Journal of Mathematical Modeling
Vol. 10, No. 4, 2022, pp. 437-452. Research Article

JMM

A posteriori error analysis for the Cahn-Hilliard equation

Ali Mesforush[†][*], **Stig Larsson**[‡]

[†]*Faculty of Mathematical Sciences, Shahrood University of Thechnology, Shahrood, Iran*
[‡]*Department of Mathematical Sciences, Chalmers University of Technology, Goteborg, Sweden*
Email(s): ali.mesforush@shahroodut.ac.ir, stig@chalmers.se

Abstract. The Cahn-Hilliard equation is discretized by a Galerkin finite element method based on continuous piecewise linear functions in space and discontinuous piecewise constant functions in time. A posteriori error estimates are proved by using the methodology of dual weighted residuals.

Keywords: Cahn-Hilliard, finite element, error estimate, a posteriori, dual weighted residuals.
AMS Subject Classification 2010: 34A34, 65L05.

1 Introduction

We consider the Cahn-Hilliard equation

$$\begin{aligned} u_t - \Delta w &= 0 \quad \text{in } \Omega \times [0,T], \\ w + \varepsilon \Delta u - f(u) &= 0 \quad \text{in } \Omega \times [0,T], \\ \frac{\partial u}{\partial \nu} = 0, \ \frac{\partial w}{\partial \nu} &= 0 \quad \text{on } \partial\Omega \times [0,T], \\ u(\cdot,0) &= g_0 \quad \text{in } \Omega, \end{aligned} \tag{1}$$

where Ω is a polygonal domain in $\mathbf{R}^d, d = 1,2,3$, $u = u(x,t), w = w(x,t)$, $\Delta = \sum_{i=1}^d \frac{\partial^2}{\partial x_i^2}$, $u_t = \frac{\partial u}{\partial t}$, ν is the exterior unit normal to $\partial\Omega$, and $\varepsilon > 0$ is a small parameter. The Cahn-Hilliard equation is a model for phase separation and spinodal decomposition [3]. The nonlinearity f is the derivative of a double-well potential. A typical example is $f(u) = u^3 - u$.

We discretize (1) by a Galerkin finite element method, which is based on continuous piecewise linear functions with respect to x and discontinuous piecewise constant functions with respect to t. This numerical method is the same as the implicit Euler time stepping combined with spatial discretization by a standard finite element method.

[*]Corresponding author.
Received: 6 May 2022 / Revised: 24 June 2022 / Accepted: 24 June 2022
DOI: 10.22124/JMM.2022.22244.1960

We perform an a posteriori error analysis within the framework of dual weighted residuals [2]. If $J(u)$ is a given goal functional, this results in an error estimate essentially of the form

$$|J(u)-J(U)| \leq \sum_{n=1}^{N} \sum_{K\in \mathbf{T}_n} \left\{ \rho_{u,K}\omega_{u,K} + \rho_{w,K}\omega_{w,K} \right\} + \mathscr{R},$$

where U denotes the numerical solution and $\mathbf{T}_n$ is the spatial mesh at time level t_n. The terms $\rho_{u,K}, \rho_{w,K}$ are local residuals from the first and second equations in (1), respectively. The weights $\omega_{u,K}, \omega_{w,K}$ are derived from the solution of the linearized adjoint problem. The remainder $\mathscr{R}$ is quadratic in the error.

There is an extensive literature on numerical methods for the Cahn-Hilliard equation; see, for example, [5] and [4] for a priori error estimates. Adaptive methods based on a posteriori estimates are presented in [1] and [6]. However, these estimates are restricted to spatial discretization. We are not aware of any completely discerete a posteriori error analysis.

2 Preliminaries

Here we present the methodology of dual weighted residuals [2] in an abstract form.

Let $A(\cdot;\cdot)$ be a semilinear form; that is, it is nonlinear in the first and linear in the second variable, and $J(\cdot)$ be an output functional, not necessarily linear, defined on some function space V. Consider the variational problem: Find $u\in V$ such that

$$A(u;\psi) = 0 \quad \forall \psi \in V, \tag{2}$$

and the corresponding finite element problem: Find $u_h \in V_h \subset V$ such that

$$A(u_h;\psi_h) = 0 \quad \forall \psi_h \in V_h. \tag{3}$$

We suppose that the derivatives of A and J with respect to the first variable u up to order three exist and are denoted by

$$A'(u;\varphi), A''(u;\psi,\varphi), A'''(u;\xi,\psi,\varphi),$$

and

$$J'(u;\varphi), J''(u;\psi,\varphi), J'''(u;\xi,\psi,\varphi),$$

respectively, for increments φ, ψ, $\xi \in V$. Here we use the convention that the forms are linear in the variables after the semicolon.

We want to estimate $J(u)-J(u_h)$. Introduce the dual variable $z\in V$ and define the Lagrange functional

$$\mathscr{L}(u;z) := J(u) - A(u;z),$$

and seek the stationary points $(u,z)\in V\times V$ of $\mathscr{L}(\cdot;\cdot)$; that is,

$$\mathscr{L}'(u;z,\varphi,\psi) = J'(u;\varphi) - A'(u;z,\varphi) - A(u;\psi) = 0 \quad \forall \varphi,\psi \in V. \tag{4}$$

By choosing $\varphi = 0$, we retrieve (2). By taking $\psi = 0$, we identify the linearized adjoint equation to find $z\in V$ such that

$$J'(u;\varphi) - A'(u;z,\varphi) = 0 \quad \forall \varphi \in V. \tag{5}$$

The corresponding finite element problem is: Find $(u_h, z_h) \in V_h \times V_h$ such that

$$\mathscr{L}'(u_h; z_h, \varphi_h, \psi_h) = J'(u_h; \varphi_h) - A'(u_h; z_h, \varphi_h) - A(u_h; \psi_h) = 0 \quad \forall \varphi_h, \psi_h \in V_h. \tag{6}$$

By choosing $\varphi_h = 0$, we retrieve (3). By taking $\psi_h = 0$, we identify the linearized adjoint equation to find $z_h \in V_h$ such that

$$J'(u_h; \varphi_h) - A'(u_h; z_h, \varphi_h) = 0 \quad \forall \varphi_h \in V_h. \tag{7}$$

We quote three propositions from [2, Ch. 6].

Proposition 1. *Let $L(\cdot)$ be a three times differentiable functional defined on a vector space X, which has a stationary point $x \in X$, that is,*

$$L'(x; y) = 0 \quad \forall y \in X.$$

Suppose that on a finite dimensional subspace $X_h \subset X$ the corresponding Galerkin approximation,

$$L'(x_h; y_h) = 0 \quad \forall y_h \in X_h,$$

has a solution, $x_h \in X_h$. Then there holds the error representation

$$L(x) - L(x_h) = \tfrac{1}{2} L'(x_h; x - y_h) + \mathscr{R} \quad \forall y_h \in X_h,$$

with a remainder term $\mathscr{R}$, which is cubic in the error $e := x - x_h$,

$$\mathscr{R} := \tfrac{1}{2} \int_0^1 L'''(x_h + se; e, e, e) s(s-1) \, \mathrm{d}s.$$

Since

$$\mathscr{L}(u; z) - \mathscr{L}(u_h; z_h) = J(u) - J(u_h),$$

at stationary points $(u, z), (u_h, z_h)$, Proposition 1 yields the following result for the Galerkin approximation (3) of the variational equation (2).

Proposition 2. *For any solutions u and u_h of equations (2) and (3) we have the error representation*

$$J(u) - J(u_h) = \tfrac{1}{2} \rho(u_h; z - \varphi_h) + \tfrac{1}{2} \rho^*(u_h; z_h, u - \psi_h) + \mathscr{R}^{(3)} \quad \forall \varphi_h, \psi_h \in V_h,$$

where z and z_h are solutions of the adjoint problems (5) and (7) and

$$\begin{aligned} \rho(u_h; \cdot) &= -A(u_h; \cdot), \\ \rho^*(u_h; z_h, \cdot) &= J'(u_h; \cdot) - A'(u_h; z_h, \cdot), \end{aligned}$$

and, with $e_u = u - u_h$, $e_z = z - z_h$, the remainder is

$$\begin{aligned} \mathscr{R}^{(3)} = \tfrac{1}{2} \int_0^1 \Big(& J'''(u_h + se_u; e_u, e_u, e_u) - A'''(u_h + se_u; z_h + se_z, e_u, e_u, e_u) \\ & - 3A''(u_h + se_u; e_u, e_u, e_z) \Big) s(s-1) \, \mathrm{d}s. \end{aligned}$$

The forms $\rho(\cdot;\cdot)$ and $\rho^*(\cdot;\cdot,\cdot)$ are the residuals of (2) and (5), respectively. The remainder $\mathscr{R}^{(3)}$ is cubic in the error. The following proposition shows that the residuals are equal up to a quadratic remainder.

Proposition 3. *With the notation from above, we have*

$$\rho^*(u_h;z_h,u-\psi_h) = \rho(u_h;z-\varphi_h) + \delta\rho \quad \forall \varphi_h, \psi_h \in V_h,$$

with

$$\delta\rho = \int_0^1 \Big(A''(u_h+se_u;z_h+se_z,e_u,e_u) - J''(u_h+se_u;e_u,e_u) \Big)\,\mathrm{d}s.$$

Moreover, we have the simplified error representation

$$J(u) - J(u_h) = \rho(u_h;z-\varphi_h) + \mathscr{R}^{(2)} \quad \forall \varphi_h \in V_h,$$

with quadratic remainder

$$\mathscr{R}^{(2)} = \int_0^1 \Big(A''(u_h+se_u;z,e_u,e_u) - J''(u_h+se_u;e_u,e_u) \Big)\,\mathrm{d}s.$$

3 Galerkin discretization and dual problem

In this section, we apply the dual weighted residuals methodology to the Cahn-Hilliard equation (1). We denote $I=[0,T],\ Q=\Omega\times I$, and

$$\langle v,w\rangle_{\mathscr{D}} = \int_{\mathscr{D}} vw\,\mathrm{d}z, \quad \|v\|^2_{\mathscr{D}} = \int_{\mathscr{D}} v^2\,\mathrm{d}z,$$

for subsets $\mathscr{D}$ of Q or Ω with the relevant Lebesgue measure dz. Let $V=H^1(\Omega)$ and $\mathscr{W}=C^1([0,T],V)$. By multiplying the first equation by $\psi_u\in V$ and the second equation by $\psi_w\in V$, integrating over Ω and using Green's formula, we obtain the weak formulation: Find $u,w\in\mathscr{W}$ such that $u(0)=g_0$ and

$$\begin{aligned} \langle u_t,\psi_u\rangle_\Omega + \langle\nabla w,\nabla\psi_u\rangle_\Omega &= 0 \quad \forall\psi_u\in V, t\in[0,T], \\ \langle w,\psi_w\rangle_\Omega - \varepsilon\langle\nabla u,\nabla\psi_w\rangle_\Omega - \langle f(u),\psi_w\rangle_\Omega &= 0 \quad \forall\psi_w\in V, t\in[0,T]. \end{aligned} \tag{8}$$

Split the interval $I=[0,T]$ into subintervals $I_n=[t_{n-1},t_n)$ of lengths $k_n=t_n-t_{n-1}$,

$$0=t_0<t_1<\cdots<t_n<\cdots<t_N=T.$$

For each time level $t_n, n\geq 1$, let $\mathscr{V}_n$ be the space of continuous piecewise linear functions with respect to regular spatial meshes $\mathbf{T}_n=\{K\}$, which may vary from time level to time level. By extending the spatial meshes $\mathbf{T}_n$ as constant in time to the time slab $\Omega\times I_n$, we obtain meshes $\mathscr{T}_k$ of the space-time domain $Q=\Omega\times I$, which consist of $(d+1)$-dimensional prisms $Q^n_K := K\times\bar{I}_n$. Define the finite element space

$$\mathscr{V} := \Big\{ \varphi\colon \bar{Q}\to\mathbf{R} : \varphi(\cdot,t)|_{\bar{\Omega}} \in \mathscr{V}_n, t\in I_n,\ \varphi(x,\cdot)|_{I_n}\in\Pi_0, x\in\bar{\Omega} \Big\}.$$

Here, Π_0 denotes the polynomials of degree 0. For functions from this space and their continuous analogues, we define

$$v_n^+ = \lim_{t\downarrow t_n} v(t), \quad v_n = v_n^- = \lim_{t\uparrow t_n} v(t), \quad [v]_n = v_n^+ - v_n^-.$$

For all $u, w, \psi_u, \psi_w \in \mathscr{V}$ or $\mathscr{W}$, consider the semilinear form

$$A(u,w;\psi_u,\psi_w) = \sum_{n=1}^{N} \int_{I_n} \Big\{ \langle u_t, \psi_u\rangle_\Omega + \langle \nabla w, \nabla \psi_u\rangle_\Omega + \langle w, \psi_w\rangle_\Omega - \varepsilon\langle \nabla u, \nabla \psi_w\rangle_\Omega - \langle f(u), \psi_w\rangle_\Omega \Big\}\, \mathrm{d}t + \sum_{n=2}^{N} \langle [u]_{n-1}, \psi_{u,n-1}^+\rangle_\Omega + \langle u_0^+ - g_0, \psi_{u,0}^+\rangle_\Omega.$$

Solutions $u, w \in \mathscr{W}$ of (1) satisfy the variational problem

$$A(u,w;\psi_u,\psi_w) = 0 \quad \forall \psi_u, \psi_w \in \mathscr{W} \tag{9}$$

and the finite element problem can formulated: Find $U, W \in \mathscr{V}$ such that

$$A(U,W;\psi_u,\psi_w) = 0 \quad \forall \psi_u, \psi_w \in \mathscr{V}. \tag{10}$$

We now show that this is a standard time-stepping method. Since

$$U(t) = U_n = U_n^- = U_{n-1}^+, W(t) = W_n,$$

for $t \in I_n$, we have

$$\begin{aligned} A(U,W;\psi_u,\psi_w) = &\sum_{n=1}^{N} \int_{I_n} \Big\{ \langle \nabla W_n, \nabla \psi_u\rangle_\Omega + \langle W_n, \psi_w\rangle_\Omega - \varepsilon\langle \nabla U_n, \nabla \psi_w\rangle_\Omega - \langle f(U_n), \psi_w\rangle_\Omega \Big\}\, \mathrm{d}t \\ &+ \sum_{n=2}^{N} \langle U_n - U_{n-1}, \psi_{u,n-1}^+\rangle_\Omega + \langle U_1 - g_0, \psi_{u,0}^+\rangle_\Omega. \end{aligned} \tag{11}$$

By taking

$$\psi_u(t) = \begin{cases} \chi_u \in \mathscr{V}_n, & t \in I_n, \\ 0, & \text{otherwise}, \end{cases} \qquad \psi_w(t) = \begin{cases} \chi_w \in \mathscr{V}_n, & t \in I_n, \\ 0, & \text{otherwise}, \end{cases}$$

we see that (10) amounts to the implicit Euler time-stepping,

$$\begin{aligned} \langle U_0 - g_0, \chi_u\rangle_\Omega &= 0 \quad \forall \chi_u \in \mathscr{V}_1, \\ k_n\langle \nabla W_n, \nabla \chi_u\rangle_\Omega + \langle U_n - U_{n-1}, \chi_u\rangle_\Omega &= 0 \quad \forall \chi_u \in \mathscr{V}_n, n \geq 1, \\ \langle W_n, \chi_w\rangle_\Omega - \varepsilon\langle \nabla U_n, \nabla \chi_w\rangle_\Omega - \langle f(U_n), \chi_w\rangle_\Omega &= 0 \quad \forall \chi_w \in \mathscr{V}_n, n \geq 1. \end{aligned}$$

Now take a goal functional $J(u)$, which depends only on u, and set

$$\mathscr{L}(v;z) = J(u) - A(v;z),$$

where $v = (u,w), z = (z_u, z_w)$. With $\varphi = (\varphi_u, \varphi_w), \psi = (\psi_u, \psi_w)$, the stationary points are given by

$$\mathscr{L}'(v;z,\varphi,\psi) = J'(u;\varphi_u) - A'(v;z,\varphi) - A(v;\psi) = 0 \qquad \forall \varphi, \psi \in \mathscr{W} \times \mathscr{W}.$$

With $\psi = 0$ we obtain $A'(v;z,\varphi) = J'(u;\varphi_u)$, the adjoint problem. So we should compute $A'(u,w;z_u,z_w,\varphi_u,\varphi_w)$ and $J'(u;\varphi_u)$. To this end we write

$$\begin{aligned}A(u,w;\psi_u,\psi_w) = \langle u_t,\psi_u\rangle_Q + \langle \nabla w, \nabla\psi_u\rangle_Q + \langle w,\psi_w\rangle_Q - \varepsilon\langle\nabla u,\nabla\psi_w\rangle_Q \\ - \langle f(u),\psi_w\rangle_Q + \langle u(0)-g_0,\psi_u(0)\rangle_\Omega.\end{aligned}$$

Hence,

$$\begin{aligned}A'(u,w;z_u,z_w,\varphi_u,\varphi_w) =& \langle\varphi_{u,t},z_u\rangle_Q + \langle\nabla\varphi_w,\nabla z_u\rangle_Q + \langle\varphi_w,z_w\rangle_Q \\ &- \varepsilon\langle\nabla\varphi_u,\nabla z_w\rangle_Q - \langle\varphi_u,z_w\rangle_Q + \langle\varphi_u(0),z_u(0)\rangle_\Omega.\end{aligned}$$

By integration by parts in t,

$$\langle\varphi_{u,t},z_u\rangle_Q = -\langle\varphi_u,z_{u,t}\rangle_Q + \langle\varphi_u(T),z_u(T)\rangle_\Omega - \langle\varphi_u(0),z_u(0)\rangle_\Omega,$$

we obtain

$$\begin{aligned}A'(u,w;z_u,z_w,\varphi_u,\varphi_w) =& -\langle\varphi_u,z_{u,t}\rangle_Q + \langle\nabla\varphi_w,\nabla z_u\rangle_Q + \langle\varphi_w,z_w\rangle_Q \\ &+ \varepsilon\langle\nabla\varphi_u,\nabla z_w\rangle_Q - \langle\varphi_u,f'(u)z_w\rangle_Q + \langle\varphi_u(T),z_u(T)\rangle_\Omega.\end{aligned}$$

The adjoint problem is thus to find $z_u,z_w\in\mathscr{W}$ such that

$$\begin{aligned}\langle\varphi_u,-z_{u,t}\rangle_Q - \varepsilon\langle\nabla\varphi_u,\nabla z_w\rangle_Q - \langle\varphi_u,f'(u)z_w\rangle_Q + \langle\varphi_u(T),z_u(T)\rangle_\Omega \\ + \langle\nabla\varphi_w,\nabla z_w\rangle_Q + \langle\varphi_w,z_w\rangle_Q = J'(u;\varphi_u) \quad \forall\varphi_u,\varphi_w\in\mathscr{W}.\end{aligned} \tag{12}$$

We now specialize to the case of a linear goal functional of the form

$$J(u) = \langle u,g\rangle_Q + \langle u(T),g_T\rangle_\Omega,$$

for some $g\in L_2(Q)$ and $g_T\in L_2(\Omega)$. Then

$$J'(u;\varphi_u) = \langle\varphi_u,g\rangle_Q + \langle\varphi_u(T),g_T\rangle_\Omega. \tag{13}$$

The adjoint problem then becomes: Find $z_u,z_w\in\mathscr{W}$ such that

$$\begin{aligned}\langle\varphi_u,-z_{u,t}-f'(u)z_w-g\rangle_Q - \varepsilon\langle\nabla\varphi_u,\nabla z_w\rangle_Q + \langle\varphi_u(T),z_u(T)-g_T\rangle_\Omega = 0 \quad \forall\varphi_u\in\mathscr{W}, \\ \langle\varphi_w,z_w\rangle_Q + \langle\nabla\varphi_w,\nabla z_u\rangle_Q = 0 \quad \forall\varphi_w\in\mathscr{W}.\end{aligned} \tag{14}$$

The strong form of this is

$$\begin{aligned}-\partial_t z_u + \varepsilon\Delta z_w - f'(u)z_w &= g \quad &&\text{in } Q, \\ z_w - \Delta z_u &= 0 \quad &&\text{in } Q, \\ \frac{\partial z_u}{\partial\nu} = 0,\ \frac{\partial z_w}{\partial\nu} &= 0 \quad &&\text{on } \partial\Omega\times I, \\ z_u(T) &= g_T \quad &&\text{in } \Omega.\end{aligned} \tag{15}$$

4 A posteriori error estimates

From Proposition 3 we have the error representation

$$J(u)-J(U) = -A(U,W;z_u-\pi z_u, z_w-\pi z_w)+\mathscr{R}^{(2)}, \tag{16}$$

where $z=(z_u,z_w)$ is the solution of the adjoint problem (12) and $\pi z_u, \pi z_w \in \mathscr{V}$ are appropriate approximations to be defined below. The remainder is quadratic in the error.

In order to write this as a sum of local contributions we must rewrite $A(U,W;\psi_u,\psi_w)$ in (11). First we compute $\int_{I_n}\langle \nabla W, \nabla\psi_u\rangle_\Omega \,\mathrm{d}t$. By using Green's formula elementwise, we have

$$\int_{I_n}\langle \nabla W,\nabla\psi_u\rangle_\Omega \,\mathrm{d}t = \int_{I_n}\sum_{K\in\mathbf{T}_n}\langle \nabla W,\nabla\psi_u\rangle_K \,\mathrm{d}t = \int_{I_n}\sum_{K\in\mathbf{T}_n} -\langle \Delta W,\psi_u\rangle_K \,\mathrm{d}t + \int_{I_n}\sum_{K\in\mathbf{T}_n}\langle \partial_\nu W,\psi_u\rangle_{\partial K}\,\mathrm{d}t,$$

where $\partial_\nu W = \nu\cdot\nabla W$. We divide the boundary $\partial K \in \mathbf{T}_n$ into two parts: internal edges, denoted by $\mathscr{E}_I^n$, and edges on the boundary $\partial\Omega$, denoted by $\mathscr{E}_{\partial\Omega}^n$. So we get, with [ ] denoting the jump across the edge,

$$\begin{aligned}\int_{I_n}\sum_{K\in\mathbf{T}_n}\langle \partial_\nu W,\psi_u\rangle_{\partial K}\,\mathrm{d}t &= \int_{I_n}\sum_{E\in\mathscr{E}_I^n}\langle \partial_\nu W,\psi_u\rangle_E\,\mathrm{d}t + \int_{I_n}\sum_{E\in\mathscr{E}_{\partial\Omega}^n}\langle \partial_\nu W,\psi_u\rangle_E\,\mathrm{d}t\\ &= \int_{I_n}\sum_{K\in\mathbf{T}_n} -\tfrac{1}{2}\langle[\partial_\nu W],\psi_u\rangle_{\partial K\setminus\partial\Omega}\,\mathrm{d}t + \int_{I_n}\sum_{K\in\mathbf{T}_n}\langle \partial_\nu W,\psi_u\rangle_{\partial K\cap\partial\Omega}\,\mathrm{d}t.\end{aligned}$$

Let ∂_x denote the spatial boundary and define $\partial_x Q = \partial\Omega\times I$ and $\partial_x Q_K^n = \partial K\times I_n$. Hence,

$$\int_{I_n}\langle \nabla W,\nabla\psi_u\rangle_\Omega\,\mathrm{d}t = \sum_{K\in\mathbf{T}_n}\Big\{-\langle \Delta W,\psi_u\rangle_{Q_K^n} - \tfrac{1}{2}\langle[\partial_\nu W],\psi_u\rangle_{\partial_x Q_K^n\setminus\partial_x Q} + \langle \partial_\nu W,\psi_u\rangle_{\partial_x Q_K^n\cap\partial_x Q}\Big\},$$

and in the same way

$$\varepsilon\int_{I_n}\langle \nabla U,\nabla\psi_w\rangle_\Omega\,\mathrm{d}t = \sum_{K\in\mathbf{T}_n}\Big\{-\varepsilon\langle \Delta U,\psi_w\rangle_{Q_K^n} - \tfrac{1}{2}\varepsilon\langle[\partial_\nu U],\psi_w\rangle_{\partial_x Q_K^n\setminus\partial_x Q} + \varepsilon\langle \partial_\nu U,\psi_w\rangle_{\partial_x Q_K^n\cap\partial_x Q}\Big\}.$$

Note that $\Delta W = \Delta U = 0$ on Q_K^n for piecewise linear functions, but we find it instructive to keep these terms. Inserting this into (11) and noting that

$$\int_{I_n}\langle W,\psi_w\rangle_\Omega\,\mathrm{d}t = \sum_{K\in\mathbf{T}_n}\langle W,\psi_w\rangle_{Q_K^n},$$

and

$$\int_{I_n}\langle f(U),\psi_w\rangle_\Omega\,\mathrm{d}t = \sum_{K\in\mathbf{T}_n}\langle f(U),\psi_w\rangle_{Q_K^n},$$

gives

$$\begin{aligned}A(U,W;\psi_u,\psi_w) = \sum_{n=1}^{N}\sum_{K\in\mathbf{T}_n}\Big\{&-\langle \Delta W,\psi_u\rangle_{Q_K^n} + \langle \varepsilon\Delta U + W - f(U),\psi_w\rangle_{Q_K^n}\\ &-\tfrac{1}{2}\langle[\partial_\nu W],\psi_u\rangle_{\partial_x Q_K^n\setminus\partial_x Q} + \tfrac{1}{2}\varepsilon\langle[\partial_\nu U],\psi_w\rangle_{\partial_x Q_K^n\setminus\partial_x Q}\\ &+\langle \partial_\nu W,\psi_u\rangle_{\partial_x Q_K^n\cap\partial_x Q} - \varepsilon\langle \partial_\nu U,\psi_w\rangle_{\partial_x Q_K^n\cap\partial_x Q} + \langle [U]_{n-1},\psi_{u,n-1}^{+}\rangle_K\Big\},\end{aligned}$$

where we have set $U_0^- = g_0$ for simplicity. Hence (16) becomes

$$\begin{aligned} J(u) - J(U) = \sum_{n=1}^{N} \sum_{K \in \mathbf{T}_n} \Big\{ & \langle R_u, z_u - \pi z_u \rangle_{Q_K^n} + \langle R_w, z_w - \pi z_w \rangle_{Q_K^n} \\ & + \langle r_u, z_u - \pi z_u \rangle_{\partial_x Q_K^n} + \langle r_w, z_w - \pi z_w \rangle_{\partial_x Q_K^n} \\ & - \langle [U]_{n-1}, (z_u - \pi z_u)_{n-1}^+ \rangle_K \Big\} + \mathscr{R}^{(2)}, \end{aligned} \tag{17}$$

with the interior residuals

$$R_u = \Delta W, \quad R_w = -\varepsilon \Delta U - W + f(U),$$

the edge residuals

$$r_w|_\Gamma = \begin{cases} -\frac{1}{2}\varepsilon[\partial_\nu U], & \Gamma \subset \partial_x Q_K^n \setminus \partial_x Q, \\ 0, & \text{otherwise}, \end{cases}$$

$$r_u|_\Gamma = \begin{cases} \frac{1}{2}[\partial_\nu W], & \Gamma \subset \partial_x Q_K^n \setminus \partial_x Q, \\ 0, & \text{otherwise}, \end{cases}$$

and the boundary residuals

$$r_w|_\Gamma = \begin{cases} \varepsilon \partial_\nu U, & \Gamma \subset \partial_x Q_K^n \cap \partial_x Q, \\ 0, & \text{otherwise}, \end{cases}$$

$$r_u|_\Gamma = \begin{cases} -\partial_\nu W, & \Gamma \subset \partial_x Q_K^n \cap \partial_x Q, \\ 0, & \text{otherwise}. \end{cases}$$

Here the subscript u refers to residuals from the first equation in (8) and the subscript w to residuals from the second equation.

We now define $\pi z_u, \pi z_w \in \mathscr{V}$. Let

$$(P_n v)(t) = \frac{1}{k_n} \int_{I_n} v(s)\, \mathrm{d}s$$

be the orthogonal projector onto constants. Let $\pi_n : C(\bar{\Omega}) \to \mathscr{V}_n$ be the nodal interpolator; that is, it is defined by

$$(\pi_n v)(a) = v(a),$$

for all nodal points a in $\mathbf{T}_n$. Then we define $\pi : C(\bar{Q}) \to \mathscr{V}$ by $\pi v|_{I_n} = P_n \pi_n v$. Since R_u, R_w, r_u, and r_w are piecewise constant in t, we have

$$\begin{aligned} J(u) - J(U) = \sum_{n=1}^{N} \sum_{K \in \mathbf{T}_n} \Big\{ & \langle R_u, P_n(z_u - \pi_n z_u) \rangle_{Q_K^n} + \langle R_w, P_n(z_w - \pi_n z_w) \rangle_{Q_K^n} \\ & + \langle r_u, P_n(z_u - \pi_n z_u) \rangle_{\partial_x Q_K^n} + \langle r_w, P_n(z_w - \pi_n z_w) \rangle_{\partial_x Q_K^n} \\ & - \langle [U]_{n-1}, (z_u - \pi z_u)_{n-1}^+ \rangle_K \Big\} + \mathscr{R}^{(2)}. \end{aligned} \tag{18}$$

Applying the Cauchy-Schwartz inequality to each term gives

$$|J(u)-J(U)| \leq \sum_{n=1}^{N}\sum_{K\in\mathbf{T}_n}\Big\{\|R_u\|_{Q_K^n}\|P_n(z_u-\pi_n z_u)\|_{Q_K^n}+h_K^{-\frac{1}{2}}\|r_u\|_{\partial_x Q_K^n}h_K^{\frac{1}{2}}\|P_n(z_u-\pi_n z_u)\|_{\partial_x Q_K^n}$$
$$+\|R_w\|_{Q_K^n}\|P_n(z_w-\pi_n z_w)\|_{Q_K^n}+h_K^{-\frac{1}{2}}\|r_w\|_{\partial_x Q_K^n}h_K^{\frac{1}{2}}\|P_n(z_w-\pi_n z_w)\|_{\partial_x Q_K^n}$$
$$+k_n^{-\frac{1}{2}}\|[U]_{n-1}\|_K k_n^{\frac{1}{2}}\|(z_u-\pi z_u)_{n-1}^{+}\|_K\Big\}+|\mathscr{R}^{(2)}|.$$

Here $h_K=\mathrm{diam}(K)$. For $a,b,c,d\geq 0$ we have

$$(ab+cd)\leq (a^2+c^2)^{\frac{1}{2}}(b^2+d^2)^{\frac{1}{2}}.$$

We use this inequality for each term in the previous inequality and set

$$\rho_{u,K}=\Big(\|R_u\|_{Q_K^n}^2+h_K^{-1}\|r_u\|_{\partial_x Q_K^n}^2\Big)^{\frac{1}{2}},$$
$$\omega_{u,K}=\Big(\|P_n(z_u-\pi_n z_u)\|_{Q_K^n}^2+h_K\|P_n(z_u-\pi_n z_u)\|_{\partial_x Q_K^n}^2\Big)^{\frac{1}{2}},$$
$$\rho_{w,K}=\Big(\|R_w\|_{Q_K^n}^2+h_K^{-1}\|r_w\|_{\partial_x Q_K^n}^2\Big)^{\frac{1}{2}},$$
$$\omega_{w,K}=\Big(\|P_n(z_w-\pi_n z_w)\|_{Q_K^n}^2+h_K\|P_n(z_w-\pi_n z_w)\|_{\partial_x Q_K^n}^2\Big)^{\frac{1}{2}},$$
$$\rho_K=\Big(k_n^{-1}\|[U]^{n-1}\|_K^2\Big)^{\frac{1}{2}},$$
$$\omega_K=\Big(k_n\|(z_u-\pi z_u)_{n-1}^{+}\|_K^2\Big)^{\frac{1}{2}}.$$

Note that, since $R_u=\Delta W=0$ for piecewise linear functions, the first term in $\rho_{u,K}$ and $\omega_{u,K}$ can actually be removed. So we have

$$|J(u)-J(U)|\leq \sum_{n=1}^{N}\sum_{K\in\mathbf{T}_n}\Big\{\rho_{u,K}\omega_{u,K}+\rho_{w,K}\omega_{w,K}+\rho_K\omega_K\Big\}+|\mathscr{R}^{(2)}|.$$

We have proved the following theorem:

Theorem 1. *We have the a posteriori error estimate*

$$|J(u)-J(U)|\leq \sum_{n=1}^{N}\sum_{K\in\mathbf{T}_n}\Big\{\rho_{u,K}\omega_{u,K}+\rho_{w,K}\omega_{w,K}+\rho_K\omega_K\Big\}+|\mathscr{R}^{(2)}|. \tag{19}$$

Note that on each space-time cell Q_K^n, the terms $\rho_{u,K}\omega_{u,K}$ and $\rho_{w,K}\omega_{w,K}$ can be used to control the spatial mesh and the term $\rho_K\omega_K$ to control the time step k_n in an adaptive algoritheorem ; see [2]. We do not pursue this here.

In the following we want to obtain a weight free a posteriori error estimate where the weights in (19) are replaced by a global stability constant. We need the following interpolation error estimate, see [2, lemma 9.4].

Lemma 1. *With π and π_n as defined as before, there holds*

$$\|P_n(z-\pi_n z)\|_{Q_K^n}+h_K^{\frac{1}{2}}\|P_n(z-\pi_n z)\|_{\partial_x Q_K^n}\leq Ch_K^2\|\mathrm{D}^2 z\|_{Q_K^n}, \tag{20}$$

$$\|z(t_{n-1})-P_n z\|_K\leq Ck_n^{\frac{1}{2}}\|\partial_t z\|_{Q_K^n}. \tag{21}$$

Here $\|\mathrm{D}^2 z\|_{Q_K^n}$ *denotes the seminorm* $\left(\sum_{|\alpha|=2}\|D^\alpha z\|^2_{Q_K^n}\right)^{\frac{1}{2}}$.

In the following we assume that $J(\cdot)$ is a linear functional given by (13) and Ω is such that we have the elliptic regularity estimate

$$\|\mathrm{D}^2 v\|_\Omega\leq C\|\Delta v\|_\Omega \quad \forall v\in H^2(\Omega) \text{ with } \frac{\partial v}{\partial \nu}\Big|_\Gamma=0. \tag{22}$$

We also assume a global bound for $f'(u)$, which is reasonable since it is known that $\|u\|_{L_\infty(Q)}\leq C$ (c.f. [5]).

In particular, with

$$g=\frac{u-U}{\|u-U\|_Q} \text{and } g_T=\frac{(u_N-U_N)}{\|u_N-U_N\|_\Omega}$$

the following theorem provides bounds for the norms of the error, $\|u-U\|_Q$ and $\|u_N-U_N\|_\Omega$.

Theorem 2. *Assume that $\|f'(u)\|_{L_\infty}\leq\beta$ and that (22) holds. Let z_u, z_w be the solutions of (15). Then there is $C=C(\beta)$ such that the following a posteriori error estimates hold.*
(i) Let $g\in L_2(Q)$ with $\|g\|_Q=1$ and $g_T=0$. Then

$$|\langle u-U,g\rangle_Q|\leq CC_S\sum_{n=1}^N\sum_{K\in\mathbf{T}_n}\left\{h_K^4(\rho_{u,K}^2+\rho_{w,K}^2)+(h_K^4+k_n^2)\rho_K^2\right\}^{\frac{1}{2}}+|\mathscr{R}^{(2)}|, \tag{23}$$

where

$$C_S=\sup_{g\in L_2(Q)}\frac{\left(\|\mathrm{D}^2 z_u\|_Q^2+\|\partial_t z_u\|_Q^2+\|\mathrm{D}^2 z_w\|_Q^2\right)^{\frac{1}{2}}}{\|g\|_Q}.$$

(ii) Let $g_T\in L_2(\Omega)$ with $\|g_T\|_\Omega=1$ and $g=0$. Then

$$|\langle u-U,g_T\rangle_\Omega|\leq CC_S\sum_{n=1}^N\sum_{K\in\mathbf{T}_n}\left\{h_K^4(\rho_{u,K}^2+\sigma_n^{-1}\rho_{w,K}^2+\sigma_n^{-1}\rho_K^2)+k_n^2\sigma^{-1}\rho_K^2\right\}^{\frac{1}{2}}+|\mathscr{R}^{(2)}|, \tag{24}$$

where $\sigma(t)=T-t$,

$$\sigma_n=\begin{cases}\sigma(t_n)=T-t_n, & n=1,\ldots,N-1,\\ k_N, & n=N,\end{cases}$$

and

$$C_S=\sup_{g_T\in L_2(\Omega)}\left(\varepsilon^{-1}\max_I\|z_u\|_\Omega^2+\varepsilon^{-1}\|z_w\|_Q^2+\|\mathrm{D}^2 z_u\|_Q^2+\|\sigma^{\frac{1}{2}}\partial_t z_u\|_Q^2+\varepsilon^2\|\sigma^{\frac{1}{2}}\mathrm{D}^2 z_w\|_Q^2\right)^{\frac{1}{2}}/\|g_T\|_\Omega.$$

Proof. Part (i). From Theorem 1 we have

$$\omega_{u,K} = \Big(\|P_n(z_u-\pi_n z_u)\|^2_{Q^n_K} + h_K\|P_n(z_u-\pi_n z_u)\|^2_{\partial_x Q^n_K}\Big)^{\frac12} \le Ch_K^2\|\mathrm{D}^2 z_u\|_{Q^n_K},$$
$$\omega_{w,K} = \Big(\|P_n(z_w-\pi_n z_w)\|^2_{Q^n_K} + h_K\|P_n(z_w-\pi_n z_w)\|^2_{\partial_x Q^n_K}\Big)^{\frac12} \le Ch_K^2\|\mathrm{D}^2 z_w\|_{Q^n_K},$$

and

$$\begin{aligned}\omega_K &= k_n^{\frac12}\|(z_u-\pi_n z_u)^+_{n-1}\|_K \le k_n^{\frac12}\|P_n(z_u-\pi_n z_u)\|_K + k_n^{\frac12}\|z_u(t_{n-1})-P_n z_u\|_K\\ &\le Ch_K^2\|\mathrm{D}^2 z_u\|_{Q^n_K} + Ck_n\|\partial_t z_u\|_{Q^n_K} + |\mathscr{R}^{(2)}|.\end{aligned}$$

Hence,

$$\begin{aligned}|\langle u-U,g\rangle_Q| &\le \sum_{n=1}^N\sum_{K\in\mathbf{T}_n}\Big\{\rho_{u,K}\omega_{u,K}+\rho_{w,K}\omega_{w,K}+\rho_K\omega_K\Big\}\\ &\le \sum_{n=1}^N\sum_{K\in\mathbf{T}_n}\Big\{Ch_K^2\rho_{u,K}\|\mathrm{D}^2 z_u\|_{Q^n_K}+Ch_K^2\rho_{w,K}\|\mathrm{D}^2 z_w\|_{Q^n_K}+\rho_K(Ch_K^2\|\mathrm{D}^2 z_u\|_{Q^n_K}+Ck_n\|\partial_t z_u\|_{Q^n_K})\Big\},\end{aligned}$$

and the desired estimate (23) follows by the Cauchy-Schwartz inequality

$$\begin{aligned}\sum_{n=1}^N\sum_{K\in\mathbf{T}_n}h_K^2\rho_{u,K}\|\mathrm{D}^2 z_u\|_{Q^n_K} &\le \Big(\sum_{n=1}^N\sum_{K\in\mathbf{T}_n}h_K^4\rho_{u,K}^2\Big)^{\frac12}\Big(\sum_{n=1}^N\sum_{K\in\mathbf{T}_n}h_K^2\rho_{u,K}\|\mathrm{D}^2 z_u\|^2_{Q^n_K}\Big)^{\frac12}\\ &= \Big(\sum_{n=1}^N\sum_{K\in\mathbf{T}_n}h_K^4\rho_{u,K}^2\Big)^{\frac12}\|\mathrm{D}^2 z_u\|_Q \le C_S\Big(\sum_{n=1}^N\sum_{K\in\mathbf{T}_n}h_K^4\rho_{u,K}^2\Big)^{\frac12}\|g\|_Q,\end{aligned}$$

and similarly for the other terms.

Part (ii). The previous bound for $\sum_{n=1}^N\sum_{K\in\mathbf{T}_n}\rho_{u,K}\omega_{u,K}$ applies here also. Consider then

$$\sum_{n=1}^N\sum_{K\in\mathbf{T}_n}\rho_{w,K}\omega_{w,K} \le \sum_{n=1}^{N-1}\sum_{K\in\mathbf{T}_n}\rho_{w,K}Ch_K^2\|\mathrm{D}^2 z_w\|_{Q^n_K} + \sum_{K\in\mathbf{T}_N}\rho_{w,K}\omega_{w,K}.$$

Here,

$$\begin{aligned}\sum_{n=1}^{N-1}\sum_{K\in\mathbf{T}_n}\rho_{w,K}Ch_K^2\|\mathrm{D}^2 z_w\|_{Q^n_K} &= \sum_{n=1}^{N-1}\sum_{K\in\mathbf{T}_n}\rho_{w,K}Ch_K^2\|\sigma^{-\frac12}\sigma^{\frac12}\mathrm{D}^2 z_w\|_{Q^n_K}\\ &\le C\sum_{n=1}^{N-1}\sum_{K\in\mathbf{T}_n}\rho_{w,K}\sigma_n^{-\frac12}h_K^2\|\sigma^{\frac12}\mathrm{D}^2 z_w\|_{Q^n_K}\\ &\le C\Big(\sum_{n=1}^{N-1}\sum_{K\in\mathbf{T}_n}\sigma_n^{-1}h_K^4\rho_{w,K}^2\Big)^{\frac12}\Big(\sum_{n=1}^{N-1}\sum_{K\in\mathbf{T}_n}\|\sigma^{\frac12}\mathrm{D}^2 z_w\|^2_{Q^n_K}\Big)^{\frac12}\\ &\le C\Big(\sum_{n=1}^{N-1}\sum_{K\in\mathbf{T}_n}\sigma_n^{-1}h_K^4\rho_{w,K}^2\Big)^{\frac12}\|\sigma_n^{\frac12}\mathrm{D}^2 z_w\|_Q\\ &\le C_S C\Big(\sum_{n=1}^{N-1}\sum_{K\in\mathbf{T}_n}\sigma_n^{-1}h_K^4\rho_{w,K}^2\Big)^{\frac12}\|g_T\|_\Omega.\end{aligned}$$

The term with $n=N$ is special. We go back to (18) and replace it by

$$\sum_{K\in\mathbf{T}_N}\langle R_w, z_w-\pi_N z_w\rangle_{Q_K^N}=\sum_{K\in\mathbf{T}_N}\Big\langle R_w,(I-\pi_N)\int_{I_N} z_w\,\mathrm{d}t\Big\rangle_K\le\sum_{K\in\mathbf{T}_N}\|R_w\|_K C h_K^2\Big\|\mathrm{D}^2\int_{I_N} z_w\,\mathrm{d}t\Big\|_K.$$

Here, by the regularity estimate (22), $\varepsilon\Delta z_w=\partial_t z_u+f'(u)z_w$ from the first equation in (15), and $\|f'(u)\|_{L_\infty}\le\beta$, we have

$$\begin{aligned}\Big\|\mathrm{D}^2\int_{I_N} z_w\,\mathrm{d}t\Big\|_K&\le C\Big\|\int_{I_N}\Delta z_w\,\mathrm{d}t\Big\|_K=C\varepsilon^{-1}\Big\|\int_{I_N}(\partial_t z_u+f'(u)z_w)\,\mathrm{d}t\Big\|_K\\&\le C\varepsilon^{-1}\Big(\|z_u(t_N)\|_K+\|z_u(t_{N-1})\|_K+\beta k_N^{\frac12}\|z_w\|_{Q_K^N}\Big).\end{aligned}$$

Hence, since $\rho_{w,K}=\|R_w\|_{Q_K^N}=k_N^{\frac12}\|R_w\|_K$, we have

$$\begin{aligned}\sum_{K\in\mathbf{T}_N}\langle R_w, z_w-\pi_N z_w\rangle_{Q_K^N}&\le\sum_{K\in\mathbf{T}_N}\|R_w\|_K C h_K^2\varepsilon^{-1}\Big(\|z_u(t_N)\|_K+\|z_u(t_{N-1})\|_K+k_N^{\frac12}\|z_w\|_{Q_K^N}\Big)\\&=C\varepsilon^{-1}\sum_{K\in\mathbf{T}_N}k_N^{-\frac12}h_K^2\rho_{w,K}\Big(\|z_u(t_N)\|_K+\|z_u(t_{N-1})\|_K+k_N^{\frac12}\|z_w\|_{Q_K^N}\Big)\\&\le C\varepsilon^{-1}\Big(\sum_{K\in\mathbf{T}_N}k_N^{-1}h_K^4\rho_{w,K}^2\Big)^{\frac12}\Big(\|z_u(t_N)\|_\Omega+\|z_u(t_{N-1})\|_\Omega+k_N^{\frac12}\|z_w\|_Q\Big)\\&\le C\varepsilon^{-1}C_S\|g_T\|_\Omega\Big(\sum_{K\in\mathbf{T}_N}\sigma_N^{-1}h_K^4\rho_{w,K}^2\Big)^{\frac12},\end{aligned}$$

where we have used $\sigma_N=k_N$. So we have

$$\sum_{n=1}^N\sum_{K\in\mathbf{T}_n}\rho_{w,K}\omega_{w,K}\le CC_S\|g_T\|_\Omega\Big(\sum_{n=1}^N\sum_{K\in\mathbf{T}_n}\sigma_n^{-1}h_K^4\rho_{w,K}^2\Big)^{\frac12}. \tag{25}$$

Now we compute $\sum_{n=1}^N\sum_{K\in\mathbf{T}_n}\rho_K\omega_K$. For $K\in\mathbf{T}_N$ we use

$$\begin{aligned}\omega_K&=k_N^{\frac12}\|(z_u-\pi z_u)_{N-1}^+\|_K\le k_N^{\frac12}\|P_N(z_u-\pi_N z_u)\|_K+k_N^{\frac12}\|z_u(t_{N-1})-P_N z_u\|_K\\&=\|P_N(z_u-\pi_N z_u)\|_{Q_K^N}+k_N^{\frac12}\|z_u(t_{N-1})-P_N z_u\|_K\le Ch_K^2\|\mathrm{D}^2 z_u\|_{Q_K^N}+k_N^{\frac12}\|z_u(t_{N-1})-P_N z_u\|_K.\end{aligned}$$

Then we have

$$\begin{aligned}\sum_{n=1}^N\sum_{K\in\mathbf{T}_n}\rho_K\omega_K&=C\sum_{n=1}^N\sum_{K\in\mathbf{T}_n}\rho_K h_K^2\|\mathrm{D}^2 z_u\|_{Q_K^n}+C\sum_{n=1}^{N-1}\sum_{K\in\mathbf{T}_n}\rho_K k_n\sigma_n^{-\frac12}\|\sigma^{\frac12}\partial_t z_u\|_{Q_K^n}\\&\quad+\sum_{K\in\mathbf{T}_N}\rho_K k_N^{\frac12}\|z_u(t_{N-1})-P_N z_u\|_K\\&\le C\Big(\sum_{n=1}^N\sum_{K\in\mathbf{T}_n}h_K^4\rho_K^2\Big)^{\frac12}\|\mathrm{D}^2 z_u\|_Q+C\Big(\sum_{n=1}^{N-1}\sum_{K\in\mathbf{T}_n}\rho_K^2k_n^2\sigma_n^{-1}\Big)^{\frac12}\|\sigma^{\frac12}\partial_t z_u\|_Q\\&\quad+C\Big(\sum_{K\in\mathbf{T}_N}k_N\rho_K^2\Big)^{\frac12}\|z_u(t_{N-1})-P_N z_u\|_\Omega.\end{aligned}$$

Using $\sigma_N = k_N$ and

$$\|z_u(t_{N-1}) - P_N z_u\|_\Omega \leq 2 \max_I \|z_u\|_\Omega \leq 2C_S \|g_T\|_\Omega,$$

gives

$$\begin{aligned}\sum_{n=1}^{N}\sum_{K\in\mathbf{T}_n} \rho_K\omega_K &\leq C\Big(\sum_{n=1}^{N}\sum_{K\in\mathbf{T}_n} h_K^4\rho_K^2\Big)^{\frac12} C_S\|g_T\|_\Omega + C\Big(\sum_{n=1}^{N-1}\sum_{K\in\mathbf{T}_n} \rho_K^2 k_n^2\sigma_n^{-1}\Big)^{\frac12} C_S\|g_T\|_\Omega\\ &\quad + C\Big(\sum_{K\in\mathbf{T}_N} k_N\rho_K^2\Big)^{\frac12} C_S\|g_T\|_\Omega\\ &= CC_S\Big(\sum_{n=1}^{N}\sum_{K\in\mathbf{T}_n} h_K^4\rho_K^2\Big)^{\frac12}\|g_T\|_\Omega + CC_S\Big(\sum_{n=1}^{N}\sum_{K\in\mathbf{T}_n} \rho_K^2 k_n^2\sigma_n^{-1}\Big)^{\frac12}\|g_T\|_\Omega.\end{aligned}$$

This completes the proof.

□

Finally, we prove a priori bounds for the stability constants C_S.

Theorem 3. *Assume that $\|f'(u)\|_{L_\infty(Q)} \leq \beta$ and $\varepsilon \in (0,1]$ and that (22) holds. Then the solution of (15) admits the following a priori bounds, where $C = C(\beta)$. If $g_T = 0$, then*

$$\|\mathrm{D}^2 z_u\|_Q^2 + \|\partial_t z_u\|_Q^2 + \varepsilon^2\|\mathrm{D}^2 z_w\|_Q^2 \leq C\|g\|_Q^2 \mathrm{e}^{C\varepsilon^{-1}T}. \tag{26}$$

If $g = 0$, then, with $\sigma(t) = T - t$,

$$\varepsilon^{-1}\max_I \|z_u\|_\Omega^2 + \|z_w\|_Q^2 + \|\mathrm{D}^2 z_u\|_Q^2 + \|\sigma^{\frac12}\partial_t z_u\|_Q^2 + \varepsilon^2\|\sigma^{\frac12}\mathrm{D}^2 z_w\|_Q^2 \leq C\varepsilon^{-1}\|g_T\|_\Omega^2 \mathrm{e}^{C\varepsilon^{-1}T}. \tag{27}$$

Proof. We first estimate $\|z_w\|_Q^2$. To this end we use $\Delta z_u = z_w$ from the second equation of (15) to get

$$\langle \Delta z_w, z_u\rangle_\Omega = \langle z_w, \Delta z_u\rangle_\Omega = \|z_w\|_\Omega^2.$$

Then we multiply the first equation of (15) by z_u, and integrate over $[t,T]$,

$$\int_t^T \langle -\partial_t z_u, z_u\rangle_\Omega \,\mathrm{d}s + \varepsilon\int_t^T \|z_w\|_\Omega^2\,\mathrm{d}s - \int_t^T \langle f'(u) z_w, z_u\rangle_\Omega\,\mathrm{d}s = \int_t^T \langle g, z_u\rangle_\Omega\,\mathrm{d}s.$$

By assumption we know that $\|f'(u)\|_{L_\infty(Q)} \leq \beta$, so we have

$$\begin{aligned}\tfrac12\|z_u(t)\|_\Omega^2 - \tfrac12\|z_u(T)\|_\Omega^2 + \varepsilon\int_t^T \|z_w\|_\Omega^2\,\mathrm{d}s &\leq \int_t^T \|f'(u)\|_{L_\infty(Q)}\|z_w\|_\Omega\|z_u\|_\Omega\,\mathrm{d}s + \int_t^T \|g\|_\Omega\|z_u\|_\Omega\,\mathrm{d}s\\ &\leq \int_t^T \big(\tfrac{\beta^2}{2\varepsilon}\|z_u\|_\Omega^2 + \tfrac{\varepsilon}{2}\|z_w\|_\Omega^2\big)\,\mathrm{d}s + \int_t^T \big(\tfrac{c}{2}\|g\|_\Omega^2 + \tfrac{1}{2c}\|z_u\|_\Omega^2\big)\,\mathrm{d}s\\ &\leq \tfrac{\beta^2}{\varepsilon}\int_t^T \|z_u\|_\Omega^2\,\mathrm{d}s + \tfrac{\varepsilon}{2}\int_t^T \|z_w\|_\Omega^2\,\mathrm{d}s + \int_t^T \big(\tfrac{c}{2}\|g\|_\Omega^2 + \tfrac{1}{2c}\|z_u\|_\Omega^2\big)\,\mathrm{d}s.\end{aligned}$$

Hence, with $z_u(T) = g_T$ and $c = \frac{\varepsilon}{\beta^2}$,

$$\begin{aligned} \|z_u(t)\|_\Omega^2 + \varepsilon \int_t^T \|z_w\|_\Omega^2 \,\mathrm{d}s &\le \frac{\varepsilon}{\beta^2}\|g\|_Q^2 + \|g_T\|_\Omega^2 + 2\beta^2\varepsilon^{-1}\int_t^T \|z_u\|_\Omega^2 \,\mathrm{d}s \\ &\le \tfrac{C}{\varepsilon}\|g\|_Q^2 + \|g_T\|_\Omega^2 + C\varepsilon^{-1}\int_t^T \|z_u\|_\Omega^2 \,\mathrm{d}s. \end{aligned}$$

Define

$$\Phi(t) = \|z_u(t)\|_\Omega^2 + \varepsilon \int_t^T \|z_w(s)\|_\Omega^2 \,\mathrm{d}s.$$

Obviously we have $\|z_u(s)\|_\Omega^2 \le \Phi(s)$, so that

$$\Phi(t) \le C\varepsilon\|g\|_Q^2 + \|g_T\|_\Omega^2 + C\varepsilon^{-1}\int_t^T \Phi(s)\,\mathrm{d}s.$$

We apply Gronwall's lemma to get

$$\Phi(t) \le C(\varepsilon\|g\|_Q^2 + \|g_T\|_\Omega^2)\mathrm{e}^{C\varepsilon^{-1}(T-t)}.$$

This means

$$\|z_u(t)\|_\Omega^2 + \varepsilon \int_t^T \|z_w\|_\Omega^2 \,\mathrm{d}s \le C(\varepsilon\|g\|_Q^2 + \|g_T\|_\Omega^2)\mathrm{e}^{C\varepsilon^{-1}(T-t)}.$$

We conclude

$$\begin{aligned} \max_I \|z_u\|_\Omega^2 &\le C(\varepsilon\|g\|_Q^2 + \|g_T\|_\Omega^2)\mathrm{e}^{C\varepsilon^{-1}T}, \\ \|z_w\|_Q^2 &\le C(\|g\|_Q^2 + \varepsilon^{-1}\|g_T\|_\Omega^2)\mathrm{e}^{C\varepsilon^{-1}T}. \end{aligned}$$

From the second equation we know $z_w = \Delta z_u$. So, by (22) and (28),

$$\|\mathrm{D}^2 z_u\|_Q^2 \le C\|\Delta z_u\|_Q^2 = C\|z_w\|_Q^2 \le C(\|g\|_Q^2 + \varepsilon^{-1}\|g_T\|_\Omega^2)\mathrm{e}^{C\varepsilon^{-1}T}. \tag{28}$$

This takes care of the first terms in (26) and (27).

Now assume that $g_T = 0$. Consider the dual problem (15) and multiply the first equation by $-\partial_t z_u$ and integrate over Q to get

$$\langle \partial_t z_u, \partial_t z_u\rangle_Q - \varepsilon\langle \Delta z_w, \partial_t z_u\rangle_Q - \langle f'(u) z_w, \partial_t z_u\rangle_Q = -\langle g, \partial_t z_u\rangle_Q. \tag{29}$$

So, by using $z_w = \Delta z_u$ from the second equation, we get

$$\langle \Delta z_w, \partial_t z_u\rangle_Q = \langle z_w, \partial_t \Delta z_u\rangle_Q = \langle \Delta z_u, \partial_t \Delta z_u\rangle_Q = \tfrac{1}{2}\int_0^T \frac{\mathrm{d}}{\mathrm{d}t}\|\Delta z_u\|_\Omega^2 \,\mathrm{d}t.$$

By putting this in (29) and using that $\|f'(u)\|_{L_\infty(Q)} \le \beta$, we have

$$\begin{aligned} \|\partial_t z_u\|_Q^2 - \tfrac{\varepsilon}{2}\|\Delta z_u(T)\|_\Omega^2 + \tfrac{\varepsilon}{2}\|\Delta z_u(0)\|_\Omega^2 &\le \|f'(u)\|_{L_\infty(Q)}\|z_w\|_Q\|\partial_t z_u\|_Q + \|g\|_Q\|\partial_t z_u\|_Q \\ &\le \tfrac{c\beta^2}{2}\|z_w\|_Q^2 + \tfrac{1}{2c}\|\partial_t z_u\|_Q^2 + \tfrac{c}{2}\|g\|_Q^2 + \tfrac{1}{2c}\|\partial_t z_u\|_Q^2. \end{aligned}$$

Put $c = 2$ and kick back $\|\partial_t z_u\|_Q^2$ to get, with $z_u(T) = g_T = 0$,

$$\tfrac{1}{2}\|\partial_t z_u\|_Q^2 + \tfrac{\varepsilon}{2}\|\Delta z_u(0)\|_\Omega^2 \le \beta^2\|z_w\|_Q^2 + \|g\|_Q^2.$$

Hence, by (28) with $C = C(\beta)$,

$$\|\partial_t z_u\|_Q^2 \le C\|z_w\|_Q^2 + C\|g\|_Q^2 \le C\|g\|_Q^2 \mathrm{e}^{-C\varepsilon^{-1}T}. \tag{30}$$

It remains to bound $\|\mathrm{D}^2 z_w\|_Q^2$. From the first equation of (15) we get

$$\varepsilon\Delta z_w = g + \partial_t z_u + f'(u) z_w.$$

Taking norms and using (22), (28), and (30) gives

$$\begin{aligned}
\varepsilon^2\|\mathrm{D}^2 z_w\|_Q^2 &\le \varepsilon^2 C\|\Delta z_w\|_Q^2 = C\|g + \partial_t z_u + f'(u) z_w\|_Q^2 \\
&\le C\Big(\|g\|_Q^2 + \|\partial_t z_u\|_Q^2 + \|f'(u)\|_{L_\infty(Q)}^2 \|z_w\|_Q^2\Big) \\
&\le C\|g\|_Q^2 \mathrm{e}^{C\varepsilon^{-1}T}.
\end{aligned}$$

This completes the proof of (26)

Now let $g = 0$ and set $\sigma(t) = T - t$. Multiply the first equation of (15) by $-\sigma\partial_t z_u$ to get

$$\langle \partial_t z_u, \sigma\partial_t z_u\rangle_Q - \varepsilon\langle \Delta z_w, \sigma\partial_t z_u\rangle_Q - \langle f'(u) z_w, \sigma\partial_t z_u\rangle_Q = 0.$$

Here, since $z_w = \Delta z_u$ and $\sigma'(t) = -1$,

$$\begin{aligned}
\langle \Delta z_w, \sigma\partial_t z_u\rangle_Q &= \langle z_w, \sigma\Delta\partial_t z_u\rangle_Q = \langle \Delta z_u, \sigma\Delta\partial_t z_u\rangle_Q = \tfrac{1}{2}\int_0^T \frac{\mathrm{d}}{\mathrm{d}t}(\sigma\|\Delta z_u\|_\Omega^2)\,\mathrm{d}t - \tfrac{1}{2}\int_0^T \sigma'\|\Delta z_u\|_\Omega^2\,\mathrm{d}t \\
&= \tfrac{1}{2}\sigma(T)\|\Delta z_u(T)\|_\Omega^2 - \tfrac{1}{2}\sigma(0)\|\Delta z_u(0)\|_\Omega^2 + \tfrac{1}{2}\int_0^T \|z_w\|_\Omega^2\,\mathrm{d}t = -\tfrac{1}{2}T\|\Delta z_u(0)\|_\Omega^2 + \tfrac{1}{2}\|z_w\|_Q^2.
\end{aligned}$$

Hence,

$$\begin{aligned}
\|\sigma^{\frac{1}{2}}\partial_t z_w\|_Q^2 + \|\Delta z_u(0)\|_\Omega^2 &\le \tfrac{\varepsilon}{2}\|z_w\|_Q^2 + \|f'(u)\|_{L_\infty}\|\sigma^{\frac{1}{2}} z_w\|_Q \|\sigma^{\frac{1}{2}}\partial_t z_u\|_Q \\
&\le \tfrac{1}{2}(\varepsilon + \beta^2 T)\|z_w\|_Q^2 + \tfrac{1}{2}\|\sigma^{\frac{1}{2}}\partial_t z_u\|_Q^2.
\end{aligned}$$

So by (28) we have

$$\|\sigma^{\frac{1}{2}}\partial_t z_u\|_Q \le (\varepsilon + \beta^2 T)\|z_w\|_Q^2 C\varepsilon^{-1}\|g_T\|_\Omega^2 \mathrm{e}^{C\varepsilon^{-1}T}.$$

Finally, from (22) and $\varepsilon\Delta z_w = \partial_t z_u + f'(u) z_w$ we get

$$\begin{aligned}
\varepsilon^2\|\sigma^{\frac{1}{2}}\mathrm{D}^2 z_w\|_Q^2 &\le \varepsilon^2 C\|\sigma^{\frac{1}{2}}\Delta z_w\|_Q^2 = C\|\sigma^{\frac{1}{2}}(\partial_t z_u + f'(u) z_w)\|_Q^2 \\
&\le C\Big(\|\sigma^{\frac{1}{2}}\partial_t z_u\|_Q^2 + T\|z_w\|_Q^2\Big) \le C\varepsilon^{-1}\|g_T\|_\Omega^2 \mathrm{e}^{C\varepsilon^{-1}T}.
\end{aligned}$$

This completes the proof of (27). □

References

[1] L. Baňas R. Nürnberg, *Adaptive finite element methods for Cahn-Hilliard equations*, J. Comput. Appl. Math. **218** (2008) 2–11.

[2] W. Bangerth, R. Rannacher, *Adaptive Finite Element Methods for Differential Equations*, Birkhäuser, 2003.

[3] J. W. Cahn, J. E. Hilliard, *Free energy of a nonuniform system I. Interfacial free energy*, J. Chem. Phys. **28** (1958) 258–267

[4] C.M. Elliott, S. Larsson, *Error estimates with smooth and nonsmooth data for a finite element method for the Cahn-Hilliard equation*, Math. Comput. **58** (1992) 603–630, S33–S36.

[5] C.M. Elliott, S. Zheng, *On the Cahn-Hilliard equation* , Arch. Rational Mech. Anal. **96** (1986) 339–357

[6] X. Feng, H. Wu, *A posteriori error estimates for finite element approximations of the Cahn-Hilliard equation and the Hele-Shaw flow*, J. Comput. Math. **26** (2008) 767–796.