

# HIGHER ORDER CONTINUOUS GALERKIN–PETROV TIME STEPPING SCHEMES FOR TRANSIENT CONVECTION-DIFFUSION-REACTION EQUATIONS\*

NAVEED AHMED<sup>1</sup> AND GUNAR MATTHIES<sup>2</sup>

**Abstract.** We present the analysis for the higher order continuous Galerkin–Petrov (cGP) time discretization schemes in combination with the one-level local projection stabilization in space applied to time-dependent convection-diffusion-reaction problems. Optimal *a priori* error estimates will be proved. Numerical studies support the theoretical results. Furthermore, a numerical comparison between continuous Galerkin–Petrov and discontinuous Galerkin time discretization schemes will be given.

**Mathematics Subject Classification.** 65M12, 65M15, 65M60.

Received August 1, 2014. Revised March 25, 2015.

Published online September 2, 2015.

## 1. INTRODUCTION

In recent years many numerical methods have been developed for the numerical solution of time-dependent convection-diffusion-reaction equations. Some of the most effective and popular algorithms for treating time-dependent problems can be defined through a process in which the spatial and temporal discretization are separated. A common approach is to first apply the Galerkin method in space to reduce the time-dependent partial differential equation into a system of ordinary differential equations. A suitable time discretization method is then applied to solve it. An alternative to this approach is a coupled space-time formulation which results in a system where all degrees inside the space-time cylinder are coupled. Hence, a  $(d + 1)$ -dimensional problem is discretized for a spatial domain with  $d$  dimensions. We will separate in this paper the discretizations in space and time.

The application of standard finite element methods to convection-dominated problems leads to spurious oscillations which spread over the whole spatial domain unless the discretization parameter is of the same order as the perturbation parameter.

---

*Keywords and phrases.* Transient convection-diffusion-reaction problem, local projection stabilization, continuous Galerkin–Petrov method, discontinuous Galerkin method.

\* *This research work was partially supported by the German research foundation (DFG) under the Grant MA 4713/2-1.*

<sup>1</sup> Weierstrass Institute for Applied Analysis and Stochastics (WIAS), Mohrenstr. 39, 10117 Berlin, Germany.  
[ahmed@wias-berlin.de](mailto:ahmed@wias-berlin.de)

<sup>2</sup> Technische Universität Dresden, Institut für Numerische Mathematik, 01062 Dresden, Germany.  
[gunar.matthies@tu-dresden.de](mailto:gunar.matthies@tu-dresden.de)

To overcome the instability while ensuring high accuracy, several stabilized techniques were proposed in the literature. One of the popular remedies is the streamline-upwind Petrov–Galerkin (SUPG) method which was introduced by Hughes and Brooks [15]. The SUPG method was originally proposed for steady state problems and has been extended to transient problem, see [8, 22]. Standard energy arguments applied to the fully discrete problem yield error estimates under conditions which couple the choice of the stabilization parameters to the length of the time step. In particular, the SUPG stabilization vanishes in the time-continuous limit. This behavior is caused by the time derivative appearing in the stabilization terms of the SUPG method to guarantee the strong consistency. The appearing non-symmetric term is difficult to handle in the numerical analysis. Using a different analysis for the case of time independent coefficients on uniform grids, optimal error estimates with the standard choice of the stabilization parameter independent of the time step length have been proved and confirmed by numerical experiments, see [22].

Comparisons of the SUPG method with other stabilization techniques can be found in [11, 26]. The stability of consistent stabilization methods for convection-diffusion and flow problems in the small time step limit has been investigated in [6, 14].

A stabilization technique which became very popular during the last decade is the local projection stabilization (LPS) scheme [5, 7, 28]. Compared to the standard Galerkin approach the local projection method provides additional control on the fluctuations of the gradient or parts of its. Although the method is weakly consistent only, the consistency error can be bounded such that the optimal order of convergence is maintained. Originally proposed for the Stokes's problem [4], the local projection method was extended successfully to transport problems [5], Oseen problems [7, 28], and convection-diffusion-reaction problems [29]. In contrast to the SUPG method, neither time derivatives nor second order spatial derivatives have to be assembled for the stabilization term of LPS. Furthermore, no coupling conditions of stabilization parameters to the length of the time step arise.

The local projection stabilization can be applied in one-level or two-level versions. In the latter ones, the projection spaces are defined on coarser meshes but standard finite element spaces can be used for approximation and projection. The one-level LPS uses the same mesh for approximation and projection but the approximation space is enriched compared to standard finite element spaces. The two-level version of LPS introduces additional couplings between degrees of freedoms which do not belong to the same finite element cell. Hence, the stencil of the stiffness matrix increases. This does not occur for the one-level variant of the LPS. The additional degrees of freedom which appear due to the enrichment can be eliminated using static condensation. In this work, we will use the one-level version of LPS.

Using stabilization techniques in space leads to a dramatic decrease of the oscillations which appear now only close to boundary and interior layers. Moreover, the amplitudes are much smaller. To remove the remaining oscillations, techniques for shock or discontinuity capturing can be applied. Numerical and analytical results are given for instance in [18–20].

Accurate numerical solutions of time-dependent problems require higher order methods both in space and time. In this paper, we will apply higher order variational type time discretization schemes. In particular, we will use continuous Galerkin–Petrov (cGP) and discontinuous Galerkin (dG) time stepping schemes. The cGP methods are a class of finite element methods using discrete solution spaces in time which consist of continuous piece-wise polynomials of degree less than or equal to  $k$  and test spaces which are built by discontinuous polynomials of degree up to order  $k - 1$ . In dG methods, both solution and test spaces are constructed by the discontinuous polynomials of degree less than or equal to  $k$ . Since the test functions in time are for both considered temporal discretizations allowed to be discontinuous at the discrete time points, the solution of cGP and dG schemes can be calculated by a time marching process.

The dG methods were first introduced by Reed and Hill [30] for neutron transport problems. The use of continuous and discontinuous finite element methods to discretize time-dependent problems has been analyzed for ordinary and partial differential equations by several authors. The time discretization by discontinuous Galerkin's methods was introduced and analyzed in [12] for the numerical solution of ordinary differential equations and is combined with continuous finite element methods in space for parabolic problem in [13, 33, 34].

The combination of LPS methods in space and dG methods in time has been analyzed for transient convection-diffusion-reaction problems in [2]. The continuous Galerkin method in time for the heat equation has been studied by Aziz and Monk in [3]. They have proved optimal error estimates as well as super-convergence results at the endpoints of the discrete time intervals. Schieweck [32] has investigated the cGP method for linear ordinary differential equations in an abstract Hilbert-space setting and for nonlinear systems of ordinary differential equations in  $d$  space dimensions. He has proved A-stability and optimal error estimates of the associated cGP method. Moreover, it was shown that this discretization method has an energy decreasing property for the gradient flow equation of an energy functional. Numerical comparisons of dG and cGP methods as time discretization of heat equations and transient Stokes problems are presented in [16,17]. A family of finite element methods for the numerical solution of nonlinear systems of ordinary differential equations was recently given in [27]. Furthermore, it was shown there that the new methods can be interpreted as pure collocation and pure variational methods with special numerical integration. In [27], simple post-processing algorithms were presented which allow to increase the obtained accuracy in time by one order in time-integrated norms.

The goal of the present paper is to combine the local projection stabilization in space with continuous Galerkin–Petrov discretizations in time. We will give a stability result and error estimates for the fully discrete scheme. Furthermore, a numerical comparison of cGP and dG time discretization schemes combined with LPS in space will be given.

This paper is organized as follows: In Section 2, a weak formulation of the time-dependent convection-diffusion-reaction equation and some basic notation are given. The semi-discretization in space and the local projection stabilization method are introduced in Section 3. Section 4 presents the fully discrete problem with a continuous Galerkin–Petrov time discretization and its error analysis. Finally, numerical results which confirm the theoretical prediction will be given in Section 5. Furthermore, a comparison to discontinuous Galerkin time discretization schemes will be presented.

## 2. MODEL PROBLEM

Let  $\Omega$  be a polygonal or polyhedral domain in  $\mathbb{R}^d$ ,  $d = 2, 3$ , with Lipschitz boundary  $\partial\Omega$ . We consider the following time-dependent convection-diffusion-reaction equation:

Find  $u : \Omega \times [0, T] \rightarrow \mathbb{R}$  such that

$$\begin{cases} u' - \varepsilon \Delta u + \mathbf{b} \cdot \nabla u + \sigma u = f & \text{in } \Omega \times (0, T), \\ u = 0 & \text{on } \partial\Omega \times (0, T), \\ u(\cdot, 0) = u_0 & \text{in } \Omega \end{cases} \quad (2.1)$$

with a small positive constant  $0 < \varepsilon \ll 1$ ,  $\mathbf{b}(x)$ , and  $\sigma(x)$  are given function,  $u_0$  the initial data, and  $T > 0$  a given final time. We assume that  $\mathbf{b}$  and  $\sigma$  are time-independent whereas  $f$  may depend on time  $t$ . Furthermore, let the data  $\mathbf{b}$ ,  $\sigma$ ,  $u_0$ , and  $f$  be sufficiently smooth in  $\Omega$  and  $\Omega \times (0, T)$ , respectively. We assume in the following that there exists a positive constant  $\sigma_0$  such that

$$\sigma(x) - \frac{1}{2} \operatorname{div} \mathbf{b}(x) \geq \sigma_0 > 0 \quad \text{in } \Omega. \quad (2.2)$$

Note that the assumption (2.2) is no restriction for time-dependent convection-diffusion-reaction problems of type (2.1). Indeed, if condition (2.2) is not fulfilled we consider the problem

Find  $v : \Omega \times [0, T] \rightarrow \mathbb{R}$  such that

$$\begin{cases} v' - \varepsilon \Delta v + \mathbf{b} \cdot \nabla v + (\sigma + M)v = e^{-Mt} f & \text{in } \Omega \times (0, T), \\ v = 0 & \text{on } \partial\Omega \times (0, T), \\ v(\cdot, 0) = u_0 & \text{in } \Omega \end{cases}$$

for the function  $v$  defined by

$$v(x, t) := e^{-Mt}u(x, t).$$

If  $M$  is chosen sufficiently large then condition (2.2) is fulfilled for the modified problem for  $v$ .

Throughout this paper, standard notation and conventions will be used. For a measurable set  $G \subset \mathbb{R}^d$ , the inner product in  $L^2(G)$  and  $L^2(G)^d$  will be denoted by  $(\cdot, \cdot)_G$ . The norm and semi-norm in  $W^{m,p}(G)$  are given by  $\|\cdot\|_{m,p,G}$  and  $|\cdot|_{m,p,G}$ , respectively. In the case  $p = 2$ , we write  $H^m(G)$ ,  $\|\cdot\|_{m,G}$ , and  $|\cdot|_{m,G}$  instead of  $W^{m,2}(G)$ ,  $\|\cdot\|_{m,2,G}$ , and  $|\cdot|_{m,2,G}$ . If  $G = \Omega$ , the index  $G$  in inner products, norms, and semi-norms will be omitted. The subspace of functions from  $H^1(\Omega)$  having zero boundary trace is denoted by  $H_0^1(\Omega)$ .

We will write shortly  $\alpha \sim \beta$  if there exist positive constants  $C_1$  and  $C_2$  such that  $\alpha \leq C_1\beta$  and  $\beta \leq C_2\alpha$  hold true.

We consider also some Bochner spaces. Let  $W$  be a Banach space with norm  $\|\cdot\|_W$  and  $I := [0, T]$ . We define

$$\begin{aligned} C(I; W) &:= \{v : I \rightarrow W, \quad v \text{ continuous}\}, \\ L^2(I; W) &:= \left\{ v : I \rightarrow W, \quad \int_0^T \|v(t)\|_W^2 dt < \infty \right\}, \\ H^m(I; W) &:= \left\{ v \in L^2(I; W) : \frac{\partial^j v}{\partial t^j} \in L^2(I; W), \quad 1 \leq j \leq m \right\}, \end{aligned}$$

where the derivatives  $\partial^j v / \partial t^j$ ,  $j = 1, \dots, m$ , are understood in the sense of distributions on  $I$ . We use in the following the short notation  $Y(W) := Y(I; W)$ . The norms in the above defined spaces are given by

$$\|v\|_{C(W)} := \sup_{t \in I} \|v(t)\|_W, \quad \|v\|_{L^2(W)} := \left( \int_0^T \|v(t)\|_W^2 dt \right)^{1/2}, \quad \|v\|_{H^m(W)} := \left( \sum_{j=0}^m \left\| \frac{\partial^j v}{\partial t^j} \right\|_{L^2(W)}^2 \right)^{1/2}.$$

Let us introduce the space  $V := H_0^1(\Omega)$ , its dual space  $V' := H^{-1}(\Omega)$ , and  $\langle \cdot, \cdot \rangle$  for the duality product between these two spaces.

Let  $u_0 \in H_0^1(\Omega)$  and  $f \in L^2(V')$ . A function  $u$  is a weak solution of problem (2.1), if

$$u \in X := \{u \in L^2(V) : u' \in L^2(V')\} \tag{2.3}$$

with

$$\langle u'(t), v \rangle + a(u(t), v) = \langle f(t), v \rangle \quad \forall v \in V \tag{2.4}$$

for almost all  $t \in I$  and

$$u(0) = u_0, \tag{2.5}$$

where the bilinear form  $a$  is given by

$$a(u, v) := \varepsilon(\nabla u, \nabla v) + (\mathbf{b} \cdot \nabla u, v) + (\sigma u, v).$$

Note that the definition of  $X$  in (2.3) implies the continuity of  $u$  as a mapping  $I \rightarrow L^2(\Omega)$  such that the initial condition (2.5) is well-defined.

Integrating (2.4) over  $[0, T]$ , we obtain the problem:

Find  $u \in X$  with  $u(0) = u_0$  and

$$\int_0^T \langle u'(t), v(t) \rangle + a(u(t), v(t)) dt = \int_0^T \langle f(t), v(t) \rangle dt \quad \forall v \in L^2(V). \tag{2.6}$$

In the following, we shall denote by  $\varphi'$ ,  $\varphi''$ , and  $\varphi^{(k)}$  the first, second, and  $k$ th order time derivative of a function  $\varphi$  which is sufficiently smooth in time.

For finite element discretizations of (2.6), let  $\{\mathcal{T}_h\}$  denote a family of shape regular triangulations of  $\Omega$  into open  $d$ -simplices, quadrilaterals, or hexahedra  $K$  such that

$$\overline{\Omega} = \bigcup_{K \in \mathcal{T}_h} \overline{K}.$$

The diameter of  $K \in \mathcal{T}_h$  will be denoted by  $h_K$  and the mesh size  $h$  is defined by  $h := \max_{K \in \mathcal{T}_h} h_K$ . Let  $V_h$  be a finite element space defined on  $\mathcal{T}_h$  such that the inverse inequality

$$\|\nabla v_h\|_{0,K} \leq c_{\text{inv}} h_K^{-1} \|v_h\|_{1,K}$$

holds for all  $K \in \mathcal{T}_h$  and all  $v_h \in V_h$  where the constant  $c_{\text{inv}}$  is independent of  $K$  and  $h$ . We set  $X_h := \mathbf{H}^1(V_h)$ . Let  $u_{0,h} \in V_h$  denote a suitable approximation of the initial condition  $u_0$ . Later on,  $u_{0,h}$  will be specified.

The standard Galerkin's method applied to problem (2.6) consists in

Find  $u_h \in X_h$  such that  $u_h(0) = u_{0,h}$  and

$$\int_0^T (u_h'(t), v_h(t)) + a(u_h(t), v_h(t)) dt = \int_0^T (f(t), v_h(t)) dt \quad \forall v_h \in L^2(V_h). \quad (2.7)$$

In the convection-dominant case  $\varepsilon \ll 1$ , it is well-known that the standard Galerkin's method (2.7) applied to problem (2.1) is unstable and leads to solutions which are globally polluted by spurious oscillations unless the discretization parameter  $h$  is of the same order as the perturbation parameter  $\varepsilon$ .

### 3. SPATIAL STABILIZATION BY THE LPS METHOD

We concentrate on the one-level local projection stabilization method in which approximation space and projection space are defined on the same mesh. Let  $D(K)$ ,  $K \in \mathcal{T}_h$ , be finite dimensional spaces and  $\pi_K : L^2(K) \rightarrow D(K)$  the local  $L^2$ -projection into  $D(K)$ . The local fluctuation operator  $\kappa_K : L^2(K) \rightarrow L^2(K)$  is given by  $\kappa_K v := v - \pi_K v$ . The stabilization term  $S_h$  is defined by

$$S_h(u_h, v_h) := \sum_{K \in \mathcal{T}_h} \mu_K (\kappa_K \nabla u_h, \kappa_K \nabla v_h)_K$$

where  $\mu_K$ ,  $K \in \mathcal{T}_h$ , are user chosen non-negative constants. The fluctuation operator is applied component-wise to vector-valued functions. The used local projection stabilization gives, compared to the standard Galerkin's methods, additional control on the fluctuation of the gradient. Note that one can replace the gradient  $\nabla w_h$  by the derivative in streamline direction  $\mathbf{b} \cdot \nabla w_h$  or (even better [24, 25]) by  $\mathbf{b}_K \cdot \nabla w_h$  where  $\mathbf{b}_K$  is a piece-wise constant approximation of  $\mathbf{b}$ .

Stability and convergence properties of local projection methods are based on the following assumptions with respect to the approximation space  $V_h$  and the local projection spaces  $D(K)$ ,  $K \in \mathcal{T}_h$ , see [28, 31].

**Assumption 1.** There is an interpolation operator  $j_h : \mathbf{H}^2(\Omega) \rightarrow V_h$  such that for all  $K \in \mathcal{T}_h$  the standard error estimate

$$\|v - j_h v\|_{0,K} + h_K |v - j_h v|_{1,K} \leq C h_K^l \|v\|_{l,K} \quad \forall v \in \mathbf{H}^l(K), \quad 2 \leq l \leq r+1, \quad (3.1)$$

and the additional orthogonality condition

$$(v - j_h v, q_h)_K = 0 \quad \forall v \in \mathbf{H}^2(\Omega), \quad \forall q_h \in D(K) \quad (3.2)$$

hold true.

**Assumption 2.** The fluctuation operators  $\kappa_K$ ,  $K \in \mathcal{T}_h$ , satisfy the following approximation property

$$\|\kappa_K q\|_{0,K} \leq Ch_K^l |q|_{l,K} \quad \forall q \in H^l(K), 0 \leq l \leq r. \tag{3.3}$$

In order to concentrate on the main aspects of temporal discretizations, we will assume in the following

$$f \in L^2(L^2), \quad u_0 \in H^2(\Omega),$$

and the discrete initial condition  $u_{0,h} = j_h u_0$  is well-defined. In the case  $u_0 \in H^1(\Omega)$  only, one could use an interpolation operator  $j_h : H^1(\Omega) \rightarrow V_h$  with the orthogonality condition (3.2) and the error estimate

$$\|v - j_h v\|_{0,K} + h_K |v - j_h v|_{1,K} \leq Ch_K^l \|v\|_{l,\omega(K)} \quad \forall v \in H^l(K), 2 \leq l \leq r + 1,$$

with a suitable neighborhood  $\omega(K)$  of  $K \in \mathcal{T}_h$ , see [28]. The upcoming analysis would more technical and is skipped for this reason.

The stabilized semi-discrete problem for  $u_h$  reads:

Find  $u_h \in X_h$  such that  $u_h(0) = j_h u_0$  and

$$\int_0^T (u_h'(t), v_h(t)) + a_h(u_h(t), v_h(t)) dt = \int_0^T (f(t), v_h(t)) dt \quad \forall v_h \in L^2(V_h) \tag{3.4}$$

where the stabilized bilinear form  $a_h$  is given by

$$a_h(u_h, v_h) := a(u_h, v_h) + S_h(u_h, v_h).$$

Thanks to (2.2) we have

$$a_h(v_h, v_h) \geq \|v_h\|^2 \quad \forall v_h \in V_h, \tag{3.5}$$

*i.e.*, the bilinear form  $a_h$  is coercive with respect to the mesh-dependent norm

$$\|v\| := \{\varepsilon |v|_1^2 + \sigma_0 \|v\|_0^2 + S_h(v, v)\}^{1/2}. \tag{3.6}$$

#### 4. TIME DISCRETIZATION

We discretize in this section problem (3.4) in time by using continuous Galerkin–Petrov’s methods and discontinuous Galerkin’s methods. To this end, we consider a partition  $0 = t_0 < t_1 < \dots < t_N = T$  of the time interval  $I$  and set  $I_n := (t_{n-1}, t_n]$ ,  $\tau_n := t_n - t_{n-1}$ ,  $n = 1, \dots, N$ , and  $\tau := \max_{1 \leq n \leq N} \tau_n$ . For a given non-negative integer  $k$ , we define the spaces

$$\begin{aligned} X^k &:= \left\{ u \in C(I, V) : u|_{I_n} \in \mathbb{P}_k(I_n, V), \quad n = 1, \dots, N \right\}, \\ Y^k &:= \left\{ v \in L^2(I, V) : v|_{I_n} \in \mathbb{P}_k(I_n, V), \quad n = 1, \dots, N \right\} \end{aligned}$$

with values in  $V$  and the fully discrete spaces

$$\begin{aligned} X_h^k &:= \left\{ u \in C(I, V_h) : u|_{I_n} \in \mathbb{P}_k(I_n, V_h), \quad n = 1, \dots, N \right\}, \\ Y_h^k &:= \left\{ v \in L^2(I, V_h) : v|_{I_n} \in \mathbb{P}_k(I_n, V_h), \quad n = 1, \dots, N \right\} \end{aligned}$$

with values in  $V_h$  where

$$\mathbb{P}_k(I_n, V) := \left\{ u : I_n \rightarrow V : u(t) = \sum_{j=0}^k U^j t^j, U^j \in V, j = 0, \dots, k \right\},$$

$$\mathbb{P}_k(I_n, V_h) := \left\{ u_h : I_n \rightarrow V_h : u_h(t) = \sum_{j=0}^k U_h^j t^j, U_h^j \in V_h, j = 0, \dots, k \right\}$$

denote the spaces of  $V$ -valued and  $V_h$ -valued polynomials of order up to  $k$  in time, respectively. The functions in the spaces  $Y^k$  and  $Y_h^k$  are allowed to be discontinuous at the nodes  $t_n, n = 1, \dots, N - 1$ . We define for functions  $v \in Y^k + Y_h^k$  by

$$v_n^- := \lim_{t \rightarrow t_n-0} v(t), \quad v_n^+ := \lim_{t \rightarrow t_n+0} v(t), \quad [v]_n := v_n^+ - v_n^-, \tag{4.1}$$

the left-sided value  $v_n^-$ , right-sided value  $v_n^+$ , and the jump  $[v]_n$  at  $t_n$ .

### 4.1. The continuous Galerkin–Petrov (cGP) method

We describe in this section the combination of the continuous Galerkin–Petrov (cGP) time discretization scheme with the LPS finite element method in space to get a fully discrete version of (3.4).

Now, let us introduce the bilinear forms  $B, B_h$  and the linear form  $l$  as

$$B(u, v) := \int_0^T \left\{ (u'(t), v(t)) + a(u(t), v(t)) \right\} dt, \tag{4.2}$$

$$B_h(u, v) := \int_0^T \left\{ (u'(t), v(t)) + a_h(u(t), v(t)) \right\} dt, \tag{4.3}$$

$$l(v) := \int_0^T (f(t), v(t)) dt. \tag{4.4}$$

Then, the continuous and fully discrete problems read:

Find  $u \in X$  such that  $u(0) = u_0$  and

$$B(u, v) = l(v) \quad \forall v \in L^2(V). \tag{4.5}$$

Find  $u_{h,\tau} \in X_h^k$  such that  $u_{h,\tau}(0) = j_h u_0$  and

$$B_h(u_{h,\tau}, v_{h,\tau}) = l(v_{h,\tau}) \quad \forall v_{h,\tau} \in Y_h^{k-1}. \tag{4.6}$$

Since the test functions are allowed to be discontinuous at the discrete time points  $t_n, n = 1, \dots, N - 1$ , we can choose test functions  $v_{h,\tau} = v_h \varphi(t)$  with a time independent  $v_h \in V_h$  and a scalar function  $\varphi : I \rightarrow \mathbb{R}$  which is zero on  $I \setminus I_n$  and a polynomial of degree less than or equal to  $k - 1$  on  $I_n$ . Then, the solution of the cGP( $k$ ) method can be determined by successively solving one local problem on each time interval  $I_n$ .

Let us denote by  $\pi_{k-1} : L^2(L^2) \rightarrow Y_h^{k-1}$  the  $L^2$ -projection into the space of  $V_h$ -valued functions which are allowed to be discontinuous at the discrete time points. Hence, we have

$$\pi_{k-1} w \in Y_h^{k-1}, \quad \int_{I_n} (w(t) - \pi_{k-1} w(t)) t^j dt = 0, \quad j = 0, \dots, k - 1. \tag{4.7}$$

We consider the mesh-dependent norm

$$\|v\|_{\text{cGP}} = \left( \int_0^T \|\pi_{k-1} v(t)\|^2 dt + \frac{1}{2} \|v(T)\|_0^2 \right)^{1/2}.$$

Note that  $\|\cdot\|_{\text{cGP}}$  is on  $X_h^k \subset Y_h^k$  not only a semi-norm but a norm. Indeed, the first term inside the definition of  $\|\cdot\|_{\text{cGP}}$  guarantees that  $\|v\|_{\text{cGP}} = 0$  results in a function  $v$  which is on each time interval  $I_n$  given by  $L_k^{(n)}(t)\varphi(x)$  where  $L_k^{(n)}$  is the transformed  $k$ th Legendre polynomial on  $I_n$  and  $\varphi \in V_h$ . Due to  $v(T) = 0$  and  $L_k^{(N)}(T) = 1$ , the function  $v$  vanishes on the last time interval  $I_N$ . The continuity of  $v$  on  $I$  gives then  $v(t_{N-1}) = 0$ . By recursion we obtain that  $v \equiv 0$  on  $I$ . Hence,  $\|\cdot\|_{\text{cGP}}$  is a norm on  $X_h^k$ .

We will show in the following lemma a property of the bilinear form  $B_h$  which will be used in the next theorem to prove existence, uniqueness, and stability of fully discrete problem (4.6).

**Lemma 4.1.** *The condition*

$$\|v_{h,\tau}\|_{\text{cGP}}^2 \leq B_h(v_{h,\tau}, \pi_{k-1}v_{h,\tau}) + \frac{1}{2}\|v_{h,\tau}(0)\|_0^2 \tag{4.8}$$

holds true for all  $v_{h,\tau} \in X_h^k$ .

*Proof.* Using the definition (4.3) of  $B_h$ , we obtain

$$B_h(v_{h,\tau}, \pi_{k-1}v_{h,\tau}) = \int_0^T (v'_{h,\tau}, \pi_{k-1}v_{h,\tau}) + a_h(v_{h,\tau}, \pi_{k-1}v_{h,\tau}) \, dt. \tag{4.9}$$

We get

$$\int_0^T (v'_{h,\tau}, \pi_{k-1}v_{h,\tau}) \, dt = \int_0^T (v'_{h,\tau}, v_{h,\tau}) \, dt = \frac{1}{2} \int_0^T \frac{d}{dt} \|v_{h,\tau}\|_0^2 \, dt = \frac{1}{2} \|v_{h,\tau}(T)\|_0^2 - \frac{1}{2} \|v_{h,\tau}(0)\|_0^2$$

for the first term on the right-hand side of (4.9) where the projection property (4.7) and  $v'_{h,\tau} \in Y_h^{k-1}$  were applied. The second term in (4.9) uses the fact that convection and reaction are independent of time. Since  $v_{h,\tau} \in X_h^k$  and  $\pi_{k-1}v_{h,\tau} \in Y_h^{k-1}$ , the definition (4.7) of  $\pi_{k-1}$  and the coercivity (3.5) of the bilinear form  $a_h$  give

$$\int_0^T a_h(v_{h,\tau}, \pi_{k-1}v_{h,\tau}) \, dt = \int_0^T a_h(\pi_{k-1}v_{h,\tau}, \pi_{k-1}v_{h,\tau}) \, dt \geq \int_0^T \|\pi_{k-1}v_{h,\tau}\|^2 \, dt.$$

Putting all these estimates into (4.9), we get

$$B_h(v_{h,\tau}, \pi_{k-1}v_{h,\tau}) \geq \int_0^T \|\pi_{k-1}v_{h,\tau}\|^2 \, dt + \frac{1}{2} \|v_{h,\tau}(T)\|_0^2 - \frac{1}{2} \|v_{h,\tau}(0)\|_0^2 = \|v_{h,\tau}\|_{\text{cGP}}^2 - \frac{1}{2} \|v_{h,\tau}(0)\|_0^2. \tag{4.10}$$

Hence, the statement of this lemma follows. □

The next theorem states the stability of the fully discrete method (4.6).

**Theorem 4.2.** *The solution  $u_{h,\tau}$  of the fully discrete problem (4.6) is uniquely determined and satisfies the stability estimate*

$$\|u_{h,\tau}\|_{\text{cGP}} \leq \frac{1}{\sqrt{\sigma_0}} \left( \int_0^T \|f\|_0^2 \, dt \right)^{1/2} + \|j_h u_0\|_0 \tag{4.11}$$

with the constant  $\sigma_0$  from (2.2).



*Proof.* Due to Lemma 4.1, we have

$$\begin{aligned}
\|u_{h,\tau}\|_{\text{cGP}}^2 &\leq B_h(u_{h,\tau}, \pi_{k-1}u_{h,\tau}) + \frac{1}{2}\|u_{0,h}\|_0^2 = \int_0^T (f, \pi_{k-1}u_{h,\tau}) \, dt + \frac{1}{2}\|u_{0,h}\|_0^2 \\
&\leq \int_0^T \|f\|_0 \|\pi_{k-1}u_{h,\tau}\|_0 \, dt + \frac{1}{2}\|u_{0,h}\|_0^2 \\
&\leq \frac{1}{\sqrt{\sigma_0}} \int_0^T \|f\|_0 \sqrt{\sigma_0} \|\pi_{k-1}u_{h,\tau}\|_0 \, dt + \frac{1}{2}\|u_{0,h}\|_0^2 \\
&\leq \frac{1}{2\sigma_0} \int_0^T \|f\|_0^2 \, dt + \frac{1}{2} \int_0^T \|\pi_{k-1}u_{h,\tau}\|^2 \, dt + \frac{1}{2}\|u_{0,h}\|_0^2 \\
&\leq \frac{1}{2\sigma_0} \int_0^T \|f\|_0^2 \, dt + \frac{1}{2}\|u_{h,\tau}\|_{\text{cGP}}^2 + \frac{1}{2}\|u_{0,h}\|_0^2
\end{aligned}$$

and the statement of this Lemma follows since  $u_{h,\tau}(0) = j_h u(0)$ .  $\square$

We define for the sufficiently smooth solution  $u$  of (4.5) its time interpolant  $\tilde{u} \in X^k$  on the interval  $I_n$  by

$$\tilde{u}(t_{n-1}) = u(t_{n-1}), \quad \tilde{u}(t_n) = u(t_n), \quad \int_{I_n} (u(t) - \tilde{u}(t), w(t)) \, dt = 0 \quad \forall w \in \mathbb{P}_{k-2}(I_n, V). \quad (4.12)$$

Hence, the standard interpolation error estimates

$$\int_{I_n} |u(t) - \tilde{u}(t)|_m \, dt \leq C\tau_n^{k+1} \int_{I_n} |u^{(k+1)}(t)|_m \, dt, \quad (4.13)$$

$$\int_{I_n} |(u - \tilde{u})'(t)|_m \, dt \leq C\tau_n^k \int_{I_n} |u^{(k+1)}(t)|_m \, dt \quad (4.14)$$

hold true for  $m \in \{0, 1\}$  and all time intervals  $I_n$ ,  $n = 1, \dots, N$ .

**Theorem 4.3.** *Assume  $\mu_K \sim h_K$  for all  $K \in \mathcal{T}_h$ . Suppose A1 and A2. Let  $u_{h,\tau}$  and  $u$  be the solutions of the fully discrete problem (4.6) and the continuous problem (4.5), respectively. Moreover, let  $u \in \mathbf{H}^1(\mathbf{H}^{r+1}) \cap \mathbf{H}^{k+1}(\mathbf{H}^1)$ . Then, there exists a positive constant  $C$  independent of  $h$ ,  $\tau$ , and  $\varepsilon$ , such that the error estimate*

$$\|u_{h,\tau} - u\|_{\text{cGP}} \leq C \left[ \tau^{k+1} \|u\|_{\mathbf{H}^{k+1}(\mathbf{H}^1)} + (\varepsilon^{1/2} + h^{1/2}) h^r \|u\|_{\mathbf{H}^1(\mathbf{H}^{r+1})} \right] \quad (4.15)$$

holds true.

*Proof.* The error analysis starts by decomposing the error  $e := u_{h,\tau} - u$  into the interpolation error  $\eta := j_h \tilde{u} - u$  and the difference  $\xi := u_{h,\tau} - j_h \tilde{u}$  between discrete solution and the interpolant of  $u$ . Hence, we have

$$u_{h,\tau} - u = e = \xi + \eta.$$

Lemma 4.1 for the discrete error function  $\xi$  provides

$$\|\xi\|_{\text{cGP}}^2 \leq B_h(\xi, v_{h,\tau}) + \frac{1}{2}\|\xi(0)\|_0^2 = B_h(\xi, \pi_{k-1}\xi) \quad (4.16)$$

since  $\xi(0) = u_{h,\tau}(0) - j_h \tilde{u}(0) = 0$  due to the choice of the discrete initial condition and the properties (4.12) of the interpolation in time.

In the following, we will bound the right-hand side of (4.16). We obtain from (4.5) and (4.6) the error equation

$$\begin{aligned}
 B_h(\xi, \pi_{k-1}\xi) &= \int_0^T (f, \pi_{k-1}\xi) dt - \int_0^T ((j_h \tilde{u})', \pi_{k-1}\xi) dt - \int_0^T a_h(j_h \tilde{u}, \pi_{k-1}\xi) dt \\
 &= \int_0^T (u', \pi_{k-1}\xi) dt + \int_0^T a(u, \pi_{k-1}\xi) dt - \int_0^T ((j_h \tilde{u})', \pi_{k-1}\xi) dt - \int_0^T a_h(j_h \tilde{u}, \pi_{k-1}\xi) dt \\
 &= - \int_0^T (\eta', \pi_{k-1}\xi) dt - \int_0^T a(\eta, \pi_{k-1}\xi) dt - \int_0^T S_h(\eta, \pi_{k-1}\xi) dt - \int_0^T S_h(u, \pi_{k-1}\xi) dt. \tag{4.17}
 \end{aligned}$$

The arising terms on the right-hand side of (4.17) will be bounded by terms depending on the solution  $u$  of the continuous problem (4.5). Using the error splitting and an integration by parts with respect to time, we get

$$\begin{aligned}
 - \int_0^T (\eta', \pi_{k-1}\xi) dt &= \int_0^T ((u - \tilde{u})', \pi_{k-1}\xi) dt + \int_0^T ((\tilde{u} - j_h \tilde{u})', \pi_{k-1}\xi) dt \\
 &= \sum_{n=1}^N \left( - \int_{I_n} (u - \tilde{u}, (\pi_{k-1}\xi)') dt + (u - \tilde{u}, \pi_{k-1}\xi) \Big|_{t_{n-1}}^{t_n} \right) + \int_0^T ((\tilde{u} - j_h \tilde{u})', \pi_{k-1}\xi) dt \\
 &= \int_0^T (\tilde{u}' - j_h \tilde{u}', \pi_{k-1}\xi) dt.
 \end{aligned}$$

Here, we have used the properties (4.12) of interpolation in time and the fact that time derivative and interpolation in space commute. Hence, using the Cauchy–Schwarz’s inequality, the interpolation properties of  $j_h$ , and the stability property of the interpolation in time, we obtain

$$\begin{aligned}
 \left| - \int_0^T (\eta', \pi_{k-1}\xi) dt \right| &\leq \sum_{n=1}^N \int_{I_n} \|\tilde{u}' - j_h \tilde{u}'\|_0 \|\pi_{k-1}\xi\|_0 dt \leq Ch^{r+1} \sum_{n=1}^N \int_{I_n} \|\tilde{u}'\|_{r+1} \|\pi_{k-1}\xi\|_0 dt \\
 &\leq Ch^{r+1} \left( \sum_{n=1}^N \int_{I_n} \|\tilde{u}'\|_{r+1}^2 dt \right)^{1/2} \left( \sum_{n=1}^N \int_{I_n} \sigma_0 \|\pi_{k-1}\xi\|_0^2 dt \right)^{1/2} \\
 &\leq Ch^{r+1} \|u\|_{H^1(H^{r+1})} \|\xi\|_{cGP}.
 \end{aligned}$$

To get an estimate for the second term on the right-hand side of (4.17), we exploit the definition of the bilinear form  $a$ . Using the Cauchy–Schwarz’s inequality, the error splitting, and the approximation properties (3.1) and (4.13), we get

$$\begin{aligned}
 &\left| \int_0^T \left\{ \varepsilon(\nabla \eta, \nabla \pi_{k-1}\xi) + (\sigma \eta, \pi_{k-1}\xi) \right\} dt \right| \\
 &\leq C \sum_{n=1}^N \int_{I_n} (\varepsilon |u - j_h \tilde{u}|_1^2 + \|u - j_h \tilde{u}\|_0^2)^{1/2} \|\pi_{k-1}\xi\| dt \\
 &\leq C \sum_{n=1}^N \int_{I_n} (\varepsilon |u - \tilde{u}|_1^2 + \varepsilon |\tilde{u} - j_h \tilde{u}|_1^2 + \|u - \tilde{u}\|_0^2 + \|\tilde{u} - j_h \tilde{u}\|_0^2)^{1/2} \|\pi_{k-1}\xi\| dt \\
 &\leq C \sum_{n=1}^N \left[ \int_{I_n} \left( (\varepsilon + h^2) h^{2r} \|\tilde{u}\|_{r+1}^2 + (\varepsilon + 1) \tau_n^{2(k+1)} \left( \|u^{(k+1)}\|_0^2 + |u^{(k+1)}|_1^2 \right) \right)^{1/2} \|\pi_{k-1}\xi\| dt \right] \\
 &\leq C \left[ (\varepsilon^{1/2} + h) h^r \|u\|_{L^2(H^{r+1})} + (\varepsilon^{1/2} + 1) \tau^{k+1} \|u\|_{H^{k+1}(H^1)} \right] \|\xi\|_{cGP}.
 \end{aligned}$$

In order to estimate the convection term in the bilinear form  $a$ , we proceed as follows. Using the error splitting, we get

$$\int_0^T (\mathbf{b} \cdot \nabla \eta, \pi_{k-1} \xi) dt = - \int_0^T (\mathbf{b} \cdot \nabla(u - \tilde{u}), \pi_{k-1} \xi) dt - \int_0^T (\mathbf{b} \cdot \nabla(\tilde{u} - j_h \tilde{u}), \pi_{k-1} \xi) dt. \quad (4.18)$$

The interpolation error estimate (4.13) gives for the first term in (4.18)

$$\begin{aligned} \left| - \int_0^T (\mathbf{b} \cdot \nabla(u - \tilde{u}), \pi_{k-1} \xi) dt \right| &\leq \frac{\|\mathbf{b}\|_{0,\infty}}{\sqrt{\sigma_0}} \sum_{n=1}^N \int_{I_n} \|\nabla(u - \tilde{u})\|_0 \sqrt{\sigma_0} \|\pi_{k-1} \xi\|_0 dt \\ &\leq C \left( \sum_{n=1}^N \tau_n^{2k+2} \int_{I_n} |u^{(k+1)}|_1^2 dt \right)^{1/2} \left( \sum_{n=1}^N \int_{I_n} \|\pi_{k-1} \xi\|^2 dt \right)^{1/2} \\ &\leq C \tau^{k+1} \|u\|_{\mathbb{H}^{k+1}(\mathbb{H}^1)} \|\xi\|_{\text{cGP}}. \end{aligned}$$

Integrating the second term in (4.18) by parts in space gives

$$\left| - \int_0^T (\mathbf{b} \cdot \nabla(\tilde{u} - j_h \tilde{u}), \pi_{k-1} \xi) dt \right| \leq \left| \int_0^T (\tilde{u} - j_h \tilde{u}, \mathbf{b} \cdot \nabla \pi_{k-1} \xi) dt \right| + \left| \int_0^T ((\pi_{k-1} \xi) \operatorname{div} \mathbf{b}, \tilde{u} - j_h \tilde{u}) dt \right|. \quad (4.19)$$

The first term in (4.19) is estimated by using the orthogonality condition (3.2) as follows

$$\begin{aligned} \left| \int_0^T (\tilde{u} - j_h \tilde{u}, \mathbf{b} \cdot \nabla \pi_{k-1} \xi) dt \right| &= \left| \int_0^T \sum_{K \in \mathcal{T}_h} (\tilde{u} - j_h \tilde{u}, \kappa_K(\mathbf{b} \cdot \nabla \pi_{k-1} \xi))_K dt \right| \\ &\leq \int_0^T \sum_{K \in \mathcal{T}_h} \|\tilde{u} - j_h \tilde{u}\|_{0,K} \|\kappa_K(\mathbf{b} \cdot \nabla \pi_{k-1} \xi)\|_{0,K} dt. \end{aligned}$$

One can obtain for  $\mathbf{b}|_K \in (W^{1,\infty}(K))^d$  the estimate

$$\|\kappa_K(\mathbf{b} \cdot \nabla \pi_{k-1} \xi)\|_{0,K} \leq C \|\mathbf{b}\|_{1,\infty,K} \|\pi_{k-1} \xi\|_{0,K} + \|\mathbf{b}\|_{0,\infty,K} \|\kappa_K(\nabla \pi_{k-1} \xi)\|_{0,K},$$

see ([28], Cor. 2.14). Using this estimate, the approximation property (3.1),  $\mu_K \sim h_K$ , and the stability property of the interpolation in time, we get

$$\begin{aligned} \left| \int_0^T (\tilde{u} - j_h \tilde{u}, \mathbf{b} \cdot \nabla \pi_{k-1} \xi) dt \right| &\leq C \int_0^T \sum_{K \in \mathcal{T}_h} h_K^{r+1} \|\tilde{u}\|_{r+1,K} \left( \|\pi_{k-1} \xi\|_{0,K} + \|\kappa_K(\nabla \pi_{k-1} \xi)\|_{0,K} \right) dt \\ &\leq C \left( \int_0^T \sum_{K \in \mathcal{T}_h} h_K^{2r+2} \|\tilde{u}\|_{r+1,K}^2 dt \right)^{1/2} \left( \int_0^T \sigma_0 \|\pi_{k-1} \xi\|_0^2 dt \right)^{1/2} \\ &\quad + C \left( \int_0^T \sum_{K \in \mathcal{T}_h} \mu_K^{-1} h_K^{2r+2} \|\tilde{u}\|_{r+1,K}^2 dt \right)^{1/2} \left( \int_0^T \sum_{K \in \mathcal{T}_h} \mu_K \|\kappa_K(\nabla \pi_{k-1} \xi)\|_{0,K}^2 dt \right)^{1/2} \\ &\leq Ch^{r+1/2} \|u\|_{L^2(\mathbb{H}^{r+1})} \|\xi\|_{\text{cGP}}. \end{aligned}$$

Using the Cauchy–Schwarz’s inequality and the properties of the spatial interpolation, we obtain for the second term on the right-hand side of (4.19)

$$\begin{aligned} \left| \int_0^T ((\pi_{k-1}\xi) \operatorname{div} \mathbf{b}, \tilde{u} - j_h \tilde{u}) \, dt \right| &\leq C \int_0^T \sum_{K \in \mathcal{T}_h} \|\tilde{u} - j_h \tilde{u}\|_0 \sigma_0^{1/2} \|\pi_{k-1}\xi\|_0 \, dt \\ &\leq Ch^{r+1} \|u\|_{L^2(\mathbb{H}^{k+1})} \|\xi\|_{\text{cGP}}. \end{aligned}$$

Inserting these estimates into (4.18), we get

$$\left| \int_0^T (\mathbf{b} \cdot \nabla \eta, \pi_{k-1}\xi) \, dt \right| \leq C \left[ h^{r+1/2} \|u\|_{L^2(\mathbb{H}^{r+1})} + \tau^{k+1} \|u\|_{\mathbb{H}^{k+1}(\mathbb{H}^1)} \right] \|\xi\|_{\text{cGP}}.$$

Hence, we have for the second term on the right-hand side of (4.17) the following estimate

$$\left| \int_0^T a(\eta, \pi_{k-1}\xi) \, dt \right| \leq C \left[ (\varepsilon^{1/2} + h^{1/2}) h^r \|u\|_{L^2(\mathbb{H}^{r+1})} + \tau^{k+1} \|u\|_{\mathbb{H}^{k+1}(\mathbb{H}^1)} \right]. \tag{4.20}$$

We rewrite the third term in (4.17) by using the error splitting as

$$- \int_0^T S_h(\eta, \pi_{k-1}\xi) \, dt = \int_0^T S_h(u - \tilde{u}, \pi_{k-1}\xi) \, dt + \int_0^T S_h(\tilde{u} - j_h \tilde{u}, \pi_{k-1}\xi) \, dt. \tag{4.21}$$

Then, the stability of the fluctuation operator  $\kappa_K$ , the boundedness of  $\mu_K$ , and the error estimates of the interpolation in time give for the first term in (4.21)

$$\begin{aligned} \left| \int_0^T S_h(u - \tilde{u}, \pi_{k-1}\xi) \, dt \right| &\leq \sum_{n=1}^N \int_{I_n} S_h^{1/2}(u - \tilde{u}, u - \tilde{u}) S_h^{1/2}(\pi_{k-1}\xi, \pi_{k-1}\xi) \, dt \\ &\leq \sum_{n=1}^N \int_{I_n} \left( \sum_{K \in \mathcal{T}_h} \mu_K \|\kappa_K \nabla(u - \tilde{u})\|_{0,K}^2 \right)^{1/2} \|\pi_{k-1}\xi\| \, dt \\ &\leq C \tau^{k+1} \|u\|_{\mathbb{H}^{k+1}(\mathbb{H}^1)} \|\xi\|_{\text{cGP}}. \end{aligned}$$

Similarly, the  $L^2$ -stability of the fluctuation operator  $\kappa_K$ , the parameter choice  $\mu_K \sim h_K$ , and (3.1) result for the second term of (4.21) in

$$\begin{aligned} \left| \int_0^T S_h(\tilde{u} - j_h \tilde{u}, \pi_{k-1}\xi) \, dt \right| &\leq \int_0^T S_h^{1/2}(\tilde{u} - j_h \tilde{u}, \tilde{u} - j_h \tilde{u}) S_h^{1/2}(\pi_{k-1}\xi, \pi_{k-1}\xi) \, dt \\ &\leq \int_0^T \left( \sum_{K \in \mathcal{T}_h} \mu_K \|\kappa_K(\nabla(\tilde{u} - j_h \tilde{u}))\|_{0,K}^2 \right)^{1/2} \|\pi_{k-1}\xi\| \, dt \\ &\leq Ch^{r+1/2} \int_0^T \|\tilde{u}\|_{r+1} \|\pi_{k-1}\xi\| \, dt \leq Ch^{r+1/2} \|u\|_{L^2(\mathbb{H}^{r+1})} \|\xi\|_{\text{cGP}}. \end{aligned}$$

Hence, we get

$$\left| - \int_0^T S_h(\eta, \pi_{k-1}\xi) dt \right| \leq C(h^{r+1/2}\|u\|_{L^2(\mathbb{H}^{r+1})} + \tau^{k+1}\|u\|_{\mathbb{H}^{k+1}(\mathbb{H}^1)})\|\xi\|_{\text{cGP}}. \quad (4.22)$$

To estimate the last term in (4.17), we use the Cauchy–Schwarz’s inequality, the approximation properties of the fluctuation operator  $\kappa_K$ , and the parameter choice  $\mu_K \sim h_K$  to get

$$\begin{aligned} \left| \int_0^T S_h(u, \pi_{k-1}\xi) dt \right| &\leq \int_0^T S_h^{1/2}(u, u) S_h^{1/2}(\pi_{k-1}\xi, \pi_{k-1}\xi) dt \leq C \int_0^T \left( \sum_{K \in \mathcal{T}_h} \mu_K \|\kappa_K(\nabla u)\|_{0,K}^2 \right)^{1/2} \|\pi_{k-1}\xi\| dt \\ &\leq Ch^{r+1/2}\|u\|_{L^2(\mathbb{H}^{r+1})}\|\xi\|_{\text{cGP}}. \end{aligned} \quad (4.23)$$

Putting the above estimates into (4.17), we obtain

$$B_h(u_{h,\tau} - j_h \tilde{u}, \pi_{k-1}\xi) \leq C \left[ (\varepsilon^{1/2} + h^{1/2})h^r\|u\|_{\mathbb{H}^1(\mathbb{H}^{r+1})} + \tau^{k+1}\|u\|_{\mathbb{H}^{k+1}(\mathbb{H}^1)} \right] \|\xi\|_{\text{cGP}}.$$

Using (4.16), it follows that

$$\|u_{h,\tau} - j_h \tilde{u}\|_{\text{cGP}} \leq C \left[ (\varepsilon^{1/2} + h^{1/2})h^r\|u\|_{\mathbb{H}^1(\mathbb{H}^{r+1})} + \tau^{k+1}\|u\|_{\mathbb{H}^{k+1}(\mathbb{H}^1)} \right]. \quad (4.24)$$

Finally, the statement of this Theorem follows by applying the triangle inequality and the interpolation error estimates in space and time.  $\square$

The next theorem gives an  $L^2(L^2)$  error estimate for the error  $u - u_{h,\tau}$  which is optimal with respect to time.

**Theorem 4.4.** *Assume  $\mu_K \sim h_K$  for all  $K \in \mathcal{T}_h$ . Suppose A1 and A2. Let  $u_{h,\tau}$  and  $u$  be the solutions of the fully discrete problem (4.6) and the continuous problem (4.5), respectively. Moreover, let  $u \in \mathbb{H}^1(\mathbb{H}^{r+1}) \cap \mathbb{H}^{k+1}(\mathbb{H}^1)$ . Then, there exists a positive constant  $C$  independent of  $h$ ,  $\tau$ , and  $\varepsilon$ , such that the error estimate*

$$\left( \int_0^T \|u(t) - u_{h,\tau}(t)\|_0^2 dt \right)^{1/2} \leq CT^{1/2} \left[ \tau^{k+1}\|u\|_{\mathbb{H}^{k+1}(\mathbb{H}^1)} + (\varepsilon^{1/2} + h^{1/2})h^r\|u\|_{\mathbb{H}^1(\mathbb{H}^{r+1})} \right] \quad (4.25)$$

holds true.

*Proof.* Let  $e_\tau = u_{h,\tau} - j_h \tilde{u}$ . Applying the ideas leading to (4.24) in the proof of Theorem 4.3 not only on  $[0, T]$  but also on  $[0, t_n]$ ,  $n = 1, \dots, N - 1$ , results in the estimate

$$\int_0^{t_n} \|\pi_{k-1}e_\tau(t)\|^2 dt + \frac{1}{2}\|e_\tau(t_n)\|_0^2 \leq C \left[ \tau^{2(k+1)}\|u\|_{\mathbb{H}^{k+1}(\mathbb{H}^1)}^2 + (\varepsilon + h)h^{2r}\|u\|_{\mathbb{H}^1(\mathbb{H}^{r+1})}^2 \right]$$

where the ranges of the time integrals on the right-hand side were extended from  $[0, t_n]$  to  $[0, T]$  due to the monotonicity of the integrals. After neglecting the non-negative integral on the left-hand side and multiplying by  $\tau_n$ , a summation over  $n = 1, \dots, N$  provides

$$\sum_{n=1}^N \tau_n \|e_\tau(t_n)\|_0^2 \leq C \left( \sum_{n=1}^N \tau_n \right) \left[ \tau^{2k+2}\|u\|_{\mathbb{H}^{k+1}(\mathbb{H}^1)}^2 + (\varepsilon + h)h^{2r}\|u\|_{\mathbb{H}^1(\mathbb{H}^{r+1})}^2 \right]. \quad (4.26)$$

Since  $e_\tau$  is a piece-wise polynomial of degree less than or equal to  $k$  in time, a norm equivalence on finite dimensional spaces gives

$$\int_{t_{n-1}}^{t_n} \|e_\tau(t)\|_0^2 dt \leq C_k \left( \int_{t_{n-1}}^{t_n} \|\pi_{k-1}e_\tau(t)\|_0^2 dt + \tau_n \|e_\tau(t_n)\|_0^2 \right) \tag{4.27}$$

where  $C_k$  depends on the polynomial degree  $k$  but is independent on  $\tau_n$  and  $h$ . Hence, we obtain

$$\begin{aligned} \int_0^T \|u_{h,\tau}(t) - j_h \tilde{u}(t)\|_0^2 dt &= \sum_{n=1}^N \int_{t_{n-1}}^{t_n} \|e_\tau(t)\|_0^2 dt \leq C_k \sum_{n=1}^N \left( \int_{t_{n-1}}^{t_n} \|\pi_{k-1}e_\tau(t)\|_0^2 dt + \tau_n \|e_\tau(t_n)\|_0^2 \right) \\ &\leq C_k \left( \int_0^T \|\pi_{k-1}e_\tau(t)\|_0^2 dt + \sum_{n=1}^N \tau_n \|e_\tau(t_n)\|_0^2 \right) \\ &\leq CT \left[ \tau^{2k+2} \|u\|_{\mathbf{H}^{k+1}(\mathbf{H}^1)}^2 + (\varepsilon + h)h^{2r} \|u\|_{\mathbf{H}^1(\mathbf{H}^{r+1})}^2 \right]. \end{aligned}$$

Here, the bound for the two terms derived from Thm. 4.3, (4.26), and the fact the time step lengths  $\tau_n$  sum up to  $T$  were used. Hence, the statement of the theorem follows by applying the triangle inequality and the interpolation error estimates.  $\square$

### 4.2. The discontinuous Galerkin (dG) method

We describe in this subsection briefly the discontinuous Galerkin (dG) time discretization in combination with the LPS method in space. Let us introduce the bilinear form  $B_h^{\text{dG}}$  by

$$B_h^{\text{dG}}(u, v) := \sum_{n=1}^N \int_{I_n} \left\{ (u'(t), v(t)) + a_h(u(t), v(t)) \right\} dt + \sum_{n=1}^{N-1} ([u]_n, v_n^+) + (u_0^+, v_0^+) \tag{4.28}$$

where the jumps  $[\varphi]_n$  and the right-sided values  $\varphi_n^+$  are defined in (4.1). The linear form  $l^{\text{dG}}$  is given by

$$l^{\text{dG}}(v) := (u_0, v_0^+) + \int_0^T (f(t), v(t)) dt. \tag{4.29}$$

The fully discrete scheme reads:

Find  $u_{h,\tau} \in Y_h^k$  such that

$$B_h^{\text{dG}}(u_{h,\tau}, v_{h,\tau}) = l^{\text{dG}}(v_{h,\tau}) \quad \forall v_{h,\tau} \in Y_h^k. \tag{4.30}$$

The mesh-dependent norm is defined by

$$\|v\|_{\text{dG}} := \left( \sum_{n=1}^N \int_{I_n} \|v(t)\|^2 dt + \frac{1}{2} \|v_0^+\|_0^2 + \frac{1}{2} \sum_{n=1}^{N-1} \|[v]_n\|_0^2 + \frac{1}{2} \|v_N^-\|_0^2 \right)^{1/2}. \tag{4.31}$$

The next theorem cites the main results from [2] where the local projection stabilization method in space has been combined with the discontinuous Galerkin method in time.

**Theorem 4.5.** *Suppose A1, A2, and  $\mu_K \sim h_K$  for all  $K \in \mathcal{T}_h$ . Let  $u_{h,\tau}$  and  $u$  be the solutions of the fully discrete scheme (4.30) and the continuous problem (2.4), respectively. Moreover, let  $u \in \mathbf{H}^1(\mathbf{H}^{r+1}) \cap \mathbf{H}^{k+1}(\mathbf{H}^1)$ . Then, there exists a positive constant  $C$  independent of  $h, \tau$ , and  $\varepsilon$ , such that the error estimate*

$$\|u_{h,\tau} - u\|_{\text{dG}} \leq C\tau^{k+1/2} \|u\|_{\mathbf{H}^{k+1}(\mathbf{H}^1)} + C(\varepsilon^{1/2} + h^{1/2})h^r \left( \|u\|_{\mathbf{H}^1(\mathbf{H}^{r+1})} + \|u\|_{\mathbf{C}(\mathbf{H}^{k+1})} \right) \tag{4.32}$$

holds true.

## 5. NUMERICAL STUDIES

We present in this section some numerical experiments to assess accuracy and performance of local projection stabilization techniques in space combined with higher order variational time discretization schemes. All computations have been performed using MooNMD [21].

In our numerical computations, we have used mapped finite element spaces [10] where the enriched spaces on the reference cell  $\widehat{K}$  are given by

$$\begin{aligned}\mathbb{P}_r^{\text{bubble}}(\widehat{K}) &:= \mathbb{P}_r(\widehat{K}) + \hat{b}_\Delta \mathbb{P}_{r-1}(\widehat{K}), & r \geq 1, \\ \mathbb{Q}_r^{\text{bubble}}(\widehat{K}) &:= \mathbb{Q}_r(\widehat{K}) + \text{span} \left\{ \hat{b}_\square \hat{x}_i^{r-1}, i = 1, 2 \right\}, & r \geq 1.\end{aligned}$$

Here,  $\hat{b}_\Delta$  and  $\hat{b}_\square$  are the cubic bubble function on the reference triangle and the biquadratic bubble function on the reference square, respectively. The mapped finite element spaces based on  $\mathbb{P}_r^{\text{bubble}}(\widehat{K})$  and  $\mathbb{Q}_r^{\text{bubble}}(\widehat{K})$  will be shortly denoted by  $\mathbb{P}_r^{\text{bubble}}$  and  $\mathbb{Q}_r^{\text{bubble}}$ , respectively.

The combinations  $V_h = \mathbb{P}_r^{\text{bubble}}$ ,  $D(K) = \mathbb{P}_{r-1}(K)$ ,  $r \geq 1$ , on triangular meshes and the combinations  $V_h = \mathbb{Q}_r^{\text{bubble}}$ ,  $D(K) = \mathbb{P}_{r-1}(K)$ ,  $r \geq 1$ , on quadrilateral meshes fulfill the assumptions A1 and A2. Further examples of approximation spaces  $V_h$  and projection spaces  $D(K)$  satisfying A1 and A2 are given in [28, 31].

The stabilization parameters  $\mu_K$  have been chosen as

$$\mu_K = \mu_0 h_K \quad \forall K \in \mathcal{T}_h$$

where  $\mu_0$  denotes a positive constant which will be given for each of the test calculations.

We define by

$$\|v\|_{L^2(L^2)} := \left\{ \int_0^T \|v(t)\|_0^2 dt \right\}^{1/2}, \quad \|v\|_\infty := \max_{1 \leq n \leq N} \|v(t_n)\|_0$$

the  $L^2(L^2)$ -norm and the discrete  $\ell^\infty(L^2)$ -norm in time.

**Example 1.** The rotating Gaussian benchmark as a pure transport problem in two space dimensions is taken from [2, 9]. Hence, we consider the problem (2.1) in

$$\Omega = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 < 1\}$$

with the data

$$\varepsilon = 0, \quad b = (-y, x)^T, \quad \sigma = f = 0, \quad T = 2\pi,$$

and the Gaussian initial condition

$$u_0(x, y) = \exp \left( -10(x - 0.3)^2 - 10(y - 0.3)^2 \right)$$

which is centered at (0.3, 0.3). Note that the solution  $u$  is at all times given by a suitably rotated initial condition.

The calculations have been performed on triangular meshes which were obtained from an initial triangulation by successive refinement with boundary adaption due to the curved boundary. The initial mesh (level 0) and the mesh on level 3 are shown in Figure 1.

In this example, we are interested in the convergence order in space. To this end, we keep the error in time small by using the methods cGP(3) and dG(2) with time step length  $\tau = 2\pi \times 10^{-3}$ . Hence, 1000 time steps correspond to one complete revolution of the Gaussian. We used the LPS discretizations with  $V_h = \mathbb{P}_3^{\text{bubble}}$ ,  $D(K) = \mathbb{P}_2(K)$ , and the stabilization parameters  $\mu_K = 0.1h_K$ . Table 1 shows for the time discretizations

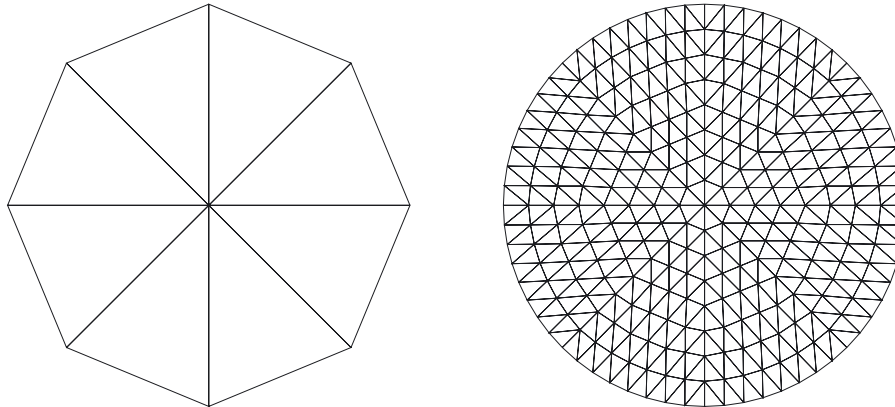


FIGURE 1. Triangular meshes for Example 1: coarsest mesh (*left*) and mesh after three refinement steps (*right*).

TABLE 1. Example 1: errors and convergence orders in space for cGP(3) and dG(2).

Level	$\ u - u_{h,\tau}\ _{\text{cGP}}$		$\ u - u_{h,\tau}\ _{\text{dG}}$	
	Error	Order	Error	Order
2	7.892-2		7.900-2	
3	8.234-3	3.26	8.238-3	3.26
4	7.001-4	3.56	7.003-4	3.56
5	6.192-5	3.50	6.192-5	3.50
6	5.471-6	3.50	5.473-6	3.50
Theory		3.5		3.5

cGP(3) and dG(2) the errors and convergence orders in the cGP-norm and the dG-norm. It is clearly to see that the convergence orders predicted by Theorem 4.3 are achieved. Moreover, the differences between the two time discretization schemes are very small. This is due to the fact that the temporal error is almost eliminated by the used higher order time discretizations with small time step length.

**Example 2.** In order to assess the effect of the applied time discretizations, we exclude in this example the spatial error. We consider problem (2.1) in  $\Omega = (0, 1)^2$  with  $\varepsilon = 10^{-8}$ ,  $\mathbf{b} = (1, 2)$ ,  $\sigma = 1$ , and  $T = 1$ . The right-hand side  $f$  and the initial condition  $u_0$  are chosen such that

$$u(x, y, t) = 1 + 2x + 3t^{200}y$$

is the solution of (2.1) equipped with non-homogeneous Dirichlet boundary conditions.

We used the LPS discretizations with  $V_h = \mathbb{Q}_1^{\text{bubble}}$ ,  $D(K) = \mathbb{P}_0(K)$ , and stabilization parameter  $\mu_K = 0.1h_K$ . A uniform mesh consisting of  $4 \times 4$  squares was considered in our calculations. Note that for any time  $t$  the solution  $u$  can be represented exactly by a function from the finite element space  $V_h$ . Hence, all occurring errors will result from the temporal discretizations. The higher order time discretization methods cGP(3), cGP(4) and dG(2), dG(3) were applied.

We report in Table 2 and 3 the errors and convergence orders for these methods. We see that all considered methods are accurate of order  $k + 1$  in the  $L^2(L^2)$ -norm as predicted by Theorem 4.4. The orders  $k + 1$  and  $k + 1/2$  in the cGP-norm and the dG-norm are achieved which also confirms the theoretical results. The values in the discrete  $\ell^\infty(L^2)$ -norms show clearly that the cGP( $k$ ) methods,  $k = 3, 4$ , are super-convergent of order  $2k$



TABLE 2. Example 2: errors and convergence orders for cGP(3) and dG(2).

$\tau$	$\ u - u_{h,\tau}\ _{L^2(L^2)}$				$\ u - u_{h,\tau}\ _\infty$				$\ u - u_{h,\tau}\ _{cGP}$		$\ u - u_{h,\tau}\ _{dG}$	
	cGP(3)		dG(2)		cGP(3)		dG(2)		cGP(3)		dG(2)	
	Error	Order	Error	Order	Error	Order	Error	Order	Error	Order	Error	Order
1/20	5.028-2		1.415-1		1.655-1		4.296-1		6.766-1		1.189	
1/40	8.130-3	2.63	3.453-2	2.04	8.821-3	4.23	4.047-2	3.40	1.164-1	2.53	4.565-1	1.38
1/80	8.336-4	3.28	6.038-3	2.52	2.230-4	1.82	1.950-3	0.87	1.282-2	3.18	1.227-1	1.90
1/160	6.340-5	3.71	8.364-4	2.85	1.350-5	7.53	2.325-4	6.57	1.003-3	3.68	2.497-2	2.29
1/320	4.205-6	3.91	1.055-4	2.99	6.531-8	7.69	2.215-6	6.71	6.704-5	3.90	4.517-3	2.47
1/640	2.670-7	3.97	1.305-5	3.00	1.030-9	5.99	6.958-8	4.99	4.265-6	3.97	7.945-4	2.50
Theory		4		3		6		5		4		2.5

TABLE 3. Example 2: errors and convergence orders for cGP(4) and dG(3).

$\tau$	$\ u - u_{h,\tau}\ _{L^2(L^2)}$				$\ u - u_{h,\tau}\ _\infty$				$\ u - u_{h,\tau}\ _{cGP}$		$\ u - u_{h,\tau}\ _{dG}$	
	cGP(4)		dG(3)		cGP(4)		dG(3)		cGP(4)		dG(3)	
	Error	Order	Error	Order	Error	Order	Error	Order	Error	Order	Error	Order
1/20	1.682-2		5.201-2		1.510-2		4.925-2		3.103-1		5.381-1	
1/40	1.659-3	3.34	8.285-3	2.65	2.247-4	6.07	1.368-3	5.17	3.276-2	3.24	1.339-1	2.00
1/80	9.414-5	4.14	8.293-4	3.32	1.599-5	3.81	1.916-4	2.84	1.934-3	4.08	2.005-2	2.74
1/160	3.704-6	4.67	6.088-5	3.77	2.201-8	9.50	5.182-6	8.53	7.705-5	4.65	2.130-3	3.23
1/320	1.240-7	4.90	3.930-6	3.95	2.683-11	9.67	1.257-9	8.69	2.587-6	4.90	1.961-4	3.44
1/640	3.945-9	4.97	2.456-7	4.00	2.295-13	6.52	9.657-11	7.02	8.240-8	4.97	1.738-5	3.49
Theory		5		4		8		7		5		3.5

TABLE 4. Example 2: errors and convergence orders of post-processed solution for cGP(3), cGP(4), dG(2), and dG(3).

$\tau$	$\ u - \Pi u_{h,\tau}\ _{L^2(L^2)}$				$\ u - \Pi u_{h,\tau}\ _{L^2(L^2)}$			
	cGP(3)		cGP(4)		dG(2)		dG(3)	
	Error	Order	Error	Order	Error	Order	Error	Order
1/20	3.834-2		1.200-2		1.049-1		3.675-2	
1/40	4.566-3	3.07	7.620-4	3.98	1.826-2	2.52	3.706-3	3.31
1/80	2.786-4	4.04	2.363-5	5.00	2.936-3	3.24	2.063-4	4.17
1/160	1.115-5	4.64	4.750-7	5.64	1.451-4	3.73	7.886-6	4.71
1/320	3.740-7	4.90	8.978-9	5.90	9.435-6	3.94	2.586-7	4.93
1/640	1.190-8	4.97	1.270-10	5.97	6.911-7	4.00	8.130-9	4.99
Theory		5		6		4		5

while the dG(k) methods,  $k = 2, 3$ , are super-convergent of order  $2k + 1$ . These superconvergence results are known for the heat equation, see [3, 35].

Table 4 shows the error  $e = u - \Pi u_{h,\tau}$  in the  $L^2(L^2)$ -norm where  $\Pi u_{h,\tau}$  denotes the post-processed solution which is obtained by means of a simple post-processing from the solution  $u_{h,\tau}$  of the original cGP method or dG method. Details of the post-processing can be found in [1, 27]. Note that the post-processed solutions of both cGP(k) and dG(k) are super-convergent of order  $k + 2$  in the  $L^2(L^2)$ -norm. This is confirmed by the error norms in Table 4.

TABLE 5. Example 3: errors and variation of solutions for cGP(2), cGP(3), dG(2), and dG(3).

$\mu_0$	$\ u - u_{h,\tau}\ _{L^2(V_h)}$				Var(0.5)			
	cGP(2)	cGP(3)	dG(2)	dG(3)	cGP(2)	cGP(3)	dG(2)	dG(3)
0.01	7.3575-3	5.5714-3	7.3575-3	5.5714-3	3.3597-2	4.3397-2	3.9573-2	4.3397-2
0.1	4.4980-3	3.2775-3	4.9799-3	1.8913-3	1.8605-2	1.7010-2	1.8605-2	1.7010-2
1.0	6.7158-3	4.7016-3	6.7158-3	4.7016-3	2.7546-2	2.9773-2	2.7546-2	2.9773-2

**Example 3.** This example is taken from [23]. The prescribed solution has the form

$$u(x, y, t) = 16 \sin(\pi t)x(1 - x)y(1 - y) \left( \frac{1}{2} + \frac{\arctan [2\varepsilon^{-1/2}w(x, y)]}{\pi} \right), \tag{5.1}$$

where

$$w(x, y) := 0.25^2 - (x - 0.5)^2 - (y - 0.5)^2.$$

This is a hump which changes its height in the course of time. The steepness of the circular internal layer depends on the diffusion parameter  $\varepsilon$ . Analogue to [23], we present the simulation for  $\Omega = (0, 1)^2$ ,  $\varepsilon = 10^{-6}$ ,  $\mathbf{b} = (2, 3)$ ,  $\sigma = 1$ , and  $T = 2$ . We use

$$\text{var}(t) := \max_{(x,y) \in \Omega} u_{h,\tau}(x, y, t) - \min_{(x,y) \in \Omega} u_{h,\tau}(x, y, t)$$

as a measure for undershoots and overshoots. For calculating the minimal and maximal values of  $u_{h,\tau}$ , we used only the values at the vertices of the underlying mesh. For evaluating the size of the spurious oscillations, we assess the variation of the discrete solutions  $u_{h,\tau}$  at  $t = 0.5$  where the hump reaches its maximal height. The exact value is  $\text{var}(0.5) = 0.997453575$ . Furthermore, some indication of the smearing of the solution is derived from

$$\|e\|_{L^2(V_h)} := \left( \int_0^T (\|e(t)\|_0^2 + \varepsilon \|e(t)\|_1^2) dt \right)^{1/2}, \quad e = u - u_{h,\tau}.$$

The time step length was chosen to be  $\tau = 10^{-3}$ . The computations were performed on a uniform grid consisting of  $128 \times 128$  squares. This leads to 98,817 degrees of freedom for the spatial discretization with  $V_h = \mathbb{Q}_2^{\text{bubble}}$ ,  $D(K) = \mathbb{P}_1(K)$  and 180,993 degrees of freedom for  $V_h = \mathbb{Q}_3^{\text{bubble}}$ ,  $D(K) = \mathbb{P}_2(K)$  including the Dirichlet degrees of freedom. Note that in the following figures only the nodal values at the mesh vertices are shown, *i.e.*, the additional bubble part of the solution will not be presented. We have performed calculations with  $\mu_0 \in \{0.01, 0.1, 1.0\}$ .

The results for cGP(2) with  $V_h = \mathbb{Q}_2^{\text{bubble}}$ ,  $D(K) = \mathbb{P}_1(K)$  and cGP(3) with  $V_h = \mathbb{Q}_3^{\text{bubble}}$ ,  $D(K) = \mathbb{P}_2(K)$  are presented in Table 5 and in Figure 2. It can be observed that small and large coefficients  $\mu_0$  inside the definition  $\mu_K = \mu_0 h_K$  lead to spurious oscillations behind the hump in the direction of convection. The solution obtained by cGP(2) with  $\mu_0 = 0.1$  possesses still small overshoots and undershoots whereas the numerical solution obtained by cGP(3) captured very well the solution. Furthermore, almost no oscillations are present.

We visualize in Figure 3 the computed profiles of the solution for dG(2) with  $V_h = \mathbb{Q}_2^{\text{bubble}}$ ,  $D(K) = \mathbb{P}_1(K)$  and dG(3) with  $V_h = \mathbb{Q}_3^{\text{bubble}}$ ,  $D(K) = \mathbb{P}_2(K)$ . The coefficient  $\mu_0$  took the values 0.01, 0.1, and 1. It can be noted that there are no significant differences between the solution obtained by dG( $k$ ) and cGP( $k$ ), see Table 5. Furthermore, the spurious oscillations vanish almost completely also for dG(3).

## 6. CONCLUSIONS

We have combined higher order variational time discretization by continuous Galerkin–Petrov’s methods and discontinuous Galerkin’s methods with the one-level local projection stabilization technique in space. Error

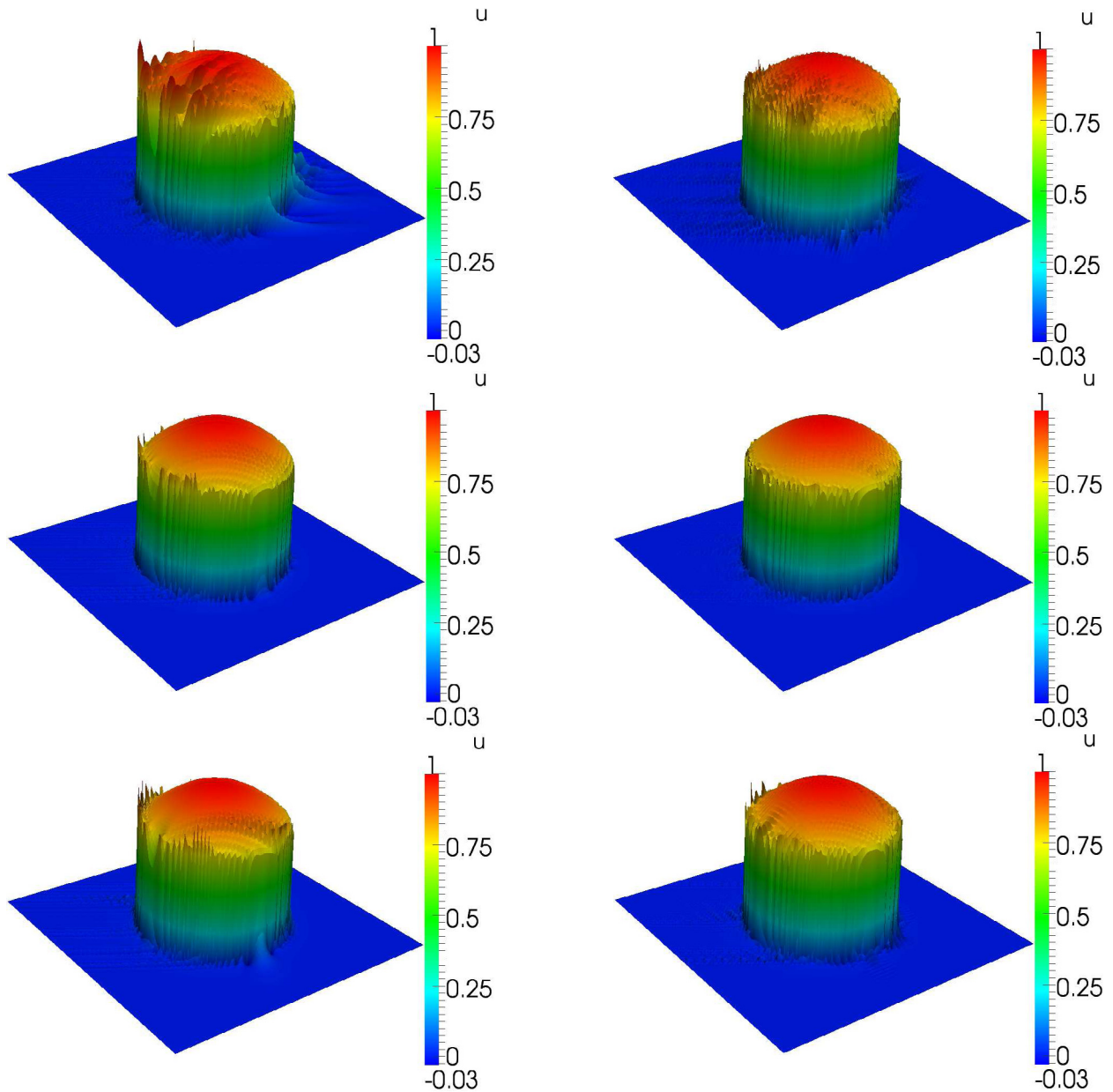


FIGURE 2. Example 3: numerical solution at  $t = 0.5$ , left: cGP(2) with  $V_h = \mathbb{Q}_2^{\text{bubble}}$ ,  $D(K) = \mathbb{P}_1(K)$ , right: cGP(3) with  $\mathbb{Q}_3^{\text{bubble}}$ ,  $D(K) = \mathbb{P}_2(K)$ ; top to bottom:  $\mu_0 = 0.01$ ,  $\mu_0 = 0.1$ ,  $\mu_0 = 1.0$ .

estimates for the cGP( $k$ ) method are given and proved. The numerical experiments support the theoretical results. Both time discretization schemes allow to get highly accurate numerical solution. Due to the applied spatial stabilization, the remaining oscillations are located near to layers. Moreover, higher order approximations in time provide solutions with less oscillations. Furthermore, the differences between the solutions obtained

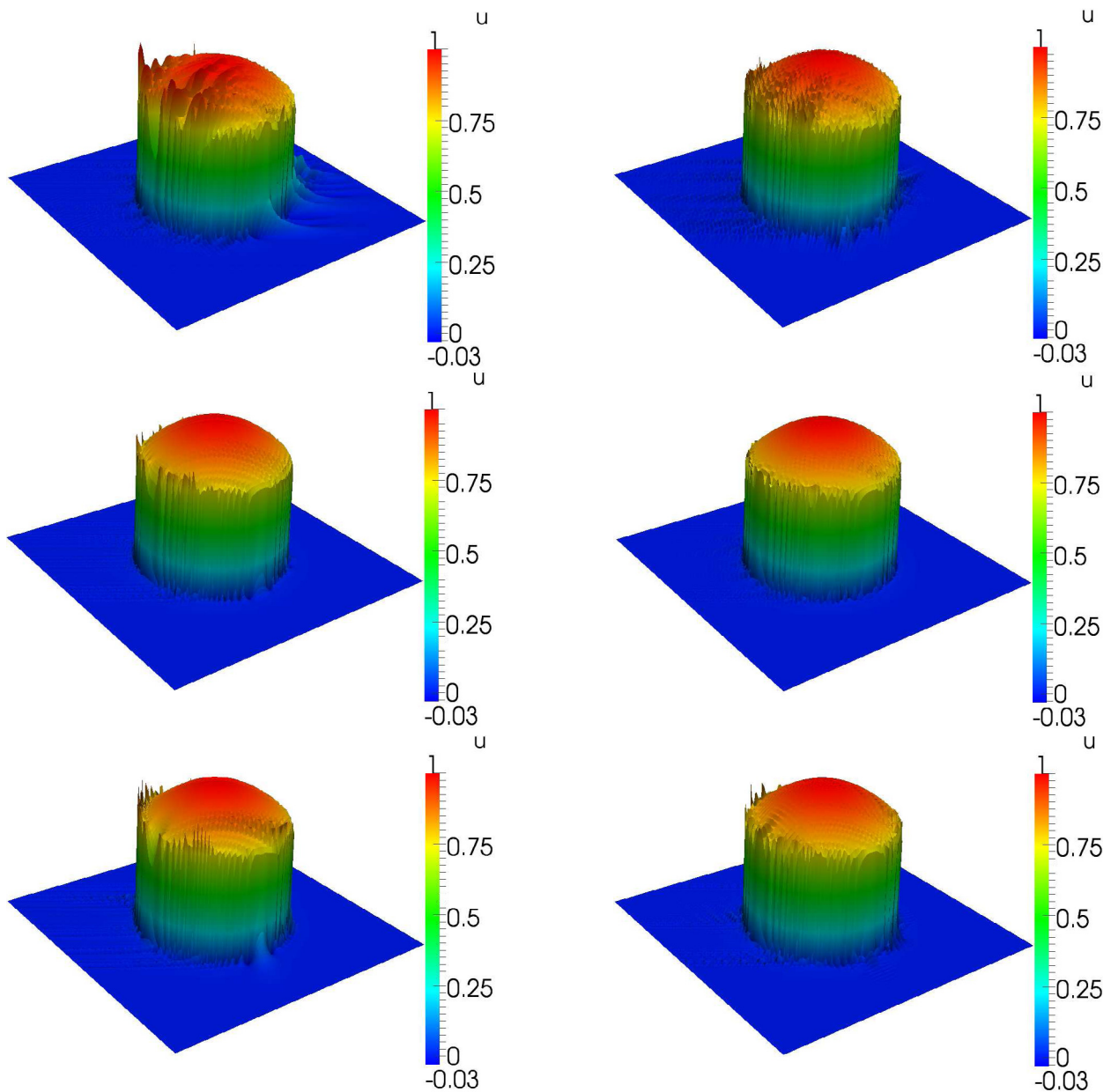


FIGURE 3. Example 3: numerical solution at  $t = 0.5$ , left: dG(2) with  $V_h = \mathbb{Q}_2^{\text{bubble}}$ ,  $D(K) = \mathbb{P}_1(K)$ , right: dG(3) with  $\mathbb{Q}_3^{\text{bubble}}$ ,  $D(K) = \mathbb{P}_2(K)$ ; top to bottom:  $\mu_0 = 0.01$ ,  $\mu_0 = 0.1$ ,  $\mu_0 = 1.0$ .

by continuous Galerkin–Petrov’s methods and discontinuous Galerkin’s methods were small in all considered examples.

*Acknowledgements.* We would like to thank Julia Novo (Universidad Autónoma de Madrid) for helpful discussions in the proof of Theorem 4.4.

## REFERENCES

- [1] N. Ahmed and G. Matthies, Numerical study of SUPG and LPS methods combined with higher order variational time discretization schemes applied to time-dependent convection-diffusion-reaction equations. Weierstraß-Institut für Angewandte Analysis und Stochastik, Berlin (2014). Preprint 1948.
- [2] 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. Engrg.* **200** (2011) 1747–1756.
- [3] A.K. Aziz and P. Monk, Continuous finite elements in space and time for the heat equation. *Math. Comput.* **52** (1989) 255–274.
- [4] R. Becker and M. Braack, A finite element pressure gradient stabilization for the Stokes equations based on local projections. *Calcolo* **38** (2001) 173–199.
- [5] R. Becker and M. Braack, A two-level stabilization scheme for the Navier-Stokes equations. In *Numer. Math. Advanced Appl.* Springer, Berlin (2004) 123–130.
- [6] P.B. Bochev, M.D. Gunzburger and J.N. Shadid, Stability of the SUPG finite element method for transient advection-diffusion problems. *Comput. Methods Appl. Mech. Engrg.* **193** (2004) 2301–2323.
- [7] M. Braack and E. Burman, Local projection stabilization for the Oseen problem and its interpretation as a variational multiscale method. *SIAM J. Numer. Anal.* **43** (2006) 2544–2566.
- [8] E. Burman, Consistent SUPG-method for transient transport problems: stability and convergence. *Comput. Methods Appl. Mech. Engrg.* **199** (2010) 1114–1123.
- [9] E. Burman and M.A. Fernández, Finite element methods with symmetric stabilization for the transient convection-diffusion-reaction equation. *Comput. Methods Appl. Mech. Engrg.* **198** (2009) 2508–2519.
- [10] P.G. Ciarlet, The finite element method for elliptic problems. Vol. 4 of *Stud. Math. Appl.* North-Holland Publishing Co., Amsterdam (1978).
- [11] R. Codina, Comparison of some finite element methods for solving the diffusion-convection-reaction equation. *Comput. Methods Appl. Mech. Engrg.* **156** (1998) 185–210.
- [12] K. Eriksson, D. Estep, P. Hansbo and C. Johnson, Computational differential equations. Cambridge University Press, Cambridge (1996).
- [13] K. Eriksson, C. Johnson and V. Thomée, Time discretization of parabolic problems by the discontinuous Galerkin method. *RAIRO: M2AN* **19** (1985) 611–643.
- [14] M.-C. Hsu, Y. Bazilevs, V.M. Calo, T.E. Tezduyar and T.J.R. Hughes, Improving stability of stabilized and multiscale formulations in flow simulations at small time steps. *Comput. Methods Appl. Mech. Engrg.* **199** (2010) 828–840.
- [15] T.J.R. Hughes and A.N. Brooks, A multidimensional upwind scheme with no crosswind diffusion. In Finite element methods for convection dominated flows (Papers, Winter Ann. Meeting Amer. Soc. Mech. Engrs., New York, 1979). Vol. 34 of *AMD. Amer. Soc. Mech. Engrs.* (ASME). New York (1979) 19–35.
- [16] 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.
- [17] S. Hussain, F. Schieweck and S. Turek, A note on accurate and efficient higher order Galerkin time stepping schemes for the nonstationary Stokes equations. *Open Numer. Methods J.* **4** (2012) 35–45.
- [18] V. John and P. Knobloch, On spurious oscillations at layers diminishing (SOLD) methods for convection-diffusion equations. I. A review. *Comput. Methods Appl. Mech. Engrg.* **196** (2007) 2197–2215.
- [19] V. John and P. Knobloch, On the performance of SOLD methods for convection-diffusion problems with interior layers. *Int. J. Comput. Sci. Math.* **1** (2007) 245–258.
- [20] V. John and P. Knobloch, On spurious oscillations at layers diminishing (SOLD) methods for convection-diffusion equations. II. Analysis for  $P_1$  and  $Q_1$  finite elements. *Comput. Methods Appl. Mech. Engrg.* **197** (2008) 1997–2014.
- [21] V. John and G. Matthies, MoonNMD – a program package based on mapped finite element methods. *Comput. Vis. Sci.* **6** (2004) 163–169.
- [22] V. John and J. Novo, Error analysis of the SUPG finite element discretization of evolutionary convection-diffusion-reaction equations. *SIAM J. Numer. Anal.* **49** (2011) 1149–1176.
- [23] V. John and E. Schmeier, Finite element methods for time-dependent convection-diffusion-reaction equations with small diffusion. *Comput. Methods Appl. Mech. Engrg.* **198** (2008) 475–494.
- [24] P. Knobloch, On the application of local projection methods to convection-diffusion-reaction problems. In BAIL 2008 – boundary and interior layers. Vol. 69 of *Lect. Notes Comput. Sci. Eng.* Springer, Berlin (2009) 183–194.
- [25] P. Knobloch, A generalization of the local projection stabilization for convection-diffusion-reaction equations. *SIAM J. Numer. Anal.* **48** (2010) 659–680.
- [26] G. Lube and D. Weiss, Stabilized finite element methods for singularly perturbed parabolic problems. *Appl. Numer. Math.* **17** (1995) 431–459.
- [27] G. Matthies and F. Schieweck, Higher order variational time discretizations for nonlinear systems of ordinary differential equations. Fakultät für Mathematik, Otto-von-Guericke-Universität Magdeburg (2011). Preprint 23/2011.
- [28] G. Matthies, P. Skrzypacz and L. Tobiska, A unified convergence analysis for local projection stabilisations applied to the Oseen problem. *ESAIM: M2AN* **41** (2007) 713–742.
- [29] G. Matthies, P. Skrzypacz and L. Tobiska, Stabilization of local projection type applied to convection-diffusion problems with mixed boundary conditions. *Electron. Trans. Numer. Anal.* **32** (2008) 90–105.
- [30] W.H. Reed and T.R. Hill, *Triangular mesh methods for the neutron transport equation*. Tech. Report LA-UR-73-479, Los Alamos Scientific Laboratory (1973).

- [31] H.-G. Roos, M. Stynes and L. Tobiska, Robust numerical methods for singularly perturbed differential equations. Convection-diffusion-reaction and flow problems. In vol. 24 of *Springer Ser. Comput. Math.* Springer-Verlag, Berlin, 2nd edition (2008).
- [32] F. Schieweck.  $A$ -stable discontinuous Galerkin-Petrov time discretization of higher order. *J. Numer. Math.* **18** (2010) 25–57.
- [33] D. Schötzau and C. Schwab, An  $hp$  a priori error analysis of the DG time-stepping method for initial value problems. *Calcolo*, **37** (2000) 207–232.
- [34] D. Schötzau and C. Schwab,  $hp$ -discontinuous Galerkin time-stepping for parabolic problems. *C. R. Acad. Sci. Paris Sér. I Math.* **333** (2001) 1121–1126.
- [35] V. Thomée, Galerkin finite element methods for parabolic problems. In vol. 25 of *Springer Ser. Comput. Math.* Springer-Verlag, Berlin (1997).