IMA Journal of Numerical Analysis (2019) **00**, 1–35
doi:10.1093/imanum/drz033

On the standard Galerkin method with explicit RK4 time stepping for the shallow water equations

D. C. Antonopoulos, V. A. Dougalis* and G. Kounadis
Department of Mathematics, University of Athens, 15784 Zographou, Greece, and Institute of Applied and Computational Mathematics, FORTH, 70013 Heraklion, Greece
*Corresponding author: doug@math.uoa.gr

[Received on 3 January 2019; revised on 3 January 2019]

We consider a simple initial-boundary-value problem for the shallow water equations in one space dimension. We discretize the problem in space by the standard Galerkin finite element method on a quasiuniform mesh and in time by the classical four-stage, fourth order, explicit Runge–Kutta scheme. Assuming smoothness of solutions, a Courant number restriction and certain hypotheses on the finite element spaces, we prove L^2 error estimates that are of fourth-order accuracy in the temporal variable and of the usual, due to the nonuniform mesh, suboptimal order in space. We also make a computational study of the numerical spatial and temporal orders of convergence, and of the validity of a hypothesis made on the finite element spaces.

Keywords: shallow water equations; standard Galerkin finite element method; fourth order four-stage explicit Runge–Kutta method; error estimates.

1. Introduction

In this paper we will consider the following initial-boundary-value problem (ibvp) for the *shallow water equations* posed on the spatial interval $[0, 1]$. For $T > 0$ we seek $\eta = \eta(x, t)$, $u = u(x, t)$, $0 \leqslant x \leqslant 1$, $0 \leqslant t \leqslant T$, such that

$$\begin{aligned} &\left.\begin{aligned} \eta_t + u_x + (\eta u)_x &= 0, \\ u_t + \eta_x + uu_x &= 0, \end{aligned}\right\} \quad 0 \le x \le 1, \ \ 0 \le t \le T, \\ &\eta(x,0) = \eta_0(x), \quad u(x,0) = u_0(x), \quad 0 \leqslant x \leqslant 1, \\ &u(0,t) = 0, \qquad u(1,t) = 0, \qquad 0 \leqslant t \leqslant T, \end{aligned} \tag{SW}$$

where η_0, u_0 are given real-valued functions defined on $[0, 1]$. The shallow water equations approximate the Euler equations of water wave theory in the case of long waves in a channel of finite depth. In (SW) the variables are nondimensional and unscaled; $x \in [0, 1]$ and $t \geqslant 0$ are proportional to position along the finite channel $[0, 1]$ and time, respectively, $\eta = \eta(x, t)$ is proportional to the elevation of the free surface above a level of rest corresponding to $\eta = 0$ and $u = u(x, t)$ is proportional to the depth-averaged horizontal velocity of the fluid. In these variables the (horizontal) bottom of the channel is at a depth equal to -1.

Even if the initial conditions η_0 and u_0 are smooth, (SW) is not expected to have global smooth solutions. There is however a local H^2 well-posedness theory; in Petcu & Temam (2011) it was proved that if $u_0 \in H^2 \cap \mathring{H}^1$, $\eta_0 \in H^2$, with $1 + \eta_0 \geqslant 2\alpha > 0$ on $[0, 1]$ for some constant α, then there exists a $T = T(\|u_0\|_2, \|\eta_0\|_2) > 0$ and a unique solution $(\eta, u) \in L^\infty(0, T; H^2 \times (H^2 \cap \mathring{H}^1))$ of (SW) such that $1 + \eta \geqslant \alpha > 0$ on $[0, 1] \times [0, T]$. (Here, and in the sequel, for integer $m \geqslant 0$ H^m, $\|\cdot\|_m$, denote the L^2-based Sobolev space of functions on $[0, 1]$ and its associated norm, and $\mathring{H}^1$ the subspace of H^1 whose elements are equal to zero at $x = 0, 1$.) For a Banach space X of functions on $[0, 1]$, $L^\infty(0, T; X)$ is the space of L^∞ maps from $[0, T]$ to X.

In the paper at hand we will approximate the solution of (SW) by a fully discrete scheme using the standard Galerkin finite element method for the discretization in space with suitable finite element spaces, whose elements are at least continuously differentiable on $[0, 1]$ and are piecewise polynomial functions of degree $r - 1$, $r \geqslant 3$, with respect to a quasiuniform partition of $[0, 1]$ of maximum meshlength h. Precise assumptions about the finite element spaces will be stated in Section 2. For the temporal integration we will use the classical, four-stage, fourth-order, explicit Runge–Kutta (RK) scheme with a uniform time step k. The main contribution of the paper is the proof that the fully discrete scheme has a temporal error of $O(k^4)$. Specifically, in Section 3 we analyze the spatial and temporal consistency and in Section 4 the convergence of the scheme, and show that if the solution of (SW) is sufficiently smooth and $1 + \eta$ is positive in $[0, 1] \times [0, T]$, there exists a positive constant λ_0 such that if the Courant number $\lambda = k/h$ satisfies $\lambda \leqslant \lambda_0$, then the L^2 norm of the error of the fully discrete approximation is of $O(k^4 + h^{r-1})$. (It is well known that the best order of spatial accuracy one may achieve for first-order hyperbolic problems using the standard Galerkin method on a nonuniform mesh is $r - 1$ in general, cf. e.g. Dupont (1973), Dupont (1974), Antonopoulos & Dougalis (2016)). In Section 5 we make a computational study of the numerical spatial and temporal orders of convergence and of the validity of a certain hypothesis made on the finite element spaces.

Explicit RK methods of higher (at least third) order of accuracy have been widely used for the temporal discretization of ode systems obtained from spatial discretizations of first-order hyperbolic equations. Such ode systems are usually only mildly stiff and may be stably integrated with such explicit RK schemes under Courant-number restrictions. Regarding rigorous error estimates for fully discrete schemes of finite element-high order RK type we mention Zhang & Shu (2004), who prove error estimates for a fully discrete DG - 3$^{\text{d}}$ order Shu–Osher RK scheme, cf. Shu & Osher (1988), for scalar conservation laws. The same authors analyze in Zhang & Shu (2014) a similar fully discrete scheme applied to a scalar linear hyperbolic equation with discontinuous initial condition. In Burman *et al.* (2010) the authors consider ibvp's for first-order linear hyperbolic problems of Friedrichs type in several space dimensions, discretized in space by a class of symmetrically stabilized finite element methods that includes DG schemes, and in time by, among other, third-order accurate, explicit RK schemes, and prove L^2-error estimates of optimal order in time and quasioptimal $(r - 1/2)$ in space. In Antonopoulos & Dougalis (2016) two of the present authors proved, among others, $O(k^3 + h^{r-1})$ L^2-error estimates for (SW) discretized by the standard Galerkin method coupled with the Shu–Osher RK scheme. As far as we know, the paper at hand presents the first error estimate of optimal temporal accuracy for a fully discrete finite element scheme with fourth-order explicit RK time stepping for a nonlinear hyperbolic system. For practical issues regarding the application of DG-high order RK schemes to nonlinear hyperbolic systems including the shallow water equations, we refer the reader to the recent review papers Qiu & Zhang (2016) and Xing (2017), and to Kubatko *et al.* (2014) and its references on the strong stability of higher order RK schemes.

In addition to previously introduced notation, in the sequel we let $C^m = C^m[0, 1]$, $m = 0, 1, 2, \ldots$, be the space of m times continuously differentiable functions on $[0, 1]$. The inner product and norm on $L^2 = L^2(0, 1)$ will be denoted by $(\cdot, \cdot)$, $\|\cdot\|$, respectively, while the norms of $L^\infty = L^\infty(0, 1)$ and of the L^∞-based Sobolev space $W^{1,\infty} = W^{1,\infty}(0,1)$ by $\|\cdot\|_\infty$, $\|\cdot\|_{1,\infty}$. We let $\mathbb{P}_r$ be the polynomials of degree at most r.

2. Approximation properties of the finite element spaces and preliminaries

Let $0 = x_1 < x_2 < \cdots < x_{N+1} = 1$ be a quasiuniform partition of $[0,1]$ with $h := \max_i(x_{i+1} - x_i)$. For integers r, μ with $r \geqslant 3$ and $1 \leqslant \mu \leqslant r-2$, let $S_h = S_h^{r,\mu} = \{\phi \in C^\mu : \phi\big|_{[x_i,x_{i+1}]} \in \mathbb{P}_{r-1}\}$, and $S_{h,0} = \{\phi \in S_h, \phi(0) = \phi(1) = 0\}$. We will assume, cf. DeBoor & Fix (1973) and Schreiber (1980), that if $w \in H^s$, $2 \leqslant s \leqslant r$, there exists a $\chi \in S_h$, such that

$$\|w - \chi\| + h\|w' - \chi'\| \leqslant Ch^s\|w^{(s)}\|, \tag{2.1a}$$

and that if $w \in H^s$, $3 \leqslant s \leqslant r$, χ satisfies in addition

$$\|w - \chi\|_2 \leqslant Ch^{s-2}\|w^{(s)}\|, \tag{2.1b}$$

for some constant C independent of h and w. We will also assume that similar properties hold for $S_{h,0}$ if w satisfies in addition $w(0) = w(1) = 0$. Well-known examples of spaces satisfying (2.1$a-b$) include the Hermite piecewise polynomial functions, for which $r = 2\mu + 2$ (Birkhoff *et al.*, 1968), and the spaces of smooth splines of even order (i.e., piecewise polynomial of odd degree), for which $r = \mu + 2$, where $\mu \geqslant 2$ is even (Schultz, 1970). (For smooth splines of any order $r = \mu + 2$, $\mu \geqslant 1$, (2.1b) holds at least for uniform meshes, cf. e.g. Bramble & Schatz (1976) and its references.)

Note that, as a consequence of the quasiuniformity of the mesh, the following inverse inequalities hold for $\chi \in S_h$ or $\chi \in S_{h,0}$

$$\begin{aligned} \|\chi\|_\alpha &\leqslant Ch^{-(\alpha-\beta)}\|\chi\|_\beta, \quad 0 \leqslant \beta \leqslant \alpha \leqslant \mu + 1, \\ \|\chi\|_{j,\infty} &\leqslant Ch^{-(j+1/2)}\|\chi\|, \quad 0 \leqslant j \leqslant \mu, \end{aligned} \tag{2.2}$$

for constants C independent of h and χ. Also, as a consequence of (2.1a–b) and the quasiuniformity of the mesh, it follows that if P is the L^2-projection operator onto S_h, then the following hold, (Douglas *et al.*, 1975; Thomée & Wahlbin, 1983),

$$\|Pv\|_1 \leqslant C\|v\|_1, \quad \forall\, v \in H^1, \tag{2.3a}$$

$$\|Pv\|_\infty \leqslant C\|v\|_\infty, \quad \forall\, v \in C^0, \tag{2.3b}$$

$$\|Pv - v\|_\infty \leqslant Ch^r\|v\|_{r,\infty}, \quad \forall\, v \in C^r, \tag{2.3c}$$

for some constants C independent of h and v. The same inequalities hold for the L^2-projection operator P_0 onto $S_{h,0}$ when, in addition, $v(0) = v(1) = 0$. (In the sequel we shall refer to the analogous results for P_0 on $S_{h,0}$ using the same formula numbers, i.e., (2.3a–c).)

The standard Galerkin method for the semidiscretization of (SW) is defined as follows: we seek $\eta_h : [0,T] \to S_h$, $u_h : [0,T] \to S_{h,0}$, such that for $t \in [0,T]$

$$\begin{aligned}(\eta_{ht},\phi) + (u_{hx},\phi) + ((\eta_h u_h)_x,\phi) = 0, \quad \forall\, \phi \in S_h,\\ (u_{ht},\chi) + (\eta_{hx},\chi) + (u_h u_{hx},\chi) = 0, \quad \forall\, \chi \in S_{h,0},\end{aligned} \tag{2.4}$$

with initial conditions

$$\eta_h(0) = P\eta_0, \quad u_h(0) = P_0 u_0. \tag{2.5}$$

In (Antonopoulos & Dougalis, 2016, Proposition 2.1) it was proved that if (η,u), the solution of (SW), is sufficiently smooth and satisfies $1+\eta > 0$ for $t \in [0,T]$, and if $r \geqslant 3$ and h is sufficiently small, then the semidiscrete ivp (2.4)–(2.5) has a unique solution (η_h,u_h) for $t \in [0,T]$, satisfying

$$\max_{0\leqslant t\leqslant T}(\|\eta(t)-\eta_h(t)\| + \|u(t)-u_h(t)\|) \leqslant Ch^{r-1}. \tag{2.6}$$

It is well known that $r-1$ is the best order of convergence in L^2 expected for the standard Galerkin method for first-order hyperbolic problems on general quasiuniform meshes; for a uniform mesh better rates of convergence may be obtained (Dupont, 1973). For uniform meshes it was proved in Antonopoulos & Dougalis (2016) that in the case of (SW), one obtains $O(h^2)$ L^2-convergence for the semidiscrete approximation with continuous, piecewise linear functions. In the case of the periodic ivp for the shallow water equations, the semidiscrete approximation with smooth splines on a uniform mesh gives an optimal-order L^2 error estimate of $O(h^r)$, cf. Antonopoulos & Dougalis (2016). The assumption that $r \geq 3$ is needed in the proof of (2.6) in order to control the $W^{1,\infty}$ norm of an error term, and was also present in the error analysis of Dupont (1974) for a close relative of the SW system. (Numerical experiments in Antonopoulos & Dougalis (2016) on quasiuniform meshes suggest that (2.6) holds for continuous, piecewise linear functions ($r=2$) as well; hence the assumption $r \geqslant 3$ may be technical.)

In the analysis of the fully discrete scheme under consideration we will assume that $r \geqslant 3$ and that the mesh is quasiuniform, so that the spatial error in L^2 will be $O(h^{r-1})$. The emphasis of the convergence proof will be placed in getting the optimal temporal-order L^2 error estimate $O(k^4 + h^{r-1})$. In the proof, the fully discrete approximations will not be compared to the semidiscrete solution, but directly to the L^2 projection of the solution of the continuous problem (SW). Thus, the semidiscretization will not be further utilized in this paper. In the sequel we will assume that (SW) possesses a unique, sufficiently smooth solution (η,u) for $0 \leqslant t \leqslant T$, such that $1+\eta \geqslant \alpha > 0$ for $(x,t) \in [0,1]\times[0,T]$ for some constant α. We will denote by C positive constants independent of the discretization parameters.

In the proofs of Sections 3 and 4 we will make use of several estimates that follow from the assumptions on the approximation and inverse properties of the finite element spaces made thus far. One of them is the following superapproximation property of $S_{h,0}$, Dupont (1974), Douglas *et al.* (1975), according to which

$$\|P_0[(1+\eta)\xi] - (1+\eta)\xi\| \leqslant Ch\|\xi\|, \quad \forall\, \xi \in S_{h,0}. \tag{2.7}$$

We will also use the following results, that we state as Lemmata.

LEMMA 2.1 Let $H = P\eta$. (i) Then

$$\|P_0[(1+H)\xi] - (1+H)\xi\| \leqslant Ch\|\xi\|, \quad \forall\, \xi \in S_{h,0}. \tag{2.8}$$

(ii) If $f \in L^2(0,1)$ and

$$((1+H)\xi, P_0 f) = ((1+H)\xi, f) + b(\xi, f), \quad \xi \in S_{h,0}, \tag{2.9}$$

then $|b(\xi,f)| \leqslant Ch\|\xi\|\,\|f\|$.

Proof. (i) We have

$$\begin{aligned} P_0[(1+H)\xi] - (1+H)\xi &= P_0[(1+H)\xi] - P_0[(1+\eta)\xi] + P_0[(1+\eta)\xi] - (1+\eta)\xi \\ &\quad + (1+\eta)\xi - (1+H)\xi \\ &= P_0[(H-\eta)\xi] + P_0[(1+\eta)\xi] - (1+\eta)\xi - (H-\eta)\xi, \end{aligned}$$

whence, from (2.3b), (2.7),

$$\|P_0[(1+H)\xi] - (1+H)\xi\| \leqslant C(\|H-\eta\|_\infty \|\xi\| + h\|\xi\|),$$

and therefore (2.8) follows from (2.3c). (ii) We have

$$b(\xi,f) = (P_0[(1+H)\xi], f) - ((1+H)\xi, f) = (P_0[(1+H)\xi] - (1+H)\xi, f),$$

and therefore, by (2.8), $|b(\xi,f)| \leqslant Ch\|\xi\|\,\|f\|$. □

LEMMA 2.2 Let η be the first component of the solution of (SW) for which we suppose that $1+\eta \geqslant \alpha > 0$, and $H = P\eta$. If $\eta \in C^r$, then for sufficiently small h we have

$$1 + H \geqslant \tfrac{\alpha}{2}.$$

In addition, if $f \in L^2(0,1)$ then

$$\tfrac{\alpha}{2}\|f\|^2 \leqslant ((1+H)f, f) \leqslant C'\|f\|^2, \tag{2.10}$$

for some constant C' depending on $\|\eta\|_{r,\infty}$.

Proof. From (2.3c) we have

$$1 + \eta - C_1 h^r \leqslant 1 + H \leqslant 1 + \eta + C_1 h^r,$$

for some constant C_1. Therefore if $h \leqslant (\alpha/(2C_1))^{1/r}$ then $\alpha/2 \leqslant 1 + H \leqslant C'$, from which (2.10) follows. □

For the purposes of the proof of convergence of the fully discrete scheme, we will also need two more properties of the L^2-projection operators P and P_0, in addition to (2.3a–c). The first one follows from the approximation and inverse properties of the finite element spaces already mentioned. It expresses the fact that P is stable in H^2, i.e., that there exists a constant C such that

$$\|Pv\|_2 \leqslant C\|v\|_2, \quad \forall\, v \in H^2. \tag{2.11}$$

In addition, the analogous stability estimate holds for P_0 on $v \in H^2 \cap \overset{\circ}{H}{}^1$. It is straightforward to check that (2.11) follows from the hypotheses on S_h made thus far. Indeed, let $R_h : H^1 \to S_h$ be the H^1-projection onto S_h defined for $w \in H^1$ by $(R_h w, \phi)_1 = (w, \phi)_1$ for all $\phi \in S_h$. (Here $(\cdot,\cdot)_1$ denotes the H^1 inner product.) Suppose $v \in H^2$ and let ψ be the interpolant of v in the space of continuous, piecewise linear functions defined with respect to the partition $\{x_i\}_{i=1}^{N+1}$. Then, by a local inverse inequality for $R_h v - \psi \in \mathbb{P}_{r-1}(x_j, x_{j+1})$ and the quasiuniformity of the mesh, we have

$$\|(R_h v)''\|^2 = \sum_{j=1}^{N} \int_{x_j}^{x_{j+1}} ((R_h v)'' - \psi'')^2 \leqslant C h^{-2} \sum_{j=1}^{N} \int_{x_j}^{x_{j+1}} ((R_h v)' - \psi')^2.$$

Hence $\|(R_h v)''\| \leqslant C h^{-1} \|(R_h v)' - \psi'\| \leqslant C h^{-1}(\|R_h v - v\|_1 + \|v - \psi\|_1)$. Since $\|R_h v - v\|_1 \leqslant C \inf_{\chi \in S_h} \|v - \chi\|_1 \leqslant Ch\|v\|_2$, by (2.1a), it follows that $\|(R_h v)''\| \leqslant C\|v\|_2$, from which the stability of R_h in H^2 follows in view of the fact that $\|R_h v\|_1 \leqslant \|v\|_1$, $v \in H^1$. Finally,

$$\begin{aligned}\|Pv\|_2 &\leqslant \|Pv - R_h v\|_2 + \|R_h v\|_2 \leqslant C h^{-2}\|P(v - R_h v)\| + C\|v\|_2 \\ &\leqslant C h^{-2}\|v - R_h v\| + C\|v\|_2 \leqslant C\|v\|_2.\end{aligned}$$

(In the final step we used (2.1a) for $s = 2$.)

In addition, in the course of the proof of the consistency estimates of the fully discrete scheme in Proposition 3.2 in Section 3, we will need the property that if $v \in H^s$, $s \geqslant 3$, is independent of h, then

$$\|Pv\|_3 \leqslant C_s(v), \tag{2.12}$$

where $C_s(v)$ is a constant depending only on v and s. We will also assume that (2.12) holds for P_0 as well, if in addition $v(0) = v(1) = 0$. This property does not follow from our hypotheses (2.1a–b), (2.2). It holds for the Hermite piecewise polynomial functions on a general nonuniform mesh, provided $\mu \geqslant 2$ (hence, for $r-1 \geqslant 5$, i.e., for at least piecewise quintic polynomials), cf. Birkhoff *et al.* (1968), and also for smooth splines if $\mu \geqslant 2$, i.e., for which $r - 1 \geqslant 3$, i.e., at least cubic splines. (If $r - 1$ is odd, this requires just a quasiuniform mesh, cf. Schultz (1970), while if $r - 1$ is even, a uniform mesh guarantees (2.12) for $\mu \geqslant 2$, cf. Bramble & Schatz (1976).)

3. The fully discrete scheme and its consistency

For a positive integer M, we let $k = T/M$, $t^n = nk$, $n = 0, 1, \ldots, M$, and using the notation established in Section 2 we let $H(t) = P\eta(t)$, $U(t) = P_0 u(t)$, $H^n = H(t^n)$, $U^n = U(t^n)$, where (η, u) is the solution of (SW). We also define

$$\Phi = U + HU, \qquad \Phi^n = \Phi(t^n), \tag{3.1}$$

$$F = H_x + UU_x, \quad F^n = F(t^n). \tag{3.2}$$

We discretize in time the ode system represented by the semidiscretization (2.4)–(2.5) by the explicit, fourth-order accurate 'classical' RK scheme (RK4), written as follows. Seek $H_h^n \in S_h$, $U_h^n \in S_{h,0}$,

$0 \leqslant n \leqslant M$ and $H_h^{n,j} \in S_h$, $U_h^{n,j} \in S_{h,0}$ for $j = 1, 2, 3$, $0 \leqslant n \leqslant M-1$, such that

$$\begin{aligned} H_h^{n,j} - H_h^n + ka_j P \Phi_{hx}^{n,j-1} &= 0, \\ U_h^{n,j} - U_h^n + ka_j P_0 F_h^{n,j-1} &= 0, \end{aligned} \tag{3.3}$$

for $j = 1, 2, 3$ and

$$\begin{aligned} H_h^{n+1} - H_h^n + kP\left[\sum_{j=1}^{4} b_j \Phi_h^{n,j-1}\right]_x &= 0, \\ U_h^{n+1} - U_h^n + kP_0\left[\sum_{j=1}^{4} b_j F_h^{n,j-1}\right] &= 0, \end{aligned} \tag{3.4}$$

where

$$\begin{aligned} \Phi_h^{n,j} &= U_h^{n,j} + H_h^{n,j} U_h^{n,j}, \\ F_h^{n,j} &= H_{hx}^{n,j} + U_h^{n,j} U_{hx}^{n,j}, \end{aligned} \tag{3.5}$$

for $j = 0, 1, 2, 3$ and

$$H_h^{n,0} = H_h^n, \quad U_h^{n,0} = U_h^n, \quad a_1 = a_2 = 1/2, \quad a_3 = 1, \quad b_1 = b_4 = 1/6, \quad b_2 = b_3 = 1/3,$$

with

$$H_h^0 = \eta_h(0) = P\eta_0, \quad U_h^0 = u_h(0) = P_0 u_0. \tag{3.6}$$

In order to study the temporal consistency of the scheme (3.3)–(3.6), we define the intermediate stages $V^{n,j} \in S_h$, $W^{n,j} \in S_{h,0}$, $j = 0, 1, 2, 3$, $0 \leqslant n \leqslant M-1$, by the equations

$$V^{n,j} - H^n + ka_j P \Phi_x^{n,j-1} = 0, \tag{3.7}$$

$$W^{n,j} - U^n + ka_j P_0 F^{n,j-1} = 0 \tag{3.8}$$

and

$$V^{n,0} = H^n, \quad W^{n,0} = U^n,$$

where

$$\Phi^{n,j} = W^{n,j} + V^{n,j} W^{n,j}, \tag{3.9}$$

$$F^{n,j} = V_x^{n,j} + W^{n,j} W_x^{n,j}, \tag{3.10}$$

for $j = 0, 1, 2, 3$, with $\Phi^{n,0} = \Phi^n$, $F^{n,0} = F^n$.

We first estimate the continuous spatial truncation error resulting from replacing η and u in (SW) by their L^2 projections on the finite element spaces. In the sequel we assume that the solution (η, u) of (SW) is sufficiently smooth for the purposes of the error estimation.

LEMMA 3.1 Let (η, u) be the solution of (SW) in $[0, T]$. Let $H(t) = P\eta(t)$, $U(t) = P_0 u(t)$ and let $\psi(t) \in S_h$, $\zeta(t) \in S_{h,0}$, for $0 \leqslant t \leqslant T$, be such that

$$H_t + P\big(U + HU\big)_x = \psi, \tag{3.11}$$

$$U_t + P_0\big(H_x + UU_x\big) = \zeta. \tag{3.12}$$

Then, for $j = 0, 1, 2, 3$, we have

$$\|\partial_t^j \psi\| + \|\partial_t^j \zeta\| \leqslant Ch^{r-1}, \tag{3.13}$$

for $0 \leqslant t \leqslant T$.

Proof. Subtracting (3.11) from the equation $P\big(\eta_t + u_x + (\eta u)_x\big) = 0$ and defining $\rho := \eta - H$, $\sigma := u - U$, we have

$$P\sigma_x + P\big((\eta u)_x - (HU)_x\big) = -\psi.$$

Since $\eta u - HU = \eta u - (\eta - \rho)(u - \sigma) = \eta\sigma + u\rho - \rho\sigma$ we see that

$$P\big(\sigma_x + (\eta\sigma)_x + (u\rho)_x - (\rho\sigma)_x\big) = -\psi,$$

from which, using the approximation properties of the spaces S_h, $S_{h,0}$, it follows that

$$\begin{aligned} \|\partial_t^j \psi\| &\leqslant \|P\partial_t^j \sigma_x\| + \|P\partial_t^j (\eta\sigma)_x\| + \|P\partial_t^j (u\rho)_x\| + \|P\partial_t^j (\rho\sigma)_x\| \\ &\leqslant C(h^{r-1} + h^{2r-1}) \leqslant Ch^{r-1}, \end{aligned}$$

for $j = 0, 1, 2, 3$. Subtracting (3.12) from the equation $P_0\big(u_t + \eta_x + uu_x\big) = 0$ gives

$$P_0\rho_x + P_0(uu_x - UU_x) = -\zeta.$$

Since $uu_x - UU_x = uu_x - (u - \sigma)(u_x - \sigma_x) = (u\sigma)_x - \sigma\sigma_x$ it follows, for $j = 0, 1, 2, 3$, that

$$\begin{aligned} \|\partial_t^j \zeta\| &\leqslant \|P_0\partial_t^j \rho_x\| + \|P_0\partial_t^j (u\sigma)_x\| + \|P_0\partial_t^j (\sigma\sigma_x)\| \\ &\leqslant C(h^{r-1} + h^{2r-1}) \leqslant Ch^{r-1}, \end{aligned}$$

and (3.13) is proved. □

In the following proposition we estimate appropriately defined local errors of the fully discrete scheme (3.3)–(3.6). The local errors $\delta_1^n \in S_h$, $\delta_2^n \in S_{h,0}$ are expressed in terms of the L^2 projections $H^n = P\eta(t^n)$, $U^n = P_0 u(t^n)$ and the quantities $\Phi^{n,i}$, $F^{n,i}$ defined by (3.9), (3.10) as nonlinear functions of the intermediate stages $V^{n,i}$, $W^{n,i}$ of a single step of the RK4 scheme with starting values H^n, U^n, cf. (3.7), (3.8). The plan of the error estimation is straightforward, but the details of the proof are rather technical. We find expansions of $\Phi^{n,i}$, $F^{n,i}$, for $i = 1, 2, 3$ in powers of k up to terms of $O(k^2)$ for $i = 1$ and up to terms of $O(k^3)$ for $i = 2$ and 3, and we estimate the remainders by bounds of $O(h^{r-1} + k^4)$ in appropriate norms. The constants in these error bounds depend polynomially on the Courant number

$\lambda = k/h$. The expressions for $\Phi^{n,i}$, $F^{n,i}$ are combined as in the final step of the RK4 scheme to yield the required estimates of the local errors after cancelation of the lower-order terms.

In bounding the remainders of $\Phi^{n,i}$, $F^{n,i}$ for $i = 1, 2$, use is made of the standard approximation and inverse properties (2.1), (2.2) of the finite element spaces, in particular of the stability and approximation estimates of the L^2 projections that follow from these properties. But in the course of bounding some terms of the remainders of $\Phi^{n,3}$ and $F^{n,3}$, we need to find L^2 bounds independent of h of third-order spatial derivatives of H_t and U_t; for this purpose we use the hypothesis that (2.12) holds.

PROPOSITION 3.2 Let (η, u) be the solution of (SW) and let $\lambda = k/h$. If δ_1^n, δ_2^n, for $0 \leqslant n \leqslant M-1$, are such that

$$\delta_1^n = H^{n+1} - H^n + kP\left[\sum_{j=1}^{4} b_j \Phi^{n,j-1}\right]_x, \tag{3.14}$$

$$\delta_2^n = U^{n+1} - U^n + kP_0\left[\sum_{j=1}^{4} b_j F^{n,j-1}\right], \tag{3.15}$$

then, there exists a constant C_λ that depends polynomially on λ such that

$$\max_{0\leqslant n\leqslant M-1}(\|\delta_1^n\| + \|\delta_2^n\|) \leqslant C_\lambda k(k^4 + h^{r-1}).$$

Proof. From (3.7) it follows that

$$V^{n,1} - H^n + a_1 kP\Phi_x^n = 0.$$

Hence, by (3.11), (3.1),

$$V^{n,1} = H^n + a_1 kH_t^n - a_1 k\psi^n. \tag{3.16}$$

In addition, by (3.8) and (3.12) we have

$$W^{n,1} - U^n + a_1 kP_0 F^n = 0,$$

and consequently, from (3.12), (3.2),

$$W^{n,1} = U^n + a_1 kU_t^n - a_1 k\zeta^n. \tag{3.17}$$

So, by (3.16) and (3.17),

$$V^{n,1}W^{n,1} = H^n U^n + a_1 k(HU)_t^n + a_1^2 k^2 H_t^n U_t^n + v_1^n, \tag{3.18}$$

where

$$v_1^n = -a_1 k(H^n \zeta^n + U^n \psi^n) - a_1^2 k^2 (H_t^n \zeta^n + U_t^n \psi^n - \psi^n \zeta^n).$$

Therefore, by (3.17), (3.18), we obtain

$$W^{n,1} + V^{n,1}W^{n,1} = U^n + H^nU^n + a_1k\big(U_t^n + (HU)_t^n\big) + a_1^2k^2H_t^nU_t^n + v_2^n,$$

so that, the desired expansion of $\Phi^{n,1}$ in powers of k is given by

$$\Phi^{n,1} = \Phi^n + a_1k\Phi_t^n + a_1^2k^2H_t^nU_t^n + v_2^n, \tag{3.19}$$

where

$$v_2^n = -a_1k\zeta^n + v_1^n.$$

Note that by (3.13) and (2.2), (2.3) it follows that

$$\|v_2^n\|_1 \leqslant C_\lambda h^{r-1}, \tag{3.20}$$

where C_λ is a first-order polynomial in λ with positive coefficients. (In the sequel we shall denote by C_λ polynomials of λ with positive coefficients without reference to their degree.) In deriving (3.20) we made use of the fact (something that we will also do in the sequel, without explicit mention) that the quantities $\|\partial_t^iH^n\|_j$, $\|\partial_t^iU^n\|_j$ for $j = 0, 1, 2$ and for each i, are bounded, uniformly in n, by constants independent of the discretization parameters k and h. This follows from (2.3a), (2.11) and the smoothness of η and u. The same holds for the quantities $\|\partial_t^iH^n\|_{j,\infty}$, $\|\partial_t^iU^n\|_{j,\infty}$ for $j = 0, 1$, as seen from (2.3b) and (2.1b). Now, from (3.17)

$$W^{n,1}W_x^{n,1} = U^nU_x^n + a_1k(UU_x)_t^n + a_1^2k^2U_t^nU_{tx}^n + w_1^n, \tag{3.21}$$

where

$$w_1^n = -a_1k(U^n\zeta^n)_x - a_1^2k^2(U_t^n\zeta^n)_x + a_1^2k^2\zeta^n\zeta_x^n.$$

Hence, from (3.16), (3.21), it follows that

$$V_x^{n,1} + W^{n,1}W_x^{n,1} = H_x^n + U^nU_x^n + a_1k(H_x + UU_x)_t^n + a_1^2k^2U_t^nU_{tx}^n + w_2^n,$$

i.e.,

$$F^{n,1} = F^n + a_1kF_t^n + a_1^2k^2U_t^nU_{tx}^n + w_2^n, \tag{3.22}$$

which is the required expansion of $F^{n,1}$. In the above

$$w_2^n = -a_1k\psi_x^n + w_1^n,$$

for which, using (3.13), the inverse inequalities and the remarks following (3.20), we obtain the estimate

$$\|w_2^n\| \leqslant C_\lambda h^{r-1}. \tag{3.23}$$

We now find the expansions of $\Phi^{n,2}$ and $F^{n,2}$ up to $O(k^3)$ terms. From (3.7), i.e.,

$$V^{n,2} = H^n - a_2kP\Phi_x^{n,1},$$

it follows, in view of (3.19), that

$$V^{n,2} = H^n - a_2kP\Phi_x^n - a_1a_2k^2P\Phi_{tx}^n - a_1^2a_2k^3P(H_t^nU_t^n)_x - a_2kPv_{2x}^n,$$

which, in view of (3.11) gives

$$V^{n,2} = H^n + a_2kH_t^n + a_1a_2k^2H_{tt}^n - a_1^2a_2k^3P(H_t^nU_t^n)_x + \psi_1^n, \tag{3.24}$$

where

$$\psi_1^n = -a_2k\psi^n - a_1a_2k^2\psi_t^n - a_2kPv_{2x}^n.$$

From (3.20) and (3.13) it follows that

$$\|\psi_1^n\| \leqslant C_\lambda kh^{r-1}, \tag{3.25}$$

and, by the inverse properties, that

$$\|\psi_{1x}^n\| \leqslant C_\lambda h^{r-1}. \tag{3.26}$$

Moreover, since from (3.8)

$$W^{n,2} = U^n - a_2kP_0F^{n,1},$$

we obtain, using (3.22),

$$W^{n,2} = U^n - a_2kP_0F^n - a_1a_2k^2P_0F_t^n - a_1^2a_2k^3P_0(U_t^nU_{tx}^n) - a_2kP_0w_2^n,$$

and finally, from (3.12),

$$W^{n,2} = U^n + a_2kU_t^n + a_1a_2k^2U_{tt}^n - a_1^2a_2k^3P_0(U_t^nU_{tx}^n) + \zeta_1^n, \tag{3.27}$$

where

$$\zeta_1^n = -a_2k\zeta^n - a_1a_2k^2\zeta_t^n - a_2kP_0w_2^n,$$

that we estimate, using (3.13) and (3.23), by

$$\|\zeta_1^n\| \leqslant C_\lambda kh^{r-1}, \tag{3.28}$$

and, using the inverse properties, by

$$\|\zeta_{1x}^n\| \leqslant C_\lambda h^{r-1}. \tag{3.29}$$

Now, from (3.24), (3.27), taking into account that $a_1 = a_2$, we have

$$\begin{aligned} V^{n,2}W^{n,2} &= H^nU^n + a_2k(HU)_t^n + a_1a_2k^2(H^nU_{tt}^n + H_t^nU_t^n + H_{tt}^nU^n) + a_1a_2^2k^3(H_t^nU_{tt}^n + H_{tt}^nU_t^n) \\ &\quad - a_1^2a_2k^3\big(H^nP_0(U_t^nU_{tx}^n) + U^nP(H_t^nU_t^n)_x\big) + v_3^n, \end{aligned} \tag{3.30}$$

where

$$v_3^n = H^n\zeta_1^n + a_2kH_t^n\zeta_1^n + a_1a_2k^2H_{tt}^n\zeta_1^n - a_1^2a_2k^3P(H_t^nU_t^n)_x\zeta_1^n - a_1^2a_2^2k^4\big(H_t^nP_0(U_t^nU_{tx}^n) + U_t^nP(H_t^nU_t^n)_x\big)$$
$$+ a_1^2a_2^2k^4H_{tt}^nU_{tt}^n - a_1^3a_2^2k^5\big(H_{tt}^nP_0(U_t^nU_{tx}^n) + U_{tt}^nP(H_t^nU_t^n)_x\big) + a_1^4a_2^2k^6P(H_t^nU_t^n)_xP_0(H_t^nU_{tx}^n) + \psi_1^nW^{n,2}.$$

Using (3.28), (3.29), (3.25), (3.13), and taking into account the inverse inequalities and the remarks following (3.20), we may estimate v_3^n as follows:

$$\|v_3^n\| \leqslant C_\lambda kh^{r-1} + Ck^4 \quad \text{and} \quad \|v_3^n\|_1 \leqslant C_\lambda h^{r-1} + Ck^4. \tag{3.31}$$

Finally, writing (3.30) in the form

$$V^{n,2}W^{n,2} = H^nU^n + a_2k(HU)_t^n + a_1a_2k^2(HU)_{tt}^n - a_1a_2k^2H_t^nU_t^n + a_1a_2^2k^3(H_tU_t)_t^n$$
$$- a_1^2a_2k^3\big(H^nP_0(U_t^nU_{tx}^n) + U^nP(H_t^nU_t^n)_x\big) + v_3^n,$$

we obtain, using (3.9) and (3.27), the desired expansion of $\Phi^{n,2}$ in powers of k:

$$\begin{aligned}\Phi^{n,2} &= \Phi^n + a_2k\Phi_t^n + a_1a_2k^2\Phi_{tt}^n - a_1a_2k^2H_t^nU_t^n + a_1a_2^2k^3(H_tU_t)_t^n\\ &\quad - a_1^2a_2k^3\big[(1 + H^n)P_0(U_t^nU_{tx}^n) + U^nP(H_t^nU_t^n)_x\big] + v_4^n,\end{aligned} \tag{3.32}$$

where $v_4^n = \zeta_1^n + v_3^n$. Hence, from (3.28), (3.29) and (3.31), we have

$$\|v_4^n\| \leqslant C_\lambda kh^{r-1}, \quad \|v_{4x}^n\| \leqslant C_\lambda h^{r-1} + Ck^4. \tag{3.33}$$

Now it follows from (3.27) that

$$W^{n,2}W_x^{n,2} = \Big[U^n + a_2kU_t^n + a_1a_2k^2U_{tt}^n - a_1^2a_2k^3P_0(U_t^nU_{tx}^n) + \zeta_1^n\Big]$$
$$\times\Big[U_x^n + a_2kU_{tx}^n + a_1a_2k^2U_{ttx}^n - a_1^2a_2k^3\big(P_0(U_t^nU_{tx}^n)\big)_x + \zeta_{1x}^n\Big],$$

and, consequently, since $a_1 = a_2$, that

$$\begin{aligned}W^{n,2}W_x^{n,2} &= U^nU_x^n + a_2k(UU_x)_t^n + a_1a_2k^2(UU_x)_{tt}^n - a_1a_2k^2U_t^nU_{tx}^n + a_1a_2^2k^3(U_t^nU_{tt}^n)_x\\ &\quad - a_1^2a_2k^3\big[U^nP_0(U_t^nU_{tx}^n)\big]_x + w_3^n,\end{aligned} \tag{3.34}$$

where

$$w_3^n = U^n\zeta_{1x}^n + W_x^{n,2}\zeta_1^n + a_2kU_t^n\zeta_{1x}^n + a_1a_2k^2U_{tt}^n\zeta_{1x}^n - a_1^2a_2k^3P_0(U_t^nU_{tx}^n)\zeta_{1x}^n$$
$$+ a_1^2a_2^2k^4\big[U_{tt}^nU_{ttx}^n - \big(U_t^nP_0(U_t^nU_{tx}^n)\big)_x\big] - a_1^3a_2^2k^5\big(U_{tt}^nP_0(U_t^nU_{tx}^n)\big)_x + a_1^4a_2^2k^6P_0(U_t^nU_{tx}^n)\big(P_0(U_t^nU_{tx}^n)\big)_x.$$

Using the approximation and inverse properties of the finite element spaces, (3.29) and observations like the ones following (3.20), we may estimate w_3^n by the inequalities

$$\|w_3^n\| \leqslant C_\lambda h^{r-1} + Ck^4, \quad k\|w_{3x}^n\| \leqslant C_\lambda (h^{r-1} + k^4). \tag{3.35}$$

Now, the definition of $F^{n,2}$ in (3.10), (3.24) and (3.34) gives

$$\begin{aligned} F^{n,2} &= F^n + a_2kF_t^n + a_1a_2k^2F_{tt}^n - a_1a_2k^2U_t^nU_{tx}^n + a_1a_2^2k^3(U_t^nU_{tt}^n)_x \\ &\quad - a_1^2a_2k^3\big[P(H_t^nU_t^n)_x + U^nP_0(U_t^nU_{tx}^n)\big]_x + w_4^n, \end{aligned} \tag{3.36}$$

where

$$w_4^n = \psi_{1x}^n + w_3^n.$$

From (3.26) and (3.35) it follows that

$$\|w_4^n\| \leqslant C_\lambda h^{r-1} + Ck^4. \tag{3.37}$$

This completes the required expansion of $F^{n,2}$ in powers of k.

We now compute the required expansions of $\Phi^{n,3}$ and $F^{n,3}$ up to $O(k^3)$ terms. In the course of estimating some of the $O(k^4)$ remainder terms, we need to find L^2 bounds independent of h of third-order spatial derivatives of U_t and H_t, and for this purpose we need the hypothesis (2.12). Since

$$V^{n,3} = H^n - a_3kP\Phi_x^{n,2} = H^n - kP\Phi_x^{n,2},$$

from (3.32) and (3.11) it follows that

$$\begin{aligned} V^{n,3} &= H^n + kH_t^n - k\psi^n + a_2k^2H_{tt}^n - a_2k^2\psi_t^n + a_1a_2k^3H_{ttt}^n - a_1a_2k^3\psi_{tt}^n + a_1a_2k^3P(H_t^nU_t^n)_x \\ &\quad - a_1a_2^2k^4P\big((H_tU_t)_{tx}^n\big) + a_1^2a_2k^4P\big[(1+H^n)P_0(U_t^nU_{tx}^n) + U^nP(H_t^nU_t^n)_x\big]_x - kPv_{4x}^n, \end{aligned}$$

which we write as

$$V^{n,3} = H^n + kH_t^n + a_2k^2H_{tt}^n + a_1a_2k^3H_{ttt}^n + a_1a_2k^3P(H_t^nU_t^n)_x + \psi_2^n, \tag{3.38}$$

where

$$\begin{aligned} \psi_2^n &= -k\psi^n - a_2k^2\psi_t^n - a_1a_2k^3\psi_{tt}^n - a_1a_2^2k^4P\big((H_tU_t)_{tx}^n\big) \\ &\quad + a_1^2a_2k^4P\big[(1+H^n)P_0(U_t^nU_{tx}^n) + U^nP(H_t^nU_t^n)_x\big]_x - kPv_{4x}^n. \end{aligned}$$

From (3.33), the approximation and inverse properties of the finite element spaces and the remarks following (3.20) we get

$$\|\psi_2^n\| \leqslant C_\lambda kh^{r-1} + Ck^4. \tag{3.39}$$

By similar considerations and using also the hypothesis (2.12) we infer in addition that

$$\|\psi_{2x}^n\| \leqslant C_\lambda(h^{r-1} + k^4). \tag{3.40}$$

Also, since

$$W^{n,3} = U^n - a_3 k P_0 F^{n,2} = U^n - kP_0 F^{n,2},$$

from (3.36) and (3.12) we obtain

$$\begin{aligned} W^{n,3} &= U^n + kU_t^n - k\zeta^n + a_2k^2U_{tt}^n - a_2k^2\zeta_t^n + a_1a_2k^3U_{ttt}^n - a_1a_2k^3\zeta_{tt}^n + a_1a_2k^3P_0(U_t^nU_{tx}^n) \\ &\quad + a_1^2a_2k^4P_0\big[P(H_t^nU_t^n)_x - U_t^nU_{tt}^n + U^nP_0(U_t^nU_{tx}^n)\big]_x - kP_0w_4^n, \end{aligned}$$

i.e.,

$$W^{n,3} = U^n + kU_t^n + a_2k^2U_{tt}^n + a_1a_2k^3U_{ttt}^n + a_1a_2k^3P_0(U_t^nU_{tx}^n) + \zeta_2^n, \tag{3.41}$$

where

$$\zeta_2^n = -k\zeta^n - a_1k^2\zeta_t^n - a_1a_2k^3\zeta_{tt}^n + a_1^2a_2k^4P_0\big[P(H_t^nU_t^n)_x - U_t^nU_{tt}^n + U^nP_0(U_t^nU_{tx}^n)\big]_x - kP_0w_4^n.$$

From (3.13), (3.37), the approximation and inverse properties of the finite element spaces and the remarks following (3.20) we may see that

$$\|\zeta_2^n\| \leqslant C_\lambda k h^{r-1} + Ck^4. \tag{3.42}$$

By similar considerations and also the hypothesis (2.12) it follows that

$$\|\zeta_{2x}^n\| \leqslant C_\lambda(h^{r-1} + k^4). \tag{3.43}$$

From (3.38) and (3.41), since $a_2 = 1/2 = a_1$, we see that

$$\begin{aligned} V^{n,3}W^{n,3} &= H^nU^n + k(HU)_t^n + a_2k^2(HU)_{tt}^n + a_1a_2k^3(HU)_{ttt}^n - a_1a_2k^3(H_tU_t)_t^n \\ &\quad + a_1a_2k^3\big[H^nP_0(U_t^nU_{tx}^n) + U^nP(H_t^nU_t^n)_x\big] + v_5^n. \end{aligned} \tag{3.44}$$

From (3.13), (3.42), (3.39), (3.40), the approximation and inverse properties of the finite element spaces, and the remarks following (3.20) it follows for the remainder term in (3.44) that

$$\|v_5^n\| \leqslant C_\lambda k h^{r-1} + Ck^4 + C_\lambda k^8. \tag{3.45}$$

Similar considerations and in addition (3.43), (3.40) lead to

$$\|v_{5x}^n\| \leqslant C_\lambda(h^{r-1} + k^4). \tag{3.46}$$

Hence, from (3.41), (3.9), (3.44) we have the desired expansion of $\Phi^{n,3}$ given by

$$\begin{aligned}\Phi^{n,3} &= \Phi^n + k\Phi_t^n + a_2k^2\Phi_{tt}^n + a_1a_2k^3\Phi_{ttt}^n - a_1a_2k^3(H_tU_t)_t^n \\ &+a_1a_2k^3\big[(1+H^n)P_0(U_t^nU_{tx}^n) + U^nP(H_t^nU_t^n)_x\big] + v_6^n,\end{aligned} \tag{3.47}$$

in which $v_6^n = \psi_2^n + v_5^n$. Hence, from (3.39), (3.40), (3.45), (3.46), it follows that

$$\|v_6^n\| \leqslant C_\lambda kh^{r-1} + Ck^4, \quad \|v_{6x}^n\| \leqslant C_\lambda(h^{r-1} + k^4). \tag{3.48}$$

From (3.41) we see that

$$\begin{aligned}W^{n,3}W_x^{n,3} = &\Big[U^n + kU_t^n + a_2k^2U_{tt}^n + a_1a_2k^3U_{ttt}^n + a_1a_2k^3P_0(U_t^nU_{tx}^n) + \zeta_2^n\Big] \\ &\times\Big[U_x^n + kU_{tx}^n + a_2k^2U_{ttx}^n + a_1a_2k^3U_{tttx}^n + a_1a_2k^3\big(P_0(U_t^nU_{tx}^n)\big)_x + \zeta_{2x}^n\Big],\end{aligned}$$

and therefore, using the fact that $a_1 = 1/2$,

$$\begin{aligned}W^{n,3}W_x^{n,3} &= U^nU_x^n + k(UU_x)_t^n + a_2k^2(UU_x)_{tt}^n + a_1a_2k^3(UU_x)_{ttt}^n \\ &-a_1a_2k^3(U_tU_{tx})_t^n + a_1a_2k^3\big[U^nP_0(U_t^nU_{tx}^n)\big]_x + w_5^n.\end{aligned} \tag{3.49}$$

For the remainder term, using (3.42), (3.43), the approximation and inverse properties of the finite element spaces, and considerations such as the ones following (3.20) we get

$$\|w_5^n\| \leqslant C_\lambda(h^{r-1} + k^4). \tag{3.50}$$

Finally, from (3.38), (3.10) and (3.49), we have the required expansion of $F^{n,3}$

$$\begin{aligned}F^{n,3} &= F^n + kF_t^n + a_2k^2F_{tt}^n + a_1a_2k^3F_{ttt}^n - a_1a_2k^3(U_tU_{tt})_x^n \\ &+ a_1a_2k^3\big[P(H_t^nU_t^n)_x + U^nP_0(U_t^nU_{tx}^n)\big]_x + w_6^n,\end{aligned} \tag{3.51}$$

where $w_6^n = \psi_{2,x}^n + w_5^n$. Therefore, by (3.40) and (3.51) we conclude that

$$\|w_6^n\| \leqslant C_\lambda(h^{r-1} + k^4). \tag{3.52}$$

We now come to the final line of the RK4 algorithm and the expansions of $\sum_{j=1}^4 b_j\Phi^{n,j-1}$, $\sum_{j=1}^4 b_jF^{n,j-1}$ that are needed in the expressions (3.14), (3.15) of the local errors. Since $\sum_{j=1}^4 b_j = 1$, from (3.19), (3.32), (3.47), it follows that

$$\begin{aligned}
&b_1\Phi^n + b_2\Phi^{n,1} + b_3\Phi^{n,2} + b_4\Phi^{n,3}\\
&\quad = \Phi^n + (b_2a_1 + b_3a_2 + b_4)k\Phi^n_t + (b_2a_1^2 - b_3a_1a_2)k^2H^n_tU^n_t + (b_3a_1a_2 + b_4a_2)k^2\Phi^n_{tt}\\
&\qquad + (b_3a_1a_2^2 - b_4a_1a_2)k^3(H_tU_t)^n_t + (-b_3a_1^2a_2 + b_4a_1a_2)k^3\big[(1+H^n)P_0(U^n_tU^n_{tx}) + U^nP(H^n_tU^n_t)_x\big]\\
&\qquad + b_4a_1a_2k^3\Phi^n_{ttt} + b_2v^n_2 + b_3v^n_4 + b_4v^n_6,
\end{aligned}$$

and, therefore, using the values of b_1, b_2, b_3, b_4, a_1, a_2,

$$b_1\Phi^n + b_2\Phi^{n,1} + b_3\Phi^{n,2} + b_4\Phi^{n,3} = \Phi^n + \tfrac{k}{2}\Phi^n_t + \tfrac{k^2}{6}\Phi^n_{tt} + \tfrac{k^3}{24}\Phi^n_{ttt} + \tfrac{1}{3}v^n_2 + \tfrac{1}{3}v^n_4 + \tfrac{1}{6}v^n_6.$$

From (3.1), (3.11), (3.14) and the above equality, we see that

$$\delta^n_1 = H^{n+1} - H^n - kH^n_t - \tfrac{k^2}{2}H^n_{tt} - \tfrac{k^3}{6}H^n_{ttt} - \tfrac{k^4}{24}H^n_{tttt} + \alpha^n,$$

where

$$\alpha^n = k\psi^n + \tfrac{k^2}{2}\psi^n_t + \tfrac{k^3}{6}\psi^n_{tt} + \tfrac{k^4}{24}\psi^n_{ttt} + \tfrac{k}{3}(Pv^n_{2x} + Pv^n_{4x}) + \tfrac{k}{6}Pv^n_{6x}.$$

Therefore, since by (3.13), (3.20), (3.33), (3.48),

$$\|\alpha^n\| \leqslant C_\lambda k(h^{r-1} + k^4),$$

it follows by Taylor's theorem that

$$\|\delta^n_1\| \leqslant C_\lambda k(h^{r-1} + k^4). \tag{3.53}$$

In addition from (3.22), (3.36), (3.51) we obtain

$$\begin{aligned}
&b_1F^n + b_2F^{n,1} + b_3F^{n,2} + b_4F^{n,3}\\
&\quad = F^n + (b_2a_1 + b_3a_2 + b_4)kF^n_t + (b_2a_1^2 - b_3a_1a_2)k^2U^n_tU^n_{tx} + (b_3a_1a_2 + b_4a_2)k^2F^n_{tt}\\
&\qquad + (b_3a_1a_2^2 - b_4a_1a_2)k^3(U^n_tU^n_{tt})_x + (-b_3a_1^2a_2 + b_4a_1a_2)k^3\big[P(H^n_tU^n_t) + U^nP_0(U^n_tU^n_{tx})\big]_x\\
&\qquad + b_4a_1a_2k^3F^n_{ttt} + b_2w^n_2 + b_3w^n_4 + b_4w^n_6,
\end{aligned}$$

and therefore

$$b_1F^n + b_2F^{n,1} + b_3F^{n,2} + b_4F^{n,3} = F^n + \tfrac{k}{2}F^n_t + \tfrac{k^2}{6}F^n_{tt} + \tfrac{k^3}{24}F^n_{ttt} + \tfrac{1}{3}w^n_2 + \tfrac{1}{3}w^n_4 + \tfrac{1}{6}w^n_6.$$

From (3.2), (3.12), (3.15) and the above, it follows that

$$\delta^n_2 = U^{n+1} - U^n - kU^n_t - \tfrac{k^2}{2}U^n_{tt} - \tfrac{k^3}{6}U^n_{ttt} - \tfrac{k^4}{24}U^n_{tttt} + \beta^n,$$

where

$$\beta^n = k\zeta^n + \tfrac{k^2}{2}\zeta_t^n + \tfrac{k^3}{6}\zeta_{tt}^n + \tfrac{k^4}{24}\zeta_{ttt}^n + \tfrac{k}{3}(P_0 w_2^n + P_0 w_4^n) + \tfrac{k}{6}P_0 w_6^n,$$

and, in view of (3.23), (3.37) and (3.52),

$$\|\beta^n\| \leqslant C_\lambda k(h^{r-1} + k^4).$$

By the above and Taylor's theorem we see that $\|\delta_2^n\| \leqslant C_\lambda k(h^{r-1} + k^4)$. This estimate and (3.53) conclude the proof of the proposition. □

4. Error estimate

In this section we analyze the convergence of the fully discrete scheme (3.3)–(3.6) to the solution of (SW) in the $L^2 \times L^2$ norm. We start with three technical Lemmata whose notation and results will be used in the course of proof of the error estimate in Theorem 4.4.

LEMMA 4.1 For $j = 1, 2, 3$, let $V^{n,j}$, $W^{n,j}$ be defined by (3.7)–(3.10) and let $\lambda = k/h$. Then, there exist constants C_λ depending polynomially on λ, such that

$$\|H^n - V^{n,j}\|_\infty + \|U^n - W^{n,j}\|_\infty \leqslant C_\lambda k, \tag{4.1}$$

$$\|H^n - V^{n,j}\|_{1,\infty} + \|U^n - W^{n,j}\|_{1,\infty} \leqslant C_\lambda. \tag{4.2}$$

Proof. From (3.7), (3.8), and (3.9), (3.10), we have, for $j = 1, 2, 3$,

$$H^n - V^{n,j} = ka_j P\Phi_x^{n,j-1} = ka_j P(W^{n,j-1} + V^{n,j-1}W^{n,j-1})_x,$$

$$U^n - W^{n,j} = ka_j P_0 F^{n,j-1} = ka_j P_0(V_x^{n,j-1} + W^{n,j-1}W_x^{n,j-1}),$$

and so, by (2.3b)

$$\|H^n - V^{n,j}\|_\infty \leqslant Ck(\|W_x^{n,j-1}\|_\infty + \|V_x^{n,j-1}\|_\infty\|W^{n,j-1}\|_\infty + \|V^{n,j-1}\|_\infty\|W_x^{n,j-1}\|_\infty),$$

$$\|U^n - W^{n,j}\|_\infty \leqslant Ck(\|V_x^{n,j-1}\|_\infty + \|W^{n,j-1}\|_\infty\|W_x^{n,j-1}\|_\infty),$$

for $j = 1, 2, 3$. From these relations, using e.g., (2.11) and the inverse properties of S_h, $S_{h,0}$, we may derive recursively (4.1) and (4.2). □

LEMMA 4.2 Let $\varepsilon^n \in S_h$, $e^n \in S_{h,0}$ and suppose that $\rho^{n,j}$, $r^{n,j}$ are functions defined for $j = 1, 2, 3$ by

$$\rho^{n,j} = (1 + H^n)P_0 r_x^{n,j-1} + U^n P\rho_x^{n,j-1}, \tag{4.3}$$

$$r^{n,j} = P\rho_x^{n,j-1} + U^n P_0 r_x^{n,j-1}, \tag{4.4}$$

with

$$\rho^{n,0} = \rho^n = (1 + H^n)e^n + U^n\varepsilon^n, \tag{4.5}$$

$$r^{n,0} = r^n = \varepsilon^n + U^n e^n. \tag{4.6}$$

Then, there exists a constant C such that

$$\|\rho^n\| + \|r^n\| \leqslant C(\|\varepsilon^n\| + \|e^n\|), \tag{4.7}$$

$$\|\rho_x^{n,j}\| + \|r_x^{n,j}\| \leqslant \tfrac{C}{h^{j+1}}(\|\varepsilon^n\| + \|e^n\|), \quad j = 1, 2, 3. \tag{4.8}$$

If, moreover, $\|\varepsilon^n\|_{1,\infty} + \|e^n\|_{1,\infty} \leqslant \widetilde{C}$ for some constant $\widetilde{C}$, then, for $j = 0, 1, 2, 3$,

$$\|\rho_x^{n,j}\|_\infty + \|r_x^{n,j}\|_\infty \leqslant \tfrac{C}{h^j}. \tag{4.9}$$

Proof. The inequality (4.7) follows from (4.5) and (4.6) and (2.3b). To prove (4.8) note that

$$\|\rho_x^n\| + \|r_x^n\| \leqslant \|H_x^n e^n\| + \|(1 + H^n)e_x^n\| + \|\varepsilon_x^n\| + \|U_x^n e^n\| + \|U^n e_x^n\|.$$

Similarly,

$$\|\rho_x^{n,j}\| + \|r_x^{n,j}\| \leqslant \tfrac{C}{h}(\|r_x^{n,j-1}\| + \|\rho_x^{n,j-1}\|),$$

for $j = 1, 2, 3$, and (4.8) follows by recursion. Since

$$\|\rho_x^n\|_\infty + \|r_x^n\|_\infty \leqslant \|H_x^n e^n\|_\infty + \|(1 + H^n)e_x^n\|_\infty + \|\varepsilon_x^n\|_\infty + \|U_x^n e^n\|_\infty + \|U^n e_x^n\|_\infty,$$

the hypothesis of the Lemma and (2.11), imply

$$\|\rho_x^n\|_\infty + \|r_x^n\|_\infty \leqslant C.$$

Moreover, since

$$\|\rho_x^{n,j}\|_\infty + \|r_x^{n,j}\|_\infty \leqslant \tfrac{C}{h}(\|r_x^{n,j-1}\|_\infty + \|\rho_x^{n,j-1}\|_\infty)$$

holds for $j = 1, 2, 3$, a recursive argument yields (4.9). □

Lemma 4.3 Given $\varepsilon^n \in S_h$, $e^n \in S_{h,0}$, let $\rho^{n,j}$, $r^{n,j}$, $0 \leqslant j \leqslant 3$, be defined as in Lemma 4.2. In addition, let $\rho^{n,-1}(x) = \int_0^x \varepsilon^n$, $r^{n,-1}(x) = \int_0^x e^n$, for $0 \leqslant x \leqslant 1$. Then, for $0 \leqslant i < j \leqslant 3$, we have

$$(P\rho_x^{n,i}, \rho_x^{n,j}) + \big((1 + H^n)P_0 r_x^{n,i}, r_x^{n,j}\big) = -(\rho_x^{n,i+1}, P\rho_x^{n,j-1}) - \big((1 + H^n)r_x^{n,i+1}, P_0 r_x^{n,j-1}\big) + \gamma_i^{n,j-1}, \tag{4.10}$$

where

$$\gamma_i^{n,j-1} = (U_x^n P\rho_x^{n,i}, P\rho_x^{n,j-1}) + \big([(1 + H^n)U_x^n - H_x^n U^n]P_0 r_x^{n,i}, P_0 r_x^{n,j-1}\big). \tag{4.11}$$

In the particular cases $j = i + 1$, $i = -1, 0, 1, 2$, (4.10) may be simplified to

$$(P\rho_x^{n,i}, \rho_x^{n,i+1}) + \big((1 + H^n)P_0 r_x^{n,i}, r_x^{n,i+1}\big) = \tfrac{1}{2}\gamma_i^{n,i}. \tag{4.12}$$

In all cases $-1 \leqslant i < j \leqslant 3$, we have the estimates

$$|\gamma_i^{n,j-1}| \leqslant \tfrac{C}{h^{i+j+1}}(\|\varepsilon^n\|^2 + \|e^n\|^2). \tag{4.13}$$

Proof. In all cases $-1 \leqslant i < j \leqslant 3$, since $\rho^{n,j}$, $P_0 r_x^{n,i}$ vanish at $x = 0, 1$, integrating by parts yields

$$(P\rho_x^{n,i}, \rho_x^{n,j}) + \big((1+H^n)P_0 r_x^{n,i}, r_x^{n,j}\big) = -\big((P\rho_x^{n,i})_x, \rho^{n,i}\big) - \big([(1+H^n)P_0 r_x^{n,i}]_x, r^{n,j}\big) \equiv A^{i,j}. \tag{4.14}$$

We first examine the cases with $0 \leqslant i < j \leqslant 3$. We have, using the definitions of the $\rho^{n,\alpha}$, $r^{n,\alpha}$ and some computation, that

$$\begin{aligned}
A^{i,j} &= -\big((P\rho_x^{n,i})_x, (1+H^n)P_0 r_x^{n,j-1} + U^n P\rho_x^{n,j-1}\big) - \big([(1+H^n)P_0 r_x^{n,i}]_x, P\rho_x^{n,j-1} + U^n P_0 r_x^{n,j-1}\big) \\
&= -\big([(1+H^n)P_0 r_x^{n,i} + U^n P\rho_x^{n,i}]_x - U_x^n P\rho_x^{n,i}, P\rho_x^{n,j-1}\big) \\
&\quad - \big((1+H^n)(P\rho_x^{n,i} + U^n P_0 r_x^{n,i})_x - (1+H^n)U_x^n P_0 r_x^{n,i} + H_x^n U^n P_0 r_x^{n,i}, P_0 r_x^{n,j-1}\big) \\
&= -(\rho_x^{n,i+1}, P\rho_x^{n,j-1}) - \big((1+H^n) r_x^{n,i+1}, P_0 r_x^{n,j-1}\big) + \gamma_i^{n,j-1},
\end{aligned}$$

where the $\gamma_i^{n,j-1}$ are defined by (4.11). The last equality above and (4.14) give (4.10). The remaining case $i = -1, j = 0$ is a special case of (4.12); the latter follows from (4.14), (4.11), similar computations as above and the identity $(\alpha v, (\beta v)_x) = ((\alpha\beta_x - \alpha_x\beta)v, v)/2$ valid for $\alpha, \beta \in H^1$, $v \in \overset{\circ}{H}{}^1$.

The estimate (4.13) for $j = 1, 2, 3$, $0 \leqslant i < j$, follows from (4.11), and (4.8), (2.3b). If $i = -1$ the proof is similar and takes account of the facts that $\rho_x^{n,-1} = \varepsilon^n$, $r_x^{n,-1} = e^n$. □

The main error estimate of the paper, which incorporates the crucial stability step applied to an error energy inequality, follows. We remark that the proof does not use the hypothesis (2.12) except in its last step, where the local error estimates of Proposition 3.2 (recall that the latter rely on the validity of (2.12)) are brought to bear.

THEOREM 4.4 Let (η, u) be the solution of (SW) in $[0,1] \times [0,T]$ with $1+\eta \geqslant \alpha > 0$, for some constant α, and let (H_h^n, U_h^n) be its fully discrete approximation defined by (3.3)–(3.6). Let $\lambda = k/h$ and let h be sufficiently small. Then there exists a constant λ_0 depending on α, and a constant C independent of k and h, such that for $\lambda \leqslant \lambda_0$,

$$\max_{0 \leqslant n \leqslant M} (\|\eta(t^n) - H_h^n\| + \|u(t^n) - U_h^n\|) \leqslant C(h^{r-1} + k^4). \tag{4.15}$$

Proof. It suffices to show that

$$\max_{0 \leqslant n \leqslant M} (\|H^n - H_h^n\| + \|U^n - U_h^n\|) \leqslant C(h^{r-1} + k^4). \tag{4.16}$$

To make the exposition easier to follow, we break up the proof in five parts.

(i) *Notation and the basic error equations*

Let

$$\varepsilon^{n,j} = V^{n,j} - H_h^{n,j}, \quad e^{n,j} = W^{n,j} - U_h^{n,j}, \quad j = 0, 1, 2, 3,$$

with $\varepsilon^{n,0} = \varepsilon^n = H^n - H_h^n$, $e^{n,0} = e^n = U^n - U_h^n$. Then, from (3.3), (3.7), (3.8) it follows for $j = 1, 2, 3$ that

$$\begin{aligned} \varepsilon^{n,j} &= \varepsilon^n - a_j kP(\Phi^{n,j-1} - \Phi_h^{n,j-1})_x, \\ e^{n,j} &= e^n - a_j kP_0(F^{n,j-1} - F_h^{n,j-1}), \end{aligned} \tag{4.17}$$

and from (3.4), (3.14), (3.15) that

$$\begin{aligned} \varepsilon^{n+1} &= \varepsilon^n - kP\left[\sum_{j=1}^{4} b_j(\Phi^{n,j-1} - \Phi_h^{n,j-1})_x\right] + \delta_1^n, \\ e^{n+1} &= e^n - kP_0\left[\sum_{j=1}^{4} b_j(F^{n,j-1} - F_h^{n,j-1})\right] + \delta_2^n. \end{aligned} \tag{4.18}$$

Also, from (3.9), (3.10), (3.5), we have for $j = 0, 1, 2, 3$

$$\begin{aligned} \Phi^{n,j} - \Phi_h^{n,j} &= W^{n,j} + V^{n,j}W^{n,j} - U_h^{n,j} - H_h^{n,j}U_h^{n,j} \\ &= e^{n,j} + V^{n,j}W^{n,j} - (V^{n,j} - \varepsilon^{n,j})(W^{n,j} - e^{n,j}), \end{aligned}$$

and

$$\begin{aligned} F^{n,j} - F_h^{n,j} &= V_x^{n,j} + W^{n,j}W_x^{n,j} - H_{hx}^{n,j} - U_h^{n,j}U_{hx}^{n,j} \\ &= \varepsilon_x^{n,j} + W^{n,j}W_x^{n,j} - (W^{n,j} - e^{n,j})(W_x^{n,j} - e_x^{n,j}), \end{aligned}$$

so that

$$\begin{aligned} \Phi^{n,j} - \Phi_h^{n,j} &= (1 + V^{n,j})e^{n,j} + W^{n,j}\varepsilon^{n,j} - \varepsilon^{n,j}e^{n,j}, \\ F^{n,j} - F_h^{n,j} &= \varepsilon_x^{n,j} + (W^{n,j}e^{n,j})_x - \tfrac{1}{2}((e^{n,j})^2)_x. \end{aligned} \tag{4.19}$$

Since now

$$\begin{aligned} (1 + V^{n,j})e^{n,j} &= \left[1 + H^n - (H^n - V^{n,j})\right]e^{n,j} = (1 + H^n)e^{n,j} - (H^n - V^{n,j})e^{n,j}, \\ W^{n,j}\varepsilon^{n,j} &= \left[U^n - (U^n - W^{n,j})\right]\varepsilon^{n,j} = U^n\varepsilon^{n,j} - (U^n - W^{n,j})\varepsilon^{n,j}, \\ W^{n,j}e^{n,j} &= \left[U^n - (U^n - W^{n,j})\right]e^{n,j} = U^n e^{n,j} - (U^n - W^{n,j})e^{n,j}, \end{aligned}$$

it follows from the equations (4.19) that for $j = 0, 1, 2, 3$

$$\begin{aligned} \Phi^{n,j} - \Phi_h^{n,j} &= (1 + H^n)e^{n,j} + U^n\varepsilon^{n,j} + v^{n,j}, \\ F^{n,j} - F_h^{n,j} &= \varepsilon_x^{n,j} + (U^n e^{n,j})_x + w_x^{n,j}, \end{aligned} \tag{4.20}$$

where

$$\begin{aligned} v^{n,j} &= -(H^n - V^{n,j})e^{n,j} - (U^n - W^{n,j})\varepsilon^{n,j} - \varepsilon^{n,j}e^{n,j}, \\ w^{n,j} &= -(U^n - W^{n,j})e^{n,j} - \tfrac{1}{2}(e^{n,j})^2. \end{aligned} \tag{4.21}$$

We note for future reference that it follows from (4.21), the inequalities (4.1) and (4.2) and the inverse inequalities on the spaces S_h, $S_{h,0}$ that

$$\|v_x^{n,j}\| + \|w_x^{n,j}\| \leqslant (C_\lambda + \|\varepsilon^{n,j}\|_{1,\infty} + \|e^{n,j}\|_{1,\infty})(\|\varepsilon^{n,j}\| + \|e^{n,j}\|), \tag{4.22}$$

$$\|v_x^{n,j}\|_\infty + \|w_x^{n,j}\|_\infty \leqslant C_\lambda(\|\varepsilon^{n,j}\|_\infty + \|e^{n,j}\|_\infty) + 3\|\varepsilon^{n,j}\|_{1,\infty}\|e^{n,j}\|_{1,\infty}, \tag{4.23}$$

where, as usual, C_λ denotes a constant depending polynomially on λ.

(ii) *Expansions of $\Phi^{n,j} - \Phi_h^{n,j}$, $F^{n,j} - F_h^{n,j}$ in powers of k*

In this part of the proof we derive suitable representations of the differences $\Phi^{n,j} - \Phi_h^{n,j}$, $F^{n,j} - F_h^{n,j}$ that will be used in the energy identities.

If $j = 0$, we have

$$\begin{aligned} \Phi^n - \Phi_h^n &= \rho^n + \rho_1^n, \\ F^n - F_h^n &= r_x^n + r_{1x}^n, \end{aligned} \tag{4.24}$$

where the ρ^n, r^n were defined in Lemma 4.2 and satisfy the inequalities (4.7) and (4.8), and

$$\begin{aligned} \rho_1^n &= v^{n,0} = -\varepsilon^n e^n, \\ r_1^n &= w^{n,0} = -\tfrac{1}{2}(e^n)^2. \end{aligned} \tag{4.25}$$

From (4.17) and (4.24) it follows that

$$\begin{aligned} \varepsilon^{n,1} &= \varepsilon^n - ka_1 P\rho_x^n - ka_1 P\rho_{1x}^n, \\ e^{n,1} &= e^n - ka_1 P_0 r_x^n - ka_1 P_0 r_{1x}^n, \end{aligned} \tag{4.26}$$

and, using the notation of Lemma 4.2, (4.3) and (4.4), that

$$\begin{aligned} (1 + H^n)e^{n,1} + U^n\varepsilon^{n,1} &= \rho^n - ka_1\rho^{n,1} - ka_1[(1 + H^n)P_0 r_{1x}^n + U^n P\rho_{1x}^n], \\ \varepsilon^{n,1} + U^n e^{n,1} &= r^n - ka_1 r^{n,1} - ka_1(P\rho_{1x}^n + U^n P_0 r_{1x}^n). \end{aligned}$$

Consequently, from the equations (4.20) we obtain

$$\begin{aligned} \Phi^{n,1} - \Phi_h^{n,1} &= \rho^n - ka_1\rho^{n,1} + \rho_1^{n,1}, \\ F^{n,1} - F_h^{n,1} &= r_x^n - ka_1 r_x^{n,1} + r_{1x}^{n,1}, \end{aligned} \tag{4.27}$$

where

$$\begin{aligned} \rho_1^{n,1} &= -ka_1[(1+H^n)P_0 r_{1x}^n + U^n P \rho_{1x}^n] + v^{n,1}, \\ r_1^{n,1} &= -ka_1(P\rho_{1x}^n + U^n P_0 r_{1x}^n) + w^{n,1}. \end{aligned} \tag{4.28}$$

Hence, from (4.17) and the equations (4.27) we get

$$\begin{aligned} \varepsilon^{n,2} &= \varepsilon^n - ka_2 P\rho_x^n + k^2 a_1 a_2 P\rho_x^{n,1} - ka_2 P\rho_{1x}^{n,1}, \\ e^{n,2} &= e^n - ka_2 P_0 r_x^n + k^2 a_1 a_2 P_0 r_x^{n,1} - ka_2 P_0 r_{1x}^{n,1}, \end{aligned} \tag{4.29}$$

and, using the notation introduced in (4.3) and (4.4),

$$\begin{aligned} (1+H^n)e^{n,2} + U^n\varepsilon^{n,2} &= \rho^n - ka_2\rho^{n,1} + k^2 a_1 a_2 \rho^{n,2} - ka_2\big[(1+H^n)P_0 r_{1x}^{n,1} + U^n P\rho_{1x}^{n,1}\big], \\ \varepsilon^{n,2} + U^n e^{n,2} &= r^n - ka_2 r^{n,1} + k^2 a_1 a_2 r^{n,2} - ka_2\big(P\rho_{1x}^{n,1} + U^n P_0 r_{1x}^{n,1}\big). \end{aligned}$$

So, from the equations (4.20) we see that

$$\begin{aligned} \Phi^{n,2} - \Phi_h^{n,2} &= \rho^n - ka_2\rho^{n,1} + k^2 a_1 a_2 \rho^{n,2} + \rho_1^{n,2}, \\ F^{n,2} - F_h^{n,2} &= r_x^n - ka_2 r_x^{n,1} + k^2 a_1 a_2 r_x^{n,2} + r_{1x}^{n,2}, \end{aligned} \tag{4.30}$$

where

$$\begin{aligned} \rho_1^{n,2} &= -ka_2[(1+H^n)P_0 r_{1x}^{n,1} + U^n P\rho_{1x}^{n,1}] + v^{n,2}, \\ r_1^{n,2} &= -ka_2(P\rho_{1x}^{n,1} + U^n P_0 r_{1x}^{n,1}) + w^{n,2}. \end{aligned} \tag{4.31}$$

Hence, from the equations (4.17), taking into account that $a_3 = 1$, we obtain

$$\begin{aligned} \varepsilon^{n,3} &= \varepsilon^n - kP\rho_x^n + k^2 a_2 P\rho_x^{n,1} - k^3 a_1 a_2 P\rho_x^{n,2} - kP\rho_{1x}^{n,2}, \\ e^{n,3} &= e^n - kP_0 r_x^n + k^2 a_2 P_0 r_x^{n,1} - k^3 a_1 a_2 P_0 r_x^{n,2} - kP_0 r_{1x}^{n,2}, \end{aligned} \tag{4.32}$$

and, according to (4.3) and (4.4),

$$\begin{aligned} (1+H^n)e^{n,3} + U^n\varepsilon^{n,3} &= \rho^n - k\rho^{n,1} + k^2 a_2\rho^{n,2} - k^3 a_1 a_2 \rho^{n,3} - k\big[(1+H^n)P_0 r_{1x}^{n,2} + U^n P\rho_{1x}^{n,2}\big], \\ \varepsilon^{n,3} + U^n e^{n,3} &= r^n - kr^{n,1} + k^2 a_2 r^{n,2} - k^3 a_1 a_2 r^{n,3} - k\big(P\rho_{1x}^{n,2} + U^n P_0 r_{1x}^{n,2}\big). \end{aligned}$$

Consequently,

$$\begin{aligned} \Phi^{n,3} - \Phi_h^{n,3} &= \rho^n - k\rho^{n,1} + k^2 a_2\rho^{n,2} - k^3 a_1 a_2 \rho^{n,3} + \rho_1^{n,3}, \\ F^{n,3} - F_h^{n,3} &= r_x^n - kr_x^{n,1} + k^2 a_2 r_x^{n,2} - k^3 a_1 a_2 r_x^{n,3} + r_{1x}^{n,3}, \end{aligned} \tag{4.33}$$

where

$$\begin{aligned} \rho_1^{n,3} &= -k\big[(1+H^n)P_0 r_{1x}^{n,2} + U^n P\rho_{1x}^{n,2}\big] + v^{n,3}, \\ r_1^{n,3} &= -k\big(P\rho_{1x}^{n,2} + U^n P_0 r_{1x}^{n,2}\big) + w^{n,3}. \end{aligned} \tag{4.34}$$

(iii) *'Inductive' hypothesis and consequent estimates*
We let now n^* be the maximal integer for which

$$\|\varepsilon^n\|_{1,\infty} + \|e^n\|_{1,\infty} \leqslant 1, \quad 0 \leqslant n \leqslant n^*. \tag{H}$$

Then, for $0 \leqslant n \leqslant n^*$, from (4.25) it follows that

$$\begin{aligned} \|\rho_{1x}^n\| + \|r_{1x}^n\| &\leqslant C(\|\varepsilon^n\| + \|e^n\|), \\ \|\rho_{1x}^n\|_\infty + \|r_{1x}^n\|_\infty &\leqslant C. \end{aligned} \tag{4.35}$$

In addition, from (4.26), (H), (4.25), (4.8) for $j = 0, 1$, (4.9) for $j = 0$, the second inequality of (4.35), the inverse properties of S_h, $S_{h,0}$ and (2.3b), it follows that

$$\|\varepsilon^{n,1}\| + \|e^{n,1}\| \leqslant C_\lambda(\|\varepsilon^n\| + \|e^n\|), \tag{4.36}$$

$$\|\varepsilon^{n,1}\|_{1,\infty} + \|e^{n,1}\|_{1,\infty} \leqslant C_\lambda. \tag{4.37}$$

Also, from (4.22), (4.23), (4.36) and (4.37), we have

$$\|v_x^{n,1}\| + \|w_x^{n,1}\| \leqslant C_\lambda(\|\varepsilon^n\| + \|e^n\|), \tag{4.38}$$

$$\|v_x^{n,1}\|_\infty + \|w_x^{n,1}\|_\infty \leqslant C_\lambda. \tag{4.39}$$

Now, for $j = 1, 2, 3$, in view of (4.28), (4.31), (4.34) and the inverse properties of S_h and $S_{h,0}$, it follows that

$$\begin{aligned} \|\rho_{1x}^{n,j}\| + \|r_{1x}^{n,j}\| &\leqslant C_\lambda\big(\|\rho_{1x}^{n,j-1}\| + \|r_{1x}^{n,j-1}\|\big) + \big(\|v_x^{n,j}\| + \|w_x^{n,j}\|\big), \\ \|\rho_{1x}^{n,j}\|_\infty + \|r_{1x}^{n,j}\|_\infty &\leqslant C_\lambda\big(\|\rho_{1x}^{n,j-1}\|_\infty + \|r_{1x}^{n,j-1}\|_\infty\big) + \big(\|v_x^{n,j}\|_\infty + \|w_x^{n,j}\|_\infty\big). \end{aligned}$$

In addition, for $j = 2, 3$, (4.7), (4.8), (4.29), (4.32), (4.9) and the inverse properties of S_h, $S_{h,0}$ give

$$\begin{aligned} \|\varepsilon^{n,j}\| + \|e^{n,j}\| &\leqslant C_\lambda(\|\varepsilon^n\| + \|e^n\|) + Ck\big(\|\rho_{1x}^{n,j-1}\| + \|r_{1x}^{n,j-1}\|\big), \\ \|\varepsilon^{n,j}\|_{1,\infty} + \|e^{n,j}\|_{1,\infty} &\leqslant C_\lambda\big(1 + \|\rho_{1x}^{n,j-1}\|_\infty + \|r_{1x}^{n,j-1}\|_\infty\big), \quad 0 \leqslant n \leqslant n^*. \end{aligned}$$

Therefore, for $0 \leqslant n \leqslant n^*$ and $j = 0, 1, 2, 3$, in view of (4.35), (4.38), (4.39), (4.22) and arguing recursively, we finally obtain

$$\|\rho_{1x}^{n,j}\| + \|r_{1x}^{n,j}\| \leqslant C_\lambda(\|\varepsilon^n\| + \|e^n\|), \tag{4.40}$$

$$\|\rho_{1x}^{n,j}\|_\infty + \|r_{1x}^{n,j}\|_\infty \leqslant C_\lambda. \tag{4.41}$$

(iv) *Basic energy identity and estimation of the terms in its right-hand side*
From (4.18), (4.24), (4.27), (4.30), (4.33) and the definitions of the constants a_j, b_j of the RK scheme,

we have

$$\begin{aligned} \varepsilon^{n+1} &= f^n + f_1^n + \delta_1^n, \\ e^{n+1} &= g^n + g_1^n + \delta_2^n, \end{aligned} \tag{4.42}$$

where

$$\begin{aligned} f^n &= \varepsilon^n - kP\rho_x^n + \tfrac{k^2}{2}P\rho_x^{n,1} - \tfrac{k^3}{6}P\rho_x^{n,2} + \tfrac{k^4}{24}P\rho_x^{n,3}, \\ g^n &= e^n - kP_0r_x^n + \tfrac{k^2}{2}P_0r_x^{n,1} - \tfrac{k^3}{6}P_0r_x^{n,2} + \tfrac{k^4}{24}P_0r_x^{n,3}, \\ f_1^n &= -\tfrac{k}{6}\big(P\rho_{1x}^n + 2P\rho_{1x}^{n,1} + 2P\rho_{1x}^{n,2} + P\rho_{1x}^{n,3}\big), \\ g_1^n &= -\tfrac{k}{6}\big(P_0r_{1x}^n + 2P_0r_{1x}^{n,1} + 2P_0r_{1x}^{n,2} + P_0r_{1x}^{n,3}\big). \end{aligned}$$

From these relations and (4.7), (4.8) it follows that

$$\|f^n\| + \|g^n\| \leqslant C_\lambda(\|\varepsilon^n\| + \|e^n\|), \tag{4.43}$$

and, moreover, for $0 \leqslant n \leqslant n^*$, from (4.40)

$$\|f_1^n\| + \|g_1^n\| \leqslant C_\lambda k(\|\varepsilon^n\| + \|e^n\|). \tag{4.44}$$

Now, by the definitions of f^n, g^n, we may obtain the basic energy identity of our scheme:

$$\|f^n\|^2 + \big((1+H^n)g^n, g^n\big) = \|\varepsilon^n\|^2 + \big((1+H^n)e^n, e^n\big) + \sum_{i=1}^{8} k^i\beta_i^n. \tag{4.45}$$

We will now identify and estimate the quantities β_i^n, $1 \leqslant i \leqslant 8$, in the right-hand side of the above. For β_1^n we have

$$\beta_1^n = -2(\varepsilon^n, P\rho_x^n) - 2\big((1+H^n)e^n, P_0r_x^n\big).$$

Since, by (4.12),

$$(\varepsilon^n, \rho_x^n) + \big((1+H^n)e^n, r_x^n\big) = \tfrac{1}{2}\gamma_{-1}^{n,-1},$$

it follows that

$$\beta_1^n = -2\big((1+H^n)e^n, P_0r_x^n - r_x^n\big) - \gamma_{-1}^{n,-1}.$$

From this relation, Lemma 2.1(ii), (4.8) and (4.13), we see that

$$|\beta_1^n| \leqslant C\big(\|\varepsilon^n\|^2 + \|e^n\|^2\big). \tag{4.46}$$

The quantity β_2^n is given by

$$\beta_2^n = (\varepsilon^n, P\rho_x^{n,1}) + \big((1+H^n)e^n, P_0r_x^{n,1}\big) + \|P\rho_x^n\|^2 + \big((1+H^n)P_0r_x^n, P_0r_x^n\big).$$

Since, by (4.10)

$$(\varepsilon^n, \rho_x^{n,1}) + \big((1+H^n)e^n, r_x^{n,1}\big) = -(\rho_x^n, P\rho_x^n) - \big((1+H^n)r_x^n, P_0 r_x^n\big) + \gamma_{-1}^{n,0},$$

we see that

$$\beta_2^n = \big((1+H^n)e^n, P_0 r_x^{n,1} - r_x^{n,1}\big) + \big((1+H^n)P_0 r_x^n, P_0 r_x^n - r_x^n\big) + \gamma_{-1}^{n,0}.$$

Hence, by Lemma 2.1(ii), (4.8) and (4.13) it follows that

$$|\beta_2^n| \leqslant \tfrac{C}{h}\big(\|\varepsilon^n\|^2 + \|e^n\|^2\big). \tag{4.47}$$

For β_3^n we find

$$\beta_3^n = -\tfrac{1}{3}\big((\varepsilon^n, P\rho_x^{n,2}) + ((1+H^n)e^n, P_0 r_x^{n,2})\big) - \big((P\rho_x^n, P\rho_x^{n,1}) + ((1+H^n)P_0 r_x^n, P_0 r_x^{n,1})\big),$$

i.e.,

$$\begin{aligned}\beta_3^n = &-\tfrac{1}{3}\big[(\varepsilon^n, P\rho_x^{n,2}) + \big((1+H^n)e^n, P_0 r_x^{n,2}\big) + (P\rho_x^n, P\rho_x^{n,1}) + \big((1+H^n)P_0 r_x^n, P_0 r_x^{n,1}\big)\big]\\ &-\tfrac{2}{3}\big[(P\rho_x^n, P\rho_x^{n,1}) + \big((1+H^n)P_0 r_x^n, P_0 r_x^{n,1}\big)\big].\end{aligned}$$

Using (4.10), (4.12) we see that

$$\begin{aligned}(\varepsilon^n, P\rho_x^{n,2}) + \big((1+H^n)e^n, r_x^{n,2}\big) &= -(\rho_x^n, P\rho_x^{n,1}) - \big((1+H^n)r_x^n, P_0 r_x^{n,1}\big) + \gamma_{-1}^{n,1},\\ (P\rho_x^n, \rho_x^{n,1}) + \big((1+H^n)P_0 r_x^n, r_x^{n,1}\big) &= \tfrac{1}{2}\gamma_0^{n,0},\end{aligned}$$

whence

$$\begin{aligned}\beta_3^n = &-\tfrac{1}{3}\big[\big((1+H^n)e^n, P_0 r_x^{n,2} - r_x^{n,2}\big) + \big((1+H^n)P_0 r_x^{n,1}, P_0 r_x^n - r_x^n\big) + \gamma_{-1}^{n,1}\big]\\ &-\tfrac{2}{3}\big[\big((1+H^n)P_0 r_x^n, P_0 r_x^{n,1} - r_x^{n,1}\big) + \tfrac{1}{2}\gamma_0^{n,0}\big],\end{aligned}$$

and, therefore, using again Lemma 2.1(ii), (4.8) and (4.13) we may estimate β_3^n as

$$|\beta_3^n| \leqslant \tfrac{C}{h^2}(\|\varepsilon^n\|^2 + \|e^n\|^2). \tag{4.48}$$

For β_4^n there holds

$$\begin{aligned}\beta_4^n = &\tfrac{1}{12}\big[(\varepsilon^n, P\rho_x^{n,3}) + \big((1+H^n)e^n, P_0 r_x^{n,3}\big)\big] + \tfrac{1}{3}\big[(P\rho_x^n, P\rho_x^{n,2}) + \big((1+H^n)P_0 r_x^n, P_0 r_x^{n,2}\big)\big]\\ &+\tfrac{1}{4}\big[\|P\rho_x^{n,1}\|^2 + \big((1+H^n)P_0 r_x^{n,1}, P_0 r_x^{n,1}\big)\big],\end{aligned}$$

or

$$\beta_4^n = \tfrac{1}{12}\big[(\varepsilon^n, P\rho_x^{n,3}) + \big((1+H^n)e^n, P_0 r_x^{n,3}\big) + (P\rho_x^n, P\rho_x^{n,2}) + \big((1+H^n)P_0 r_x^n, P_0 r_x^{n,2}\big)\big]$$
$$+ \tfrac{1}{4}\big[(P\rho_x^n, P\rho_x^{n,2}) + \big((1+H^n)P_0 r_x^n, P_0 r_x^{n,2}\big) + \|P\rho_x^{n,1}\|^2 + \big((1+H^n)P_0 r_x^{n,1}, P_0 r_x^{n,1}\big)\big].$$

Since, in view of (4.10),

$$(\varepsilon^n, P\rho_x^{n,3}) + \big((1+H^n)e^n, r_x^{n,3}\big) = -(\rho_x^n, P\rho_x^{n,2}) - \big((1+H^n)r_x^n, P_0 r_x^{n,2}\big) + \gamma_{-1}^{n,2},$$
$$(P\rho_x^n, \rho_x^{n,2}) + \big((1+H^n)P_0 r_x^n, r_x^{n,2}\big) = -(\rho_x^{n,1}, P\rho_x^{n,1}) - \big((1+H^n)r_x^{n,1}, P_0 r_x^{n,1}\big) + \gamma_0^{n,1},$$

it follows from Lemma 2.1(ii), (4.8) and (4.13) that

$$\beta_4^n = \tfrac{1}{12}\big[\big((1+H^n)e^n, P_0 r_x^{n,3} - r_x^{n,3}\big) + \big((1+H^n)P_0 r_x^{n,2}, P_0 r_x^n - r_x^n\big) + \gamma_{-1}^{n,2}\big]$$
$$+ \tfrac{1}{4}\big[\big((1+H^n)P_0 r_x^n, P_0 r_x^{n,2} - r_x^{n,2}\big) + \big((1+H^n)P_0 r_x^{n,1}, P_0 r_x^{n,1} - r_x^{n,1}\big) + \gamma_0^{n,1}\big],$$

and, consequently, that

$$|\beta_4^n| \leqslant \tfrac{C}{h^3}(\|\varepsilon^n\|^2 + \|e^n\|^2). \tag{4.49}$$

The quantity β_5^n is given by

$$\beta_5^n = -\tfrac{1}{12}\big[(P\rho_x^n, P\rho_x^{n,3}) + \big((1+H^n)P_0 r_x^n, P_0 r_x^{n,3}\big)\big] - \tfrac{1}{6}\big[(P\rho_x^{n,1}, P\rho_x^{n,2}) + \big((1+H^n)P_0 r_x^{n,1}, P_0 r_x^{n,2}\big)\big],$$

or by

$$\beta_5^n = -\tfrac{1}{12}\big[(P\rho_x^n, P\rho_x^{n,3}) + \big((1+H^n)P_0 r_x^n, P_0 r_x^{n,3}\big) + (P\rho_x^{n,1}, P\rho_x^{n,2}) + \big((1+H^n)P_0 r_x^{n,1}, P_0 r_x^{n,2}\big)\big]$$
$$- \tfrac{1}{12}\big[(P\rho_x^{n,1}, P\rho_x^{n,2}) + \big((1+H^n)P_0 r_x^{n,1}, P_0 r_x^{n,2}\big)\big].$$

However, in view of (4.10) and (4.12), we have

$$(P\rho_x^n, \rho_x^{n,3}) + \big((1+H^n)P_0 r_x^n, r_x^{n,3}\big) = -(\rho_x^{n,1}, P\rho_x^{n,2}) - \big((1+H^n)r_x^{n,1}, P_0 r_x^{n,2}\big) + \gamma_0^{n,2},$$
$$(P\rho_x^{n,1}, \rho_x^{n,2}) + \big((1+H^n)P_0 r_x^{n,1}, r_x^{n,2}\big) = \tfrac{1}{2}\gamma_1^{n,1},$$

whence, from Lemma 2.1(ii), (4.8) and (4.13)

$$\beta_5^n = -\tfrac{1}{12}\big[\big((1+H^n)P_0 r_x^n, P_0 r_x^{n,3} - r_x^{n,3}\big) + \big((1+H^n)P_0 r_x^{n,2}, P_0 r_x^{n,1} - r_x^{n,1}\big) + \gamma_0^{n,2}\big]$$
$$- \tfrac{1}{12}\big[\big((1+H^n)P_0 r_x^{n,1}, P_0 r_x^{n,2} - r_x^{n,2}\big) + \tfrac{1}{2}\gamma_1^{n,1}\big],$$

and so

$$|\beta_5^n| \leqslant \tfrac{C}{h^4}(\|\varepsilon^n\|^2 + \|e^n\|^2). \tag{4.50}$$

For β_6^n we have

$$\beta_6^n = \tfrac{1}{24}\big[(P\rho_x^{n,1}, P\rho_x^{n,3}) + \big((1+H^n)P_0r_x^{n,1}, P_0r_x^{n,3}\big)\big] + \tfrac{1}{36}\big[\|P\rho_x^{n,2}\|^2 + \big((1+H^n)P_0r_x^{n,2}, P_0r_x^{n,2}\big)\big],$$

which gives

$$\begin{aligned}\beta_6^n = {} & \tfrac{1}{24}\big[(P\rho_x^{n,1}, P\rho_x^{n,3}) + \big((1+H^n)P_0r_x^{n,1}, P_0r_x^{n,3}\big) + \|P\rho_x^{n,2}\|^2 + \big((1+H^n)P_0r_x^{n,2}, P_0r_x^{n,2}\big)\big] \\ & - \tfrac{1}{72}\big[\|P\rho_x^{n,2}\|^2 + \big((1+H^n)P_0r_x^{n,2}, P_0r_x^{n,2}\big)\big].\end{aligned}$$

Since now

$$(P\rho_x^{n,1}, \rho_x^{n,3}) + \big((1+H^n)P_0r_x^{n,1}, r_x^{n,3}\big) = -(\rho_x^{n,2}, P\rho_x^{n,2}) - \big((1+H^n)r_x^{n,2}, P_0r_x^{n,2}\big) + \gamma_1^{n,2},$$

we write

$$\beta_6^n = \beta_6^{n,1} + \beta_6^{n,2}, \tag{4.51}$$

where

$$\begin{aligned}\beta_6^{n,1} &= \tfrac{1}{24}\big[\big((1+H^n)P_0r_x^{n,1}, P_0r_x^{n,3} - r_x^{n,3}\big) + \big((1+H^n)P_0r_x^{n,2}, P_0r_x^{n,2} - r_x^{n,2}\big) + \gamma_1^{n,2}\big], \\ \beta_6^{n,2} &= -\tfrac{1}{72}\big[\|P\rho_x^{n,2}\|^2 + \big((1+H^n)P_0r_x^{n,2}, P_0r_x^{n,2}\big)\big].\end{aligned}$$

From (4.8), (4.13) and Lemma 2.1(ii) we see that

$$|\beta_6^{n,1}| \leqslant \tfrac{C}{h^5}(\|\varepsilon^n\|^2 + \|e^n\|^2). \tag{4.52}$$

Now, from Lemma 2.2 for sufficiently small h we infer that

$$\beta_6^{n,2} \leqslant -\tfrac{C_\alpha}{72}(\|P\rho_x^{n,2}\|^2 + \|P_0r_x^{n,2}\|^2), \tag{4.53}$$

where $C_\alpha = \min(1, \alpha/2)$. The quantity β_7^n is given by

$$\beta_7^n = -\tfrac{1}{72}\big[(P\rho_x^{n,2}, P\rho_x^{n,3}) + \big((1+H^n)P_0r_x^{n,2}, P_0r_x^{n,3}\big)\big],$$

and since, by (4.12),

$$(P\rho_x^{n,2}, \rho_x^{n,3}) + \big((1+H^n)P_0r_x^{n,2}, r_x^{n,3}\big) = \tfrac{1}{2}\gamma_2^{n,2},$$

we have

$$\beta_7^n = -\tfrac{1}{72}\big[\big((1+H^n)P_0r_x^{n,2}, P_0r_x^{n,3} - r_x^{n,3}\big) + \tfrac{1}{2}\gamma_2^{n,2}\big],$$

and, therefore, in view of Lemma 2.1(ii) and (4.13),

$$|\beta_7^n| \leqslant \tfrac{C}{h^6}(\|\varepsilon^n\|^2 + \|e^n\|^2). \tag{4.54}$$

Finally, β_8^n is given by

$$\beta_8^n = \tfrac{1}{24^2}\big[\|P\rho_x^{n,3}\|^2 + \big((1+H^n)P_0 r_x^{n,3}, P_0 r_x^{n,3}\big)\big].$$

Hence

$$\beta_8^n \leqslant \tfrac{1}{24^2}(\|\rho_x^{n,3}\|^2 + C'\|r_x^{n,3}\|^2),$$

and from (4.3), (4.4) and the inverse properties of S_h, $S_{h,0}$, we get that

$$\beta_8^n \leqslant \tfrac{C_0}{h^2}(\|P\rho_x^{n,2}\|^2 + \|P_0 r_x^{n,2}\|^2), \tag{4.55}$$

where C_0 is a constant independent of h and k. We conclude therefore from (4.45)–(4.55) that

$$\begin{aligned}\|f^n\|^2 + \big((1+H^n)g^n, g^n\big) \leqslant \|\varepsilon^n\|^2 + \big((1+H^n)e^n, e^n\big) + C_\lambda k(\|\varepsilon^n\|^2 + \|e^n\|^2)\\ + k^6\big(\lambda^2 C_0 - \tfrac{C_\alpha}{72}\big)(\|P\rho_x^{n,2}\|^2 + \|P_0 r_x^{n,2}\|^2).\end{aligned}$$

(v) *Stability, use of local error estimates and completion of the proof*
From the last inequality above, for $\lambda \leqslant \lambda_0 = \sqrt{C_\alpha/(72C_0)}$ it follows that

$$\|f^n\|^2 + \big((1+H^n)g^n, g^n\big) \leqslant \|\varepsilon^n\|^2 + \big((1+H^n)e^n, e^n\big) + C_\lambda k(\|\varepsilon^n\|^2 + \|e^n\|^2). \tag{4.56}$$

Therefore, using the equations (4.42), we see that

$$\begin{aligned}\|\varepsilon^{n+1}\|^2 + \big((1+H^{n+1})e^{n+1}, e^{n+1}\big) &= \|f^n\|^2 + 2(f^n, f_1^n + \delta_1^n) + \|f_1^n + \delta_1^n\|^2 + \big((1+H^{n+1})g^n, g^n\big)\\ &\quad + 2\big((1+H^{n+1})g^n, g_1^n + \delta_2^n\big) + \big((1+H^{n+1})(g_1^n + \delta_2^n), g_1^n + \delta_2^n\big).\end{aligned} \tag{4.57}$$

From (4.43), (4.44) for $0 \leqslant n \leqslant n^*$, we obtain

$$\|f^n\|\,\|f_1^n\| + \|g^n\|\,\|g_1^n\| \leqslant C_\lambda k(\|\varepsilon^n\|^2 + \|e^n\|^2),$$

and from Proposition 3.2, and (4.43), (4.44) that

$$\|f^n\|\,\|\delta_1^n\| + \|g^n\|\,\|\delta_2^n\| \leqslant C_\lambda k\big(\|\varepsilon^n\|^2 + \|e^n\|^2 + (h^{r-1} + k^4)^2\big).$$

Moreover, taking into account that

$$\big((1+H^{n+1})g^n, g^n\big) \leqslant \big((1+H^n)g^n, g^n\big) + Ck\|g^n\|^2,$$

we get from (4.57) in view of (4.56) and (2.10),

$$\|\varepsilon^{n+1}\|^2 + \big((1+H^{n+1})e^{n+1}, e^{n+1}\big) \leqslant \big(1 + \tfrac{C_\lambda}{C_\alpha}k\big)\Big(\|\varepsilon^n\|^2 + \big((1+H^n)e^n, e^n\big)\Big) + C'_\lambda k(h^{r-1} + k^4)^2,$$

for $0 \leqslant n \leqslant n^*$. Therefore from Gronwall's lemma it follows that

$$\|\varepsilon^n\|^2 + \big((1+H^n)e^n, e^n\big) \leqslant C_1\Big((\|\varepsilon^0\|^2 + \big((1+H^0)e^0, e^0\big)\Big) + C_2(h^{r-1} + k^4)^2,$$

where C_1, C_2 do not depend on n^*. Therefore, by (2.10)

$$\|\varepsilon^n\|^2 + \|e^n\|^2 \leqslant C_1(\|\varepsilon^0\|^2 + \|e^0\|^2) + C_2(h^{r-1} + k^4)^2,$$

i.e.,

$$\|\varepsilon^n\| + \|e^n\| \leqslant C(h^{r-1} + k^4),$$

for $0 \leqslant n \leqslant n^* + 1$, where the constant C does not depend on n^*. From the inverse inequalities of S_h, $S_{h,0}$ and the fact that $r \geqslant 3$ it follows that n^* was not maximal. Hence, we may take $n^* = M - 1$ and obtain the result of the theorem. □

5. Computational remarks

In this section we present results of numerical experiments that we performed in order to determine computationally the spatial and temporal rates of convergence of fully discrete schemes of the type analyzed in the previous sections. We also report on some computational results on the validity of the property (2.12) in the case of cubic and quartic splines.

(i) *Spatial rates of convergence*

As previously mentioned it is well known that in the case of first-order hyperbolic problems, the standard Galerkin method on a general quasiuniform mesh converges in L^2 with a spatial rate of $r-1$. We illustrate this for the problem at hand in the case of C^2 cubic splines ($r = 4$) defined on the quasiuniform mesh $0 = x_1 < x_2 < \ldots < x_{N+1} = 1$, where $x_{i+1} = x_i + h_i$, $1 \leqslant i \leqslant N$, N even, and $h_i = 0.8h$ if $i \equiv 0 \bmod 2$, $h_i = 1.2h$ if $i \equiv 1 \bmod 2$, and $h = 1/N$. We solve the system of shallow water equations (SW) with the addition of a suitable right-hand side and initial conditions, so that its exact solution is $\eta(x,t) = \exp(2t)(x + \cos(\pi x) + 2)$, $u(x,t) = \exp(-xt)\sin(\pi x)$. We integrate the semidiscrete problem in time for $0 \leqslant t \leqslant 1$ by the classical RK4 scheme taking small enough time steps so that the temporal error is negligible in comparison with the spatial one. Table 1 shows the numerical rates of convergence at $t = 1$ in the L^2 and L^∞ norms, and the H^1 seminorm as N increases, when $k/h = 1/20$. The L^2 and L^∞ rates are practically equal to 3, while the H^1 seminorm rate is practically 2. (The analogous experiment with C^4 quintic splines ($r = 6$) yielded numerical rates of convergence in L^2, L^∞ and H^1 approximately equal to 5, 5 and 4, respectively.)

In the case of uniform spatial mesh the numerical experiments suggest that the L^2 rate of convergence is $O(h^r)$, i.e., optimal. This was proved in Antonopoulos & Dougalis (2016) for the finite element space of continuous piecewise linear functions ($r = 2$) for (SW), and for general r in the case of periodic boundary conditions. Table 4.2 in Antonopoulos & Dougalis (2016) suggests that the numerical L^2 rates of convergence for C^2 cubic splines are also optimal, i.e., equal to 4. Here we illustrate this property in the case of C^4 quintic splines. Table 2 shows the associated numerical rates with $h = 1/N$, $k = 10^{-4}$ for the same test problem at $t = 1$. The L^2, L^∞, H^1 rates are observed to be close to 6, 6 and 5, for both components.

TABLE 1 *Spatial rates of convergence, cubic splines, quasiuniform mesh, $T = 1$, $\frac{k}{h} = \frac{1}{20}$, (a): η, (b): u*

N	L^2 error	Rate	L^∞ error	Rate	H^1 seminorm error	Rate
160	1.1057e−06	—	2.4537e−06	—	5.8016e−04	—
200	5.6700e−07	2.993	1.2600e−06	2.987	3.6898e−04	2.028
240	3.2848e−07	2.994	7.2837e−07	3.006	2.5514e−04	2.024
280	2.0700e−07	2.996	4.5857e−07	3.002	1.8686e−04	2.020
320	1.3875e−07	2.996	3.0686e−07	3.008	1.4273e−04	2.017
360	9.7479e−08	2.998	2.1566e−07	2.995	1.1256e−04	2.017
400	7.1102e−08	2.995	1.5726e−07	2.998	9.1058e−05	2.012
440	5.3431e−08	2.998	1.1834e−07	2.983	7.5161e−05	2.013

(*a*)

N	L^2 error	Rate	L^∞ error	Rate	H^1 seminorm error	Rate
160	2.3101e−08	—	4.9500e−08	—	1.1641e−05	—
200	1.1881e−08	2.980	2.4909e−08	3.078	7.4840e−06	1.980
240	6.8975e−09	2.983	1.4296e−08	3.046	5.2139e−06	1.982
280	4.3513e−09	2.989	8.8425e−09	3.116	3.8374e−06	1.989
320	2.9189e−09	2.990	5.8042e−09	3.153	2.9420e−06	1.990
360	2.0516e−09	2.994	4.0743e−09	3.005	2.3263e−06	1.994
400	1.4972e−09	2.990	2.9627e−09	3.024	1.8863e−06	1.990
440	1.1255e−09	2.994	2.2195e−09	3.030	1.5597e−06	1.994

(*b*)

TABLE 2 *Spatial rates of convergence, quintic splines $T = 1$, uniform mesh, $h = 1/N$, $k = 10^{-4}$, (a): η, (b): u*

N	L^2 error	Rate	L^∞ error	Rate	H^1 seminorm error	Rate
12	5.5379e−07	—	1.4510e−06	—	4.2901e−05	—
18	4.7013e−08	6.083	1.2278e−07	6.091	4.7221e−06	5.442
24	8.2765e−09	6.038	2.1654e−08	6.032	1.0096e−06	5.362
30	2.1511e−09	6.038	5.6238e−09	6.042	3.0752e−07	5.327
36	7.1581e−10	6.035	1.8738e−09	6.028	1.1680e−07	5.310

(*a*)

N	L^2 error	Rate	L^∞ error	Rate	H^1 seminorm error	Rate
12	9.2535e−09	—	2.1916e−08	—	4.4551e−07	—
18	7.8813e−10	6.075	1.8705e−09	6.070	5.7648e−08	5.043
24	1.4005e−10	6.005	3.3366e−10	5.992	1.3670e−08	5.003
30	3.6090e−11	6.077	8.7567e−11	5.995	4.4472e−09	5.032
36	1.1975e−11	6.051	2.9254e−11	6.014	1.7807e−09	5.020

(*b*)

(ii) *Temporal rates of convergence*

We turn now to the computational determination of the temporal accuracy, which is a harder exercise. We follow the technique proposed in Bona *et al.* (1995). We select a test problem with known exact solution and, for a fixed spatial grid (i.e., fixed h), we compute the numerical solutions up to $t = T$ with decreasing values of $k = T/M$ satisfying the stability condition. The L^2 error $E = E(T)$ ceases to decrease of course after a certain k when the temporal error becomes much smaller than the spatial one. Denote by $V^{M_{ref}}(h, k_{ref})$ the numerical solution (here $V = \eta$ or u) computed with a time step $k_{ref} = T/M_{ref}$, which is taken well below the threshold, after which E stabilizes. Therefore, the error of the approximation $V^{M_{ref}}(h, k_{ref})$ is almost purely spatial. We then compute a modified L^2 error for values of k much larger than k_{ref}, which is defined by

$$E^* = E^*(T) = \|V^M(h,k) - V^{M_{ref}}(h, k_{ref})\|,$$

where $T = Mk$. It is reasonable to expect that the subtraction $V^M(h,k) - V^{M_{ref}}(h, k_{ref})$ will essentially cancel the spatial error of $V^M(h,k)$ for a range of values of k; thus the temporal order of accuracy of the scheme may emerge from a sequence of computations of E^* with decreasing k in that range. The success of this procedure depends of course on finding an appropriate range of time steps depending on the chosen spatial grid, the solution of the test problem, k_{ref}, and the order of magnitude of the errors. For scalar problems and time-stepping schemes with weak stability conditions, such as those considered in Bona *et al.* (1995), this technique works rather well. In the case of systems of pde's and a high-order conditionally stable scheme, such as the one at hand, one has to experiment considerably; for example, we found that the test problems should be chosen so that the errors of all components of the system (here η and u) are of the same order of magnitude. The results of our experiments are shown in Table 3 for cubic and quintic splines on uniform and quasiuniform spatial meshes. The exact solution was taken now to be $\eta = \exp(-4t^2)(x + \cos(\pi x))$, $u = \exp(-tx)\sin(\pi x)$ and corresponding right-hand sides and initial conditions were found. The errors and temporal rates at $T = 1$ were computed with uniform mesh with $h = 1/N$ and the quasiuniform mesh defined in part (i) of this section. For each $M = T/k$ we show the modified L^2 error E^* and the corresponding numerical temporal rate of convergence. In all cases we took $M_{ref} = 600$; the L^2 error of $V^{M_{ref}}(h, k_{ref})$ is denoted by E_{ref}. The fourth-order temporal convergence emerges in all cases. (The experiments also gave fourth-order temporal convergence when RK4 was coupled with continuous, piecewise linear spatial discretizations. In all cases the spatial grid was taken coarse enough so that the spatial errors not be too small.)

(iii) *Remarks on the validity of* (2.12)

In closing, we report on a few numerical experiments we performed in order to check the validity of the hypothesis (2.12) in the case of C^2 cubic ($r = 4$) and C^3 quartic ($r = 5$) splines. To this effect we computed the $H^3(0,1)$ error $\|Pv - v\|_3$ for a C^∞ function, v and a function that was C^2 and piecewise C^3, i.e., so that $v \in H^3$ but $v \notin H^4$, and found its numerical rate of convergence as $h \to 0$ in the case of uniform and quasiuniform meshes. In all cases we found that $\|Pv - v\|_3$ was of $O(h^\alpha)$ with $\alpha > 0$, which suggests that $\|Pv\|_3 \leqslant C(v)$ for a function v that is at least in H^3. In the case of a smooth v (we took $v(x) = \sin(\pi x/2 + 1)$) the numerical rate of convergence of $\|Pv - v\|_3$ was found to be optimal, i.e., equal to one for cubic and to two for quartic splines, for uniform and quasiuniform meshes. (Results not shown.)

We then experimented with the C^2 function whose third derivative is given by

$$v'''(x) = \begin{cases} \exp(x), & 0 \leqslant x < 1/4 \\ \sin(\pi x), & 1/4 \leqslant x < 1/2, \\ \exp(-x), & 1/2 \leqslant x < 3/4, \\ \cos(\pi x), & 3/4 \leqslant x \leqslant 1. \end{cases}$$

TABLE 3 *Temporal rates of convergence, $T = 1$*

	η		u	
M	E^*	Rate	E^*	Rate
110	2.5095e−08	—	2.3825e−08	—
115	2.1068e−08	3.934	1.9943e−08	4.001
120	1.7814e−08	3.942	1.6825e−08	3.994
125	1.5163e−08	3.947	1.4296e−08	3.990
130	1.2987e−08	3.950	1.2225e−08	3.990
135	1.1188e−08	3.950	1.0515e−08	3.992
140	9.6915e−09	3.949	9.0931e−09	3.996
145	8.4378e−09	3.948	7.9024e−09	4.000
150	7.3808e−09	3.948	6.8997e−09	4.002
E_{ref}				
600	7.6301e-09	—	4.9031e-09	—

(a) Cubic splines, uniform mesh, $N = 60$

	η		u	
M	E^*	Rate	E^*	Rate
105	3.0062e−08	—	2.8915e−08	—
110	2.4975e−08	3.985	2.4051e−08	3.959
115	2.0905e−08	4.002	2.0182e−08	3.946
120	1.7619e−08	4.018	1.7073e−08	3.930
E_{ref}				
600	2.2953e-06	—	8.2945e-07	—

(b) Cubic splines, quasiuniform mesh, $N = 60$

	η		u	
M	E^*	Rate	E^*	Rate
60	2.7218e−07	—	2.6114e−07	—
65	1.9786e−07	3.984	1.9000e−07	3.974
70	1.4716e−07	3.995	1.4164e−07	3.963
75	1.1167e−07	3.999	1.0776e−07	3.963
80	8.6261e−08	4.001	8.3416e−08	3.967
85	6.7679e−08	4.002	6.5559e−08	3.973
95	4.3353e−08	4.004	4.2123e−08	3.977
100	3.5312e−08	4.000	3.4364e−08	3.969
E_{ref}				
600	2.2956e-09	—	6.7454e-10	—

(c) Quintic splines, uniform mesh, $N = 20$

	η		u	
M	E^*	Rate	E^*	Rate
80	8.5189e−08	—	8.6138e−08	—
85	6.6830e−08	4.004	6.7616e−08	3.994
90	5.3130e−08	4.013	5.3949e−08	3.950
95	4.2745e−08	4.023	4.3591e−08	3.943
100	3.4762e−08	4.031	3.5620e−08	3.937
105	2.8548e−08	4.036	2.9397e−08	3.935
E_{ref}				
600	1.3611e-08	—	4.6738e-09	—

(d) Quintic splines, quasiuniform grid, $N = 30$

TABLE 4 *Errors $Pv - v$ and orders of convergence, nonsmooth v, cubic splines, uniform mesh*

N	$\|Pv - v\|$	Order	$\|Pv - v\|_1$	Order	$\|Pv - v\|_2$	Order	$\|Pv - v\|_3$	Order	$\|Pv - v\|_3$	Order	$\|Pv - v\|_\infty$	Order
9	3.38e−06	—	7.35e−06	—	1.30e−04	—	5.45e−03	—	2.66e−01	—	3.63e−01	—
17	3.56e−07	3.542	2.25e−05	2.751	1.67e−03	1.856	1.76e−01	0.649	2.17e−01	0.811	1.10e−06	2.984
33	3.39e−08	3.546	4.20e−06	2.532	6.09e−04	1.525	1.25e−01	0.513	1.45e−01	0.607	1.48e−07	3.030
65	3.16e−09	3.498	7.76e−07	2.492	2.21e−04	1.493	8.91e−02	0.500	9.88e−02	0.564	1.93e−08	3.000
129	2.88e−10	3.496	1.40e−07	2.496	7.93e−05	1.498	6.32e−02	0.501	6.80e−02	0.545	2.47e−09	2.998
257	2.58e−11	3.498	2.51e−08	2.498	2.82e−05	1.499	4.48e−02	0.500	4.71e−02	0.532	3.13e−10	2.999
513	2.30e−12	3.499	4.45e−09	2.499	1.00e−05	1.499	3.17e−02	0.500	3.29e−02	0.522	3.94e−11	2.999
1025	2.04e−13	3.499	7.90e−10	2.499	3.54e−06	1.500	2.24e−02	0.500	2.30e−02	0.515	4.94e−12	3.000
2049	1.81e−14	3.500	1.40e−10	2.500	1.25e−06	1.500	1.59e−02	0.500	1.61e−02	0.511	6.18e−13	3.000
4097	1.60e−15	3.499	2.47e−11	2.500	4.44e−07	1.500	1.12e−02	0.500	1.14e−02	0.508	7.73e−14	3.001

(a) Nonsmooth v, cubic splines, uniform mesh

N	$\|Pv - v\|$	Order	$\|Pv - v\|_1$	Order	$\|Pv - v\|_2$	Order	$\|Pv - v\|_3$	Order	$\|Pv - v\|_3$	Order	$\|Pv - v\|_\infty$	Order
9	1.32e−06	—	3.24e−06	—	7.11e−05	—	5.86e−03	—	3.39e−01	—	4.46e−01	—
17	1.71e−07	3.216	1.17e−05	2.841	1.19e−03	2.505	1.31e−01	1.497	1.60e−01	1.614	5.18e−07	2.881
33	1.63e−08	3.541	1.96e−06	2.692	2.68e−04	2.250	5.84e−02	1.216	6.73e−02	1.304	6.95e−08	3.029
65	1.46e−09	3.554	3.48e−07	2.546	9.31e−05	1.559	4.05e−02	0.539	4.47e−02	0.603	8.81e−09	3.046
129	1.31e−10	3.526	6.20e−08	2.519	3.30e−05	1.513	2.86e−02	0.505	3.07e−02	0.548	1.11e−09	3.022
257	1.16e−11	3.512	1.10e−08	2.509	1.17e−05	1.506	2.03e−02	0.502	2.13e−02	0.532	1.39e−10	3.010
513	1.03e−12	3.506	1.95e−09	2.504	4.14e−06	1.503	1.43e−02	0.501	1.48e−02	0.522	1.75e−11	3.005
1025	9.10e−14	3.503	3.45e−10	2.502	1.46e−06	1.501	1.01e−02	0.501	1.04e−02	0.515	2.19e−12	3.003
2049	8.05e−15	3.501	6.09e−11	2.501	5.17e−07	1.501	7.17e−03	0.500	7.29e−03	0.511	2.74e−13	3.001
4097	7.19e−16	3.486	1.08e−11	2.499	1.83e−07	1.500	5.07e−03	0.500	5.13e−03	0.507	3.42e−14	2.999

(b) Nonsmooth v, quartic splines, quasiuniform mesh

The grids that we considered were uniform with $h = 1/N$ and quasiuniform with $x_{i+1} = x_i + h_i$, $1 \leqslant i \leqslant N$, $h_i = 3h/2$, if $i \equiv 0 \bmod 2$, and $h_i = h/2$, if $i \equiv 1 \bmod 2$, N odd and $h = 2/(2N-1)$. (We mainly took N odd so that the discontinuities of v did not occur at meshpoints. For N even we took $h = 1/N$.)

In Table 4(a) we show the results obtained in the case of cubic splines on a uniform grid with odd N. The order of convergence α was found to be approximately equal to 0.5. (The table also shows the errors and rates of convergence for a variety of other norms and seminorms.) The same rates of convergence were found (results not shown) in the case of the quasiuniform grid with odd N. (In the case of even N optimal-order results were found, i.e., $\alpha = 1$, for both uniform and quasiuniform meshes.)

In the case of quartic splines in all cases of uniform and quasiuniform meshes with odd or even N we observed, as expected, that α was approximately equal to 0.5, due to the restricted regularity of v. We just show in Table 4(b) the results on the quasiuniform grid with N odd.

Funding

Work supported by IACM-FORTH as part of the project 'Innovative Actions in Environmental Research and Development (PErAn)' (MIS 5002358), implemented under the 'Action for the strategic development of the Research and Technological sector' funded by the Operational Program 'Competitiveness, and Innovation' (NSRF 2014-2020) and cofinanced by Greece and the EU (European Regional Development Fund).

References

Antonopoulos, D. C. & Dougalis, V. A. (2016) Error estimates for the standard Galerkin-finite element method for the shallow water equations. *Math. Comp.*, **85**, 1143–1182.

Birkhoff, G., Schultz, M. H. & Varga, R. S. (1968) Piecewise Hermite interpolation in one and two variables with applications to partial differential equations. *Numer. Math.*, **11**, 232–256.

Bona, J. L., Dougalis, V. A., Karakashian, O. A. & McKinney, W. R. (1995) Conservative, high-order numerical schemes for the generalized Korteweg–de Vries equation. *Phil. Trans. R. Soc. Lond. A*, **351**, 107–164.

Bramble, J. H. & Schatz, A. H. (1976) Estimates for spline projections. *RAIRO Anal. Numér.*, **10**, 5–37.

Burman, E., Ern, A. & Fernandez, M. A. (2010) Explicit Runge–Kutta schemes and finite elements with symmetric stabilization for first-order linear pde systems. *SIAM J. Numer. Anal.*, **48**, 2019–2042.

DeBoor, C. & Fix, G. J. (1973) Spline approximation by quasiinterpolants. *J. Approx. Theory*, **8**, 19–45.

Douglas Jr., J., Dupont, T. & Wahlbin, L. (1975) Optimal L^∞ error estimates for Galerkin approximations to solutions of two-point boundary value problems. *Math. Comp.*, **29**, 475–483.

Dupont, T. (1973) Galerkin methods for first order hyperbolics: an example. *SIAM J. Numer. Anal.*, **10**, 890–899.

Dupont, T. (1974) Galerkin Methods for Modeling Gas Pipelines. Constructive and Computational Methods for Differential and Integral Equations (Sympos. Indiana Univ., Bloomington, Ind., 1974). Lecture Notes in Math., vol. 430, 112–130. Berlin: Springer.

Kubatko, E. J., Yeager, B. A. & Ketcheson, D. I. (2014) Optimal strong-stability preserving Runge–Kutta time discretizations for discontinuous Galerkin methods. *J. Sci. Computing*, **60**, 313–344.

Petcu, M. & Temam, R. (2011) The one dimensional shallow water equations with Dirichlet boundary conditions. *Discr. Cont. Dyn. Syst., Series S*, **4**, 209–222.

Qiu, J. & Zhang, Q. (2016) Stability, error estimate and limiters of discontinuous Galerkin methods. *Handbook of Numerical Analysis* (R. Abgrall & C.-W. Shu eds), vol. 17. Elsevier **34**, 147–171.

Schreiber, R. (1980) Finite element methods of high-order accuracy for singular two-point boundary value problems with nonsmooth solutions. *SIAM J. Numer. Anal.*, **17**, 547–566.

SCHULTZ, M. H. (1970) Error bounds for polynomial spline interpolation. *Math. Comp.*, **24**, 507–515.

SHU, C.-W. & OSHER, S. (1988) Efficient implementation of essentially non-oscillatory shock-capturing schemes. *J. Comp. Phys.*, **77**, 439–471.

THOMÉE, V. & WAHLBIN, L. B. (1983) Maximum-norm stability and error estimates in Galerkin methods for parabolic equations in one space variable. *Numer. Math.*, **41**, 345–371.

XING, Y. (2017) Numerical methods for the nonlinear shallow water equations. *Handbook of Numerical Analysis* (R. Abgrall & C.-W. Shu eds), vol. 18. Elsevier, 361–384.

ZHANG, Q. & SHU, C.-W. (2004) Error estimates to smooth solutions of Runge–Kutta discontinuous Galerkin methods for scalar conservation laws. *SIAM J. Numer. Anal.*, **42**, 641–646.

ZHANG, Q. & SHU, C.-W. (2014) Error estimates for the third order explicit Runge–Kutta discontinuous Galerkin method for a linear hyperbolic equation in one-dimension with discontinuous initial data. *Numer. Math.*, **126**, 703–740.