 4.2. Error estimates for $U_{\tau,h}$

This section is devoted to the desired norm estimates for the error $E(t) := U(t) - U_{\tau,h}(t)$ where $U_{\tau,h}$ is the solution of Problem 3.3. For our error analysis we consider the decomposition

$$E(t) = \Theta(t) + E_{\tau,h}(t) \quad \text{with} \quad \Theta(t) := U(t) - \mathcal{R}_h R^k_\tau U(t) \quad \text{and} \quad E_{\tau,h} := \mathcal{R}_h R^k_\tau U(t) - U_{\tau,h} \tag{4.13}$$

for all $t \in \overline{I}$ and define the components $E_{\tau,h}(t) = \left(e^0_{\tau,h}(t), e^1_{\tau,h}(t)\right)$. We observe that both $\Theta$ and $E_{\tau,h}$ are continuously differentiable in time on $\overline{I}$ if the exact solution $U$ is sufficiently smooth. Moreover, $\Theta$ and $E_{\tau,h}$ are smooth enough to be used as arguments in the bilinear form $B^{\mathrm{GL}}_n$.

The following estimates of the interpolation error $\Theta$ in (4.13) can be found in Lemma 5.7 of [16]. They rely on the approximation properties of $\mathcal{R}_h$ and $R^k_\tau$; *cf.*, Section 4.1 of [16] for details.

**Lemma 4.6** (Estimation of the interpolation error)**.** *Let $m \in \{0, 1\}$. Then, the error estimates*

$$\|\Theta(t)\|_m \lesssim h^{r+1-m} + \tau_n^{k+1}, \quad t \in \overline{I}_n, \tag{4.14}$$

$$\|\partial_t \Theta(t)\|_m \lesssim h^{r+1-m} + \tau_n^{k}, \quad t \in \overline{I}_n, \tag{4.15}$$

*hold for all $n = 1, \ldots, N$ where $\|\cdot\|_0 := \|\cdot\|$.*

Next, we address the discrete error $E_{\tau,h}$ of the decomposition (4.13) between the interpolation $\mathcal{R}_h R_\tau^k U$ and the fully discrete solution $U_{\tau,h}$. We start with some auxiliary results.

**Lemma 4.7** (Consistency error). *Assume that $U \in C^1\left(\overline{I}; V\right) \times C^1\left(\overline{I}; H\right)$. Then, for all $n = 1, \dots, N$ the identity*

$$B_n^{\mathrm{GL}}\left(E, V_{\tau,h}\right) = Q_n^{\mathrm{GL}}\left(\langle\!\langle I_\tau^{\mathrm{GL}} F - I_\tau^{\mathrm{H}} F, V_{\tau,h}\rangle\!\rangle\right) = Q_n^{\mathrm{GL}}\left(\langle\!\langle F - I_\tau^{\mathrm{H}} F, V_{\tau,h}\rangle\!\rangle\right)$$

*is satisfied for all $V_{\tau,h} \in \left(Y_{\tau,h}^{k-2}\left(V_h\right)\right)^2$.*

*Proof.* We recall from Lemma 3.5 that for all $n = 1, \dots, N$ the identity

$$B_n^{\mathrm{GL}}\left(U_{\tau,h}, V_{\tau,h}\right) = Q_n^{\mathrm{GL}}\left(\langle\!\langle I_\tau^{\mathrm{H}} F, V_{\tau,h}\rangle\!\rangle\right) \tag{4.16}$$

holds for all $V_{\tau,h} \in \left(\mathbb{P}_{k-2}\left(I_n; V_h\right)\right)^2$. We have under sufficient smoothness assumptions on the exact solution that

$$\partial_t U\left(t_{n,\mu}^{\mathrm{GL}}\right) + \mathcal{A} U\left(t_{n,\mu}^{\mathrm{GL}}\right) = F\left(t_{n,\mu}^{\mathrm{GL}}\right), \quad \mu = 1, \dots, k. \tag{4.17}$$

By the consistency (2.4) of $\mathcal{A}_h$, the identity (4.17) implies

$$\begin{aligned} B_n^{\mathrm{GL}}\left(U, V_{\tau,h}\right) &= Q_n^{\mathrm{GL}}\left(\langle\!\langle \partial_t U + \mathcal{A}_h U, V_{\tau,h}\rangle\!\rangle\right) \\ &= Q_n^{\mathrm{GL}}\left(\langle\!\langle \partial_t U + \mathcal{A} U, V_{\tau,h}\rangle\!\rangle\right) = Q_n^{\mathrm{GL}}\left(\langle\!\langle I_\tau^{\mathrm{GL}} F, V_{\tau,h}\rangle\!\rangle\right). \end{aligned} \tag{4.18}$$

Combining (4.16) with (4.18) and recalling that $E = U - U_{\tau,h}$ prove the assertion. $\square$

The following lemma is a minor generalization of Lemma 5.9 from [16].

**Lemma 4.8.** *Let $p \in \mathbb{P}_k\left(I_n\right)$ be a polynomial of degree less than or equal to $k$. Then, the relation*

$$\partial_t p\left(t_{n,\mu}^{\mathrm{G}}\right) = \partial_t I_\tau^{\mathrm{GL}} p\left(t_{n,\mu}^{\mathrm{G}}\right)$$

*holds for all Gauss points $t_{n,\mu}^G \in I_n$, $\mu = 1, \dots, k-1$.*

Next, we give a stability result for the form $B_n^{\mathrm{GL}}(\cdot,\cdot)$. It is more general than Lemma 5.10 of [16]. This enables a wider applicability, for example, in Theorem 5.3.

**Lemma 4.9** (Stability of $B_n^{\mathrm{GL}}$). *For any function $V_{\tau,h} = \left(v_{\tau,h}^0, v_{\tau,h}^1\right) \in \left(\mathbb{P}_k\left(I_n; V_h\right)\right)^2$ there holds that*

$$\begin{aligned} &B_n^{\mathrm{GL}}\big(\left(v_{\tau,h}^0, v_{\tau,h}^1\right), \left(\Pi_n^{k-2} A_h I_\tau^{\mathrm{GL}} v_{\tau,h}^0, \Pi_n^{k-2} I_\tau^{\mathrm{GL}} v_{\tau,h}^1\right)\big) \\ &\qquad = \frac{1}{2}\left(\|\nabla v_{\tau,h}^0\left(t_n\right)\|^2 - \|\nabla v_{\tau,h}^0\left(t_{n-1}\right)\|^2 + \|v_{\tau,h}^1\left(t_n\right)\|^2 - \|v_{\tau,h}^1\left(t_{n-1}\right)\|^2\right) \end{aligned} \tag{4.19}$$

*for all $n = 1, \dots, N$.*

We note that (4.19) also holds true if the second argument in $B_n^{\mathrm{GL}}$ is $\left(-\partial_t I_\tau^{\mathrm{GL}} v_{\tau,h}^1, \partial_t I_\tau^{\mathrm{GL}} v_{\tau,h}^0\right)$. This can be concluded from the proof of Lemma 6.1 from [7].

Exploiting the correspondence of $U_{\tau,h}$ in this paper to $L_\tau U_{\tau,h}$ in [16] and keeping in mind that $k$ here is related to $k+1$ there, we can recall from [16] the following result of boundedness (*cf.* [16], Lem. 5.11).

**Lemma 4.10** (Boundedness). *Let $V_{\tau,h} = \left(\Pi_n^{k-2} A_h I_\tau^{\mathrm{GL}} e_{\tau,h}^0, \Pi_n^{k-2} I_\tau^{\mathrm{GL}} e_{\tau,h}^1\right)$. Then, the bound*

$$\left|B_n^{\mathrm{GL}}\left(\Theta, V_{\tau,h}\right)\right| \lesssim \tau_n^{1/2}\left(\tau_n^{k+1} + h^{r+1}\right)\left\{\tau_n \|E_{\tau,h}\left(t_{n-1}\right)\|^2 + \tau_n^2 Q_n^{\mathrm{G}}\left(\|\partial_t E_{\tau,h}\|^2\right)\right\}^{1/2}$$

*holds for all $n = 1, \dots, N$.*

We proceed with estimating the consistency error given in Lemma 4.7.

**Lemma 4.11** (Estimates on right-hand side term)**.** *Let* $V_{\tau,h} = \left(v_{\tau,h}^0, v_{\tau,h}^1\right) = \left(\Pi_n^{k-2} A_h I_\tau^{\mathrm{GL}} e_{\tau,h}^0, \Pi_n^{k-2} I_\tau^{\mathrm{GL}} e_{\tau,h}^1\right)$. *Then, the estimate*

$$Q_n^{\mathrm{GL}}\left(\langle\langle (0, f - I_\tau^{\mathrm{H}} f), (v_{\tau,h}^0, v_{\tau,h}^1)\rangle\rangle\right) \lesssim \tau_n^{1/2}\tau_n^{k+1}\left\{\tau_n \|E_{\tau,h}(t_{n-1})\|^2 + \tau_n^2 Q_n^{\mathrm{G}}\left(\|\partial_t E_{\tau,h}\|^2\right)\right\}^{1/2}$$

*holds for all* $n = 1, \ldots, N$.

*Proof.* The Cauchy–Schwarz inequality and standard error estimates of Hermite interpolation show that

$$\begin{aligned} Q_n^{\mathrm{GL}}\left(\langle\langle (0, f - I_\tau^{\mathrm{H}} f), (v_{\tau,h}^0, v_{\tau,h}^1)\rangle\rangle\right) &= Q_n^{\mathrm{GL}}\left(\langle f - I_\tau^{\mathrm{H}} f, \Pi_n^{k-2} I_\tau^{\mathrm{GL}} e_{\tau,h}^1\rangle\right) \\ &\le \left(Q_n^{\mathrm{GL}}\left(\|f - I_\tau^{\mathrm{H}} f\|^2\right)\right)^{1/2}\left(Q_n^{\mathrm{GL}}\left(\|\Pi_n^{k-2} I_\tau^{\mathrm{GL}} e_{\tau,h}^1\|^2\right)\right)^{1/2} \\ &\lesssim \tau_n^{1/2}\tau_n^{k+1}\left(Q_n^{\mathrm{GL}}\left(\|\Pi_n^{k-2} I_\tau^{\mathrm{GL}} e_{\tau,h}^1\|^2\right)\right)^{1/2}. \end{aligned}$$

Using the exactness of $Q_n^{\mathrm{GL}}$ for polynomials up to degree $2k-3$, the stability of the $L^2$-projection $\Pi_n^{k-2}$, the standard norm bound $\|v\|_{L^2(I_n;L^2(\Omega))}^2 \lesssim \tau_n \|v(t_{n-1})\|^2 + \tau_n^2 \|\partial_t v\|_{L^2(I_n;L^2(\Omega))}^2$ for $v \in H^1(I_n; H)$ (*cf.* [16], Lem. 4.6), and Lemma 4.8, we finally conclude that

$$\begin{aligned} Q_n^{\mathrm{GL}}\left(\|\Pi_n^{k-2} I_\tau^{\mathrm{GL}} e_{\tau,h}^1\|^2\right) &= \int_{I_n} \|\Pi_n^{k-2} I_\tau^{\mathrm{GL}} e_{\tau,h}^1\|^2 \,\mathrm{d}t \le \int_{I_n} \|I_\tau^{\mathrm{GL}} e_{\tau,h}^1\|^2 \,\mathrm{d}t \\ &\lesssim \tau_n \|I_\tau^{\mathrm{GL}} e_{\tau,h}^1(t_{n-1})\|^2 + \tau_n^2 \int_{I_n} \|\partial_t I_\tau^{\mathrm{GL}} e_{\tau,h}^1\|^2 \,\mathrm{d}t \\ &= \tau_n \|e_{\tau,h}^1(t_{n-1})\|^2 + \tau_n^2 Q_n^{\mathrm{G}}\left(\|\partial_t I_\tau^{\mathrm{GL}} e_{\tau,h}^1\|^2\right) \\ &= \tau_n \|e_{\tau,h}^1(t_{n-1})\|^2 + \tau_n^2 Q_n^{\mathrm{G}}\left(\|\partial_t e_{\tau,h}^1\|^2\right). \end{aligned}$$

Combining both estimates, the assertion of the lemma follows directly. □

**Lemma 4.12** (Estimates on $E_{\tau,h}$)**.** *Let* $U_{0,h} := (R_h u_0, R_h u_1)$. *Then, the estimate*

$$\|e_{\tau,h}^0(t_n)\|_1^2 + \|e_{\tau,h}^1(t_n)\|^2 \lesssim \left(\tau^{k+1} + h^{r+1}\right)^2 \tag{4.20}$$

*is satisfied for all* $n = 1, \ldots, N$. *Moreover, for all* $t \in \overline{I}$ *we have that*

$$\|\nabla e_{\tau,h}^0(t)\| \lesssim \tau^{k+1} + h^r, \tag{4.21}$$

$$\|e_{\tau,h}^0(t)\| + \|e_{\tau,h}^1(t)\| \lesssim \tau^{k+1} + h^{r+1}. \tag{4.22}$$

*Proof.* We conclude from Lemma 4.7 that

$$B_n^{\mathrm{GL}}(E_{\tau,h}, V_{\tau,h}) = -B_n^{\mathrm{GL}}(\Theta, V_{\tau,h}) + Q_n^{\mathrm{GL}}\left(\langle\langle F - I_\tau^{\mathrm{H}} F, V_{\tau,h}\rangle\rangle\right)$$

is satisfied for all $V_{\tau,h} \in \left(Y_{\tau,h}^{k-2}(V_h)\right)^2$. Choosing here $V_{\tau,h} = \left(\Pi_n^{k-2} A_h I_\tau^{\mathrm{GL}} e_{\tau,h}^0, \Pi_n^{k-2} I_\tau^{\mathrm{GL}} e_{\tau,h}^1\right)$ and using Lemmas 4.10 and 4.11 yield that

$$\begin{aligned} &B_n^{\mathrm{GL}}\left((e_{\tau,h}^0, e_{\tau,h}^1), (\Pi_n^{k-2} A_h I_\tau^{\mathrm{GL}} e_{\tau,h}^0, \Pi_n^{k-2} I_\tau^{\mathrm{GL}} e_{\tau,h}^1)\right) \\ &\quad = -B_n^{\mathrm{GL}}\left((\theta^0, \theta^1), (\Pi_n^{k-2} A_h I_\tau^{\mathrm{GL}} e_{\tau,h}^0, \Pi_n^{k-2} I_\tau^{\mathrm{GL}} e_{\tau,h}^1)\right) \\ &\qquad + Q_n^{\mathrm{GL}}\left((0, f - I_\tau^{\mathrm{H}} f), (\Pi_n^{k-2} A_h I_\tau^{\mathrm{GL}} e_{\tau,h}^0, \Pi_n^{k-2} I_\tau^{\mathrm{GL}} e_{\tau,h}^1)\right) \\ &\quad \lesssim \tau_n^{1/2}\left(\tau_n^{k+1} + h^{r+1}\right)\left\{\tau_n \|E_{\tau,h}(t_{n-1})\|^2 + \tau_n^2 Q_n^{\mathrm{G}}\left(\|\partial_t E_{\tau,h}\|^2\right)\right\}^{1/2}. \end{aligned} \tag{4.23}$$

Since the upper bound in (4.23) coincides with that in equation (5.46) of [16] and our $E_{\tau,h}$ can be identified with $\widetilde{E}_{\tau,h}$ of [16], we present here just a short summary of the proof of Lemma 5.12 in [16].

Next, we combine (4.23) with the stability result of Lemma 4.9 for the choice $V_{\tau,h} = \left(e^0_{\tau,h}, e^1_{\tau,h}\right)$. Further, we apply the Cauchy–Schwarz inequality to the last term on the right-hand side of (4.23). Summing up the resulting inequality leads to a telescoping sum for the contributions of (4.19). It implies that

$$\begin{aligned}\|\nabla e^0_{\tau,h}(t_n)\|^2 + \|e^1_{\tau,h}(t_n)\|^2 &\lesssim \|\nabla e^0_{\tau,h}(t_0)\|^2 + \|e^1_{\tau,h}(t_0)\|^2 + \sum_{s=1}^{n} \tau_s \left(\tau_s^{k+1} + h^{r+1}\right)^2 \\ &\quad + \sum_{s=1}^{n} \tau_s^2 Q_s^{\mathrm{G}}\left(\|\partial_t E_{\tau,h}\|^2\right) + \sum_{s=1}^{n} \tau_s \|E_{\tau,h}(t_{s-1})\|^2.\end{aligned}$$

Using

$$\|\partial_t E_{\tau,h}(t)\| \le \|\partial_t U(t) - \partial_t U_{\tau,h}(t)\| + \|-\partial_t \Theta(t)\| \lesssim \tau^k + h^{r+1}, \quad t \in \overline{I}, \tag{4.24}$$

together with the estimates (4.11) and (4.15), we obtain that

$$\begin{aligned}\|\nabla e^0_{\tau,h}(t_n)\|^2 + \|e^1_{\tau,h}(t_n)\|^2 &\lesssim \|\nabla e^0_{\tau,h}(t_0)\|^2 + \|e^1_{\tau,h}(t_0)\|^2 + \left(\tau^{k+1} + h^{r+1}\right)^2 \\ &\quad + \sum_{s=0}^{n-1} \tau_{s+1}\left(\|\nabla e^0_{\tau,h}(t_s)\|^2 + \|e^1_{\tau,h}(t_s)\|^2\right),\end{aligned}$$

where we also used the definition of the Gauss quadrature and the Poincaré inequality. Applying the discrete Gronwall lemma (*cf.* [43], p. 14) results in

$$\|\nabla e^0_{\tau,h}(t_n)\|^2 + \|e^1_{\tau,h}(t_n)\|^2 \lesssim \|\nabla e^0_{\tau,h}(t_0)\|^2 + \|e^1_{\tau,h}(t_0)\|^2 + \left(\tau^{k+1} + h^{r+1}\right)^2.$$

Exploiting $e^i_{\tau,h}(t_0) = 0$, $i \in \{0,1\}$, which holds due to the choice $U_{0,h} = (R_h u_0, R_h u_1)$ of the discrete initial value, this estimate along with the Poincaré inequality proves the assertion (4.20).

To show (4.21) and (4.22), we start for the error component $e^i_{\tau,h} \in \mathbb{P}_k(I_n, V_h)$, $i \in \{0,1\}$, with

$$\|e^i_{\tau,h}(t)\|_m \le \|e^i_{\tau,h}(t_n)\|_m + \tau_n \max\nolimits_{s\in\overline{I}_n} \|\partial_t e^i_{\tau,h}(s)\|_m, \quad t \in I_n, \tag{4.25}$$

that is deduced from the fundamental theorem of calculus. Applying (4.20) and (4.24), we get from (4.25) with $m = 0$ that

$$\|e^i_{\tau,h}(t)\| \lesssim \left(\tau^{k+1} + h^{r+1}\right) + \tau_n\left(\tau^k + h^{r+1}\right) \lesssim \tau^{k+1} + h^{r+1}, \quad t \in \overline{I}, \quad i \in \{0,1\},$$

which proves (4.22).

Similarly to (4.24), we get for the $H^1$-norm that

$$\|\partial_t e^0_{\tau,h}(t)\|_1 \le \|\partial_t u^0(t) - \partial_t u^0_{\tau,h}(t)\|_1 + \|-\partial_t \theta^0(t)\|_1 \lesssim \tau^k + h^r, \quad t \in \overline{I}, \tag{4.26}$$

where we used (4.12) along with the Poincaré inequality and (4.15). Applying (4.20) and (4.26), we get from (4.25) with $m = 1$ that

$$\|e^0_{\tau,h}(t)\|_1 \lesssim \left(\tau^{k+1} + h^{r+1}\right) + \tau_n\left(\tau^{k+1} + h^r\right) \lesssim \tau^{k+1} + h^r, \quad t \in \overline{I},$$

which proves (4.21). $\square$

We are now able to derive our final error estimates for the proposed Galerkin–collocation approximation of the solution to (1.1).

**Theorem 4.13** (Error estimate for $U_{\tau,h}$)**.** *Let $U = (u, \partial_t u)$ be the solution of the problem (1.1) and let $U_{\tau,h}$ be the fully discrete solution of Problem 3.3 with initial value $U_{0,h} = (R_h u_0, R_h u_1)$. Then, the error $E(t) = \big(e^0(t), e^1(t)\big) = U(t) - U_{\tau,h}(t)$ can be bounded for all $t \in \overline{I}$ by*

$$\|e^0(t)\| + \|e^1(t)\| \lesssim \tau^{k+1} + h^{r+1}, \tag{4.27}$$

$$\|\nabla e^0(t)\| \lesssim \tau^{k+1} + h^r. \tag{4.28}$$

*Moreover, the estimates*

$$\|e^0\|_{L^2(I;H)} + \|e^1\|_{L^2(I;H)} \lesssim \tau^{k+1} + h^{r+1}, \tag{4.29}$$

$$\|\nabla e^0\|_{L^2(I;H)} \lesssim \tau^{k+1} + h^r \tag{4.30}$$

*hold true.*

*Proof.* Recalling the error decomposition

$$E(t) = U(t) - U_{\tau,h}(t) = \Theta(t) + E_{\tau,h}(t), \tag{4.31}$$

we conclude assertion (4.27) by applying the triangle inequality along with estimate (4.14) with $m = 0$ and (4.22) to the terms on the right-hand-side of (4.31). Similarly we conclude (4.28) using the estimate (4.14) with $m = 1$ and (4.21). The assertions (4.29) and (4.30) follow from the definition of the $L^2(I;H)$-norm together with the estimates (4.27) and (4.28). □

**Remark 4.14.** We note that the estimates (4.27)–(4.30) are of optimal order in space and time.

## 5. Stability estimates and energy conservation principle

In this section we address the issue of *a priori* stability estimates for the discrete solution and energy conservation for the considered space-time finite element scheme.

### 5.1. *A priori* stability estimates

We now investigate stability of the cGP-C$^1(k)$–cG$(r)$ scheme of Problem 3.3. Hereinafter, let $\|\cdot\|_{V'}$ denote the $V' = H^{-1}$-norm. Further, let a discrete analogue be given by

$$\|v\|_{V_h'} := \sup_{w_h \in V_h \setminus \{0\}} \frac{\langle v, w_h \rangle}{\|\nabla w_h\|} \quad \forall v \in V'.$$

It can be easily seen that $\|v\|_{V_h'} \leq \|v\|_{V'}$ for all $v \in V'$ and that $\|A_h v\|_{V_h'} \leq \|\nabla v\|$ for all $v \in V$. Moreover, for $v_h \in V_h$ we even have that $\|A_h v_h\|_{V_h'} = \|\nabla v_h\|$.

*A priori* stability estimates for the fully discrete solution are given in the following theorem.

**Theorem 5.1** (Stability estimate for $U_{\tau,h}$)**.** *Let the initial value be given by $U_{0,h} = (u_{0,h}, u_{1,h})$. Then, for the fully discrete solution $U_{\tau,h} = \left(u^0_{\tau,h}, u^1_{\tau,h}\right)$ of Problem 3.3 it holds for all $t \in [0,T]$ that*

$$\|u^0_{\tau,h}(t)\| + \|u^1_{\tau,h}(t)\|_{V_h'} \lesssim \|u_{0,h}\|_1 + \|u_{1,h}\| + \|f(0)\|_{V'} + \left(\int_0^T \|\partial_t I_\tau^{\mathrm{H}} f(s)\|_{V'}^2 \,\mathrm{d}s\right)^{1/2}, \tag{5.1a}$$

$$\|\nabla u^0_{\tau,h}(t)\| + \|u^1_{\tau,h}(t)\| \lesssim \|A_h u_{0,h}\| + \|u_{1,h}\|_1 + \|f(0)\| + \left(\int_0^T \|\partial_t I_\tau^{\mathrm{H}} f(s)\|^2 \,\mathrm{d}s\right)^{1/2}. \tag{5.1b}$$

*Proof.* Since for both estimates in (5.1) the arguments are the same, we shall only prove the first one. Using that $U_{\tau,h}$ is continuously differentiable, it follows with the fundamental theorem of calculus that

$$\|u^0_{\tau,h}(t)\| + \|u^1_{\tau,h}(t)\|_{V_h'} \leq \|u^0_{\tau,h}(0)\| + \|u^1_{\tau,h}(0)\|_{V_h'} + \int_0^t \|\partial_t u^0_{\tau,h}(s)\| + \|\partial_t u^1_{\tau,h}(s)\|_{V_h'}\,\mathrm{d}s \quad \forall t \in [0,T]. \tag{5.2}$$

Thus, it remains to derive a suitable stability estimate for $\partial_t U_{\tau,h}$. From Remark 4.2 we already know that $\partial_t U_{\tau,h}$ can be regarded as the cGP($k$-1)–cG($r$) approximation of an evolution problem with adapted right-hand side and initial value. Thus, noting that all terms in (4.1) and (4.6), respectively, can be integrated exactly, the stability estimates of Section 2 from [9] become applicable, showing that

$$\max_{s\in[0,T]} \left(\|\partial_t u^0_{\tau,h}(s)\| + \|\partial_t u^1_{\tau,h}(s)\|_{V_h'}\right) \lesssim \|\partial_t u^0_{\tau,h}(0)\| + \|\partial_t u^1_{\tau,h}(0)\|_{V_h'} + \left(\int_0^T \|\partial_t I_\tau^{\mathrm{H}} f(s)\|^2_{V_h'}\,\mathrm{d}s\right)^{1/2}.$$

Furthermore, because of $\partial_t U_{\tau,h}(0) = -\mathcal{A}_h U_{0,h} + \mathcal{P}_h F(0)$ we obtain that

$$\|\partial_t u^0_{\tau,h}(0)\| = \|u_{1,h}\|, \quad \|\partial_t u^1_{\tau,h}(0)\|_{V_h'} = \|P_h f(0) - A_h u_{0,h}\|_{V_h'} \leq \|f(0)\|_{V'} + \|\nabla u_{0,h}\|.$$

Finally, combining the above estimates and identities, we gain assertion (5.1a). □

**Remark 5.2** (Stability estimate for $I_\tau^{\mathrm{GL}} U_{\tau,h}$)**.** We already noted that $U_{\tau,h}$ can be interpreted as the post-processed solution of Problem 3.2 with $F$ replaced by $I_\tau^{\mathrm{H}} F$. Moreover, $I_\tau^{\mathrm{GL}} U_{\tau,h}$ can be interpreted as cGP($k$-1)–cG($r$) approximation, *i.e.*, solution of Problem 3.1 with $I_\tau^{\mathrm{GL}} I_\tau^{\mathrm{H}} F$ instead of $F$; *cf.*, also Proposition 4.5 and Section 5.1 of [17]. Therefore, the stability estimates of Lemma 2.1 from [9] directly imply that

$$\|\nabla I_\tau^{\mathrm{GL}} u^0_{\tau,h}(t)\| + \|I_\tau^{\mathrm{GL}} u^1_{\tau,h}(t)\| \lesssim \|\nabla u_{0,h}\| + \|u_{1,h}\| + \left(\int_0^T \|I_\tau^{\mathrm{GL}} I_\tau^{\mathrm{H}} f(s)\|^2\,\mathrm{d}s\right)^{1/2} \quad \forall t \in [0,T].$$

By means of (3.6) and (5.2), this would also imply stability estimates for $\|u^0_{\tau,h}(t)\| + \|u^1_{\tau,h}(t)\|_{V_h'}$.

## 5.2. Energy conservation principle for $f \equiv 0$

For vanishing right-hand side term $f \equiv 0$ it is well-known that the solution $u$ of the initial-boundary value problem (1.1) satisfies the energy conservation

$$\|u^1(t)\|^2 + \|\nabla u^0(t)\|^2 = \|u_1\|^2 + \|\nabla u_0\|^2, \quad t \in I.$$

We will prove that the space-time finite element discretization $U_{\tau,h}$ of Problem 3.3 also satisfies the energy conservation principle at the discrete time nodes $t_n$. Preserving this fundamental property of the solution of (1.1) is an important quality criterion for discretization schemes of (1.1).

**Theorem 5.3** (Energy conservation for $U_{\tau,h}$)**.** *Suppose that $f \equiv 0$. Let the initial value be given by $U_{0,h} = (u_{0,h}, u_{1,h})$. Then, the fully discrete solution $U_{\tau,h} = \left(u^0_{\tau,h}, u^1_{\tau,h}\right)$ defined by Problem 3.3 satisfies the energy conservation property*

$$\|u^1_{\tau,h}(t_n)\|^2 + \|\nabla u^0_{\tau,h}(t_n)\|^2 = \|u_{1,h}\|^2 + \|\nabla u_{0,h}\|^2 \tag{5.3}$$

*for all $n = 1, \ldots, N$.*

*Proof.* Let $f \equiv 0$. We recall that the fully discrete solution $U_{\tau,h} = \left(u^0_{\tau,h}, u^1_{\tau,h}\right)$ defined by Problem 3.3 satisfies the variational equation (3.3). Now, choosing $V_{\tau,h} = \left(\Pi_n^{k-2} A_h I_\tau^{\mathrm{GL}} u^0_{\tau,h}, \Pi_n^{k-2} I_\tau^{\mathrm{GL}} u^1_{\tau,h}\right) \in (\mathbb{P}_{k-2}(I_n; V_h))^2$ as test function, we find from (4.19) that

$$
\begin{aligned}
0 &= Q_n^{\mathrm{GL}}\left(\langle\!\langle \partial_t U_{\tau,h}, V_{\tau,h}\rangle\!\rangle + \langle\!\langle \mathcal{A}_h U_{\tau,h}, V_{\tau,h}\rangle\!\rangle\right)\\
&= \frac{1}{2}\left(\|\nabla u^0_{\tau,h}(t_n)\|^2 - \|\nabla u^0_{\tau,h}(t_{n-1})\|^2 + \|u^1_{\tau,h}(t_n)\|^2 - \|u^1_{\tau,h}(t_{n-1})\|^2\right).
\end{aligned}
$$

Hence, changing the index $n$ to $m$ and summing up from $m=1$ to $n$, give the desired identity. □

Finally, we note that to derive (5.3) also $V_{\tau,h} = \left(-\partial_t I_\tau^{\mathrm{GL}} u^1_{\tau,h}, \partial_t I_\tau^{\mathrm{GL}} u^0_{\tau,h}\right) \in \left(\mathbb{P}_{k-2}(I_n; V_h)\right)^2$ could be chosen as test function in (3.3). The detailed proof for this choice is given in a preprint Lemma 6.1 of [7].

## 6. $\mathrm{C}^2$-regular Galerkin–collocation approximation and its relation to post-processed cGP-$\mathrm{C}^1$

In this section, let $k \geq 5$ be satisfied. Firstly, we propose a family of Galerkin–collocation time discretization schemes with twice continuously differentiable in time discrete solutions, that are referred to as cGP-C$^2(k)$–cG$(r)$ schemes. Similarly to the cGP-C$^1(k)$–cG$(r)$ approach of Problem 3.3, the higher order regularity in time is ensured by collocation conditions that are imposed in the endpoints $t_{n-1}$ and $t_n$ of the subinterval $I_n$. This construction principle can be generalized to discrete solutions of even higher order regularity in time. For this generalization we also refer to [17, 18] where the Galerkin–collocation approximation of first-order ordinary differential equations systems is studied in detail. Secondly, we show how the cGP-C$^2(k+1)$–cG$(r)$ approximation can be computed efficiently in a simple and computationally cheap post-processing step from the cGP-C$^1(k)$–cG$(r)$ approach. The post-processing introduced in [41] and generalized in [17] was recently applied in [16] to the cGP$(k)$–cG$(r)$ family of schemes given in Problem 3.1. There the post-processing is used to lift continuous in time discrete solutions to continuously differentiable ones. Moreover, an optimal order error analysis is provided for the post-processed solution.

**Problem 6.1** (Local, numerically integrated, fully discrete problem of cGP-C$^2(k)$–cG$(r)$ on $I_n$)**.**
Given $U_{\tau,h}\left(t^-_{n-1}\right)$ for $n>1$ and $U_{\tau,h}\left(t^-_0\right) = U_{0,h}$ for $n=1$, find $U_{\tau,h}|_{I_n} \in \left(\mathbb{P}_k(I_n; V)\right)^2$ such that

$$U_{\tau,h}\left(t^+_{n-1}\right) = U_{\tau,h}\left(t^-_{n-1}\right), \tag{6.1a}$$
$$\partial_t U_{\tau,h}\left(t^+_{n-1}\right) = -\mathcal{A}_h U_{\tau,h}\left(t^+_{n-1}\right) + \mathcal{P}_h F\left(t^+_{n-1}\right), \tag{6.1b}$$
$$\partial_t^2 U_{\tau,h}\left(t^+_{n-1}\right) = -\mathcal{A}_h \partial_t U_{\tau,h}\left(t^+_{n-1}\right) + \mathcal{P}_h \partial_t F\left(t^+_{n-1}\right), \tag{6.1c}$$
$$\partial_t U_{\tau,h}\left(t^-_{n}\right) = -\mathcal{A}_h U_{\tau,h}\left(t^-_{n}\right) + \mathcal{P}_h F\left(t^-_{n}\right), \tag{6.1d}$$
$$\partial_t^2 U_{\tau,h}\left(t^-_{n}\right) = -\mathcal{A}_h \partial_t U_{\tau,h}\left(t^-_{n}\right) + \mathcal{P}_h \partial_t F\left(t^-_{n}\right), \tag{6.1e}$$

and

$$Q^{\mathrm{H}}_{n,k}\left(\langle\!\langle \partial_t U_{\tau,h}, V_{\tau,h}\rangle\!\rangle + \langle\!\langle \mathcal{A}_h U_{\tau,h}, V_{\tau,h}\rangle\!\rangle\right) = Q^{\mathrm{H}}_{n,k}\left(\langle\!\langle F, V_{\tau,h}\rangle\!\rangle\right) \tag{6.1f}$$

for all $V_{\tau,h} \in \left(\mathbb{P}_{k-5}(I_n; V_h)\right)^2$.

We note that an Hermite-type quadrature formula with $k$ evaluations of function values is used in (6.1f). This differs from $Q_n^{\mathrm{H}}$ in (2.5) that is used in the cGP-C$^1(k)$–cG$(r)$ family of schemes of Problem 3.3 and is based on $k-1$ evaluations of function values only. In both cases the derivatives of the integrand are evaluated additionally in the endpoints of the subinterval $I_n$. Further, the cGP-C$^2(k)$ approach presented here differs from that in [18] by the applied quadrature formula.

**Remark 6.2.** A careful inspection of the conditions on $U_{\tau,h}$ shows that

$$\partial_t U_{\tau,h}\left(t^+_{n-1}\right) = \partial_t U_{\tau,h}\left(t^-_{n-1}\right) \quad \text{and} \quad \partial_t^2 U_{\tau,h}\left(t^+_{n-1}\right) = \partial_t^2 U_{\tau,h}\left(t^-_{n-1}\right),$$

where the discrete initial conditions are determined using

$$\partial_t U_{\tau,h}(0) = -\mathcal{A}_h U_{\tau,h}(0) + \mathcal{P}_h F(0), \quad \partial_t^2 U_{\tau,h}(0) = -\mathcal{A}_h \partial_t U_{\tau,h}(0) + \mathcal{P}_h \partial_t F(0).$$

Hence, the obtained trajectory in time is twice continuously differentiable on $\overline{I}$.

Compared to Problem 3.3, the test space of condition (6.1f) is decreased from $(\mathbb{P}_{k-3}(I_n; V_h))^2$ to $(\mathbb{P}_{k-5}(I_n; V_h))^2$ while the number of collocation conditions is increased from two to four. For $k = 5$ this results in a test space which consists of piecewise constant functions only and to two additional collocation conditions in both endpoints of the time subinterval $I_n$.

Finally we address the link between the cGP-C$^1$ and cGP-C$^2$ families of Galerkin–collocation schemes.

**Theorem 6.3.** *Let $U_{\tau,h}$ denote the solution of the* cGP-C$^1(k)$–cG$(r)$ *method given in Problem 3.3. For $n = 1, \ldots, N$ we put*

$$\widetilde{U}_{\tau,h}|_{I_n} := U_{\tau,h}|_{I_n} - K_n \vartheta_n,$$

*where $\vartheta_n \in \mathbb{P}_{k+1}(I_n; \mathbb{R})$ is uniquely determined by*

$$I_n^{\mathrm{H}} \vartheta_n \equiv 0 \quad \text{and} \quad \partial_t^2 \vartheta_n \left(t_{n-1}^+\right) = 1.$$

*If the correction coefficient $K_n$ is chosen as*

$$K_n := \begin{cases} \partial_t^2 U_{\tau,h}\left(t_0^+\right) - \partial_t^2 U_{\tau,h}(t_0), & n = 1, \\ \partial_t^2 U_{\tau,h}\left(t_{n-1}^+\right) - \partial_t^2 \widetilde{U}_{\tau,h}(t_{n-1}^-), & n > 1, \end{cases}$$

*then $\widetilde{U}_{\tau,h} \in \left(X_\tau^{k+1}(V_h)\right)^2$ is the solution of the* cGP-C$^2(k+1)$–cG$(r)$ *method given in Problem 6.1.*

The post-processing or lifting operator that is introduced in Theorem 6.3 is similar to the lifting operator of [16] that is studied there in the context of the cGP$(k)$–cG$(r)$ approach of Problem 3.2. Both post-processing procedures provide the correction as a product of a scalar polynomial $\vartheta_n$ and a coefficient $K_n \in V_h^2$ that are, however, different for the two procedures. In particular, the lifting in [16] is based on the difference of first derivatives while our post-processing uses the difference of second order derivatives. We refer to [17] for details on post-processing techniques for general nonlinear systems of ordinary differential equations and the proof of the analogue to Theorem 6.3.

## 7. Numerical studies

In this section we investigate numerically the Galerkin–collocation approximation schemes introduced in Problems 3.3 and 6.1, respectively. In particular, we aim to illustrate the error estimates given in Theorem 4.13 for the cGP-C$^1(k)$–cG$(r)$ Galerkin–collocation approximation of Problem 3.3 and the additional post-processing of the discrete solution introduced in Theorem 6.3. For further details regarding the block structure of the algebraic system of the cGP-C$^1(k)$–cG$(r)$ family of schemes along with its iterative solution and preconditioning as well as for additional numerical experiments we refer to [6,12]. The implementation of the numerical schemes was done in the high-performance `DTM++/awave` frontend solver (*cf.* [35]) for the `deal.II` library [11].

### 7.1. Convergence test for cGP-C$^1$(3)–cG(3)

In our first numerical experiment we study the convergence behavior of the Galerkin–collocation approximation $U_{\tau,h} \in \left(X_\tau^3\left(V_h^{(3)}\right) \cap C^1\left(I; V_h^{(3)}\right)\right)^2$ of Problem 3.3 for the prescribed solution

$$u(\boldsymbol{x}, t) := \sin(4\pi t) \cdot \sin(2\pi x_1) \cdot \sin(2\pi x_2) \tag{7.1}$$

TABLE 1. Calculated errors $E = \left(e^0, e^1\right)$ with $E(t) = U(t) - U_{\tau,h}(t)$ and corresponding experimental orders of convergence (EOC) for the solution $U = (u, \partial_t u)$ of (7.1) and the Galerkin–collocation approximation $U_{\tau,h} \in \left(X_\tau^3\left(V_h^{(3)}\right) \cap C^1\left(I; V_h^{(3)}\right)\right)^2$ of Problem 3.3.

| $\tau$ | $h$ | $\\|e^0\\|_{L^\infty(L^2)}$ | $\\|e^1\\|_{L^\infty(L^2)}$ | $\|\|\|E\|\|\|_{L^\infty}$ | $\\|e^0\\|_{L^2(L^2)}$ | $\\|e^1\\|_{L^2(L^2)}$ | $\|\|\|E\|\|\|_{L^2}$ |
|---|---|---|---|---|---|---|---|
| $\tau_0/2^0$ | $h_0/2^0$ | 2.834e-02 | 2.862e-01 | 6.122e-01 | 2.099e-02 | 2.234e-01 | 4.808e-01 |
| $\tau_0/2^1$ | $h_0/2^1$ | 1.383e-03 | 1.755e-02 | 5.343e-02 | 9.773e-04 | 1.186e-02 | 3.989e-02 |
| $\tau_0/2^2$ | $h_0/2^2$ | 9.261e-05 | 1.075e-03 | 6.750e-03 | 6.064e-05 | 7.140e-04 | 4.835e-03 |
| $\tau_0/2^3$ | $h_0/2^3$ | 5.911e-06 | 6.690e-05 | 8.466e-04 | 3.812e-06 | 4.446e-05 | 6.005e-04 |
| $\tau_0/2^4$ | $h_0/2^4$ | 3.714e-07 | 4.186e-06 | 1.059e-04 | 2.387e-07 | 2.777e-06 | 7.495e-05 |
| $\tau_0/2^5$ | $h_0/2^5$ | 2.325e-08 | 2.616e-07 | 1.324e-05 | 1.492e-08 | 1.735e-07 | 9.364e-06 |
| EOC | | 4.00 | 4.00 | 3.00 | 4.00 | 4.00 | 3.00 |

of the wave problem (1.1) on the space-time domain $\Omega \times I = (0,1)^2 \times (0,1)$. For the piecewise polynomial order in space and time of the finite element approach the choice $k = 3$ and $r = 3$ is thus made; *cf.*, (2.2) and (2.3). Beyond the norms of $L^\infty\left(I; L^2(\Omega)\right)$ and $L^2\left(I; L^2(\Omega)\right)$ the convergence behavior is studied further with respect to the energy quantities

$$|||E|||_{L^\infty} = \max_{t\in\mathbb{I}} \left(\|\nabla e^0(t)\|^2 + \|e^1(t)\|^2\right)^{1/2} \quad \text{and} \quad |||E|||_{L^2} = \left(\int_I \left(\|\nabla e^0(t)\|^2 + \|e^1(t)\|^2\right) \mathrm{d}t\right)^{1/2} \tag{7.2}$$

with $E(t) = U(t) - U_{\tau,h}(t)$. Throughout, the $L^\infty$-norms in time are computed on the discrete time grid

$$\mathbb{I} = \left\{t_n^j \,:\, t_n^j = t_{n-1} + j \cdot k_n \cdot \tau_n,\, k_n = 0.001,\, j = 0, \dots, 999,\, n = 1, \dots, N\right\} \cup \{t_N\}\,.$$

In the numerical experiments the domain $\Omega$ is decomposed into a sequence of successively refined meshes $\Omega_h^l$, with $l = 0, \dots, 5$, of quadrilateral finite elements. On the coarsest level, we use a uniform decomposition of $\Omega$ into 4 cells, corresponding to the mesh size $h_0 = 1/\sqrt{2}$, and of the time interval $I$ into $N = 10$ subintervals which amounts to the time step size $\tau_0 = 0.1$. In the experiments the temporal and spatial mesh sizes are successively refined by a factor of two in each refinement step.

In Table 1 we summarize the calculated results for this experiment. The experimental order of convergence (EOC) was calculated using the results from the two finest meshes. The numerical results of Table 1 nicely confirm our error estimates (4.27) and (4.29) by depicting the expected optimal fourth order rate of convergence in space and time. The third order convergence of the energy errors (7.2) is in agreement with the error estimates (4.28) and (4.30). Increasing the piecewise polynomial order in space to $r = 4$ and thus calculating an approximation $U_{\tau,h} \in \left(X_\tau^3\left(V_h^{(4)}\right) \cap C^1\left(I; V_h^{(4)}\right)\right)^2$ leads a fourth order convergence behavior in time and space which is not shown here for the sake of brevity.

### 7.2. Convergence test for cGP-$\mathbf{C^1}$(4)–cG(5) and post-processing

In the second numerical experiment we study the Galerkin–collocation scheme of Problem 3.3 for $k = 4$ to obtain a fully discrete solution $U_{\tau,h} \in \left(X_\tau^4\left(V_h^{(5)}\right) \cap C^1\left(I; V_h^{(5)}\right)\right)^2$. In addition, we will apply the post-processing considered in Section 6 and obtain a solution $\widetilde{U}_{\tau,h}$ belonging to $\left(X_\tau^5\left(V_h^{(5)}\right) \cap C^2\left(I; V_h^{(5)}\right)\right)^2$. The numerical study is done for the prescribed solution

$$u(\boldsymbol{x}, t) := \sin(4\pi t)\, x_1\,(1 - x_1)\, x_2\,(1 - x_2) \tag{7.3}$$

TABLE 2. Error $E = \left(e^0, e^1\right) = U - U_{\tau,h}$ and error $\widetilde{E} = \left(\tilde{e}^0, \tilde{e}^1\right) = U - \widetilde{U}_{\tau,h}$ of the post-processed solution $\widetilde{U}_{\tau,h}$ of Theorem 6.3, both with the corresponding experimental orders of convergence (EOC), for the solution $U = (u, \partial_t u)$ of (7.3) and the Galerkin–collocation approximation $U_{\tau,h} \in \left(X_\tau^4\left(V_h^{(5)}\right) \cap C^1\left(I; V_h^{(5)}\right)\right)^2$ of Problem 3.3.

| $\tau$ | $h$ | $\Vert e^0\Vert_{L^\infty(L^2)}$ | $\Vert e^1\Vert_{L^\infty(L^2)}$ | $\vert\vert\vert E\vert\vert\vert_{L^\infty}$ | $\Vert e^0\Vert_{L^2(L^2)}$ | $\Vert e^1\Vert_{L^2(L^2)}$ | $\vert\vert\vert E\vert\vert\vert_{L^2}$ |
|---|---|---|---|---|---|---|---|
| $\tau_0/2^0$ | $h_0$ | 8.457e-06 | 9.634e-05 | 9.637e-05 | 4.787e-06 | 5.392e-05 | 5.806e-05 |
| $\tau_0/2^1$ | $h_0$ | 2.497e-07 | 3.018e-06 | 3.022e-06 | 1.360e-07 | 1.654e-06 | 1.763e-06 |
| $\tau_0/2^2$ | $h_0$ | 7.608e-09 | 9.368e-08 | 9.372e-08 | 4.127e-09 | 5.141e-08 | 5.463e-08 |
| $\tau_0/2^3$ | $h_0$ | 2.353e-10 | 2.936e-09 | 2.936e-09 | 1.280e-10 | 1.604e-09 | 1.703e-09 |
| $\tau_0/2^4$ | $h_0$ | 7.323e-12 | 9.175e-11 | 9.175e-11 | 3.991e-12 | 5.012e-11 | 5.321e-11 |
| EOC | | 5.01 | 5.00 | 5.00 | 5.00 | 5.00 | 5.00 |
| $\tau$ | $h$ | $\Vert\tilde{e}^0\Vert_{L^\infty(L^2)}$ | $\Vert\tilde{e}^1\Vert_{L^\infty(L^2)}$ | $\vert\vert\vert\widetilde{E}\vert\vert\vert_{L^\infty}$ | $\Vert\tilde{e}^0\Vert_{L^2(L^2)}$ | $\Vert\tilde{e}^1\Vert_{L^2(L^2)}$ | $\vert\vert\vert\widetilde{E}\vert\vert\vert_{L^2}$ |
| $\tau_0/2^0$ | $h_0$ | 2.906e-06 | 1.711e-05 | 1.791e-05 | 1.936e-06 | 1.519e-05 | 1.764e-05 |
| $\tau_0/2^1$ | $h_0$ | 4.717e-08 | 2.802e-07 | 2.841e-07 | 3.150e-08 | 2.418e-07 | 2.824e-07 |
| $\tau_0/2^2$ | $h_0$ | 7.513e-10 | 4.507e-09 | 4.537e-09 | 4.972e-10 | 3.797e-09 | 4.440e-09 |
| $\tau_0/2^3$ | $h_0$ | 1.180e-11 | 7.085e-11 | 7.133e-11 | 7.788e-12 | 5.940e-11 | 6.949e-11 |
| $\tau_0/2^4$ | $h_0$ | 1.851e-13 | 1.113e-12 | 1.120e-12 | 1.216e-13 | 9.282e-13 | 1.086e-12 |
| EOC | | 6.00 | 6.00 | 6.00 | 6.00 | 6.00 | 6.00 |

of problem (1.1) on the space-time domain $\Omega \times I = (0,1)^2 \times (0,1)$. For the piecewise polynomial order in space and time of the finite element approach the choice $k = 4$ and $r = 5$ is thus made. Since this work focuses the temporal discretization, the polynomial degree $r$ in space is chosen such that the spatial approximation is exact. Hence, the convergence behavior in time can be illustrated on a fixed spatial grid that consists of $4 \times 4$ congruent squares with $h_0 = 0.25\sqrt{2}$. The largest time length is $\tau_0 = 0.1$.

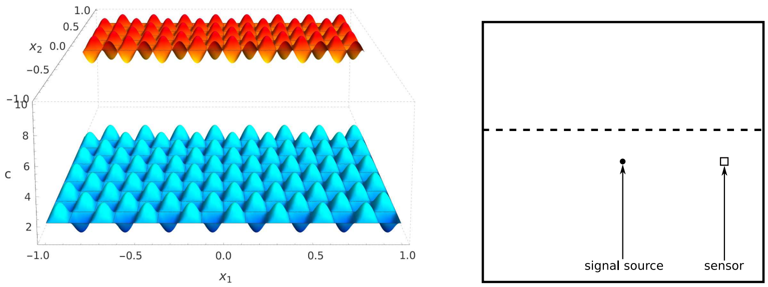
In Table 2 we summarize the calculated results for this experiment. The experimental order of convergence (EOC) was determined from the results on the two finest meshes. The numerical results of Table 2 nicely confirm the fifth order rate of convergence of the cGP-C$^1$(4) time discretization. The application of the post-processing presented in Theorem 6.3 increased all convergence rates from 5 to 6. This order can at most be expected for a polynomial approximation in time with piecewise polynomials of fifth order. By means of Theorem 6.3, Table 2 thus underlines the optimal order approximation properties of the cGP-C$^2$(5) member of the family of Galerkin–collocation schemes of Problem 6.1.

If the cGP-C$^2$(5)–cG(5) method of Problem 6.1 is directly applied for the computation, instead of using the post-processing of Theorem 6.3, then exactly the same errors as shown in Table 2 for $\widetilde{E}$ are obtained. However, using the post-processing has certain computational advantages. Since the cGP-C$^1(k)$ approach leads to system matrices of simpler block structure compared to the cGP-C$^2(k+1)$ method, the construction of efficient preconditioners simplifies; *cf.* [6] for details.

### 7.3. Sophisticated test problem

Here we illustrate the performance properties of the proposed cGP-C$^1(k)$ scheme for a more sophisticated test problem and provide a comparative study to the standard continuous approximation cGP($k$) of Problem 3.2. For this, we consider instead of (1.1) the more general initial-boundary value problem

$$\begin{aligned} \partial_t^2 u - \nabla \cdot (c\nabla u) &= f \quad \text{in} \quad \Omega \times (0,T], \\ u &= 0 \quad \text{on} \quad \partial\Omega \times (0,T], \\ u(\cdot,0) = u_0, \quad \partial_t u(\cdot,0) &= u_1 \quad \text{in} \quad \Omega, \end{aligned} \tag{7.4}$$




(A) (B)

FIGURE 1. Problem setting. (A) Coefficient $c$ of (7.5) for (7.4). (B) Geometry.

with a coefficient function $c \in L^\infty(\Omega)$ satisfying $c(\boldsymbol{x}) \geq c_0 > 0$ for almost every $\boldsymbol{x} \in \Omega$. Without effort, the results of Sections 4 and 5 can be generalized to the system (7.4). In our numerical experiment the coefficient $c$ is assumed to exhibit multi-scale properties that often arise in modern material sciences, *e.g.*, in lightweight composite material design, or also in applications related to the subsurface, *e.g.*, in seismic wave propagation. Further, $c$ has a jump discontinuity at $x_2 = 0.2$. Precisely, $c$ is defined by

$$c(\boldsymbol{x}) = \begin{cases} 2 + \sin\left(\dfrac{2\pi x_1}{\epsilon}\right) \cdot \sin\left(\dfrac{2\pi x_2}{\epsilon}\right), & x_2 < 0.2, \\ 9 + \sin\left(\dfrac{2\pi x_1}{\epsilon}\right) \cdot \sin\left(\dfrac{2\pi x_2}{\epsilon}\right), & x_2 \geq 0.2, \end{cases} \tag{7.5}$$

with $\epsilon = 0.25$. An illustration of the coefficient function $c$ is given in Figure 1A. We consider $\Omega \times I = (-1,1)^2 \times (0,1]$, let $f = 0$. For the initial value we prescribe a regularized Dirac impulse by

$$u_0(\boldsymbol{x}) = e^{-|\boldsymbol{x}_s|^2}\left(1 - |\boldsymbol{x}_s|^2\right)\Theta(1 - |\boldsymbol{x}_s|), \quad \boldsymbol{x}_s = 100\boldsymbol{x},$$

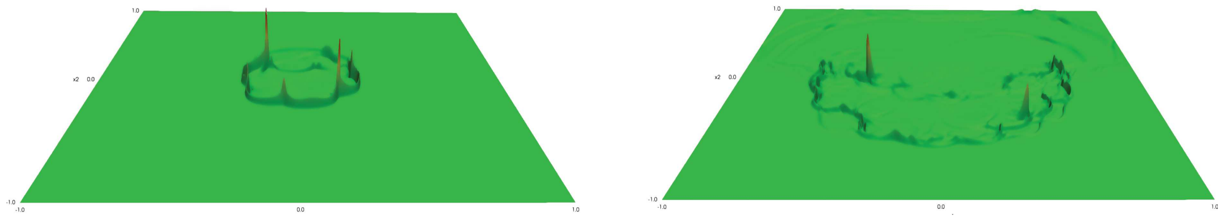
where $\Theta$ is the Heaviside function. We put $u_1 = 0$ for the second initial value. We define the control region $\Omega_c = (0.75 - h_c, 0.75 + h_c) \times (-h_c, h_c)$, with $h_c = 0.125$, where we calculate the signal arrival by

$$u_c(t) = \int_{\Omega_c} u_{\tau,h}(\boldsymbol{x}, t)\,\mathrm{d}\boldsymbol{x}. \tag{7.6}$$

Hereby, we mimic a typical problem setting of non-destructive lightweight material inspection by ultrasonic waves which is currently investigated as an intelligent future technique for structural health monitoring of lightweight material. The problem setting is sketched in Figure 1B.

We choose a spatial mesh of $256 \times 256$ cells and $\mathbb{Q}_8$ elements. In each time step, this leads to more than $4.2 \times 10^6$ degrees of freedom in space for each of the solution vectors. The solution profile computed by the cGP-C$^1$(3)–cG(8) approach of Problem 3.3 is visualized in Figures 2A and 2B, respectively.

In Figure 3 we compare the cGP-C$^1$(3)–cG(8) Galerkin–collocation approach of Problem 3.3 with the more standard cGP(2)–cG(8) continuous Galerkin approach of Problem 3.2. The goal quantity (7.6) is visualized for $t \in [0,1]$ and equidistant time step sizes $\tau_n = \tau_0 := 2 \times 10^{-5}$. The red-coloured cGP-C$^1$(3)–cG(8) solution shows the fully converged approximation. A further refinement of the step sizes in space and time does not lead to further significant variations. We note that the cGP(2)–cG(8) scheme of Problem 3.2 converges, due to some effect of superconvergence, of fourth order in the discrete time nodes $t_n$ which is proved in [16]. Therefore, both schemes admit the same asymptotic convergence rate in the discrete time nodes $t_n$. Moreover, after condensation of internal degrees of freedom by the continuity and collocation constraints the algebraic



(A) (B)

FIGURE 2. Solution profile computed by cGP-C$^1$(3)–cG(8) scheme. (A) Time $t = 0.2$. (B) Time $t = 0.45$.

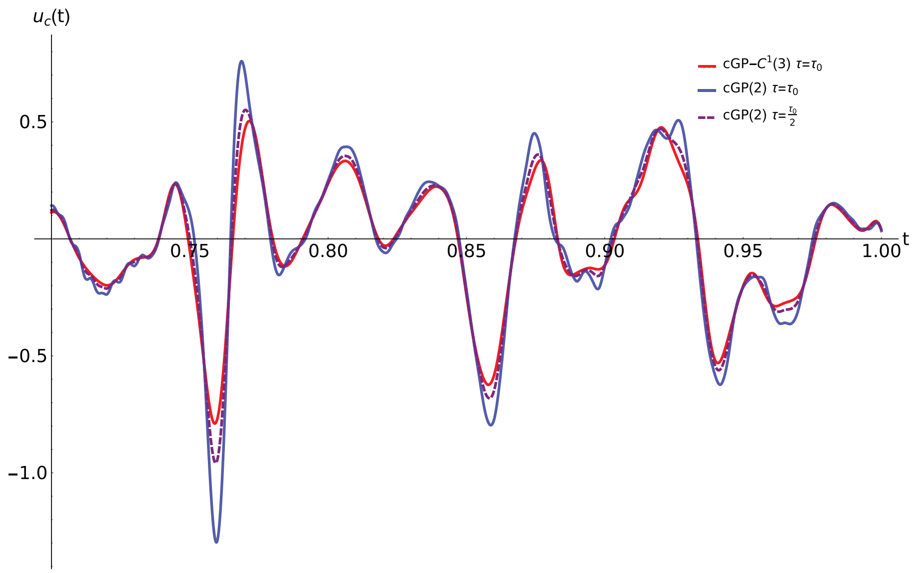



FIGURE 3. Signal (7.6) arrival at control (sensor) position (*cf.*, Fig. 1B) for cGP-C$^1$(3) and cGP(2) solutions with different time step sizes.

systems of both approaches, cGP-C$^1$(3)–cG(8) and cGP(2)–cG(8), lead to system matrices of the same size and almost the same sparsity pattern, precisely

$$\boldsymbol{A}_{\text{cGP-C}^1(3)}=\begin{pmatrix} \boldsymbol{A} & \boldsymbol{0} & \boldsymbol{0} & \frac{1}{\tau_n^2}\boldsymbol{M} \\ \boldsymbol{0} & \boldsymbol{A} & \frac{1}{\tau_n}\boldsymbol{M} & \boldsymbol{0} \\ \boldsymbol{M} & -\frac{1}{2}\boldsymbol{M} & -\frac{\tau_n}{120}\boldsymbol{M} & \frac{1}{10}\boldsymbol{M} \\ \frac{\tau_n}{2}\boldsymbol{A} & \frac{1}{\tau_n}\boldsymbol{M}-\frac{\tau_n}{10}\boldsymbol{A} & \boldsymbol{0} & \frac{\tau_n}{120}\boldsymbol{A} \end{pmatrix}, \quad \boldsymbol{A}_{\text{cGP}(2)}=\begin{pmatrix} -\frac{\tau_n}{3}\boldsymbol{M} & \boldsymbol{0} & \frac{2}{3}\boldsymbol{M} & \frac{1}{6}\boldsymbol{M} \\ \frac{2}{3}\boldsymbol{M} & \frac{1}{6}\boldsymbol{M} & \frac{\tau_n}{3}\boldsymbol{A} & \boldsymbol{0} \\ -\frac{\tau_n}{3}\boldsymbol{M} & -\frac{\tau_n}{6}\boldsymbol{M} & -\frac{2}{3}\boldsymbol{M} & \frac{5}{6}\boldsymbol{M} \\ -\frac{2}{3}\boldsymbol{M} & \frac{5}{6}\boldsymbol{M} & \frac{\tau_n}{3}\boldsymbol{A} & \frac{\tau_n}{6}\boldsymbol{A} \end{pmatrix},$$

where $\boldsymbol{M}$ and $\boldsymbol{A}$ denote the usual mass and stiffness matrix of the spatial discretization. The cGP-C$^1$(3)–cG(8) approach even admits a less dense system matrix. For a derivation and further details of $\boldsymbol{A}_{\text{cGP-C}^1(3)}$ we refer to [6], $\boldsymbol{A}_{\text{cGP}(2)}$ can be derived along the same lines. Thus, the iterative solution of both linear systems is almost of the same complexity; *cf.* [6].

Figure 3 shows that the cGP-C$^1$(3) Galerkin–collocation approach admits (without loss of accuracy) larger step sizes compared to the cGP(2) continuous Galerkin approach which makes the cGP-C$^1$(3) scheme much more efficient in terms of compute time; *cf.*, also [6]. The cGP-C$^1$(3) method with step size $\tau_0$ and the cGP(2) scheme with halved step size $\tau_0/2$ lead to almost the same numerical results for the goal quantity (7.6) whereas the cGP(2) scheme with step size $\tau_0$ shows severe perturbations. This numerical experiment nicely demonstrates the superiority of the Galerkin–collocation approach of Problem 3.3 over the standard continuous Galerkin approximation of Problem 3.2. Of course, our experiment gives no evidence that this potential of the cGP-C$^1(k)$ schemes applies to all applications.

## 8. Summary

In this work we presented a family of space-time finite element methods for wave problems. The schemes combine the concepts of collocation methods and Galerkin approximation. Continuously differentiable in time fully discrete solutions were obtained. An optimal order error analysis was provided for this class of methods. By a straightforward extension of the construction principle a further class of schemes with twice continuously differentiable in time discrete solutions was presented. A theorem regarding the connection of the two classes of schemes to each other by means of a post-processing was given. The proven error estimates and the expected convergence rates for the second class of schemes were illustrated by numerical experiments. The construction of the methods can be transferred to further classes of non-stationary partial differential equations. In addition, the presented post-processing can nicely be exploited for *a posteriori* error control and adaptive refinement of the temporal mesh.

*Acknowledgements.* The authors wish to thank the anonymous reviewers for their help to improve the paper.

## References

[1] N. Ahmed, G. Matthies, L. Tobiska and H. Xie, Discontinuous Galerkin time stepping with local projection stabilization for transient convection-diffusion-reaction problems. *Comput. Methods Appl. Mech. Eng.* **200** (2011) 1747–1756.

[2] N. Ahmed, S. Becher and G. Matthies, Higher-order discontinuous Galerkin time stepping and local projection stabilization techniques for the transient Stokes problem. *Comput. Methods Appl. Mech. Eng.* **313** (2017) 28–52.

[3] M. Ainsworth, P. Monk and W. Muniz, Dispersive and dissipative properties of discontinuous Galerkin finite element methods for the second-order wave equation. *J. Sci. Comput.* **27** (2006) 5–40.

[4] G. Akrivis, C. Makridakis and R.H. Nochetto, Optimal order a posteriori error estimates for a class of Runge–Kutta and Galerkin methods. *Numer. Math.* **114** (2009) 133–160.

[5] G. Akrivis, C. Makridakis and R.H. Nochetto, Galerkin and Runge–Kutta methods: unified formulation, a posteriori error estimates and nodal superconvergence. *Numer. Math.* **118** (2011) 429–456.

[6] M. Anselmann and M. Bause, Numerical study of Galerkin–collocation approximation in time for the wave equation. Preprint `arXiv:1905.00606` (2019). To appear in *Mathematics of Wave Phenomena*, edited by W. Dörfler *et al. Trends in Mathematics*. Birkhäuser (2019).

[7] M. Anselmann, M. Bause, S. Becher and G. Matthies, Galerkin–collocation approximation in time for the wave equation and its post-processing. Preprint `arXiv:1908.08238v1` (2019).

[8] A.K. Aziz and P. Monk, Continuous finite elements in space and time for the heat equation. *Math. Comput.* **52** (1989) 255–274.

[9] L. Bales and I. Lasiecka, Continuous finite elements in space and time for the nonhomogeneous wave equation. *Comput. Math. Appl.* **27** (1994) 91–102.

[10] W. Bangerth, M. Geiger and R. Rannacher, Adaptive Galerkin finite element methods for the wave equation. *Comput. Methods Appl. Math.* **10** (2010) 3–48.

[11] W. Bangerth, T. Heister and G. Kanschat, `deal.II`, Differential equations analysis library. Technical reference. `http://www.dealii.org` (2013).

[12] M. Bause and A. Anselmann, Comparative study of continuously differentiable Galerkin time discretizations for the wave equation. *PAMM* **19** (2019) e201900144.

[13] M. Bause and U. Köcher, Variational time discretization for mixed finite element approximations of nonstationary diffusion problems. *J. Comput. Appl. Math.* **289** (2015) 208–224.

[14] M. Bause, F.A. Radu and U. Köcher, Space-time finite element approximation of the Biot poroelasticity system with iterative coupling. *Comput. Methods Appl. Mech. Eng.* **320** (2017) 745–768.

[15] M. Bause, F.A. Radu and U. Köcher, Error analysis for discretizations of parabolic problems using continuous finite elements in time and mixed finite elements in space. *Numer. Math.* **137** (2017) 773–818.

[16] M. Bause, U. Köcher, F.A. Radu and F. Schieweck, Post-processed Galerkin approximation of improved order for wave equations. *Math. Comput.* **89** (2020) 595–627.

[17] S. Becher and G. Matthies, Variational time discretizations of higher order and higher regularity. Preprint arXiv:2003.04056 (2020).

[18] S. Becher, G. Matthies and D. Wenzel, Variational methods for stable time discretization of first-order differential equations. In: Advanced Computing in Industrial Mathematics, edited by K. Georgiev, M. Todorov and I. Georgiev. BGSIAM. Springer, Cham (2018) 63–75.

[19] J. Česenek and M. Feistauer, Theory of the space-time discontinuous Galerkin method for nonstationary parabolic problems with nonlinear convection and diffusion. *SIAM J. Numer. Anal.* **50** (2012) 1181–1206.

[20] J.D. De Basabe, M.K. Sen and M.F. Wheeler, The interior penalty discontinuous Galerkin method for elastic wave propagation: grid dispersion. *Geophys. J. Int.* **175** (2008) 83–95.

[21] L.F. Demkowicz and J. Gopalakrishnan, An overview of the discontinuous Petrov–Galerkin method. In: Recent Developments in Discontinuous Galerkin Finite Element Methods for Partial Differential Equations, edited by X. Feng, O. Karakashian and Y. Xing. Springer, Cham (2014) 149–180.

[22] W. Dörfler, S. Findeisen and C. Wieners, Space-time discontinuous Galerkin discretizations for linear first-order hyperbolic evolution systems. *Comput. Methods Appl. Math.* **16** (2016) 409–428.

[23] A. Ern and F. Schieweck, Discontinuous Galerkin method in time combined with a stabilized finite element method in space for linear first-order PDEs. *Math. Comput.* **85** (2016) 2099–2129.

[24] L.C. Evans, Partial Differential Equations. American Mathematical Society, Providence, Rhode Island (2010).

[25] S.M. Findeisen, *A parallel and adaptive space-time method for Maxwell's equations*. Ph.D. thesis, KIT, Karlruhe (2016).

[26] D.A. French and T.E. Peterson, A continuous space-time finite element method for the wave equation. *Math. Comput.* **65** (1996) 491–506.

[27] M.J. Gander and M. Neumüller, Analysis of a new space-time parallel multigrid algorithm for parabolic problems. *SIAM J. Sci. Comput.* **38** (2016) A2173–A2208.

[28] B.L. Hulme, One-step piecewise polynomial Galerkin methods for initial value problems. *Math. Comput.* **26** (1972) 416–426.

[29] B.L. Hulme, Discrete Galerkin and related one-step methods for ordinary differential equations. *Math. Comput.* **26** (1972) 881–891.

[30] S. Hussain, F. Schieweck and S. Turek, Higher order Galerkin time discretizations and fast multigrid solvers for the heat equation. *J. Numer. Math.* **19** (2011) 41–61.

[31] S. Hussain, F. Schieweck and S. Turek, An efficient and stable finite element solver of higher order in space and time for nonstationary incompressible flow. *Int. J. Numer. Methods Fluids* **73** (2013) 927–952.

[32] C. Johnson, Discontinuous Galerkin finite element methods for second order hyperbolic problems. *Comput. Methods Appl. Mech. Eng.* **107** (1993) 117–129.

[33] H. Joulak and B. Beckermann, On Gautschi's conjecture for generalized Gauss–Radau and Gauss–Lobatto formulae. *J. Comput. Appl. Math.* **233** (2009) 768–774.

[34] O. Karakashian and C. Makridakis, Convergence of a continuous Galerkin method with mesh modification for nonlinear wave equations. *Math. Comput.* **74** (2004) 85–102.

[35] U. Köcher, *Variational space-time methods for the elastic wave equation and the diffusion equation*. Ph.D. thesis, Helmut Schmidt University, Hamburg (2015).

[36] U. Köcher and M. Bause, Variational space-time methods for the wave equation. *J. Sci. Comput.* **61** (2014) 424–453.

[37] U. Langer, S. Moore and M. Neumüller, Space-time isogeometric analysis of parabolic evoluation equations. *Comput. Methods Appl. Mech. Eng.* **306** (2016) 342–363.

[38] L. Lions, Optimal Control of Systems Governed by Partial Differential Equations. Springer, Berlin (1971).

[39] L. Lions and E. Magenes, Non-Homogeneous Boundary Value Problems and Applications. Springer, Berlin (1972).

[40] C. Makridakis and R.H. Nochetto, A posteriori error analysis for higher order dissipative methods for evolution problems. *Numer. Math.* **104** (2006) 489–514.

[41] G. Matthies and F. Schieweck, Higher order variational time discretizations for nonlinear systems of ordinary differential equations. Preprint 23/2011, Fakultät für Mathematik, Otto-von-Guericke-Universität Magdeburg (2011).

[42] A. Mikelić and M.F. Wheeler, Theory of the dynamic Biot–Allard equations and their link to the quasi-static Biot system. *J. Math. Phys.* **53** (2012) 123702.

[43] A. Quarteroni and A. Valli, Numerical Approximation of Partial Differential Equations. Springer, Berlin (2008).

[44] O. Steinbach, Space-time finite element methods for parabolic problems. *Comput. Methods Appl. Math.* **15** (2015) 551–566.

[45] O. Steinbach and H. Yang, Comparison of algebraic multigrid methods for an adaptive space-time finite-element discretization of the heat equation in 3D and 4D. *Numer. Linear Algebra App.* **25** (2018) e2143.

[46] V. Thomeé, Galerkin Finite Element Methods for Parabolic Problems. Springer, Berlin (2006).

[47] M. Vlasak and F. Roskovec, On Runge–Kutta, collocation and discontinuous Galerkin methods: mutual connections and resulting consequences to the analysis. *Programs Algorithms Numer. Math.* **17** (2015) 231–236.