MATHEMATICS OF COMPUTATION
VOLUME 59, NUMBER 200
OCTOBER 1992, PAGES 383-401

CONVERGENCE OF A FINITE ELEMENT METHOD FOR THE DRIFT-DIFFUSION SEMICONDUCTOR DEVICE EQUATIONS: THE ZERO DIFFUSION CASE

BERNARDO COCKBURN AND IOANA TRIANDAF

ABSTRACT. In this paper a new explicit finite element method for numerically solving the drift-diffusion semiconductor device equations is introduced and analyzed. The method uses a mixed finite element method for the approximation of the electric field. A finite element method using discontinuous finite elements is used to approximate the concentrations, which may display strong gradients. The use of discontinuous finite elements renders the scheme for the concentrations trivially conservative and fully parallelizable. In this paper we carry out the analysis of the model method (which employs a continuous piecewise-linear approximation to the electric field and a piecewise-constant approximation to the electron concentration) in a model problem, namely, the so-called unipolar case with the diffusion terms neglected. The resulting system of equations is equivalent to a conservation law whose flux, the electric field, depends globally on the solution, the concentration of electrons. By exploiting the similarities of this system with classical scalar conservation laws, the techniques to analyze the monotone schemes for conservation laws are adapted to the analysis of the new scheme. The scheme, considered as a scheme for the electron concentration, is shown to satisfy a maximum principle and to be total variation bounded. Its convergence to the unique weak solution is proven. Numerical experiments displaying the performance of the scheme are shown.

1. INTRODUCTION

This is the first paper of a series in which we introduce and analyze a new finite element method for numerically solving the equations of the drift-diffusion model for semiconductor devices, [39]:

$$-\varepsilon\Delta\psi = q(C - n + p), \tag{1.1a}$$

$$qn_t - \operatorname{div} J_n = -qR, \tag{1.1b}$$

$$qp_t + \operatorname{div} J_p = -qR, \tag{1.1c}$$

where ψ is the electric potential, n is the electron concentration, p is the hole concentration, ε is the dielectric constant, q is the electronic elementary

Received by the editor October 29, 1990 and, in revised form, December 9, 1991.

1991 *Mathematics Subject Classification*. Primary 65N30, 65N12, 35L60, 35L65.

Key words and phrases. Semiconductor devices, conservation laws, finite elements, convergence.

The first author was partly supported by the National Science Foundation (GrantDSM-91003997) and by the University of Minnesota Army High Performance Computing Research Center.

The second author was supported by a Fellowship of the University of Minnesota Army High Performance Computing Research Center.

charge, C is the doping profile, $R = R(n, p)$ is the carrier recombination-generation rate, and J_n and J_p are the current densities. They are given by

$$J_n = q\mu_n (U_T \operatorname{grad} n - n \operatorname{grad} \psi), \tag{1.1d}$$

$$J_p = -q\mu_p (U_T \operatorname{grad} p + p \operatorname{grad} \psi), \tag{1.1e}$$

where μ_n and μ_p are the electron and hole mobilities, and U_T is the thermal voltage; see [29, pp. 7–13].

The main ideas of our method are as follows. Following [16], we discretize Poisson's equation (1.1a) by using a mixed finite element method, [33], [5]. This method has the advantage of directly giving a numerical approximation of the electric field, $-\operatorname{grad} \psi$, which is the only quantity depending on ψ that appears in the convection-diffusion equations (1.1b) and (1.1c). To discretize the latter equations, we use an extension of the Runge-Kutta Discontinuous Galerkin (RKDG) method, which is a fully parallelizable method initially devised for numerically solving nonlinear conservation laws. In the scalar case, the RKDG method can be proven to satisfy maximum principles, even when the approximate solution is locally a polynomial of total degree $k > 0$. Moreover, extensive numerical simulations show that the RKDG method can capture discontinuities within a couple of elements without producing spurious oscillations; see [9, 10, 11, 12]. Thus, the RKDG method is a natural choice in this framework, since the concentrations may present strong gradients.

The main computational advantages of our method are the following. Since the use of Lagrange multipliers, see [5] and the bibliography therein, renders the matrix of the mixed element method symmetric and positive definite, the computation of the approximation to grad ψ is very much facilitated. Also, since the RKDG method uses discontinuous approximations, the 'mass' matrix turns out to be a blockdiagonal matrix whose entries can be inverted by hand (in fact, the order of the blockdiagonal matrices is exactly equal to the number of degrees of freedom of the approximate solution u_h on the corresponding element). Moreover, since the Runge-Kutta method used is explicit, the scheme for the convection-diffusion equations is fully parallelizable. Finally, no nonlinear equation is required to be solved at each timestep.

For the sake of clarity, the analysis of our finite element method will be done on a one-dimensional model problem. We set $R \equiv 0$, and scale the equations (see [29, pp. 26–28], [38] and [3] for details) to obtain

$$-\phi_{xx} = c - u + v, \qquad \tau > 0,\ x \in (-1, 1), \tag{1.2a}$$

$$u_\tau + (u\phi_x)_x - \lambda^2 u_{xx} = 0, \qquad \tau > 0,\ x \in (-1, 1), \tag{1.2b}$$

$$v_\tau - (v\phi_x)_x - \lambda^2 v_{xx} = 0, \qquad \tau > 0,\ x \in (-1, 1), \tag{1.2c}$$

where

$$\lambda^2 = \frac{\varepsilon U_T}{q \| C \|_{L^\infty} l^2}, \tag{1.2d}$$

and l is the typical diameter of the semiconductor device. Since typical values of λ range from 10^{-3} to 10^{-5}, [29, p. 28; 37], it seems reasonable to neglect the second-order terms in (1.2b) and (1.2c). The resulting equations give a good approximation of the initial system in the so-called 'fast' time scale, see [37, 30, 34]. See also the singular perturbation analyses carried out in [4, 6]. By using

a symmetry assumption, [37, 34], we can decouple equations (1.2b) and (1.2c) and obtain the following equations for the scaled electron concentration u and the scaled electric potential ϕ:

$$\begin{aligned} -\phi_{xx} &= 1-u, && x \in (0,1),\ \tau \geq 0, \\ u_\tau + (u\phi_x)_x &= 0, && x \in (0,1),\ \tau > 0, \end{aligned}$$

where we have assumed, for simplicity, that the scaled doping profile c is identically equal to 1 on $(0,1)$. All our results also hold for c in $\mathrm{BV}(0,1)$. These are the equations of our model problem. To complete it, we have to impose the boundary conditions

$$\begin{aligned} \phi(\tau,0) &= 0, && \text{for } \tau \geq 0, \\ \phi(\tau,1) &= \phi_1(\tau), && \text{for } \tau \geq 0, \end{aligned}$$

and

$$\begin{aligned} u(\tau,0) &= u_0(\tau), && \text{if } \phi_x(\tau,0) > 0,\ \tau \geq 0, \\ u(\tau,1) &= u_1(\tau), && \text{if } \phi_x(\tau,1) < 0,\ \tau \geq 0, \end{aligned}$$

and the initial condition

$$u(0,x) = u_i(x), \qquad x \in (0,1).$$

The solution of this problem has been proven to be the limit as λ goes to zero of the corresponding 'viscous' solutions, see (2.15), in [8]. The problem of how close these solutions are will be addressed in this paper numerically only. It will be considered analytically in a forthcoming paper.

To study the above problem, we prefer to rewrite it as the following conservation law:

$$\begin{aligned} &(1.3a) && u_\tau + (u\beta)_x = 0, && \tau > 0,\ x \in (0,1), \\ &(1.3b) && u(\tau,0) = u_0(\tau), && \text{if } \beta(\tau,0) > 0,\ \tau \geq 0, \\ &(1.3c) && u(\tau,1) = u_1(\tau), && \text{if } \beta(\tau,1) < 0,\ \tau \geq 0, \\ &(1.3d) && u(0,x) = u_i(x), && x \in (0,1), \end{aligned}$$

where

$$\begin{aligned} &(1.4a) && -\beta_x = 1-u, && x \in (0,1),\ \tau \geq 0, \\ &(1.4b) && \beta = \phi_x, && x \in (0,1),\ \tau \geq 0, \\ &(1.4c) && \phi(\tau,0) = 0, && \text{for } \tau \geq 0, \\ &(1.4d) && \phi(\tau,1) = \phi_1(\tau), && \text{for } \tau \geq 0, \end{aligned}$$

since written in this form, it is easier to compare it with a classical conservation law:

(i) Notice that the equation (1.3a) would be a classical nonlinear conservation law if the operator β were an evaluation operator, i.e., if $\beta = \beta(u)$. However, in our case the value of β at a single point $(\tau, x) \in (0,T) \times (0,1)$ contains the information of all the values of the function $u(\tau,\cdot)$ on $(0,1)$. Hence a perturbation of the function u at any given point of the domain does have a

global effect immediately. This is in sharp contrast with the classical conservation laws, for which local perturbations of the solution have a local effect in finite time.

(ii) The smoothness of $\beta(\tau, \cdot)$ guarantees the uniqueness of the weak solution, [31], and so the worry about convergence to the so-called entropy solution that pervades the numerical analysis of schemes for classical conservation laws is not present here. Nevertheless, the continuity of β also guarantees that the characteristics of our system never intersect each other. Numerically, this means that there is no natural mechanism that would help the scheme to 'sharpen' the discontinuities, as happens in classical conservation laws when the nonlinearity is convex (for example).

(iii) Note that $u = 0$ and $u = 1$ are equilibrium points of the equation of u along the characteristics. In fact, if $x = x(\tau)$ denotes a given characteristic, and if we set $u = u(\tau, x(\tau))$, then

$$\frac{d}{d\tau} u = (1 - u)\, u. \tag{1.5}$$

From this equation it is clear that $u = 0$ is an unstable equilibrium point whereas $u = 1$ is an asymptotically stable point. (This situation never occurs for classical conservation laws.) The instability of $u = 0$ indicates that the numerical approximation, u_h, has to be prevented from taking negative values, however small. It also indicates that the numerical approximation of the points $u = 0$ might be a delicate matter. The asymptotic stability of $u = 1$ suggests the existence of a maximum principle for the solution u. Notice that for the very simple case in which $\beta = \beta(x) = 1/2 - x$, $u_0 \equiv 0$, $u_1 \equiv 1$, and $u_i(x) = x$, the solution of the conservation law (1.3) does not remain bounded. In this case, $\| u(\tau) \|_{L^\infty(0,1)}$ goes to infinity as τ goes to infinity; see Figure 1.

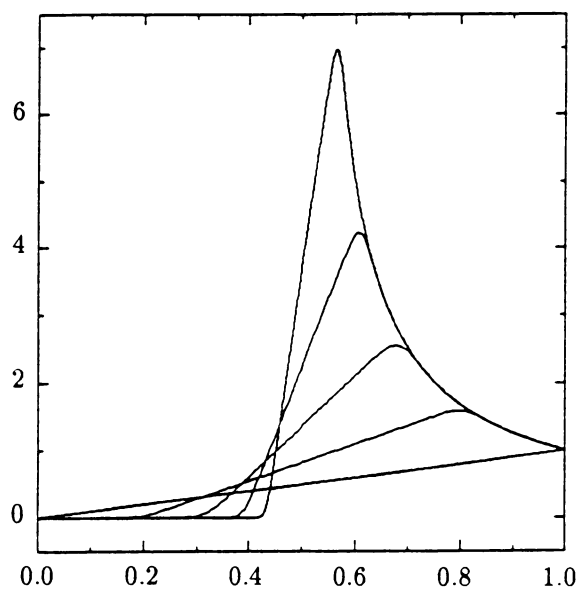

FIGURE 1. Blowing up of the solution u of (1.3) when $\beta(x) = 1/2 - x$, $u_0 \equiv 0$, $u_1 \equiv 1$, and $u_i(x) = x$

(iv) Finally, since the solution of equation (1.5) is

$$u(\tau, x(\tau)) = \left[\frac{u_i(x(0))}{1 + (e^\tau - 1)u_i(x(0))}\right] e^\tau$$
$$\approx u_i(x(0))\, e^\tau, \qquad \text{for } 0 < \tau \ll 1,$$

we expect the total variation of the concentration to grow in time as fast as e^τ, without the boundary data being responsible for such a growth. This is also in strong contrast to what happens in classical conservation laws, where the total variation does not increase in such a situation.

The finite element method we consider in this paper takes the approximation to grad ϕ to be a continuous piecewise linear function and the approximation to u to be a piecewise constant function. The resulting scheme can be considered to be a counterpart of the monotone schemes for conservation laws. With this fact in mind, the techniques for analyzing monotones schemes, [15, 18], will be suitably extended to our present setting. The main objective of this paper is to establish the stability of the new method and to prove its convergence in the case of the model problem (1.3), (1.4). Particular care has to be taken to relate the behavior of the piecewise linear continuous approximate electric field given by the mixed finite element method, with the behavior of the piecewise constant convected approximate electron concentration. Our analysis of the boundary conditions is inspired by the ideas introduced in [25]. A forthcoming paper will be devoted to obtaining error estimates.

No finite element method seems to have been proposed and analyzed for the hyperbolic problem (1.3), (1.4). Most methods are defined on the full equations (1.1), and take advantage of the parabolicity of equations for the concentrations. In order to do that, a widespread practice consists in using the so-called Boltzmann statistics as variables. This change of variables allows the currents to be expressed in the form $\{c \text{ grad } z\}$ in an effort to render the equations (1.1b) and (1.1c) 'naturally' parabolic; see, e.g., [40, 2]. However, this practice 'hides' the convective character of the equations. This is why we do not use Boltzmann statistics as variables. The same point of view is taken in [16, 32], where a direct hold of the convection phenomenon is attempted through the modified method of characteristics. We want to emphasize that, unlike the methods in [16, 32], our scheme is conservative.

Our proposed method can also be used to compute the stationary solution of (1.1) by simply letting the final time T be large enough. Of course, other more efficient ways to do so can be devised, but we shall not pursue this matter in this paper. Methods for computing stationary solutions of (1.1) can be found in, e.g., [1, 17, 22, 35]. The stationary solutions of (1.1) may be considered to be the fixed points of the so-called Gummel map; see, e.g., [19]. The iterative procedures of the above-mentioned papers try to construct numerical approximations to the Gummel map (and to its fixed points). A rigorous analysis of the convergence of such numerical approximations can be found in [20]. These iterative procedures can also be applied to solve the transient problem; see [21] and [14]. We finally point out that a new discretization technique that generalizes to the two-dimensional case the (one-dimensional) Sharfetter-Gummel method has recently been introduced in [7].

The paper is organized as follows. In §2a, we define the set of data with which we deal in this paper and define the weak solution of problem (1.3),

(1.4). In §2b, we present our numerical method. In §2c, we state and briefly discuss our results on stability (Theorem 2.1), continuity with respect to the data (Theorem 2.2), and convergence to the weak solution (Theorem 2.3). We also include therein an additional convergence result (Theorem 2.4) that will be used in a forthcoming paper to obtain error estimates. In §2d, we state a result (Theorem 2.5) concerning the smoothness with which the boundary conditions are satisfied by the approximate solution. This result requires reasonable additional restrictions on the data. Finally, in §2e, we show several numerical results illustrating the performance of the scheme. The proofs of our theorems are contained in the Supplement section at the end of this issue.

2. The numerical method and the main results

2a. **The weak solution.** We shall assume that the initial data u_i and the boundary data u_0, u_1 and ϕ_1 satisfy the following regularity conditions:

$$u_0(\tau),\, u_1(\tau),\, u_i(x) \in [0,\, u^*], \quad \tau \in [0,\, T],\ x \in [0,\, 1], \tag{2.1a}$$

$$u_0,\, u_1 \in \mathrm{BV}(0,\, T),\ \text{and } u_i(x) \in \mathrm{BV}(0,\, 1), \tag{2.1b}$$

$$\phi_1(\tau) \in [0,\, \phi_1^*], \quad \tau \in [0,\, T], \tag{2.1c}$$

$$\phi_1 \in \mathrm{BV}(0,\, T), \tag{2.1d}$$

where, see remark (iii) in §1, we assume that

$$u^* \geq 1. \tag{2.1e}$$

The weak solution of (1.3) and (1.4) is defined to be a function $(u,\, \beta,\, \phi) \in L^\infty(0,\, T;\, BV(0,\, 1)) \times L^\infty(0,\, T;\, W^{1,1}(0,\, 1)) \times L^\infty(0,\, T;\, BV(0,\, 1))$ satisfying the following weak formulation:

$$\begin{aligned} &-\int_0^T \int_0^1 \{ u(\tau,\, x)\, \varphi_\tau(\tau,\, x) + u(\tau,\, x)\, \beta(\tau,\, x)\, \varphi_x(\tau,\, x) \}\, dx\, d\tau \\ &-\int_0^1 u_i(x)\, \varphi(0,\, x)\, dx \\ &+\int_0^T f(u(\tau,\, 1-),\, u_1(\tau);\ \beta(\tau,\, 1))\, \varphi(\tau,\, 1)\, d\tau \\ &-\int_0^T f(u_0(\tau),\, u(\tau,\, 0+);\ \beta(\tau,\, 0))\, \varphi(\tau,\, 0)\, d\tau = 0, \\ &\qquad\qquad\qquad\qquad\qquad\qquad \forall \varphi \in \mathscr{C}_0^1([0,\, T) \times [0,\, 1]), \end{aligned} \tag{2.2a}$$

where the flux f is defined by

$$f(u_{\text{left}},\, u_{\text{right}};\, \beta) = u_{\text{left}}\, \beta^+ + u_{\text{right}}\, \beta^-, \tag{2.2b}$$

with $\beta^- = \min\{\beta,\, 0\}$, and $\beta^+ = \max\{\beta,\, 0\}$, and

$$\begin{aligned} &-\int_0^1 \beta_x(\tau,\, x)\, w(x)\, dx \\ &\qquad = \int_0^1 (1 - u(\tau,\, x))\, w(x)\, dx, \quad \forall w \in L^2(0,\, 1), \end{aligned} \tag{2.3a}$$

$$\begin{aligned}(2.3b)\qquad &\int_0^1 \beta(\tau, x)\, v(x)\, dx \\ &= -\int_0^1 \phi(\tau, x)\, v_x(x)\, dx + \phi_1(\tau)\, v(1), \quad \forall v \in H^1(0, 1).\end{aligned}$$

Notice that any function in $W^{1,1}(0, 1)$ is continuous, and that any function in BV$(0, 1)$ has well-defined limits from the left and from the right. Thus, the last two integrals of (2.2a) have a meaning.

Notice also that the role of the flux f is to select the correct boundary value for u according to (1.3b) and (1.3c). To see this, assume that u is a very smooth solution of (2.2). Take functions $\varphi \in \mathscr{C}_0^1((0, T) \times [0, 1))$, and integrate by parts in (2.2a). Since u satisfies equation (1.3a) pointwise, we get

$$\int_0^T (u(\tau, 0+) - u_0(\tau))\, \beta^+(\tau, 0)\, \varphi(\tau, 0)\, d\tau = 0,$$

and since we are assuming u to be smooth, this implies

$$u(\tau, 0+)\, \beta^+(\tau, 0) = u_0(\tau)\, \beta^+(\tau, 0), \qquad \tau \in (0, T),$$

or

$$u(\tau, 0+) = u_0(\tau) \quad \text{if } \beta(\tau, 0) > 0, \ \tau \in (0, T),$$

which is nothing but condition (1.3b). Condition (1.3a) can be obtained in a similar way.

2b. **The numerical scheme.** We first introduce some notation. Let $\{x_{i+1/2}\}_{i=0,\dots,n_x}$ be a partition of $[0, 1]$ with $x_{1/2} = 0$ and $x_{n_x+1/2} = 1$. Similarly, let $\{\tau^n\}_{n=0,\dots,n_T}$ be a partition of $[0, T]$, with $\tau^0 = 0$ and $\tau^{n_T} = T$. We set $I_i = (x_{i-1/2}, x_{i+1/2})$, $\Delta x_i = x_{i+1/2} - x_{i-1/2}$, and $J^n = [\tau^n, \tau^{n+1})$, $\Delta\tau^n = \tau^{n+1} - \tau^n$. Define $\Delta x = \max_{i=1,\cdots,n_x}\{\Delta x_i\}$ and $\Delta\tau = \max_{n=0,\cdots,n_T-1}\{\Delta\tau^n\}$. We associate with these partitions the following spaces:

$$(2.4a)\quad V_{\Delta x} = \{v_{\Delta x} \in \mathscr{C}^0(0, 1) : v_{\Delta x}|_{I_i} \in P^1(I_i),\ i = 1, \dots, n_x\},$$

$$(2.4b)\quad W_{\Delta x} = \{w_{\Delta x} \in L^\infty(0, 1) : w_{\Delta x}|_{I_i} \in P^0(I_i),\ i = 1, \dots, n_x\},$$

$$(2.4c)\quad W_{\Delta\tau} = \{w_{\Delta\tau} \text{ right-continuous: } w_{\Delta\tau}|_{J^n} \in P^0(J^n),\ n = 0, \dots, n_T - 1\}.$$

If $v_{\Delta x} \in V_{\Delta x}$, then $v_{i+1/2}$ denotes $v_{\Delta x}(x_{i+1/2})$ for $i = 0, \cdots, n_x$. If $w_{\Delta x} \in W_{\Delta x}$, then w_i denotes the constant value $w_{\Delta x}(x)$, $x \in I_i$, for $i = 1, \dots, n_x$; the values w_0 and w_{n_x+1} denote the exterior trace at $x = 0$, $w_{\Delta x}(0-)$, and at $x = 1$, $w_{\Delta x}(1+)$, respectively. Finally, if $w_{\Delta\tau} \in W_{\Delta\tau}$, then w^n denotes the constant value $w_{\Delta\tau}(\tau)$, $\tau \in J^n$.

To discretize (1.3), (1.4), we first discretize the data by setting

$$(2.5a)\qquad \phi_{1,\Delta\tau} = \mathbb{P}_{W_{\Delta\tau}}(\phi_1),$$

$$(2.5b)\qquad u_{0,\Delta\tau} = \mathbb{P}_{W_{\Delta\tau}}(u_0),$$

$$(2.5c)\qquad u_{1,\Delta\tau} = \mathbb{P}_{W_{\Delta\tau}}(u_1),$$

$$(2.5d)\qquad u_{i,\Delta x} = \mathbb{P}_{W_{\Delta x}}(u_i),$$

where $\mathbb{P}_{\mathscr{V}}$ denotes the L^2-projection into the space $\mathscr{V}$. The approximate solution u_h is taken to be in the space $W_{\Delta\tau} \otimes W_{\Delta x}$ and is required to satisfy the equation

$$(2.6a)\qquad (u_i^{n+1} - u_i^n)/\Delta\tau^n + (f_{i+1/2}^n - f_{i-1/2}^n)/\Delta x_i = 0,$$

where the numerical flux $f^n_{i+1/2} = f(u^n_i\,,\, u^n_{i+1}\,;\, \beta^n_{i+1/2})$ is given by (2.2b), i.e., by

$$f^n_{i+1/2} = u^n_i\, \beta^{n^+}_{i+1/2} + u^n_{i+1}\, \beta^{n^-}_{i+1/2}. \tag{2.6b}$$

The function $(\beta_h\,,\, \phi_h) \in W_{\Delta\tau} \otimes V_{\Delta x} \times W_{\Delta\tau} \otimes W_{\Delta x}$ is defined by the following mixed finite element method:

$$\begin{aligned} &-\int_0^1 (\beta_h)_x(\tau^n\,,\, x)\, w_{\Delta x}(x)\, dx \\ &\quad = \int_0^1 (1 - u_h(\tau^n\,,\, x))\, w_{\Delta x}(x)\, dx\,, \quad \forall w_{\Delta x} \in W_{\Delta x}\,, \end{aligned} \tag{2.7a}$$

$$\begin{aligned} &\int_0^1 \beta_h(\tau^n\,,\, x)\, v_{\Delta x}(x)\, dx \\ &\quad = -\int_0^1 \phi_h(\tau^n\,,\, x)\, (v_{\Delta x})_x(x)\, dx + \phi_{1,\Delta\tau}(\tau^n)\, v_{\Delta x}(1)\,, \quad \forall v_{\Delta x} \in V_{\Delta x}. \end{aligned} \tag{2.7b}$$

Note that, for a given function β_h, the scheme (2.6) is nothing but the well-known upwinding scheme (which coincides with the Godunov scheme in this case). Since this is a monotone scheme under a suitable CFL condition, it is reasonable to expect to have for this scheme convergence properties similar to the convergence properties of monotone schemes for scalar conservation laws, [15]. We shall see below that this is indeed the case.

Notice that for the upwinding numerical flux $f^n_{i+1/2}$, (2.6b), to be well defined, the function $\beta_h(\tau\,,\, \cdot)$ has to be continuous. This requirement is naturally taken into account by the mixed finite element method used to compute it. This is an important advantage of using the mixed method (2.7).

Thus the algorithm of our numerical method is:

(2.8a) Compute the functions $u_{0,\Delta\tau}$, $u_{1,\Delta\tau}$, $u_{i,\Delta x}$, and $\phi_{1,\Delta\tau}$ by (2.5);

(2.8b) Set $u_h(0\,,\, \cdot) = u_{i,\Delta x}(\cdot)$;

(2.8c) For $n = 0\,,\, \dots\,,\, n_T - 1$ compute $u_h(\tau^{n+1}\,,\, \cdot)$ as follows:

- (i) Compute $(\beta_h(\tau^n\,,\, \cdot)\,,\, \phi_h(\tau^n\,,\, \cdot))$ by using the mixed finite element method (2.7);
- (ii) Set $u_h(\tau^n\,,\, 0-) = u_{0,\Delta\tau}(\tau^n)$ and $u_h(\tau^n\,,\, 1+) = u_{1,\Delta\tau}(\tau^n)$;
- (iii) Compute $u_h(\tau^{n+1}\,,\, x)$ for $x \in (0\,,\, 1)$ by using the scheme (2.6).

2c. **Stability and convergence results.** In this section we state and briefly discuss the stability properties of the scheme (2.8), Theorem 2.1, its property of continuity with respect to the data, Theorem 2.2, and its convergence property to the weak solution of the original problem, Theorem 2.3. We also obtain an estimate which will be used elsewhere to obtain error estimates for the scheme under consideration, Theorem 2.4.

Theorem 2.1 (Stability). *Suppose that for* $n = 0\,,\, \dots\,,\, n_T - 1$ *the following* CFL *condition is satisfied*:

$$\Delta\tau^n \le \min\left\{ \frac{1}{u^*}\,,\, \frac{\Delta x_i}{(2\,u^* - 1)\,\Delta x_i + \phi_1^* + \frac{1}{2}\max\{1\,,\, u^* - 1\}} \right\}. \tag{2.9}$$

Then the following stability properties hold:

(2.10a) $$u_h(\tau, x) \in [0, u^*], \qquad (\tau, x) \in (0, T) \times (0, 1),$$

(2.10b) $$\| \beta_h \|_{L^\infty(0,T;L^\infty(0,1))} \le \phi_1^* + \tfrac{1}{2}\max\{1, u^* - 1\},$$

(2.10c) $$\| \beta_{hx} \|_{L^\infty(0,T;L^1(0,1))} \le \max\{1, u^* - 1\},$$

(2.10d) $$\| \phi_h \|_{L^\infty(0,T;L^\infty(0,1))} \le \phi_1^* + \tfrac{1}{2}\max\{1, u^* - 1\},$$

(2.10e) $$\| \phi_h \|_{L^\infty(0,T;\mathrm{BV}(0,1))} \le \phi_1^* + \tfrac{1}{2}\max\{1, u^* - 1\}.$$

Moreover, there is a constant C_1, *depending solely on the data and* T, *such that*

(2.10f) $$\| u_h \|_{L^\infty(0,T;\mathrm{BV}(0,1))} \le C_1.$$

The scheme for u_h is thus a total variation bounded scheme that satisfies a sharp maximum principle. Notice that $u^* \ge 1$, by (2.1e). The maximum principle is not true for $u^* < 1$. This reflects the fact that along the characteristics the point $u = 1$ is an asymptotically stable equilibrium point; see remark (iii) in §1.

In the following result, (v_h, γ_h, ψ_h) stands for the approximate solution of (1.3), (1.4), with data v_i, v_0, v_1, ψ_1 satisfying (2.1).

Theorem 2.2 (Continuity with respect to the data). *Suppose that the* CFL *condition* (2.9) *is satisfied. Then,*

$$\begin{aligned}
&\| u_h(\tau) - v_h(\tau) \|_{L^1(0,1)} \\
&\quad \le e^{C_1 \tau}\{\| u_i - v_i \|_{L^1(0,1)} + 2\, C_1 \| \phi_1 - \psi_1 \|_{L^1(0,\tau)} \\
&\qquad + (\phi_1^* + \tfrac{1}{2}\max\{1, u^* - 1\})(\| u_1 - v_1 \|_{L^1(0,\tau)} + \| u_0 - v_0 \|_{L^1(0,\tau)})\}.
\end{aligned}$$

Theorem 2.3 (Convergence). *Suppose that the* CFL *condition* (2.9) *is satisfied. Then the sequence* $\{(u_h, \beta_h, \phi_h)\}_{h>0}$ *generated by the scheme* (2.8) *converges in* $L^\infty(0, T; L^1(0, 1)) \times L^1(0, T; W^{1,1}(0, 1)) \times L^1(0, T; BV(0, 1))$ *to the unique weak solution of* (2.2), (2.3), (u, β, ϕ). *Moreover,*

$$u \in L^\infty(0, T; \mathrm{BV}(0, 1)) \cap \mathscr{C}^0(0, T; L^1(0, 1)).$$

In [31], the uniqueness of the weak solution of (2.2) and (2.3) was proven for ϕ_1 constant only. However, the argument used therein can be easily extended to the case we consider in this paper. The uniqueness of the weak solution can also be deduced from the first inequality of Theorem 2.4 below.

To state our following result, we need to define the 'entropy' form $E^{\varepsilon_0,\varepsilon}(u, v; \beta)$. Kruzhkov [23] introduced this form in his study of classical conservation laws. Later, Kuznetsov [24] used it to obtain an approximation theory; see also [36, 26, 27, 28, 13]. Let u be the entropy solution of a classical conservation law, and let u_h be the approximate solution given by a monotone scheme. It was proven, in [23] and [24] respectively, that

$$\begin{aligned}
&E^{\varepsilon_0,\varepsilon}(u, v) \le 0, \\
&E^{\varepsilon_0,\varepsilon}(u_h, u) \le C\,\left(\Delta x/\varepsilon + \Delta\tau/\varepsilon_0\right),
\end{aligned}$$

where v is a reasonably general function (the function β does not appear in these entropy forms E, since it does not appear in the framework of classical conservation laws). Using these key results, Kuznetsov [24] obtained a bound for the error $\| u_h - u \|_{L^\infty(0,T;L^1)}$. In a forthcoming paper, we prove that error

estimates can be obtained for the scheme under consideration provided that similar results are obtained. Our next result contains those results.

Let ε_0 and ε be arbitrary positive real numbers, and let $w\colon \mathbb{R} \to \mathbb{R}$ be an even nonnegative function in $\mathscr{C}^\infty(\mathbb{R})$ with support included in $[-1, 1]$, and such that $\int_{-1}^{1} w = 1$. We set

$$\varphi(\tau, x; \tau', x') = w_{\varepsilon_0}(\tau - \tau')\, w_\varepsilon(x - x'),$$

where $w_\nu(s) = w(s/\nu)/\nu$, $\forall s \in \mathbb{R}$. Let us denote by U an arbitrary even convex function with Lipschitz second derivative, such that $U(0) = 0$.

The entropy form, $E^{\varepsilon_0,\varepsilon}(u, v; \beta)$, is defined as follows:

$$E^{\varepsilon_0,\varepsilon}(u, v; \beta) = \int_0^T\!\!\int_0^1 \Theta(u, v(\tau, x); \beta; \varphi(\tau, x; \cdot, \cdot))\, dx\, d\tau, \tag{2.11a}$$

where

$$\begin{aligned}
\Theta(u, c; \beta; \varphi) = &-\int_0^T \int_0^1 \{ U(u(\tau, x) - c)\, \varphi_\tau(\tau, x) \\
&\qquad + U(u(\tau, x) - c)\, \beta(\tau, x)\, \varphi_x(\tau, x) \}\, dx\, d\tau \\
&+ \int_0^1 U(u(T, x) - c)\, \varphi(T, x)\, dx \\
&- \int_0^1 U(u_i(x) - c)\, \varphi(0, x)\, dx \\
&+ \int_0^T G(u(\tau, 1-) - c, u_1(\tau) - c;\ \beta(\tau, 1))\, \varphi(\tau, 1)\, d\tau \\
&- \int_0^T G(u_0(\tau) - c, u(\tau, 0+) - c;\ \beta(\tau, 0))\, \varphi(\tau, 0)\, d\tau \\
&- \int_0^T \int_0^1 \{ \beta_x(\tau, x)\, V(u(\tau, x), c)\, \varphi(\tau, x) \}\, dx\, d\tau,
\end{aligned} \tag{2.11b}$$

where the 'entropy' flux G and the function V are defined by

$$G(u_{\text{left}}, u_{\text{right}}; \beta) = U(u_{\text{left}})\, \beta^+ + U(u_{\text{right}})\, \beta^-, \tag{2.11c}$$

$$V(u, c) = U(u - c) - u\, U'(u - c), \tag{2.11d}$$

and where β is obtained from u by (2.3).

Theorem 2.4 (on the form $E^{\varepsilon_0,\varepsilon}$). *Suppose that the CFL condition* (2.9) *is satisfied. Then*

$$E^{\varepsilon_0,\varepsilon}(u, v; \beta) = \lim_{h \downarrow 0} E^{\varepsilon_0,\varepsilon}(u_h, v; \beta_h) \le 0.$$

Moreover, there exist two constants C_2 and C_3, depending only on the data and T, such that

$$E^{\varepsilon_0,\varepsilon}(u_h, u; \beta_h) \le L C_2 (\Delta x/\varepsilon + \Delta\tau/\varepsilon_0) + M C_3 \Delta\tau,$$

where $L = \sup_{u \in \mathbb{R}} |U'(u)|$, $M = \sup_{u \in \mathbb{R}} |U''(u)|$.

2d. **A result about the continuity at the boundary.** In this section we state a result concerning the behavior at the boundary of the approximate solution.

Our result will be stated in terms of the following quantities

$$\nu^+_{x,0}(\varepsilon, u; \beta) = \sup_{0\le\Delta\le\varepsilon} \int_0^T |u(\tau, \Delta) - u(\tau, 0-)|\, \beta^+(\tau, 0)\, d\tau,$$

$$\nu^-_{x,1}(\varepsilon, u; \beta) = \sup_{0\le\Delta\le\varepsilon} -\int_0^T |u(\tau, 1-\Delta) - u(\tau, 1+)|\, \beta^-(\tau, 1)\, d\tau,$$

which give a measure of the smoothness with which the boundary conditions are satisfied. To see this, suppose that $\nu^+_{x,0}(\varepsilon, u; \beta) \le C\varepsilon$ and that $\beta^+(\tau, 0) \ge \beta_{\min} > 0$ for $\tau \in [0, T]$. Then

$$\begin{aligned}\int_0^T |u(\tau, \varepsilon) - u(\tau, 0-)|\, d\tau &\le \int_0^T |u(\tau, \varepsilon) - u(\tau, 0-)|\, \beta^+(\tau, 0)\, d\tau/\beta_{\min}\\ &\le \nu^+_{x,0}(\varepsilon, u; \beta)/\beta_{\min}\\ &\le \varepsilon/\beta_{\min}.\end{aligned}$$

In other words, the mapping $x \mapsto u(x)$ is Lipschitz continuous in $L^1(0, T)$ at $x = 0$ provided that the inflow velocity, $\beta^+(\tau, 0)$, is bounded away from zero uniformly. However, this continuity property does not hold if $\beta^+(\tau, 0) = 0$ for some $\tau \in [0, T]$ (think of the extreme case in which $\beta \equiv 0$ and a discontinuity is sitting precisely at $x = 0$).

If $\beta^+_h(\tau, 0) \ge \beta_{\min} > 0$ for $\tau \in [0, T]$, it is possible to obtain an estimate of $\nu^+_{x,0}(\varepsilon, u_h; \beta_h)$. (Such an estimate follows from a uniform bound on the total variation in time of $u_h(x)$, for x near $x = 0$; see §3e of the Supplement.) However, if $\beta^+_h(\tau, 0)$ is very close to zero, the estimate is much harder to get. In the technique we use, we need to control the number of disjoint intervals on which $\beta^+_h(\tau, 0)$ can be uniformly bounded away from zero. (Similarly, we also need to control the number of disjoint intervals on which $\beta^-_h(\tau, 1)$ can be uniformly bounded away from zero.) For this purpose, we introduce the set $\{K_l\}_{l=1}^N$ of disjoint intervals such that $(0, T) = \bigcup_{l=1}^N K_l$, and consider the following two important cases:

(2.12a) $u^* = 1$, $u_0 \equiv 0$, and $u_1|_{K_l}$, $\phi_1|_{K_l}$ are constant, $\quad l = 1, \dots, N$,

(2.12b) $\phi_1|_{K_l} \in W^{1,\infty}(K_l)$, $\quad l = 1, \dots, N$.

In the first case, it can be proven that the above-mentioned numbers can be uniformly bounded. In the second case, we allow those numbers to increase unboundedly as Δx goes to zero. However, their growth can be controlled by using the regularity condition on ϕ_1.

Theorem 2.5 (Continuity at the boundary). *Suppose that the* CFL *condition* (2.9) *is satisfied. If the hypothesis* (2.12a) *is satisfied, then there exist constants* C_4 *and* C_5, *depending solely on the data and* T, *such that*

(2.13a) $$\nu^+_{x,0}(\varepsilon, u_h; \beta_h) \le C_4(\varepsilon + \Delta x),$$

(2.13b) $$\nu^-_{x,1}(\varepsilon, u_h; \beta_h) \le C_5(\varepsilon + \Delta x).$$

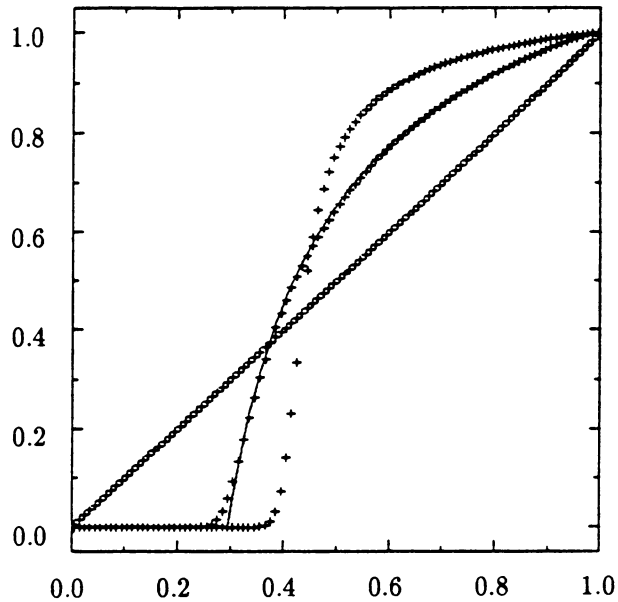

FIGURE 2. Example 1: the approximate solution u_h, with $\Delta x_i \equiv 1/100$, at $T = 0, 1, 2$. The '+' represents the value of the approximate solution at the middle of the elements. The solid line represents the exact solution at time $T = 1$

If the hypothesis (2.12b) *is satisfied, then there exist constants* $\widehat{C}_4$ *and* $\widehat{C}_5$, *depending solely on the data and* T, *such that*

$$\nu^{+}_{x,0}(\varepsilon, u_h; \beta_h) \le \widehat{C}_4 (\varepsilon + \Delta x)^{1/2}, \tag{2.14a}$$

$$\nu^{-}_{x,1}(\varepsilon, u_h; \beta_h) \le \widehat{C}_5 (\varepsilon + \Delta x)^{1/2}, \tag{2.14b}$$

provided that $\varepsilon + \Delta x$ *is small enough.*

2e. **Numerical results.** The numerical experiments we describe in this section have been designed (i) to show the qualitative behavior of the approximate solution, (ii) to display the convergence properties of the scheme, and (iii) to show that our approximate solution approximates well the viscous solution (2.15) for the specified range of λ-values, even for large values of T. In all the examples the CFL condition (2.9) has been used.

Example 1. $u_0 \equiv 0$, $u_1 \equiv 1$, $u_i(x) = x$, $\phi_1 \equiv 1/8$.

In Figure 2 we display the approximate solution u_h at different times (the exact solution at $T = 1$ is also included); compare with Figure 1. Notice that the monotonicity of the solution is preserved, as stated in Corollary 3.7 of the Supplement.

In Tables 1 and 2 we display the errors and their respective order of convergence at time $T = 1$. From Table 1 we see that the scheme is first-order accurate in the electric field and in the electric potential. We also see that in this case the scheme is first-order accurate in L^1 but only half-order accurate

TABLE 1. Example 1: Convergence in $(0, 1)$ at $T = 1$ and $\Delta x = 1/100$

	L^∞		L^1	
	$10^2 \cdot$ error	order	$10^2 \cdot$ error	order
u	5.7713	0.545	0.4045	0.937
β	0.1176	0.924	0.0279	0.997
ϕ	0.1820	0.986	0.0286	1.003

TABLE 2. Example 1: Convergence in $(0, 1) \setminus [.25, .35]$ at $T = 1$ and $\Delta x = 1/100$

	L^∞		L^1	
	$10^2 \cdot$ error	order	$10^2 \cdot$ error	order
u	2.2075	0.890	0.2142	1.121
β	0.1005	1.055	0.0211	1.112
ϕ	0.1819	0.986	0.0244	1.051

in L^∞ in the concentration. This discrepancy is caused by the discontinuity of the derivative of the exact electron concentration in the interval $(.25, .35)$. To see this, it is enough to study the errors outside this interval. Indeed, in Table 2 we see that the scheme is first-order accurate (in the electron concentration) in both L^1 and L^∞ on $[0, .25] \cup [.35, 1]$. These observations indicate that the scheme is first-order accurate when the solution is smooth enough. They also indicate that the influence of the discontinuity in the derivative of the electron concentration on the quality of the approximation has only a local effect.

Example 2. $u_0 \equiv 0$, $u_1 \equiv 1$, $u_i \equiv 0$, $\phi_1 \equiv 1/8$.

In Figure 3 (see next page) we display the approximate and exact solutions at $T = 2$. Notice how the biggest error in the approximation of u and of β occurs around the location of the discontinuity $x \approx .6398$ of u. Notice also that the biggest error in ϕ occurs at the critical point of ϕ, not at the location of the discontinuity of u.

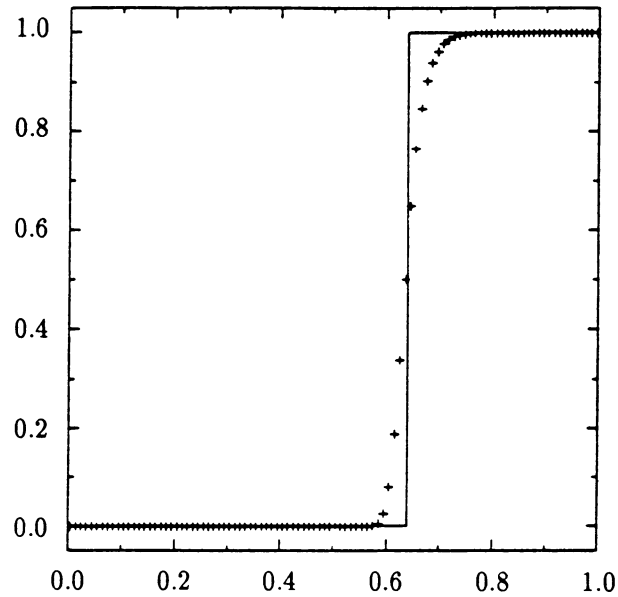

FIGURE 3a. Example 2: the approximate concentration u_h at $T = 2$ with $\Delta x_i \equiv 1 / 100$. The '+' represents the value of u_h at the middle of the elements. The solid line represents the exact solution

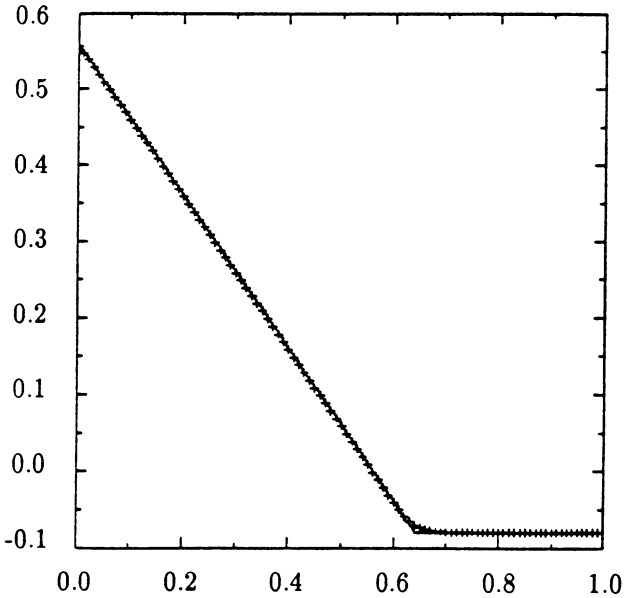

FIGURE 3b. Example 2: the approximate negative electric field β_h at $T = 2$ with $\Delta x_i \equiv 1/100$. The '+' represents the value of β_h at the extremes of the elements. The solid line represents the exact solution

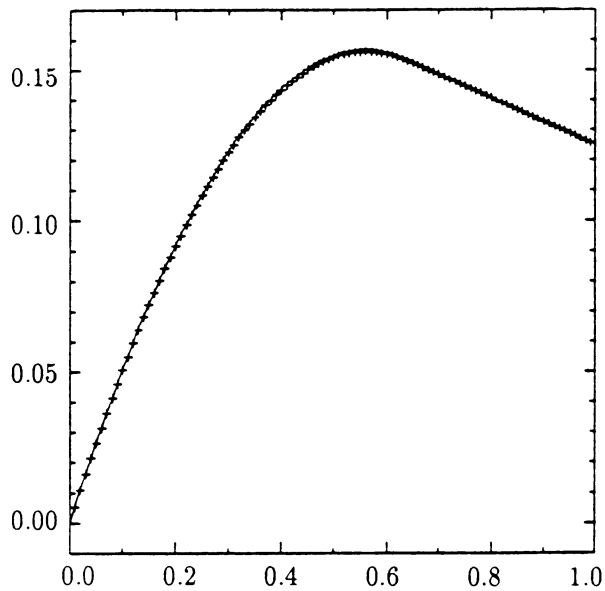

FIGURE 3c. Example 2: the approximate electric potential ϕ_h at $T = 2$ with $\Delta x_i \equiv 1/100$. The '+' represents the value of ϕ_h at the middle of the elements. The solid line represents the exact solution

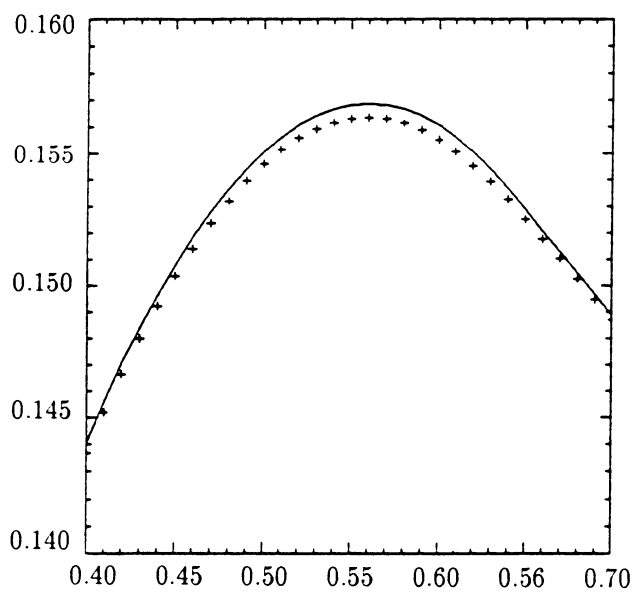

FIGURE 3d. Example 2: zoom on the previous figure. The biggest error in the approximation of ϕ by ϕ_h does not occur at the location $x \approx 0.6398$ of the discontinuity of u

TABLE 3. Example 2: Convergence in $(0, 1)$ at $T = 2$ and $\Delta x = 1/800$

	L^∞		L^1	
	$10^2 \cdot$ error	order	$10^2 \cdot$ error	order
u	-	-	0.7886	0.487
β	0.3877	0.489	0.0146	0.964
ϕ	0.0271	0.998	0.0068	0.996

In Table 3 we show the errors and their orders of convergence. Notice that the orders of convergence, in L^1, to u and β are $\sim 1/2$ and ~ 1, respectively. This indicates that the presence of the discontinuity in u has only a local effect in the approximation of β.

Example 3. $u_0 \equiv 0$, $u_1 \equiv 1$, $u_i \equiv 0$, $\phi_1 \equiv 1/8$. In this experiment we want to compare the approximate solution given by the method (2.8) with the 'viscous' solution $(u^\lambda, \beta^\lambda, \phi^\lambda)$ defined by the following equations:

$$u_\tau^\lambda + (u^\lambda \beta^\lambda)_x = \lambda^2 u_{xx}^\lambda, \qquad \tau > 0,\ x \in (0, 1), \tag{2.15a}$$

$$u^\lambda(\tau, 0) = u_0(\tau), \qquad \tau \geq 0, \tag{2.15b}$$

$$u^\lambda(\tau, 1) = u_1(\tau), \qquad \tau \geq 0, \tag{2.15c}$$

$$u^\lambda(0, x) = u_i(x), \qquad x \in (0, 1), \tag{2.15d}$$

where

$$-\beta_x^\lambda = 1 - u^\lambda, \qquad x \in (0, 1),\ \tau \geq 0, \tag{2.15e}$$

$$\beta^\lambda = \phi_x^\lambda, \qquad x \in (0, 1),\ \tau \geq 0, \tag{2.15f}$$

$$\phi^\lambda(\tau, 0) = 0, \qquad \text{for } \tau \geq 0, \tag{2.15g}$$

$$\phi^\lambda(\tau, 1) = \phi_1(\tau), \qquad \text{for } \tau \geq 0. \tag{2.15h}$$

We take the same discretization parameters as in the previous example. In Figure 4 the approximate concentration, the 'nonviscous' and the 'viscous' concentrations are compared at time $T = 10$ around the boundary layer. (If those concentrations would have been compared on the full interval $[0, 1]$, no difference between them would have been detected.) At that time all the solutions are very close to the stationary solutions. We see that the size of the boundary layer created by the diffusion is approximately equal to $\Delta x = 1/100$ for the biggest value of λ^2, i.e., for $\lambda^2 = 10^{-6}$. We see that the approximate solution provides an excellent approximation to the 'viscous' concentration.

Our numerical results thus indicate that the method (2.8) is a very robust method, which is uniformly first-order accurate in L^∞ for the electron concentration, the electric field, and the electric potential when the exact solution is

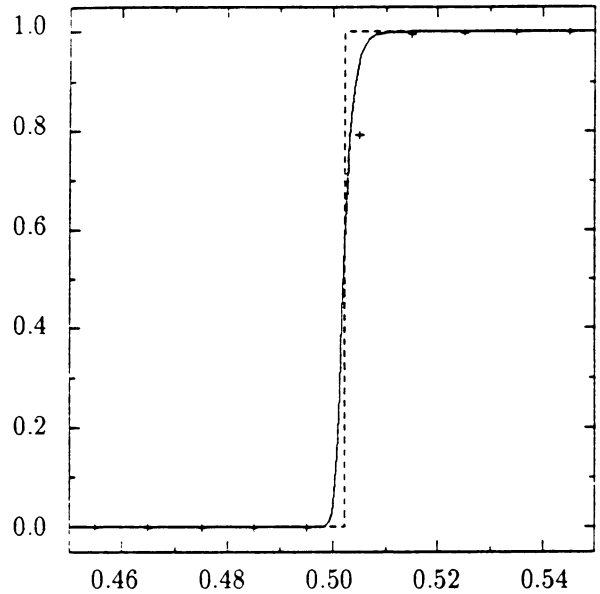

FIGURE 4. Example 3: detail of the concentrations u_h, u and u^λ at $T = 10$. In this case $\Delta x_i \equiv 1/100$ and $\lambda = 10^{-3}$. The '+' represent the values of u_h at the midpoints of the elements, the dashed line represents u, and the solid line u^λ

smooth. If the electron concentration is discontinuous, the scheme is half-order accurate in L^1 for the electron concentration and half-order accurate in L^∞ for the electric field. We have also verified that in this case, the presence of the discontinuity has only a local effect on the quality of the approximation. Finally, we have verified that the scheme provides a good approximation to the 'viscous' solution, see (2.15), even near the stationary state.

ACKNOWLEDGMENTS

The authors want to thank Franco Brezzi, Donatella Marini and Huanan Yang for fruitful discussions. The authors also want to thank the reviewer, the Managing Editor, and Zhangxin Chen for their remarks leading to an improved presentation of this paper.

BIBLIOGRAPHY

1. R. E. Bank, D. J. Rose and W. Fichtner, *Numerical methods for the semiconductor device simulation*, IEEE Trans. Electron Devices **ED-30** (1983), 1031–1041.
2. R. E. Bank, W. M. Coughran, Jr., W. Fichtner, E. H. Grosse, D. J. Rose, and R. K. Smith, *Transient simulation of silicon devices and circuits*, IEEE Trans. Computer-Aided Design **CAD-4** (1985), 436–451.
3. F. Brezzi, *Theoretical and numerical problems in reverse biased semiconductor devices*, Computing methods in Applied Sciences and Engineering, VII (R. Glowinski and J. L. Lions, eds.), Elsevier, Amsterdam, 1986, pp. 45–58.
4. F. Brezzi, A. Capelo, and L. D. Marini, *A singular perturbation analysis for semiconductor devices*, SIAM J. Math. Anal. **20** (1989), 372–387.
5. F. Brezzi, J. Douglas, Jr., and L. D. Marini, *Two families of mixed finite elements for second order elliptic problems*, Numer. Math. **47** (1985), 217–235.

6. F. Brezzi and L. Gastaldi, *Mathematical properties of one-dimensional semiconductors*, Mat. Apl. Comput. **5** (1986), 123–138.

7. F. Brezzi, L. D. Marini, and P. Pietra, *Two-dimensional exponential fitting and applications to drift-diffusion models*, SIAM J. Numer. Anal. **26** (1989), 1324–1355.

8. F. Brezzi and P. A. Markowich, *A convection-diffusion problem with small diffusion coefficient arising in semiconductor physics*, Boll. Un. Mat. Ital. **2B** (1988), 903–930.

9. B. Cockburn and C. W. Shu, *The Runge-Kutta local projection P^1-discontinuous Galerkin method for scalar conservation laws*, M^2AN **25** (1991), 337–361.

10. ——, *TVB Runge-Kutta local projection discontinuous Galerkin finite element method for scalar conservation laws* II: *General framework*, Math. Comp. **52** (1989), 411–435.

11. B. Cockburn, S. Y. Lin, and C. W. Shu, *TVB Runge-Kutta local projection discontinuous Galerkin finite element method for conservation laws* III: *One dimensional systems*, J. Comput. Phys. **84** (1989), 90–113.

12. B. Cockburn, S. Hou, and C. W. Shu, *TVB Runge-Kutta local projection discontinuous Galerkin finite element method for conservation laws* IV: *The multidimensional case*, Math. Comp. **54** (1990), 545–581.

13. B. Cockburn, *The quasi-monotone schemes for scalar conservation laws. Part* I, SIAM J. Numer. Anal. **26** (1989), 1325–1341.

14. W. M. Coughran, Jr. and J. W. Jerome, *Modular algorithms for transient semiconductor device simulation, Part* I: *Analysis of the outer iteration*, Lectures in Appl. Math., vol. 25, Amer. Math. Soc., Providence, R.I., 1990, pp. 107–149.

15. M. Crandall and A. Majda, *Monotone difference approximations for scalar conservation laws*, Math. Comp. **34** (1980), 1–21.

16. J. Douglas, Jr., I. Martínez-Gamba, and M. C. J. Squeff, *Simulation of the transient behavior of a one-dimensional semiconductor device*, Mat. Apl. Comput. **5** (1986), 103–122.

17. W. Fichtner, D. J. Rose, and R. E. Bank, *Semiconductor device simulation*, IEEE Trans. Electron Devices **ED-30** (1983), 1018–1030.

18. A. Harten, *On a class of high-resolution total-variation-stable finite-difference schemes*, SIAM J. Numer. Anal. **21** (1984), 1–23.

19. J. W. Jerome, *Consistency of semiconductor modeling: An existence/stability analysis for the stationary Van Roosbroeck system*, SIAM J. Appl. Math. **45** (1985), 565–590.

20. J. W. Jerome and T. Kerkhoven, *A finite element approximation theory for the drift diffusion semiconductor model*, SIAM J. Numer. Anal. **28** (1991), 403–422.

21. J. W. Jerome, *Evolution systems in semiconductor device modeling: A cyclic uncoupled line analysis for the Gummel map*, Math. Methods Appl. Sci. **9** (1987), 455–492.

22. T. Kerkhoven, *A proof of convergence of Gummel's algorithm for realistic device problems*, SIAM J. Numer. Anal. **23** (1986), 1121–1137.

23. S. N. Kruzhkov, *First order quasilinear equations in several independent variables*, Math. USSR Sb. **10** (1970), 217–243.

24. N. N. Kuznetsov, *Accuracy of some approximate methods for computing the weak solutions of a first-order quasi-linear equation*, USSR Comput. Math. and Math. Phys. **16** (1976), 105–119.

25. A. Y. LeRoux, *Etude du problème mixte pour une équation quasilinéaire du premier ordre*, C. R. Acad. Sci. Paris, Sér. A **285** (1977), 351–354.

26. B. J. Lucier, *Error bounds for the methods of Glimm, Godunov and LeVeque*, SIAM J. Numer. Anal. **22** (1985), 1074–1081.

27. ——, *A moving mesh numerical method for hyperbolic conservation laws*, Math. Comp. **46** (1986), 59–69.

28. ——, *On nonlocal monotone difference schemes for scalar conservation laws*, Math. Comp. **47** (1986), 19–36.

29. P. A. Markovich, *The stationary semiconductor equations*, Springer-Verlag, New York, 1986.

30. ——, *Spatial-temporal structure of solutions of the semiconductor problem, computational aspects of VLSI Design with an emphasis on semiconductor device simulation*, Lectures in Appl. Math., vol. 25, Amer. Math. Soc., Providence, R.I., 1990, pp. 1–26.

31. P. A. Markowich and P. Szmolyan, *A system of convection-diffusion equations with small diffusion coefficient arising in semiconductor physics*, preprint.

32. I. Martínez-Gamba and M. C. J. Squeff, *Simulation of the transient behavior of a one-dimensional semiconductor device,* II, SIAM J. Numer. Anal. **26** (1989), 539–552.

33. P. A. Raviart and J. M. Thomas, *A mixed finite element method for second order elliptic problems*, Lecture Notes in Math., Springer-Verlag, 1977.

34. C. Ringhofer, *An asymptotic analysis of a transient pn-junction model*, SIAM J. Appl. Math. **47** (1987), 624–642.

35. C. Ringhofer and C. Schmeiser, *An approximate Newton method for the solution of the basic semiconductor device equations*, SIAM J. Numer. Anal. **26** (1989), 507–516.

36. R. Sanders, *On convergence of monotone finite difference schemes with variable spatial differencing*, Math. Comp. **40** (1983), 91–106.

37. P. Szmolyan, *Initial transients of solutions of the semiconductor equations*, preprint.

38. ——, *Asymptotic methods for transient semiconductor device equations*, COMPEL **8** (1989), 113–122.

39. V. W. Van Roosbroeck, *Theory of flow of electrons and holes in germanium and other semiconductors*, Bell Syst. Tech. J. **29** (1950), 560–607.

40. M. Zlámal, *Finite element solution of the fundamental equations of semiconductor devices*, Math. Comp. **46** (1986), 27–43.

School of Mathematics, University of Minnesota, Minneapolis, Minnesota 55455

Current address, I. Triandaf: U.S. Naval Research Laboratory, Code 4700.3, Washington, D.C. 20375-5000

MATHEMATICS OF COMPUTATION
VOLUME 59, NUMBER 200
OCTOBER 1992, PAGES S29–S46

Supplement to

CONVERGENCE OF A FINITE ELEMENT METHOD FOR THE DRIFT-DIFFUSION SEMICONDUCTOR DEVICE EQUATIONS: THE ZERO DIFFUSION CASE

BERNARDO COCKBURN AND IOANA TRIANDAF

3. PROOFS OF THEOREMS 2.1, 2.2 AND 2.5

In this section we prove Theorems 2.1, 2.2 and 2.5. In §3a we obtain a maximum principle for u_h, and in §3b a bound on the total variation in space. In §3c the important continuity result (with respect to the data) is obtained. In §3d we end the proof of Theorems 2.1 and 2.2. In §3e we prove several results that will allow us to prove Theorem 2.5 in §3f.

3a. A maximum principle for u_h. We begin by showing that under a suitable condition the approximate concentration u_h satisfies a maximum principle. We show in particular that $u_h \geq 0$ provided $u_i, u_0, u_1 \geq 0$, as we expect from the physical point of view. In this section we prove a maximum principle for the approximate solution u_h. We stress the fact that in the framework of conservation laws it is very well known that any maximum principle for explicit schemes is valid only under a certain Courant-Friedrichs-Levy (CFL) condition. Roughly speaking, this condition asks for the speed of the propagation of the information allowed by the numerical scheme to be bigger than the maximum speed of propagation of the information inherent in the conservation law. Since in our case the latter speed is approximated by the quantity $\| \beta_h \|_{L^\infty(0,1)}$ it is clear that in order to get a maximum principle for u_h, we must obtain an upper bound for that quantity. The next three lemmas are devoted to obtain the CFL condition. The bound on $\| \beta_h \|_{L^\infty(0,1)}$ is obtained in Proposition 3.4. The maximum principle is given in Proposition 3.5.

The following simple but important result is the essential link between the mixed finite element method (2.7) and the monotone scheme (2.6). It is a restatement of equation (2.7a) with $w_{\Delta x}(x) = \begin{cases} 1, & x \in I_i, \\ 0, & \text{otherwise.} \end{cases}$

LEMMA 3.1. *For $n = 0, \cdots, n_T - 1$ we have*

$$\beta^n_{i+\frac{1}{2}} = \beta^n_{i-\frac{1}{2}} - (1 - u^n_i)\Delta x_i, \quad i = 1, \ldots, n_x.$$

The next result displays the CFL condition under which a maximum principle is satisfied.

LEMMA 3.2 (The CFL condition). *Let u^n_h be such that*

$$u^n_h(0-), u^n_h(x), u^n_h(1+) \in [0, u^*], \quad x \in (0,1).$$

Suppose that

$$1 - \frac{\Delta\tau^n}{\Delta x_i}(\beta^{n+}_{i+\frac{1}{2}} - \beta^{n-}_{i-\frac{1}{2}}) \geq \Delta\tau^n, \quad i = 1, \cdots, n_x. \tag{3.1}$$

Then

$$u^{n+1}_h(x) \in [0, u^*], \quad x \in (0,1).$$

The assumption $u^* \geq 1$, see (2.1e), is crucial; the result is simply not true if $u^* < 1$. This reflects the fact that $u = 1$ is an asymptotically stable equilibrium point for the equation satisfied by u along the characteristics, (1.5); see remark (iii) in §1.

Proof. Pick $i \in \{1, \cdots, n_x\}$. By (2.6),

$$\begin{aligned} u_i^{n+1} &= u_i^n - \frac{\Delta\tau^n}{\Delta x_i}(\beta^{n-}_{i+\frac{1}{2}} u^n_{i+1} + \beta^{n+}_{i+\frac{1}{2}} u^n_i - \beta^{n-}_{i-\frac{1}{2}} u^n_i - \beta^{n+}_{i-\frac{1}{2}} u^n_{i-1}) \\ &= D^n_i\,[A^n_{i+1}\,u^n_{i+1} + B^n_i\,u^n_i + C^n_{i-1}\,u^n_{i-1}], \end{aligned}$$

where

$$\begin{aligned} A^n_{i+1} &= -\frac{\Delta\tau^n}{\Delta x_i}\frac{\beta^{n-}_{i+\frac{1}{2}}}{D^n_i}, \\ B^n_i &= (1 - \frac{\Delta\tau^n}{\Delta x_i}(\beta^{n+}_{i+\frac{1}{2}} - \beta^{n-}_{i-\frac{1}{2}}))/D^n_i, \\ C^n_{i-1} &= \frac{\Delta\tau^n}{\Delta x_i}\frac{\beta^{n+}_{i-\frac{1}{2}}}{D^n_i}, \\ D^n_i &= 1 - \frac{\Delta\tau^n}{\Delta x_i}(\beta^n_{i+\frac{1}{2}} - \beta^n_{i-\frac{1}{2}}). \end{aligned}$$

Note that by definition and by (3.1),

$$D^n_i \geq 1 - \frac{\Delta\tau^n}{\Delta x_i}(\beta^{n+}_{i+\frac{1}{2}} - \beta^{n-}_{i-\frac{1}{2}}) \geq \Delta\tau^n, \tag{3.2a}$$

$$A^n_{i+1} + B^n_i + C^n_{i-1} = 1, \tag{3.2b}$$

$$A^n_{i+1},\ B^n_i,\ C^n_{i-1} \geq 0. \tag{3.2c}$$

It is thus clear that $u_i^{n+1} \geq 0$ by the hypothesis on u_h^n.

To prove that $u_i^{n+1} \leq u^*$, we proceed as follows. Since

$$\begin{aligned} u_i^{n+1} &= D^n_i\,[u^* - A^n_{i+1}\,(u^* - u^n_{i+1}) - B^n_i\,(u^* - u^n_i) - C^n_{i-1}\,(u^* - u^n_{i-1})] \\ &= \{D^n_i\,[1 - A^n_{i+1}\frac{u^* - u^n_{i+1}}{u^*} - B^n_i\frac{u^* - u^n_i}{u^*} - C^n_{i-1}\frac{u^* - u^n_{i-1}}{u^*}]\}\,u^* \\ &= \{D^n_i(1 - E^n_i)\}\,u^* \\ &= \{(1 + \Delta\tau^n\,(1 - u^n_i))(1 - E^n_i)\}\,u^*, \quad \text{by Lemma 3.1,} \end{aligned}$$

where

$$E^n_i = A^n_{i+1}\frac{u^* - u^n_{i+1}}{u^*} + B^n_i\frac{u^* - u^n_i}{u^*} + C^n_{i-1}\frac{u^* - u^n_{i-1}}{u^*},$$

it is evident that $u_i^{n+1} \leq u^*$ if and only if $E^n_i \geq \Delta\tau^n\,(1 - u^n_i)/(1 + \Delta\tau^n\,(1 - u^n_i))$. (Notice that the denominator in the above fraction is nonnegative thanks to Lemma 3.1 and condition (3.1).) If $u_i > 1$, the above inequality is trivially verified. If $u_i \leq 1$, we have, by the hypothesis on u^n_h and (3.2),

$$\begin{aligned} E^n_i &\geq B^n_i\frac{u^* - u^n_i}{u^*} \\ &\geq B^n_i(1 - u^n_i), \quad \text{since } u^* \geq 1, \\ &= \frac{1 - \frac{\Delta\tau^n}{\Delta x_i}(\beta^{n+}_{i+\frac{1}{2}} - \beta^{n-}_{i-\frac{1}{2}})}{1 + \Delta\tau^n\,(1 - u^n_i)}\,(1 - u^n_i) \\ &\geq \frac{\Delta\tau^n(1 - u^n_i)}{1 + \Delta\tau^n\,(1 - u^n_i)}, \end{aligned}$$

provided condition (3.1) is satisfied. This completes the proof. □

We need now to rewrite condition (3.1) in terms of u^n_h and $\phi_1(\tau^n)$ only.

LEMMA 3.3. *Suppose that the following condition is satisfied:*

$$\Delta\tau^n \leq \min\{1, \frac{1}{\|\,u^n_h\,\|_{L^\infty(0,1)}}, \frac{\Delta x_i}{\Delta x_i + \|\,\beta^n_h\,\|_{L^\infty(0,1)}}\}. \tag{3.3}$$

Then condition (3.1) *holds.*

Proof. We have

$$\begin{aligned} \Theta^n_i &:= 1 - \frac{\Delta\tau^n}{\Delta x_i}(\beta^{n+}_{i+\frac{1}{2}} - \beta^{n-}_{i-\frac{1}{2}}) \\ &= \begin{cases} 1 - \frac{\Delta\tau^n}{\Delta x_i}(\beta^n_{i+\frac{1}{2}} - \beta^n_{i-\frac{1}{2}}) & \text{if } \beta^n_{i+\frac{1}{2}} \geq 0 \text{ and } \beta^n_{i-\frac{1}{2}} \leq 0, \\ 1 - \frac{\Delta\tau^n}{\Delta x_i}|\beta^n_{i+\frac{1}{2}}| & \text{if } \beta^n_{i+\frac{1}{2}} \geq 0 \text{ and } \beta^n_{i-\frac{1}{2}} \geq 0, \\ 1 - \frac{\Delta\tau^n}{\Delta x_i}|\beta^n_{i-\frac{1}{2}}| & \text{if } \beta^n_{i+\frac{1}{2}} \leq 0 \text{ and } \beta^n_{i-\frac{1}{2}} \leq 0, \\ 1 & \text{otherwise} \end{cases} \\ &\geq \begin{cases} 1 + \Delta\tau^n\,(1 - u^n_i) & \text{if } \beta^n_{i+\frac{1}{2}} \geq 0 \text{ and } \beta^n_{i-\frac{1}{2}} \leq 0, \\ 1 - \frac{\Delta\tau^n}{\Delta x_i}\|\,\beta^n_h\,\|_{L^\infty(0,1)} & \text{if } \beta^n_{i+\frac{1}{2}} \geq 0 \text{ and } \beta^n_{i-\frac{1}{2}} \geq 0, \\ 1 - \frac{\Delta\tau^n}{\Delta x_i}\|\,\beta^n_h\,\|_{L^\infty(0,1)} & \text{if } \beta^n_{i+\frac{1}{2}} \leq 0 \text{ and } \beta^n_{i-\frac{1}{2}} \leq 0, \\ 1 & \text{otherwise,} \end{cases} \end{aligned}$$

where we have used Lemma 3.1 in the last line. Thus, $\Theta^n_i \geq \Delta\tau^n$ if

$$\Delta\tau^n \leq \begin{cases} \frac{1}{\|\,u^n_h\,\|_{L^\infty(0,1)}} & \text{if } \beta^n_{i+\frac{1}{2}} \geq 0 \text{ and } \beta^n_{i-\frac{1}{2}} \leq 0, \\ \frac{\Delta x_i}{\Delta x_i + \|\,\beta^n_h\,\|_{L^\infty(0,1)}} & \text{if } \beta^n_{i+\frac{1}{2}} \geq 0 \text{ and } \beta^n_{i-\frac{1}{2}} \geq 0, \\ \frac{\Delta x_i}{\Delta x_i + \|\,\beta^n_h\,\|_{L^\infty(0,1)}} & \text{if } \beta^n_{i+\frac{1}{2}} \leq 0 \text{ and } \beta^n_{i-\frac{1}{2}} \leq 0, \\ 1 & \text{otherwise.} \end{cases}$$

This completes the proof. □

Finally, we have to estimate $\|\,\beta^n_h\,\|_{L^\infty(0,1)}$.

PROPOSITION 3.4 (Estimate of $\| \beta_h \|_{L^\infty(0,1)}$). *For $n = 0, \cdots, n_T$ we have*

$$\| \beta_h^n - \phi_{1,\Delta\tau}(\tau^n) \|_{L^\infty(0,1)} \le \frac{1}{2} \| 1 - u_h^n \|_{L^\infty(0,1)}.$$

Proof. From equations (2.7) we easily obtain that

$$-(\beta_h^n - \phi_{1,\Delta\tau}(\tau^n))_x = 1 - u_h^n,$$
$$(\beta_h^n - \phi_{1,\Delta\tau}(\tau^n), 1) = 0.$$

The result follows immediately. □

We are now ready to state our maximum principle.

PROPOSITION 3.5 (A maximum principle for u_h). *Let u_0, u_1 and u_i be such that*

$$u_0(\tau), u_1(\tau), u_i(x) \in [0, u^*], \quad \tau \in [0, T],\ x \in (0,1),$$

for some $u^* \ge 1$. Suppose that for $i = 1, \cdots, n_x$, and $n = 0, \cdots, n_T - 1$ we have

$$\Delta\tau^n \le \min\{\frac{1}{u^*}, \frac{\Delta x_i}{\Delta x_i + \phi_{1,\Delta\tau}(\tau^n) + \frac{1}{2}\max\{1, u^* - 1\}}\}. \tag{3.4}$$

Then the following maximum principle holds:

$$u_h^n(x) \in [0, u^*], \quad x \in (0,1),\ n = 0, \cdots, n_T.$$

Proof. We proceed by induction on n. The result is trivially true for $n = 0$ by (2.5). Let us assume that it is true for $n = m$, and let us prove it is also true for $n = m + 1$. By the inductive hypothesis and (2.5) we have that

$$u_h^m(0-), u_h^m(x), u_h^m(1+) \in [0, u^*], \quad x \in (0,1).$$

Thus, $\| u_h^m \|_{L^\infty(0,1)} \le u^*$ and $\| 1 - u_h^m \|_{L^\infty(0,1)} \le \max\{1, u^* - 1\}$. Since $\phi_{1,\Delta\tau}(\tau^m) \ge 0$, the latter inequality implies that $\| \beta_h^n \|_{L^\infty(0,1)} \le \phi_{1,\Delta\tau}(\tau^m) + \frac{1}{2}\max\{1, u^* - 1\}$ by Proposition 3.4. Hence, by (3.4),

$$\begin{aligned}\Delta\tau^m &\le \min\{\frac{1}{u^*}, \frac{\Delta x_i}{\Delta x_i + \phi_{1,\Delta\tau}(\tau^m) + \frac{1}{2}\max\{1, u^* - 1\}}\} \\ &\le \min\{1, \frac{1}{\| u_h^m \|_{L^\infty(0,1)}}, \frac{\Delta x_i}{\Delta x_i + \| \beta_h^m \|_{L^\infty(0,1)}}\},\end{aligned}$$

and by Lemma 3.3, condition (3.1) is satisfied for $n = m$. By Lemma 3.2, we have that $u_h^{m+1}(x) \in [0, u^*]$, $\forall x \in (0,1)$. This completes the proof. □

3b. Bound on the total variation of u_h. In this section we show that under a suitable CFL condition the total variation of u_h^n in $(0,1)$ is uniformly bounded. Lemma 3.6 is a preliminary result. Corollary 3.7 is a monotonicity result for a special (but important) case. Lemma 3.8 relates the total variation of u_h^{n+1} to the total variation of u_h^n. Proposition 3.9 contains our total variation boundedness result.

LEMMA 3.6. *We have, for $i = 0$,*

$$\begin{aligned}u_1^{n+1} - u_0^{n+1} = &-\frac{\Delta\tau^n}{\Delta x_1}\beta_{\frac{3}{2}}^{n-}(u_2^n - u_1^n) \\ &+ (1 + \Delta\tau^n(1 - u_1^n - u_0^n) - \frac{\Delta\tau^n}{\Delta x_1}\beta_{\frac{1}{2}}^{n+})(u_1^n - u_0^n) \\ &+ ((1 + \Delta\tau^n(1 - u_0^n))u_0^n - u_0^{n+1}),\end{aligned}$$

for $i \in \{1, \cdots, n_x - 1\}$,

$$\begin{aligned}u_{i+1}^{n+1} - u_i^{n+1} = &-\frac{\Delta\tau^n}{\Delta x_{i+1}}\beta_{i+\frac{3}{2}}^{n-}(u_{i+2}^n - u_{i+1}^n) \\ &+ (1 + \Delta\tau^n(1 - u_{i+1}^n - u_i^n) - \frac{\Delta\tau^n}{\Delta x_{i+1}}\beta_{i+\frac{1}{2}}^{n+} + \frac{\Delta\tau^n}{\Delta x_i}\beta_{i+\frac{1}{2}}^{n-})(u_{i+1}^n - u_i^n) \\ &+ \frac{\Delta\tau^n}{\Delta x_i}\beta_{i-\frac{1}{2}}^{n+}(u_i^n - u_{i-1}^n),\end{aligned}$$

and for $i = n_x$

$$\begin{aligned}u_{n_x+1}^{n+1} - u_{n_x}^{n+1} = &(u_{n_x+1}^{n+1} - (1 + \Delta\tau^n(1 - u_{n_x+1}^n))u_{n_x+1}^n) \\ &+ (1 + \Delta\tau^n(1 - u_{n_x+1}^n - u_{n_x}^n) + \frac{\Delta\tau^n}{\Delta x_{n_x}}\beta_{n_x+\frac{1}{2}}^{n-})(u_{n_x}^n - u_{n_x+1}^n) \\ &+ \frac{\Delta\tau^n}{\Delta x_{n_x}}\beta_{n_x-\frac{1}{2}}^{n+}(u_{n_x}^n - u_{n_x-1}^n).\end{aligned}$$

Proof. By (2.6a) ,

$$\begin{aligned}u_{i+1}^{n+1} &= -\frac{\Delta\tau^n}{\Delta x_{i+1}}\beta_{i+\frac{3}{2}}^{n-}u_{i+2}^n + \frac{\Delta\tau^n}{\Delta x_{i+1}}\beta_{i+\frac{1}{2}}^{n+}u_i^n \\ &\quad + (1 - \frac{\Delta\tau^n}{\Delta x_{i+1}}(\beta_{i+\frac{3}{2}}^{n+} - \beta_{i+\frac{1}{2}}^{n-}))u_{i+1}^n \\ &= -\frac{\Delta\tau^n}{\Delta x_{i+1}}\beta_{i+\frac{3}{2}}^{n-}(u_{i+2}^n - u_{i+1}^n) + \frac{\Delta\tau^n}{\Delta x_{i+1}}\beta_{i+\frac{1}{2}}^{n+}(u_i^n - u_{i+1}^n) \\ &\quad + (1 - \frac{\Delta\tau^n}{\Delta x_{i+1}}(\beta_{i+\frac{3}{2}}^{n} - \beta_{i+\frac{1}{2}}^{n}))u_{i+1}^n \\ &= -\frac{\Delta\tau^n}{\Delta x_{i+1}}\beta_{i+\frac{3}{2}}^{n-}(u_{i+2}^n - u_{i+1}^n) + \frac{\Delta\tau^n}{\Delta x_{i+1}}\beta_{i+\frac{1}{2}}^{n+}(u_i^n - u_{i+1}^n) \\ &\quad + (1 + \Delta\tau^n(1 - u_{i+1}^n))u_{i+1}^n,\end{aligned}$$

by Lemma 3.1. The result follows after a few very simple algebraic manipulations. □

An immediate consequence of this result is the following monotonicity result.

COROLLARY 3.7 (A monotonicity result). *Let u_i, u_0 and u_1 be such that*

$$u_i(x) \in [0,1], x \in (0,1), \text{ is a nondecreasing function,}$$
$$u_0(\tau) = 0, \tau \in (0,T),$$
$$u_1(\tau) = 1, \tau \in (0,T),$$

or such that

$$u_i(x) \in [0,1], x \in (0,1), \text{ is a nonincreasing function,}$$
$$u_0(\tau) = 1, \tau \in (0,T),$$
$$u_1(\tau) = 0, \tau \in (0,T).$$

Suppose that for $i = 1, \cdots, n_x$ and $n = 0, \cdots, n_T - 1$ we have

$$\Delta\tau^n \le \min\{1, \frac{\Delta x_i}{\Delta x_i + \phi_{1,\Delta\tau}(\tau^n) + \frac{1}{2}}\}. \tag{3.5}$$

Then $u_h(\tau^n, \cdot)$ is a monotone function for $n = 0, \cdots, n_T$, and

$$\| u_h^n \|_{BV[0,1]} = \sum_{i=0}^{n_x} | u_{i+1}^n - u_i^n | = 1, \quad n = 0, \cdots, n_T.$$

Proof. We proceed by induction on n. The result is trivially true for $n = 0$ by (2.5d). Now, assume that the result is true for $n = m$ and let us prove that it is true for $n = m+1$. Pick $i \in \{1, \cdots, n_x - 1\}$. Consider the coefficient

$$\begin{aligned} C_{i+\frac{1}{2}}^m &:= 1 + \Delta\tau^m (1 - u_{i+1}^m - u_i^m) - \frac{\Delta\tau^m}{\Delta x_{i+1}} \beta_{i+\frac{1}{2}}^{m+} + \frac{\Delta\tau^m}{\Delta x_i} \beta_{i+\frac{1}{2}}^{m-} \\ &\ge 1 - \Delta\tau^m - \frac{\Delta\tau^m}{\Delta x_{i+1}} \beta_{i+\frac{1}{2}}^{m+} + \frac{\Delta\tau^m}{\Delta x_i} \beta_{i+\frac{1}{2}}^{m-}, \text{ by the inductive hypothesis,} \\ &\ge 1 - \Delta\tau^m - \frac{\Delta\tau^m}{\min\{\Delta x_i, \Delta x_{i+1}\}} \| \beta_h^m \|_{L^\infty(0,1)} \\ &\ge 1 - \Delta\tau^m - \frac{\Delta\tau^m}{\min\{\Delta x_i, \Delta x_{i+1}\}} (\phi_{1,\Delta\tau}(\tau^m) + \frac{1}{2}), \text{ by Lemma 3.4.} \end{aligned}$$

By (3.5), $C_{i+\frac{1}{2}}^m \ge 0$ for $i \in \{1, \cdots, n_x - 1\}$. In a similar way it can be proven that $C_{\frac{1}{2}}^m \cdot C_{n_x+\frac{1}{2}}^m \ge 0$. Thus, all the coefficients of the terms of the form $(u_{j+1} - u_j)$ in Lemma 3.6 are nonnegative, and the result follows. □

Notice that this result implies the maximum principle $u_h \in [0,1]$. In fact, since in our case $u^* = 1$, condition (3.5) is nothing but condition (3.4), and hence Proposition 3.5 is valid. This result is the discrete version of the corresponding result for the continuous problem proven in [8].

LEMMA 3.8 (Local total variation boundedness). *Suppose that for $i = 1, \cdots, n_x$, we have*

$$\Delta\tau^n \le \min\{1, \frac{\Delta x_i}{\Delta x_i \max\{1, 2\| u_h^n \|_{L^\infty(0,1)} - 1\} + \phi_{1,\Delta\tau}(\tau^n) + \frac{1}{2}\max\{1, \| u_h^n \|_{L^\infty(0,1)} - 1\}}\}. \tag{3.6}$$

Then

$$\begin{aligned} \| u_h^{n+1} \|_{BV[0,1]} &= \sum_{i=0}^{n_x} | u_{i+1}^{n+1} - u_i^{n+1} | \\ &\le \Delta\tau^n \left| \frac{u_0^{n+1} - u_0^n}{\Delta\tau^n} - (1 - u_0^n) u_0^n \right| \\ &\quad + \sum_{i=0}^{n_x} (1 + \Delta\tau^n (1 - u_{i+1}^n - u_i^n)) | u_{i+1}^n - u_i^n | \\ &\quad + \Delta\tau^n \left| \frac{u_{n_x+1}^{n+1} - u_{n_x+1}^n}{\Delta\tau^n} - (1 - u_{n_x+1}^n) u_{n_x+1}^n \right|. \end{aligned}$$

Proof. First, notice that condition (3.6) ensures the nonnegativity of the coefficients of the terms of the form $(u_{j+1}^n - u_j^n)$ in the expressions of Lemma 3.6. Taking absolute values in the expressions of Lemma 3.6, we get, for $i = 0$,

$$\begin{aligned} | u_1^{n+1} - u_0^{n+1} | &\le -\frac{\Delta\tau^n}{\Delta x_1} \beta_{\frac{3}{2}}^{n-} | u_2^n - u_1^n | \\ &\quad + (1 + \Delta\tau^n (1 - u_1^n - u_0^n) - \frac{\Delta\tau^n}{\Delta x_1} \beta_{\frac{1}{2}}^{n+}) | u_1^n - u_0^n | \\ &\quad + | (1 + \Delta\tau^n (1 - u_0^n)) u_0^n - u_0^{n+1} |, \end{aligned}$$

for $i = 1, \cdots, n_x - 1$,

$$\begin{aligned} | u_{i+1}^{n+1} - u_i^{n+1} | &\le -\frac{\Delta\tau^n}{\Delta x_{i+1}} \beta_{i+\frac{3}{2}}^{n-} | u_{i+2}^n - u_{i+1}^n | \\ &\quad + (1 + \Delta\tau^n (1 - u_{i+1}^n - u_i^n) - \frac{\Delta\tau^n}{\Delta x_{i+1}} \beta_{i+\frac{1}{2}}^{n+} + \frac{\Delta\tau^n}{\Delta x_i} \beta_{i+\frac{1}{2}}^{n-}) | u_{i+1}^n - u_i^n | \\ &\quad + \frac{\Delta\tau^n}{\Delta x_i} \beta_{i-\frac{1}{2}}^{n+} | u_i^n - u_{i-1}^n |, \end{aligned}$$

and for $i = n_x$,

$$\begin{aligned} | u_{n_x+1}^{n+1} - u_{n_x}^{n+1} | &\le | u_{n_x+1}^{n+1} - (1 + \Delta\tau^n (1 - u_{n_x+1}^n)) u_{n_x+1}^n | \\ &\quad + (1 + \Delta\tau^n (1 - u_{n_x+1}^n - u_{n_x}^n) + \frac{\Delta\tau^n}{\Delta x_{n_x}} \beta_{n_x+\frac{1}{2}}^{n-}) | u_{n_x}^n - u_{n_x+1}^n | \\ &\quad + \frac{\Delta\tau^n}{\Delta x_{n_x}} \beta_{n_x-\frac{1}{2}}^{n+} | u_{n_x}^n - u_{n_x-1}^n |. \end{aligned}$$

Hence we obtain

$$
\begin{aligned}
\| u_h^{n+1} \|_{\mathrm{BV}[0,1]} &= \sum_{i=0}^{n_x} | u_{i+1}^{n+1} - u_i^{n+1} | \\
&\le | (1 + \Delta\tau^n (1 - u_0^n)) u_0^n - u_0^{n+1} | \\
&\quad + (1 + \Delta\tau^n (1 - u_1^n - u_0^n) - \frac{\Delta\tau^n}{\Delta x_1} \beta_{\frac{1}{2}}^{n+}) | u_1^n - u_0^n | \\
&\quad - \sum_{i=1}^{n_x} \frac{\Delta\tau^n}{\Delta x_i} \beta_{i+\frac{1}{2}}^{n-} | u_{i+1}^n - u_i^n | \\
&\quad + \sum_{i=1}^{n_x - 1} (1 + \Delta\tau^n (1 - u_{i+1}^n - u_i^n) - \frac{\Delta\tau^n}{\Delta x_{i+1}} \beta_{i+\frac{1}{2}}^{n+} + \frac{\Delta\tau^n}{\Delta x_i} \beta_{i+\frac{1}{2}}^{n-}) | u_{i+1}^n - u_i^n | \\
&\quad + \sum_{i=0}^{n_x - 1} \frac{\Delta\tau^n}{\Delta x_{i+1}} \beta_{i+\frac{1}{2}}^{n+} | u_{i+1}^n - u_i^n | \\
&\quad + (1 + \Delta\tau^n (1 - u_{n_x+1}^n - u_{n_x}^n) + \frac{\Delta\tau^n}{\Delta x_{n_x}} \beta_{n_x+\frac{1}{2}}^{n-}) | u_{n_x}^n - u_{n_x+1}^n | \\
&\quad + | u_{n_x+1}^{n+1} - (1 + \Delta\tau^n (1 - u_{n_x+1}^n)) u_{n_x+1}^n |,
\end{aligned}
$$

and so,

$$
\begin{aligned}
\| u_h^{n+1} \|_{\mathrm{BV}[0,1]} &\le | (1 + \Delta\tau^n (1 - u_0^n)) u_0^n - u_0^{n+1} | \\
&\quad + \sum_{i=0}^{n_x} (1 + \Delta\tau^n (1 - u_{i+1}^n - u_i^n)) | u_{i+1}^n - u_i^n | \\
&\quad + | u_{n_x+1}^{n+1} - (1 + \Delta\tau^n (1 - u_{n_x+1}^n)) u_{n_x+1}^n |,
\end{aligned}
$$

and the result follows. □

Notice that if the segment $\{x = 0\} \times (\tau^n, \tau^{n+1})$ lies on a characteristic, then the expression

$$
\frac{u_0^{n+1} - u_0^n}{\Delta\tau^n} - (1 - u_0^n) u_0^n
$$

can be considered as a discretization of the equation (1.5). If the above quantity is equal to zero, then nothing flows across the boundary $x = 0$ during the time interval (τ^n, τ^{n+1}), and hence the influence of the boundary condition on the total variation of u_h should be none, as our result confirms. If $u_0 \equiv 0$, or if $u_0 \equiv 1$, this is precisely what happens. A similar remark can be made for the corresponding expression at the boundary $x = 1$.

Notice also that in the case of Corollary 3.7 it can be easily verified that this result reads as follows:

$$
\| u_h^{n+1} \|_{\mathrm{BV}[0,1]} \le \| u_h^n \|_{\mathrm{BV}[0,1]},
$$

as expected. However, this is a particular case that is far from being typical. By remark (iv) in §1 we expect the total variation of the exact solution to (locally) increase with a rate equal to e^t. This explains the factor e^T in our following result.

PROPOSITION 3.9 (Total variation boundedness). *Suppose that for $i = 1, \cdots, n_x$ and $n = 0, \cdots, n_T - 1$ we have*

$$
\Delta\tau^n \le \min\Big\{ \frac{1}{u^*}, \frac{\Delta x_i}{(2u^* - 1)\Delta x_i + \phi_{1,\Delta\tau}(\tau^n) + \frac{1}{2}\max\{1, u^* - 1\}} \Big\}. \tag{3.7}
$$

Then

$$
\| u_h \|_{L^\infty(0,T;\mathrm{BV}[0,1])} \le C_1,
$$

where

$$
\begin{aligned}
C_1 = e^T \big\{ &\| u_0 \|_{\mathrm{BV}[0,T]} + \| 1 - u_0 \|_{L^\infty(0,T)} \| u_0 \|_{L^1(0,T)} \\
&+ (\Delta\tau^0 \| u_0 \|_{\mathrm{BV}[0,\Delta\tau^0]} + |u_0(0+) - u_i(0+)| \\
&\quad + \| u_i \|_{\mathrm{BV}(0,1)} \\
&\quad + \Delta\tau^0 \| u_1 \|_{\mathrm{BV}[0,\Delta\tau^0]} + |u_1(0+) - u_i(1-)|) \\
&+ \| u_1 \|_{\mathrm{BV}[0,T]} + \| 1 - u_1 \|_{L^\infty(0,T)} \| u_1 \|_{L^1(0,T)} \big\}.
\end{aligned}
$$

Notice that u_0 and u_1 are the boundary data and that u_i is the initial condition.

Proof. Since condition (3.7) implies condition (3.4), Proposition 3.5 is satisfied. Since condition (3.7) is stronger than condition (3.6), Lemma 3.8 is also satisfied, and so

$$
\begin{aligned}
\| u_h^{n+1} \|_{\mathrm{BV}[0,1]} &\le \Delta\tau^n \Big| \frac{u_0^{n+1} - u_0^n}{\Delta\tau^n} - (1 - u_0^n) u_0^n \Big| \\
&\quad + (1 + \Delta\tau^n) \| u_h^n \|_{\mathrm{BV}[0,1]} \\
&\quad + \Delta\tau^n \Big| \frac{u_{n_x+1}^{n+1} - u_{n_x+1}^n}{\Delta\tau^n} - (1 - u_{n_x+1}^n) u_{n_x+1}^n \Big|.
\end{aligned}
$$

This implies that

$$
\begin{aligned}
\| u_h^{n+1} \|_{\mathrm{BV}[0,1]} &\le \sum_{\ell=0}^{n} e^{\tau^{n+1} - \tau^\ell} \Big| \frac{u_0^{\ell+1} - u_0^\ell}{\Delta\tau^\ell} - (1 - u_0^\ell) u_0^\ell \Big| \Delta\tau^\ell \\
&\quad + e^{\tau^{n+1}} \| u_h^0 \|_{\mathrm{BV}[0,1]} \\
&\quad + \sum_{\ell=0}^{n} e^{\tau^{n+1} - \tau^\ell} \Big| \frac{u_{n_x+1}^{\ell+1} - u_{n_x+1}^\ell}{\Delta\tau^\ell} - (1 - u_{n_x+1}^\ell) u_{n_x+1}^\ell \Big| \Delta\tau^\ell,
\end{aligned}
$$

and the result follows from (2.5). □

3c. Continuity with respect to the data. In this section we obtain a result concerning the continuous dependence of the approximate solution u_h with respect to the initial and boundary data. The Lemmas 3.10, 3.11 and 3.12 contain preliminary results. The continuity result is given in Proposition 3.13. Proposition 3.14 is an equicontinuity-in-time result which will be used in §4.1.

In what follows, we denote by (v_h, σ_h, ψ_h) the approximate solution (u_h, β_h, ϕ_h) of (1.3), (1.4) with u_i, u_0, u_1 and ϕ_1 replaced by v_i, v_0, v_1 and ψ_1, respectively.

LEMMA 3.10. *For $n = 0, ..., n_T - 1$ and $i = 1, ..., n_x$ we have the following identity:*

$$
\begin{aligned}
u_i^{n+1} - v_i^{n+1} = &\left\{-\frac{1}{2}\frac{\Delta\tau^n}{\Delta x_i}(\beta^{n-}_{i+\frac{1}{2}} + \sigma^{n-}_{i+\frac{1}{2}})\right\}(u^n_{i+1} - v^n_{i+1}) \\
&+ \left\{1 - \frac{1}{2}\Delta\tau^n(u_i^n + v_i^n) - \frac{1}{2}\frac{\Delta\tau^n}{\Delta x_i}((\beta^{n+}_{i+\frac{1}{2}} + \sigma^{n+}_{i+\frac{1}{2}}) - (\beta^{n-}_{i-\frac{1}{2}} + \sigma^{n-}_{i-\frac{1}{2}}))\right\}(u_i^n - v_i^n) \\
&+ \left\{\frac{1}{2}\frac{\Delta\tau^n}{\Delta x_i}(\beta^{n+}_{i-\frac{1}{2}} + \sigma^{n+}_{i-\frac{1}{2}})\right\}(u^n_{i-1} - v^n_{i-1}) \\
&+ \left(\frac{1}{2}\frac{\Delta\tau^n}{\Delta x_i}(-\beta^{n-}_{i+\frac{1}{2}} + \sigma^{n-}_{i+\frac{1}{2}})\right)((u^n_{i+1} - u^n_i) + (v^n_{i+1} - v^n_i)) \\
&+ \left(\frac{1}{2}\frac{\Delta\tau^n}{\Delta x_i}(-\beta^{n+}_{i-\frac{1}{2}} + \sigma^{n+}_{i-\frac{1}{2}})\right)((u^n_i - u^n_{i-1}) + (v^n_i - v^n_{i-1})).
\end{aligned}
$$

The proof of this lemma is obtained by using (2.6), Lemma (3.1) and some straightforward algebraic manipulations.

LEMMA 3.11. *We have, for $n = 0, ..., n_T$,*

$$
\|\beta_h^n - \sigma_h^n\|_{L^\infty(0,1)} \le \|u_h^n - v_h^n\|_{L^1(0,1)} + |\phi^n_{1,\Delta\tau} - \psi^n_{1,\Delta\tau}|.
$$

Proof. From equations (2.7) we have

$$
\begin{aligned}
&((\beta_h^n - \sigma_h^n) - (\phi^n_{1,\Delta\tau} - \psi^n_{1,\Delta\tau}), 1) = 0, \\
&(\beta_h^n - \sigma_h^n)_x = u_h^n - v_h^n,
\end{aligned}
$$

and hence the result follows. □

LEMMA 3.12. *Suppose that the* CFL *condition* (3.7) *is satisfied (for both sets of data). Then*

$$
\begin{aligned}
\|u_h^{n+1} - v_h^{n+1}\|_{L^1(0,1)} \le &\frac{1}{2}\Delta\tau^n\left(\|\beta_h^n\|_{L^\infty(0,1)} + \|\sigma_h^n\|_{L^\infty(0,1)}\right)\left(|u^n_{n_x+1} - v^n_{n_x+1}| + |u_0^n - v_0^n|\right) \\
&+ \frac{1}{2}\Delta\tau^n\left(\|u_h^n\|_{BV(0,1)} + \|v_h^n\|_{BV(0,1)}\right)|\phi^n_{1,\Delta\tau} - \psi^n_{1,\Delta\tau}| \\
&+ \left(1 + \Delta\tau^n\frac{1}{2}\left(\|u_h^n\|_{BV(0,1)} + \|v_h^n\|_{BV(0,1)}\right)\right)\|u_h^n - v_h^n\|_{L^1(0,1)}.
\end{aligned}
$$

Proof. Under the hypothesis of this lemma the terms between braces in Lemma 3.10 are nonnegative. Hence,

$$
\begin{aligned}
|u_i^{n+1} - v_i^{n+1}|\Delta x_i \le &-\frac{1}{2}\Delta\tau^n(\beta^{n-}_{i+\frac{1}{2}} + \sigma^{n-}_{i+\frac{1}{2}})|u^n_{i+1} - v^n_{i+1}| \\
&- \frac{1}{2}\Delta\tau^n(\beta^{n+}_{i+\frac{1}{2}} + \sigma^{n+}_{i+\frac{1}{2}})|u_i^n - v_i^n| \\
&+ (1 - \frac{1}{2}\Delta\tau^n(u_i^n + v_i^n))|u_i^n - v_i^n|\Delta x_i \\
&+ \frac{1}{2}\Delta\tau^n(\beta^{n-}_{i-\frac{1}{2}} + \sigma^{n-}_{i-\frac{1}{2}})|u_i^n - v_i^n| \\
&- \frac{1}{2}\Delta\tau^n(\beta^{n+}_{i-\frac{1}{2}} + \sigma^{n+}_{i+\frac{1}{2}})|u^n_{i-1} - v^n_{i-1}| \\
&+ \frac{1}{2}\Delta\tau^n|\beta^{n-}_{i+\frac{1}{2}} - \sigma^{n-}_{i+\frac{1}{2}}|(|u^n_{i+1} - u^n_i| + |v^n_{i+1} - v^n_i|) \\
&+ \frac{1}{2}\Delta\tau^n|\beta^{n+}_{i-\frac{1}{2}} - \sigma^{n+}_{i-\frac{1}{2}}|(|u_i^n - u^n_{i-1}| + |v_i^n - v^n_{i-1}|).
\end{aligned}
$$

Summing over i and rearranging terms, we get

$$
\begin{aligned}
\|u_h^{n+1} - v_h^{n+1}\|_{L(0,1)} \le &-\frac{1}{2}\Delta\tau^n(\beta^{n-}_{n_x+\frac{1}{2}} + \sigma^{n-}_{n_x+\frac{1}{2}})|u^n_{n_x+1} - v^n_{n_x+1}| \\
&+ \sum_{i=1}^{n_x}(1 - \frac{1}{2}\Delta\tau^n(u_i^n + v_i^n))|u_i^n - v_i^n|\Delta x_i \\
&+ \frac{1}{2}\Delta\tau^n(\beta^{n+}_{\frac{1}{2}} + \sigma^{n+}_{\frac{1}{2}})|u_0^n - v_0^n| \\
&+ \frac{1}{2}\Delta\tau^n\sum_{i=0}^{n_x}(|\beta^{n-}_{i+\frac{1}{2}} - \sigma^{n-}_{i+\frac{1}{2}}| + |\beta^{n+}_{i+\frac{1}{2}} - \sigma^{n+}_{i+\frac{1}{2}}|)(|u^n_{i+1} - u^n_i| + |v^n_{i+1} - v^n_i|),
\end{aligned}
$$

and since $|a^- - b^-| + |a^+ - b^+| = |a - b|$, we obtain

$$
\begin{aligned}
\|u_h^{n+1} - v_h^{n+1}\|_{L(0,1)} \le &-\frac{1}{2}\Delta\tau^n(\beta^{n-}_{n_x+\frac{1}{2}} + \sigma^{n-}_{n_x+\frac{1}{2}})|u^n_{n_x+1} - v^n_{n_x+1}| \\
&+ \sum_{i=1}^{n_x}(1 - \frac{1}{2}\Delta\tau^n(u_i^n + v_i^n))|u_i^n - v_i^n|\Delta x_i \\
&+ \frac{1}{2}\Delta\tau^n(\beta^{n+}_{\frac{1}{2}} + \sigma^{n+}_{\frac{1}{2}})|u_0^n - v_0^n| \\
&+ \frac{1}{2}\Delta\tau^n\sum_{i=0}^{n_x}|\beta^n_{i+\frac{1}{2}} - \sigma^n_{i+\frac{1}{2}}|(|u^n_{i+1} - u^n_i| + |v^n_{i+1} - v^n_i|) \\
\le &\frac{1}{2}\Delta\tau^n\left(\|\beta_h^n\|_{L^\infty(0,1)} + \|\sigma_h^n\|_{L^\infty(0,1)}\right)|u^n_{n_x+1} - v^n_{n_x+1}| \\
&+ \sum_{i=1}^{n_x}(1 - \frac{1}{2}\Delta\tau^n(u_i^n + v_i^n))|u_i^n - v_i^n|\Delta x_i \\
&+ \frac{1}{2}\Delta\tau^n\left(\|\beta_h^n\|_{L^\infty(0,1)} + \|\sigma_h^n\|_{L^\infty(0,1)}\right)|u_0^n - v_0^n| \\
&+ \frac{1}{2}\Delta\tau^n\|\beta_h^n - \sigma_h^n\|_{L^\infty(0,1)}\left(\|u_h^n\|_{BV(0,1)} + \|v_h^n\|_{BV(0,1)}\right).
\end{aligned}
$$

Using Lemma 3.11, we obtain

$$\begin{aligned}\|u_h^{n+1} - v_h^{n+1}\|_{L^1(0,1)} \le & \frac{1}{2}\Delta\tau^n \left(\|\beta_h^n\|_{L^\infty(0,1)} + \|\sigma_h^n\|_{L^\infty(0,1)}\right) |u_{n_x+1}^n - v_{n_x+1}^n| \\ & + \frac{1}{2}\Delta\tau^n \left(\|\beta_h^n\|_{L^\infty(0,1)} + \|\sigma_h^n\|_{L^\infty(0,1)}\right) |u_0^n - v_0^n| \\ & + \frac{1}{2}\Delta\tau^n \left(\|u_h^n\|_{\mathrm{BV}(0,1)} + \|v_h^n\|_{\mathrm{BV}(0,1)}\right) |\phi_{1,\Delta\tau}^n - \psi_{1,\Delta\tau}^n| \\ & + \sum_{i=1}^{n_x} (1 - \frac{1}{2}\Delta\tau^n(u_i^n + v_i^n))|u_i^n - v_i^n|\Delta x_i \\ & + \frac{1}{2}\Delta\tau^n \left(\|u_h^n\|_{\mathrm{BV}(0,1)} + \|v_h^n\|_{\mathrm{BV}(0,1)}\right) \|u_h^n - v_h^n\|_{L^1(0,1)}.\end{aligned}$$

The result follows from the fact that $u_i^n, v_i^n \ge 0$, see Proposition 3.5. □

The following result is a direct consequence of the preceding lemma.

PROPOSITION 3.13 (Continuity with respect to the data). *Suppose that the* CFL *condition* (3.7) *is satisfied (for both sets of data). Then, for* $n = 0, \ldots, n_T$, *we have*

$$\begin{aligned}\|u_h^n - v_h^n\|_{L^1(0,1)} \le & e^{\nu\tau^n}\|u_i - v_i\|_{L^1(0,1)} \\ & + \frac{1}{2}\left(\|\beta_h\|_{L^\infty(L^\infty)} + \|\sigma_h\|_{L^\infty(L^\infty)}\right) \cdot e^{\nu\tau^n} \\ & \cdot \left(\|u_0 - v_0\|_{L^1(0,\tau^n)} + \|u_1 - v_1\|_{L^1(0,\tau^n)}\right) \\ & + 2\nu e^{\nu\tau^n}\|\phi_1 - \psi_1\|_{L^1(0,\tau^n)},\end{aligned}$$

where

$$\nu = \frac{1}{2}\left(\|u_h\|_{L^\infty(0,T;\mathrm{BV}(0,1))} + \|v_h\|_{L^\infty(0,T;\mathrm{BV}(0,1))}\right).$$

The following result will be needed in §3.d and §4.1.

PROPOSITION 3.14 (Equicontinuity in time). *Suppose that the* CFL *condition* (3.7) *is satisfied. Then, for* $n = 0, \ldots, n_T - 1$ *we have*

$$\|u_h^{n+1} - u_h^n\|_{L^1(0,1)} \le C_6\Delta\tau^n,$$

where

$$C_6 = C_1\left(\|\phi_1\|_{L^\infty(0,T)} + \frac{1}{2}u^*\right) + \max\{\frac{1}{4}(u^* - 1)u^*\}.$$

The constant C_1 is given in Proposition 3.9.

Proof. From (2.6) and Lemma 3.1, we easily obtain that

$$(u_i^{n+1} - u_i^n)\Delta x_i = -\left\{\beta_{i+1/2}^{n-}(u_{i+1}^n - u_i^n) - (1 - u_i^n)u_i^n\Delta x_i - \beta_{i-1/2}^{n+}(u_{i-1}^n - u_i^n)\right\}\Delta\tau^n.$$

Taking absolute values, summing over i and making some simple algebraic manipulations, we obtain

$$\begin{aligned}\|u_h^{n+1} - u_h^n\|_{L^1(0,1)} &\le \left\{\sum_{i=0}^{n_x} |\beta_{i+1/2}^n|\,|u_{i+1}^n - u_i^n| + |1 - u_i^n|\,|u_i^n\,\Delta x_i|\right\}\Delta\tau^n \\ &\le \left\{C_1\|\beta_h\|_{L^\infty(0,T;L^\infty(0,1))} + \max\{\frac{1}{4},(u^* - 1)u^*\}\right\}\Delta\tau^n \\ &\le \left\{C_1(\|\phi_1\|_{L^\infty(0,T)} + u^*/2) + \max\{\frac{1}{4},(u^* - 1)u^*\}\right\}\Delta\tau^n,\end{aligned}$$

by Propositions 3.4, 3.5 and (2.5). This completes the proof. □

3d. Proofs of Theorems 2.1 and 2.2. First, consider Theorem 2.1. Since the CFL condition (2.9) implies condition (3.4), (2.10a) holds by Proposition 3.5. The inequality (2.10b) follows from (2.10a) and Proposition 3.4. The inequality (2.10c) follows from (2.10a) and equation (2.7a). To obtain (2.10d) and (2.10e), we first note that, by (2.7b),

$$\begin{aligned}\|\phi_h^n\|_{\mathrm{BV}[0,1]} &\le \|\beta_h^n\|_{L^1(0,1)}, \\ \|\phi_h^n\|_{L^\infty(0,1)} &\le \|\beta_h^n\|_{L^1(0,1)}.\end{aligned}$$

Thus, (2.10d) and (2.10e) follow from (2.5) and (2.10b). Finally, since the CFL condition (2.9) implies the condition (3.7), the uniform bound (2.10f) follows from Proposition 3.9. This completes the proof of Theorem 2.1.

Theorem 2.2 follows from Proposition 3.13 and Theorem 2.1.

3e. Bounds for $\nu_{x,0}^+$ and $\nu_{x,1}^-$. In this section we show that under the CFL condition (3.7), the quantities $\nu_{x,0}^+(\epsilon, u_h; \beta_h)$ and $\nu_{x,1}^-(\epsilon, u_h; \beta_h)$ can be estimated as stated in Theorem 2.5. We proceed in several steps.

First step: The case $\beta_h(x = 0) \ge \epsilon + \Delta x$. We begin by considering the case in which

$$\beta_{1/2}^n \ge \epsilon + \Delta x, \text{ for } n = 0, \cdots, n_T. \tag{3.8}$$

In what follows we let i_0 be such that $x_{i_0-1/2} \le \epsilon < x_{i_0+1/2}$. Condition (3.8) ensures that the flow of the electrons does not change sign in a neighborhood of $\{x = 0\} \times [0, T]$ uniformly with respect to Δx.

LEMMA 3.15. *Suppose that condition* (3.8) *is satisfied. Then for* $n = 0, \cdots, n_T$ *and* $i = 0, \cdots, i_0$,

$$\beta_{i+1/2}^n \ge 0.$$

Proof. For $i > 0$ we have

$$\begin{aligned}
\beta^n_{i+1/2} &= \beta^n_{1/2} - \sum_{j=1}^{i} (1 - u^n_j)\,\Delta x_j, && \text{by Lemma 3.1,}\\
&= \beta^n_{1/2} - x_{i+1/2} + \sum_{j=1}^{i} u^n_j\,\Delta x_j\\
&\geq \beta^n_{1/2} - x_{i+1/2}, && \text{by Proposition 3.5,}\\
&\geq \epsilon - x_{i-1/2}, && \text{by condition (3.8),}\\
&\geq 0, && \text{for } i = 0, \cdots, i_0,
\end{aligned}$$

by the definition of i_0. This completes the proof. □

LEMMA 3.16. *Suppose that condition* (3.8) *is satisfied. Then*

$$\nu^+_{x,0}(\epsilon, u_h; \beta_h) \leq \left\{ (u^*)^2\, T + \sup_{1\leq j\leq i_0} \sum_{n=0}^{n_T-1} |\, u^{n+1}_j - u^n_j \,| \right\} (\epsilon + \Delta x).$$

Proof. By definition of $\nu^+_{x,0}(\epsilon, u_h; \beta_h)$ and of i_0, we have

$$\begin{aligned}
\nu^+_{x,0}(\epsilon, u_h; \beta_h) &= \sup_{0\leq\Delta\leq\epsilon} \int_0^T |\, u_h(\tau, \Delta) - u_{0h}(\tau) \,|\, \beta^+_h(\tau, 0)\, d\tau\\
&= \sup_{0\leq i\leq i_0} \sum_{n=0}^{n_T-1} |\, u^n_i - u^n_0 \,|\, \beta^n_{1/2}{}^+\, \Delta\tau^n\\
&\leq \sup_{0\leq i\leq i_0} \sum_{n=0}^{n_T-1} \left\{ |\beta^n_{1/2} - \beta^n_{i+1/2}|\, u^n_i + |\beta^n_{i+1/2}\, u^n_i - \beta^n_{1/2}\, u^n_0 \,| \right\} \Delta\tau^n\\
&\leq \sup_{0\leq i\leq i_0} \sum_{n=0}^{n_T-1} \left\{ |\beta^n_{1/2} - \beta^n_{i+1/2}|\, u^n_i + \sum_{j=1}^{i} |\, \beta^n_{j+1/2}\, u^n_j - \beta^n_{j-1/2}\, u^n_{j-1} \,| \right\} \Delta\tau^n\\
&\leq (u^*)^2\, T\,(\epsilon + \Delta x) + \sum_{n=0}^{n_T-1} \sum_{j=1}^{i_0} |\, \beta^n_{j+1/2}\, u^n_j - \beta^n_{j-1/2}\, u^n_{j-1} \,|\, \Delta\tau^n,
\end{aligned}$$

by Lemma 3.1 and Proposition 3.5. By Lemma 3.15, we can rewrite the scheme (2.6), for $j \leq i_0$, as follows

$$(\beta^n_{j+1/2}\, u^n_j - \beta^n_{j-1/2}\, u^n_{j-1})\,\Delta\tau^n = -(u^{n+1}_j - u^n_j)\,\Delta x_j,$$

and hence

$$\begin{aligned}
\nu^+_{x,0}(\epsilon, u_h; \beta_h) &\leq (u^*)^2\, T\,(\epsilon + \Delta x) + \sum_{n=0}^{n_T-1} \sum_{j=1}^{i_0} |u^{n+1}_j - u^n_j|\, \Delta x_i\\
&\leq (u^*)^2\, T\,(\epsilon + \Delta x) + \sup_{1\leq j\leq i_0} \left\{ \sum_{n=0}^{n_T-1} |u^{n+1}_j - u^n_j| \right\} x_{i_0+1/2}\\
&\leq (u^*)^2\, T\,(\epsilon + \Delta x) + \sup_{1\leq j\leq i_0} \left\{ \sum_{n=0}^{n_T-1} |u^{n+1}_j - u^n_j| \right\} (\epsilon + \Delta x),
\end{aligned}$$

by definition of i_0. This completes the proof. □

We now turn to estimate the variation in time of $u_h(x_j)$, $\sum_{n=0}^{n_T-1} |u^{n+1}_j - u^n_j|$. We first obtain a basic estimate from which the wanted estimate will follow.

LEMMA 3.17 (Basic Estimate). *Suppose that the condition* (3.8) *and the* CFL *condition* (3.7) *are satisfied. Then, for* $i = 1, \ldots, i_0$ (*and* $n_T \geq 2$),

$$\begin{aligned}
\sum_{n=0}^{n_T-1} |u^{n+1}_i - u^n_i - (1 - u^n_i)u^n_i \Delta\tau^n| &\leq |u^0_i - u^0_{i-1}|\\
&\quad + \sum_{n=0}^{n_T-2} |u^{n+1}_{i-1} - u^n_{i-1} - (1 - u^n_{i-1})u^n_{i-1}\Delta\tau^n|\\
&\quad - \sum_{n=0}^{n_T-1} (1 - u^n_i - u^n_{i-1})|u^n_i - u^n_{i-1}|\Delta\tau^n.
\end{aligned}$$

Proof. Using Lemma 3.15, we can rewrite (2.6) as follows:

$$u^{n+1}_i = (1 - \frac{\Delta\tau^n}{\Delta x_i}\beta^n_{i+\frac{1}{2}})\, u^n_i + (\frac{\Delta\tau^n}{\Delta x_i}\beta^n_{i-\frac{1}{2}})\, u^n_{i-1}, \qquad i = 1, \ldots, i_0.$$

If we use Lemma 3.1 to write $\beta^n_{i-\frac{1}{2}}$ in terms of $\beta^n_{i+\frac{1}{2}}$, and set

$$a^n_i = \frac{\Delta\tau^n}{\Delta x_i}\beta^n_{i+\frac{1}{2}} + \Delta\tau^n, \tag{3.9a}$$

$$\Gamma^n_i = 1 + (1 - u^n_{i-1})\Delta\tau^n, \tag{3.9b}$$

we can rewrite the equation for the numerical scheme as follows:

$$u^{n+1}_i = (\Gamma^n_i - a^n_i)\, u^n_i + a^n_i\, u^n_{i-1}, \qquad i = 1, \cdots, i_0.$$

If we set

$$c^{n,j}_i = \begin{cases} 1 & \text{for } j = n+1,\\ \Pi^n_{\ell=j}(\Gamma^\ell_i - a^\ell_i) & \text{for } j \leq n, \end{cases} \tag{3.9c}$$

a simple computation leads us to the following expression:

$$u_i^{n+1} = c_i^{n,0} u_i^0 + \sum_{j=0}^{n} c_i^{n,j+1} a_i^j u_{i-1}^j.$$

Since by definition

$$c_i^{n,j+1} a_i^j = \Gamma_i^j c_i^{n,j+1} - c_i^{n,j}, \qquad j = 0, \ldots, n,$$

we have

$$\begin{aligned} u_i^{n+1} &= c_i^{n,0}\, u_i^0 + \sum_{j=0}^{n} (\Gamma_i^j\, c_i^{n,j+1} - c_i^{n,j})\, u_{i-1}^j \\ &= c_i^{n,0}(u_i^0 - u_{i-1}^0) + \sum_{j=0}^{n-1} c_i^{n,j+1}(\Gamma_i^j u_{i-1}^j - u_{i-1}^{j+1}) + \Gamma_i^n u_{i-1}^n, \end{aligned}$$

and hence, for $n \geq 1$,

$$\begin{aligned} u_i^{n+1} - u_i^n &= (c_i^{n,0} - c_i^{n-1,0})(u_i^0 - u_{i-1}^0) \\ &\quad - \sum_{j=0}^{n-2} (c_i^{n,j+1} - c_i^{n-1,j+1})(u_{i-1}^{j+1} - \Gamma_i^j u_{i-1}^j) \\ &\quad - c_i^{n,n}(u_{i-1}^n - \Gamma_i^{n-1} u_{i-1}^{n-1}) + \Gamma_i^n\, u_{i-1}^n - \Gamma_i^{n-1} u_{i-1}^{n-1} \\ &= (c_i^{n,0} - c_i^{n-1,0})(u_i^0 - u_{i-1}^0) \\ &\quad - \sum_{j=0}^{n-1} (c_i^{n,j+1} - c_i^{n-1,j+1})(u_{i-1}^{j+1} - \Gamma_i^j u_{i-1}^j) \\ &\quad - (u_{i-1}^n - \Gamma_i^{n-1} u_{i-1}^{n-1}) + \Gamma_i^n u_{i-1}^n - \Gamma_i^{n-1}\, u_{i-1}^{n-1}. \end{aligned}$$

The form of this expression suggests to consider the expression $(u_i^{n+1} - \Gamma_{i+1}^n u_i^n)$ instead of the expression $(u_i^{n+1} - u_i^n)$. Thus,

$$\begin{aligned} (u_i^{n+1} - \Gamma_{i+1}^n u_i^n) &= (c_i^{n,0} - c_i^{n-1,0})(u_i^0 - u_{i-1}^0) \\ &\quad - \sum_{j=0}^{n-1} (c_i^{n,j+1} - c_i^{n-1,j+1})(u_{i-1}^{j+1} - \Gamma_i^j\, u_{i-1}^j) \\ &\quad + (1 - \Gamma_{i+1}^n)\, u_i^n - (1 - \Gamma_i^n)\, u_{i-1}^n \\ &= (c_i^{n,0} - c_i^{n-1,0})(u_i^0 - u_{i-1}^0) \\ &\quad - \sum_{j=0}^{n-1} (c_i^{n,j+1} - c_i^{n-1,j+1})(u_{i-1}^{j+1} - \Gamma_i^j\, u_{i-1}^j) \\ &\quad - \Delta\tau^n (1 - u_i^n - u_{i-1}^n)(u_i^n - u_{i-1}^n). \end{aligned}$$

Notice that this formula holds for $n = 0$, if we define $c_i^{-1,0} = 1$. Notice also that by the CFL condition (3.7) and the equations (3.9) we have, for $j = 0, \ldots, n$,

$$c_i^{n,j} \in [0,1], \tag{3.10a}$$

$$\begin{aligned} c_i^{n,j} - c_i^{n-1,j} &= (\Gamma_i^n - a_i^n - 1)\, c_i^{n-1,j} \\ &= -\left(\frac{\Delta\tau^n}{\Delta x_i} \beta_{i+\frac{1}{2}}^n + \Delta\tau^n u_{i-1}^n\right) c_i^{n-1,j} \\ &\leq 0. \end{aligned} \tag{3.10b}$$

We can thus take absolute values and obtain, by (3.10b),

$$\begin{aligned} |u_i^{n+1} - \Gamma_{i+1}^n u_i^n + \Delta\tau^n (1 - u_i^n - u_{i-1}^n)(u_i^n - u_{i-1}^n)| &\leq (c_i^{n-1,0} - c_i^{n,0}) |u_i^0 - u_{i-1}^0| \\ &\quad + \sum_{j=0}^{n-1} (c_i^{n-1,j+1} - c_i^{n,j+1}) |u_{i-1}^{j+1} - \Gamma_i^j u_{i-1}^j|. \end{aligned}$$

Let us consider the term in the left-hand side. Since

$$\begin{aligned} u_i^{n+1} - \Gamma_{i+1}^n u_i^n &= -\left\{\frac{\Delta\tau^n}{\Delta x_i} \beta_{i+\frac{1}{2}}^n + \Delta\tau^n (1 - u_i^n)\right\}(u_i^n - u_{i-1}^n) \\ &= -\left\{\frac{\Delta\tau^n}{\Delta x_i} \beta_{i-\frac{1}{2}}^n\right\}(u_i^n - u_{i-1}^n), \text{ by Lemma 3.1,} \end{aligned}$$

and

$$\begin{aligned} (u_i^{n+1} - \Gamma_{i+1}^n u_i^n) + \{\Delta\tau^n (1 - u_i^n - u_{i-1}^n)(u_i^n - u_{i-1}^n)\} \\ = -\left\{\frac{\Delta\tau^n}{\Delta x_i} \beta_{i+\frac{1}{2}}^n + \Delta\tau^n u_{i-1}^n\right\}(u_i^n - u_{i-1}^n), \end{aligned}$$

and since, by the CFL condition (3.7), the terms between braces are nonnegative, it is clear that

$$\begin{aligned} &|(u_i^{n+1} - \Gamma_{i+1}^n u_i^n) + \Delta\tau^n (1 - u_i^n - u_{i-1}^n)(u_i^n - u_{i-1}^n)| \\ &\quad = |u_i^{n+1} - \Gamma_{i+1}^n u_i^n| + \Delta\tau^n (1 - u_i^n - u_{i-1}^n)|u_i^n - u_{i-1}^n|, \end{aligned}$$

and thus

$$\begin{aligned} |u_i^{n+1} - \Gamma_{i+1}^n u_i^n| &\leq (c_i^{n-1,0} - c_i^{n,0})|u_i^0 - u_{i-1}^0| \\ &\quad + \sum_{j=0}^{n-1} (c_i^{n-1,j+1} - c_i^{n,j+1})|u_{i-1}^{j+1} - \Gamma_i^j u_{i-1}^j| \\ &\quad - \Delta\tau^n (1 - u_i^n - u_{i-1}^n)\, |u_i^n - u_{i-1}^n|. \end{aligned}$$

Summing over n, we get

$$\begin{aligned} \sum_{n=0}^{n_T-1} |u_i^{n+1} - \Gamma_{i+1}^n u_i^n| &\leq (c_i^{-1,0} - c_i^{n_T-1,0})|u_i^0 - u_{i-1}^0| \\ &\quad + \sum_{n=1}^{n_T-1} \sum_{j=1}^{n} (c_i^{n-1,j} - c_i^{n,j})|u_{i-1}^j - \Gamma_i^{j-1}\, u_{i-1}^{j-1}| \\ &\quad - \sum_{n=0}^{n_T-1} \Delta\tau^n (1 - u_i^n - u_{i-1}^n)|u_i^n - u_{i-1}^n|. \end{aligned}$$

Rearranging, we get

$$\begin{aligned}\sum_{n=0}^{n_T-1} |u_i^{n+1} - \Gamma_{i+1}^n u_i^n| &\le \left(c_i^{-1,0} - c_i^{n_T-1,0}\right) |u_i^0 - u_{i-1}^0| \\ &+ \sum_{j=1}^{n_T-1} \Big[\sum_{n=j}^{n_T-1} (c_i^{n-1,j} - c_i^{n,j}) \Big] |u_{i-1}^j - \Gamma_i^{j-1} u_{i-1}^{j-1}| \\ &- \sum_{n=0}^{n_T-1} \Delta\tau^n (1 - u_i^n - u_{i-1}^n)|u_i^n - u_{i-1}^n| \\ &\le \left(c_i^{-1,0} - c_i^{n_T-1,0}\right) |u_i^0 - u_{i-1}^0| \\ &+ \sum_{j=1}^{n_T-1} \left(c_i^{j-1,j} - c_i^{n_T-1,j}\right) |u_{i-1}^j - \Gamma_i^{j-1} u_{i-1}^{j-1}| \\ &- \sum_{n=0}^{n_T-1} \Delta\tau^n (1 - u_i^n - u_{i-1}^n)|u_i^n - u_{i-1}^n|.\end{aligned}$$

Finally, using (3.10a), we get

$$\begin{aligned}\sum_{n=0}^{n_T-1} |u_i^{n+1} - \Gamma_{i+1}^n u_i^n| &\le |u_i^0 - u_{i-1}^0| + \sum_{n=0}^{n_T-2} |u_{i-1}^{n+1} - \Gamma_i^n u_{i-1}^n| \\ &- \sum_{n=0}^{n_T-1} \Delta\tau^n (1 - u_i^n - u_{i-1}^n)|u_i^n - u_{i-1}^n|.\end{aligned}$$

This completes the proof. □

When the solution has the monotonicity property of Corollary 3.7, a sharp estimate on the total variation in time can be obtained.

COROLLARY 3.18. *Let u_i, u_0 and u_1 be such that*

$$\begin{aligned}&u_i(x) \in [0,1], x \in (0,1), \text{ is a nondecreasing function,}\\ &u_0(\tau) = 0, \tau \in (0,T),\\ &u_1(\tau) = 1, \tau \in (0,T),\end{aligned}$$

and suppose that the CFL *condition* (3.7) *is satisfied. Then, for $\Delta x \le \inf_{\tau\in[0,T]} \phi_1(\tau)/2$, we have*

$$\sum_{n=0}^{n_T-1} |u_i^{n+1} - u_i^n| \le 1, \quad \text{for } x_i \le \inf_{\tau\in[0,T]} \phi_1(\tau)/4.$$

Proof. It is easy to verify that under the hypotheses on the data, condition (3.8) is satisfied for ϵ and Δx such that $\inf_{\tau\in[0,T]} \phi_1(\tau) \ge \Delta x + \epsilon$. Set $\epsilon = \inf_{\tau\in[0,T]} \phi_1(\tau)/2$. Then the condition (3.8) is satisfied for $\Delta x \le \inf_{\tau\in[0,T]} \phi_1(\tau)/2$ and, by Lemma 3.17, we have, for $i = 1, \cdots, i_0$,

$$\begin{aligned}\sum_{n=0}^{n_T-1} |u_i^{n+1} - u_i^n| &\le \sum_{n=0}^{n_T-1} |u_i^{n+1} - u_i^n - (1-u_i^n)u_i^n \Delta\tau^n| + \sum_{n=0}^{n_T-1} |(1-u_i^n)u_i^n|\Delta\tau^n \\ &\le \sum_{j=1}^{i} |u_j^0 - u_{j-1}^0| + \sum_{n=0}^{n_T-1} |u_0^{n+1} - u_0^n - (1-u_0^n)u_0^n \Delta\tau^n| \\ &- \sum_{n=0}^{n_T-1} \sum_{j=1}^{i} (1 - u_j^n - u_{j-1}^n)|u_j^n - u_{j-1}^n|\Delta\tau^n \\ &+ \sum_{n=0}^{n_T-1} |(1-u_i^n)u_i^n|\Delta\tau^n.\end{aligned}$$

By (2.5) and the hypothesis on the initial data, $\sum_{j=1}^{i} |u_j^0 - u_{j-1}^0| \le 1$. Since $u_0^n = 0$, the second term of the right-hand side is zero; also, by Corollary 3.7, u_k^n is nondecreasing, and we have

$$\begin{aligned}\sum_{n=0}^{n_T-1} \sum_{j=1}^{i} (1 - u_j^n - u_{j-1}^n)|u_j^n - u_{j-1}^n|\Delta\tau^n &= \sum_{n=0}^{n_T-1} \sum_{j=1}^{i} (1 - u_j^n - u_{j-1}^n)(u_j^n - u_{j-1}^n)\Delta\tau^n \\ &= \sum_{n=0}^{n_T-1} \sum_{j=1}^{i} (u_i^n - (u_i^n)^2)\Delta\tau^n \\ &= \sum_{n=0}^{n_T-1} (1-u_i^n)u_i^n \Delta\tau^n \\ &= \sum_{n=0}^{n_T-1} |(1-u_i^n)u_i^n|\Delta\tau^n.\end{aligned}$$

This proves the inequality for $x_i \le x_{i_0}$. Finally, if $x_i \le \inf_{\tau\in[0,T]} \phi_1(\tau)/4 = \epsilon/2$, then $x_{i+1/2} \le \epsilon/2 + \Delta x/2 = \epsilon \le x_{i_0+1/2}$, and so $x_i \le x_{i_0}$. This completes the proof. □

In the general case the estimate of the total variation in time takes a more complicated form.

LEMMA 3.19 (Estimate of the total variation in time). *Suppose that the condition* (3.8) *and the* CFL *condition* (3.7) *are satisfied. Then, for $i = 0, ..., i_0$,*

$$\sum_{n=0}^{n_T-1} |u_i^{n+1} - u_i^n| \le C_1 + C_7(0,T),$$

where

$$\begin{aligned}C_7(\tau_1,\tau_2) &= \|u_0\|_{BV(\tau_1,\tau_2)} + \|1-u_0\|_{L^\infty(\tau_1,\tau_2)}\|u_0\|_{L^1(\tau_1,\tau_2)} \\ &+ (2u^*-1)(\tau_2-\tau_1)C_1 + (\tau_2-\tau_1)\max\{1/4, (u^*-1)u^*\}.\end{aligned}$$

Proof. From Lemma 3.17 and Proposition 3.5 we get

$$\begin{aligned}\sum_{n=0}^{n_T-1} |u_i^{n+1} - u_i^n| \le \|u_h(\tau=0)\|_{\mathrm{BV}(0,1)} &\\ + \sum_{n=0}^{n_T-1} |u_0^{n+1} - u_0^n - (1-u_0^n)\, u_0^n\, \Delta\tau^n| &\\ + (2u^* - 1)\, T\, \|u_h\|_{L^\infty(0,T;\mathrm{BV}(0,1))} &\\ + T \max\{\tfrac{1}{4}, (u^*-1)\, u^*\}.&\end{aligned}$$

The result follows from (2.5), Proposition 3.9, and (2.5c). □

From Lemmas 3.16 and 3.19 we immediately obtain the following result.

COROLLARY 3.20. *Suppose that condition* (3.8) *and the* CFL *condition* (3.7) *are satisfied. Then*

$$\nu^+_{x,0}(\epsilon, u_h; \beta_h) \le ((u^*)^2\, T + C_1 + C_7(0,T))\,(\epsilon + \Delta x).$$

Second step: The general case. To treat the general case, we divide the interval $[0,T)$ into intervals on which we have information about the size of the flux $\beta_h(x=0)$. We need to introduce some notation. For a given nonnegative number δ and a given partition of the computational domain we introduce the following sets:

$$\begin{aligned}\mathcal{J}_h(\delta) &= \{n : \beta^n_{1/2} \le \delta\},\\ \mathcal{L}_h(\delta) &= \cup_{m\in\mathcal{J}_h(\delta)} [\tau^m, \tau^{m+1})\\ &= \cup_{j=1}^{N_{h,\delta,T}-1} [\tau^{n_{2j-1}}_{h,\delta}, \tau^{n_{2j}}_{h,\delta}),\\ [0,T)\setminus\mathcal{L}_h(\delta) &= \cup_{j=1}^{N_{h,\delta,T}} [\tau^{n_{2j-2}}_{h,\delta}, \tau^{n_{2j-1}}_{h,\delta}).\end{aligned}$$

Notice that $\beta(\tau, x=0) \le \delta$ if and only if $\tau \in \mathcal{L}_h(\delta)$. The complement of $\mathcal{L}_h(\delta)$ is made of $N_{h,\delta,T}$ disjoint intervals on which $\beta(x=0) > \delta$.

LEMMA 3.21. *Suppose that the* CFL *condition* (3.7) *is satisfied. Then, for every* $\delta \ge \epsilon + \Delta x$,

$$\nu^+_{x,0}(\epsilon, u_h; \beta_h) \le C_8\, \delta + N_{h,\delta,T}\, C_1\, (\epsilon + \Delta x),$$

where $C_8 = u^*\, T + (u^*)^2\, T + C_7(0,T)$.

Proof. By definition of $\nu^+_{x,0}(\epsilon, u_h; \beta_h)$ we have

$$\nu^+_{x,0}(\epsilon, u_h; \beta_h) = \sup_{0\le\Delta\le\epsilon} \{\Theta_1(\Delta) + \Theta_2(\Delta)\},$$

where

$$\begin{aligned}\Theta_1(\Delta) &= \int_{\mathcal{L}_h(\delta)} |\, u_h(\tau,\Delta) - u_{0,\Delta x}(\tau)\,|\, \beta^+_h(\tau,0)\, d\tau,\\ \Theta_2(\Delta) &= \int_{(0,T)\setminus\mathcal{L}_h(\delta)} |\, u_h(\tau,\Delta) - u_{0,\Delta x}(\tau)\,|\, \beta^+_h(\tau,0)\, d\tau.\end{aligned}$$

Since, by construction, $\beta_h(\tau, x=0) \le \delta$, for $\tau^n \in \mathcal{L}_h(\delta)$, we have that

$$\Theta_1(\Delta) \le u^*\, T\, \delta.$$

On the other hand, for $\tau \in (0,T)\setminus\mathcal{L}_h(\delta)$, $\beta_h(\tau, x=0) > \delta \ge (\epsilon + \Delta x)$, and we have, by Corollary 3.20,

$$\begin{aligned}\Theta_2(\Delta) &= \int_{(0,T)\setminus\mathcal{L}_h(\delta)} |\, u_h(\tau,\Delta) - u_{0,\Delta x}(\tau)\,|\, \beta^+_h(\tau,0)\, d\tau\\ &= \sum_{j=1}^{N_{h,\delta,T}} \int_{\tau^{n_{2j-2}}_{h,\delta}}^{\tau^{n_{2j-1}}_{h,\delta}} |\, u_h(\tau,\Delta) - u_{0,\Delta x}(\tau)\,|\, \beta^+_h(\tau,0)\, d\tau\\ &\le \left\{ \sum_{j=1}^{N_{h,\delta,T}} ((u^*)^2\, |\tau^{n_{2j-1}}_{h,\delta} - \tau^{n_{2j-2}}_{h,\delta}| + C_1 + C_7(\tau^{n_{2j-2}}_{h,\delta}, \tau^{n_{2j-1}}_{h,\delta})) \right\} (\epsilon + \Delta x).\end{aligned}$$

Finally,

$$\Theta_2(\Delta) \le \left\{ (u^*)^2\, T + C_7(0,T) + N_{h,\delta,T}\, C_1 \right\} (\epsilon + \Delta x),$$

by the definition of C_7 in Lemma 3.19. This completes the proof. □

Third step: the case (2.12a).

PROPOSITION 3.22 (Estimate of $\nu^+_{x,0}(\epsilon, u_h; \beta_h)$). *Suppose that the* CFL *condition* (3.7) *is satisfied. If the hypothesis* (2.12a) *is satisfied, then,*

$$\nu^+_{x,0}(\epsilon, u_h; \beta_h) \le C_4\, (\epsilon + \Delta x),$$

where the constant $C_4 = C_8 + C_1 \sup_{h>0, 0<\delta\le\epsilon+\Delta x} N_{h,\delta,T}$ *is independent of* ϵ *and* Δx.

Proof. From Lemma 3.21, it is clear that we only have to prove that the number $N_{h,\delta,T}$ is uniformly bounded. Since the number of intervals on which the boundary data is constant, N, is finite, it is enough to prove the result for the data being constant on the whole interval $(0,T)$.

From (2.7), we have

$$\beta_h^n(0) = \phi_1 + \int_0^1 (u_h^n(s) - 1)(s-1)\,ds$$
$$\geq \phi_1,$$

by the hypothesis (2.12a) and the maximum principle (2.10a) of Theorem 2.1. Thus, $N_{h,\delta,T} \leq 1$ provided $\phi_1 \neq 0$.

To analyze the case $\phi_1 = 0$, we have to use a different argument. For the sake of clarity, let us consider the continuous equations (1.3) and (1.4) instead of the discrete equations (2.6) and (2.7). From (1.3) and (1.4) we easily obtain the following equations:

$$\frac{d\beta(0)}{dt} = \eta^2/2 + \eta\,\beta(0),$$
$$\frac{d\eta}{dt} = -u_1(\eta + \beta(0)),$$

where $\eta = \int_0^1 (u(s) - 1)\,ds$. Notice that to obtain these equations, we have to use the fact that $u_0 = 0$ and that

$$\beta(1) = \int_0^1 (u_h^n(s) - 1)\,s\,ds \leq 0,$$

since $u^* = 1$. A simple analysis of the above dynamical system allows us to conclude that the number of times $\frac{\beta(0)}{dt} = 0$ is at most one. Thus, $\beta(0) = \epsilon > 0$ at most twice for ϵ small enough. The discrete equations (2.6) and (2.7) lead to a similar dynamical system which, for Δx small enough, displays the same property. Thus $N_{h,\delta,T}$ can be uniformly bounded. □

Fourth step: the case (2.12b). Next we prove that when ϕ_1 satisfies the hypothesis (2.12b), the number $N_{h,\delta,T}$ in Lemma 3.22 can be estimated in terms of $(\epsilon + \Delta x)$ only. It is enough to obtain the bound assuming $N = 1$.

We use the following auxiliary result that carries the information of the negative electric field β_h being a uniformly Lipschitz continuous function in time.

We recall that, by definition, the set $\mathcal{L}_h(\delta)$ is the union of intervals $[\tau_{h,\delta}^{n_{2j-1}}, \tau_{h,\delta}^{n_{2j}})$ on which $\beta_{1/2} \leq \delta$. Let us consider a subset $\tilde{\mathcal{L}}_h(\delta)$ of $\mathcal{L}_h(\delta)$ which is the union of those intervals $[\tau_{h,\delta}^{n_{2j-1}}, \tau_{h,\delta}^{n_{2j}})$ on which the minimum of $\beta_{1/2}$ is smaller than, or equal to, $\delta/2$. Let us denote by $\tilde{N}_{h,\delta,T} - 1$ the number of those intervals.

LEMMA 3.23. *We have that*

$$\tilde{N}_{h,\delta,T} \leq 2\,C_9\,T/\delta,$$

where $C_9 = (C_6 + \|(\phi_1)_\tau\|_{L^\infty(0,T)})$.

Proof. By Lemma 3.11 and Proposition 3.14, we have that

$$\|\beta_h(\tau^m) - \beta_h(\tau^n)\|_{L^\infty(0,1)} \leq C_9\,|\tau^m - \tau^n|.$$

Let $\tau^n \in \tilde{\mathcal{L}}_h(\delta)$ be such that $\beta_{1/2}(\tau^n) \leq \delta/2$, and let τ^m be in $[0,T) \setminus \tilde{\mathcal{L}}_h(\delta)$, i.e., such that $\beta_{1/2}(\tau^m) > \delta$. Then

$$|\tau^m - \tau^n| \geq \|\beta_h(\tau^m) - \beta_h(\tau^n)\|_{L^\infty(0,1)}/C_9$$
$$\geq (|\beta_{1/2}(\tau^m)| - |\beta_{1/2}(\tau^n)|)/C_9$$
$$\geq \delta/2\,C_9.$$

Hence, $T \geq \tilde{N}_{h,\delta,T}\,\delta/2\,C_9$. This completes the proof. □

COROLLARY 3.24. *We have that*

$$\nu_{x,0}^+(\epsilon, u_h; \beta_h) \leq \hat{C}_4\,(\epsilon + \Delta x)^{1/2},$$

where $\hat{C}_4 = 2\,\{2\,C_1\,C_8\,C_9\,T\}^{1/2}$, *provided that* $(\epsilon + \Delta x) \leq 2\,C_1\,C_9\,T/C_8$.

Proof. If we proceed as in Lemma 3.21 with $\tilde{\mathcal{L}}_h(\delta)$ instead of $\mathcal{L}_h(\delta)$, we easily obtain that, for $\delta \geq (\epsilon + \Delta x)$,

$$\nu_{x,0}^+(\epsilon, u_h; \beta_h) \leq C_8\,\delta + \tilde{N}_{h,\delta,T}\,C_1\,(\epsilon + \Delta x)$$
$$\leq C_8\,\delta + 2\,C7\,C_1\,T\,(\epsilon + \Delta x)/\delta,$$

by Lemma 3.23. Minimizing over δ, we get

$$\nu_{x,0}^+(\epsilon, u_h; \beta_h) \leq \{8\,C_1\,C_8\,C_9\,T\}^{1/2}\,(\epsilon + \Delta x)^{1/2},$$

provided that $\delta_{\min} = \{2\,C_1\,C_9\,T/C_8\}^{1/2}\,(\epsilon + \Delta x)^{1/2} \geq (\epsilon + \Delta x)$, i.e., provided that $(\epsilon + \Delta x) \leq 2\,C_1\,C_9\,T/C_8$. This completes the proof. □

3f. Proof of Theorem 2.5. The inequality (2.13a) follows directly from Proposition 3.22 and the inequality (2.14a) from Corollary 3.24. The inequalities (2.13b) and (2.14b) can be obtained by completely similar arguments. This ends the proof of Theorem 2.5.

4. PROOFS OF THEOREMS 2.3 AND 2.4

In this section we prove Theorems 2.3 and 2.4. In §4a, we prove that there is a subsequence $\{(u_{h'}, \beta_{h'}, \phi_{h'})\}_{h'>0}$ converging to a limit (u, β, ϕ) satisfying the weak equations

(2.3). In §4b, §4c, and §4d, we obtain results that allow us to prove that (u, β) satisfies the weak equations (2.2). In §4b, we prove that for $\varphi \in \mathcal{C}^1([0,T) \times [0,1])$ and $c \in \mathbf{R}$

$$\lim_{h' \to 0} \Theta(u_{h'}, c; \beta_{h'}; \varphi) = \Theta(u, c; \beta; \varphi),$$

provided that

$$\lim_{h' \to 0} \Theta(u_{h'}, c; \beta_{h'}; \varphi) \leq 0.$$

In §4c and §4d we obtain a bound of the form

$$\Theta(u_{h'}, c; \beta_{h'}; \varphi) \leq L\, C_2\, (\Delta x/\epsilon + \Delta\tau/\epsilon_0) + M\, C_3\, \Delta\tau.$$

With these results, Theorems 2.3 and 2.4 can be easily proven. In §4e we show how to do that.

A delicate point in this procedure is how to take the limit in the boundary terms of the form Θ. This was done in [25] in the framework of classical conservation laws. Another point we want to stress is the necessity of using entropies U which are smoother than the classical entropy $U(u) = |u|$; see [13]. Notice how the second derivative of the entropy function U is needed in the second inequality of Theorem 2.4.

4a. Existence of a converging subsequence. First, we reduce the problem of finding a converging subsequence $\{(u_{h'}, \beta_{h'}, \phi_{h'})\}_{h'>0}$ to the problem of finding a converging subsequence $\{\, u_{h'} \,\}_{h'>0}$.

LEMMA 4.1. *Suppose there is a subsequence* $\{\, u_{h'} \,\}_{h'>0}$ *converging in* $L^\infty(0,T; L^1(0,1))$ *to a limit* $u \in L^\infty(0,T; \mathrm{BV}(0,1)) \cap \mathcal{C}^0(0,T; L^1(0,1))$. *Then, the subsequence* $\{(u_{h'}, \beta_{h'}, \phi_{h'})\}_{h'>0}$ *converges in*

$$L^\infty(0,T; L^1(0,1)) \times L^1(0,T; W^{1,1}(0,1)) \times L^1(0,T; BV(0,1))$$

to a limit (u, β, ϕ) *satisfying the equations* (2.3).

Proof. By (2.7a), and by (2.7b) with $v_h \equiv 1$, we have

$$(\beta_{h'_1} - \beta_{h'_2})_x = u_{h'_1} - u_{h'_2},$$
$$\int_0^1 (\beta_{h'_1} - \beta_{h'_2}) = \mathbf{P}_{\Delta\tau'_1}\phi_1 - \mathbf{P}_{\Delta\tau'_2}\phi_1.$$

This implies that $\{\, \beta_{h'} \,\}_{h'>0}$ is a Cauchy sequence in $L^1(0,T; W^{1,1}(0,1))$. (Notice that since ϕ_1 belongs to $\mathrm{BV}(0,T)$, ϕ_1 can have discontinuities. This forces us to consider convergence in $\mathrm{L}^1(0,T)$ instead of in $\mathrm{L}^\infty(0,T)$.) Denote by β the limit of the above sequence. By (2.7a), and by (2.7b) with $v_h \equiv 1$, we easily see that β satisfies (2.3a) and (2.3b) with $v \equiv 1$ a.e. in $(0,T)$.

By using (2.7b) and a similar argument, we can see that $\{\phi_{h'}\}_{h'>0}$ is a Cauchy sequence in $L^1(0,T; BV(0,1))$ and that its limit ϕ satisfies (2.3b) a.e. in $(0,T)$. □

Now we prove that a converging subsequence $\{\, u_{h'} \,\}_{h'>0}$ exists.

LEMMA 4.2. *Suppose that the* CFL *condition* (2.9) *is satisfied. Then there exists a subsequence* $\{\, u_{h'} \,\}_{h'>0}$ *converging in* $L^\infty(0,T; L^1(0,1))$ *to a function* u *in* $L^\infty(0,T; BV(0,1)) \cap \mathcal{C}^0(0,T; L^1(0,1))$.

Proof. To prove this result, we only have to follow Crandall and Majda [15], who used a discrete version of the Ascoli-Arzelà Theorem. The equicontinuity in time is provided by Proposition 3.19 and the compactness of the range is given by Proposition 3.9. □

Thus we have the following immediate result.

COROLLARY 4.3. *Suppose that the* CFL *condition* (2.9) *is satisfied. Then there exists a subsequence* $\{(u_{h'}, \beta_{h'}, \phi_{h'})\}_{h'>0}$ *converging as indicated in Theorem 2.3 to a limit* (u, β, ϕ) *satisfying* (2.3). *Moreover,* $u \in L^\infty(0,T; BV(0,1)) \cap \mathcal{C}^0(0,T; L^1(0,1))$.

4b. Passing to the limit in $\Theta(u_{h'}, c; \beta_{h'}; \varphi)$. In this section we prove the following result.

LEMMA 4.4. *Suppose that for every* $c \in \mathbf{R}$ *and for every nonnegative* $\varphi \in \mathcal{C}^1([0,T] \times [0,1])$,

$$\lim_{h' \to 0} \Theta(u_{h'}, c; \beta_{h'}; \varphi) \leq 0.$$

Then

$$\lim_{h' \to 0} \Theta(u_{h'}, c; \beta_{h'}; \varphi) = \Theta(u, c; \beta; \varphi).$$

Proof. By a standard argument, we have that for $c \in \mathbf{R}$ and for every nonnegative $\varphi \in \mathcal{C}_0^1([0,T] \times (0,1))$,

$$\lim_{h' \to 0} \Theta(u_{h'}, c; \beta_{h'}; \varphi) = \Theta(u, c; \beta; \varphi).$$

Now, consider the case in which $\varphi \in \mathcal{C}_0^1((0,T) \times [0,1))$; it is enough to consider $\varphi(\tau, x)$ of the form $\omega(\tau)\eta(x)$. Notice that since, by Lemma 4.1, $\{\beta_{h'}\}_{h'>0}$ converges to β in $L^1(0,T; W^{(1,1)}(0,1))$, $\{\beta_{h'}(\cdot,0)\}_{h'>0}$ converges to $\beta(\cdot,0)$ in $L^1(0,T)$. Also, notice that since, by Theorem 2.1, the sequence $\{u_{h'}(\cdot,0+)\}_{h'>0}$ is included in $L^\infty(0,T)$, there is a subsequence $\{u_{h''}(\cdot,0+)\}_{h''>0}$ converging in $L^\infty(0,T)$ weak* to a limit $\tilde{u}$. Let γ_τ be the Young measure associated with $\tilde{u}$. Thus,

$$\begin{aligned}\lim_{h'' \to 0} \Theta(u_{h''}, c; \beta_{h''}; \varphi) = &- \int_0^T \int_0^1 U(u(\tau,x) - c)\omega'(\tau)\eta(x)\, dx d\tau \\ &- \int_0^1 g(x)\eta'(x) dx - g_0\eta(0) \\ &- \int_0^T \int_0^1 (\beta_x) V(u(\tau,x), c)\omega(\tau)\eta(x)\, dx d\tau,\end{aligned}$$

where

$$g(x) = \int_0^T U(u(\tau,x)-c)\beta(\tau,x)\omega(\tau)d\tau, \tag{4.1a}$$

$$g_0 = \int_0^T \{\beta^+(\tau,0)U(u_0(\tau)-c) + \beta^-(\tau,0)v(\tau)\}\,\omega(\tau)d\tau, \tag{4.1b}$$

and

$$v(\tau) = \int_0^{u^*} U(\lambda - c)\, d\gamma_\tau(\lambda).$$

To prove

$$\lim_{h''\to 0} \Theta(u_{h''}, c; \beta_{h''}; \varphi) = \Theta(u, c; \beta; \varphi),$$

we only have to prove that $g_0 = g_0^*$, where

$$g_0^* = \int_0^T \{\beta^+(\tau,0)U(u_0(\tau)-c) + \beta^-(\tau,0)U(u(\tau,0+)-c)\}\,\omega(\tau)d\tau. \tag{4.2}$$

Take η such that its support is contained in $[0,\varepsilon]$. By our hypothesis and Theorem 2.1,

$$-\int_0^1 g(x)\eta'(x)dx - g_0\,\eta(0) \le C\,\varepsilon\,\|\eta\|_{L^1(0,1)}.$$

Since ε is arbitrary, this implies that

$$g(0+) - g_0 \le 0. \tag{4.1c}$$

Notice that since $g \in \mathrm{BV}(0,1)$, by (4.1a) and Theorem 2.1, the limit $g(0+)$ exits. Next, take U and c such that $U(u-c) = \alpha\,(u-c)$ for $u \in [0,u^*]$ and $\alpha \in \mathbf{R}$. Then

$$v(\tau) = \alpha\,(\tilde{u} - c),$$

and so, by (4.1a,b), the last inequality reads as follows:

$$\alpha\left\{\int_0^T (u(\tau,0+) - u_0(\tau))\beta^+(\tau,0)\omega(\tau)d\tau + \int_0^T (u(\tau,0+) - \tilde{u}(\tau))\beta^-(\tau,0)\omega(\tau)d\tau\right\} \le 0.$$

Since the sign of α is arbitrary and this holds for every nonnegative $\omega \in C_0^1(0,T)$, this implies that for τ a.e. in $[0,T]$,

$$\beta^+(\tau,0)\big(u(\tau,0+) - u_0(\tau)\big) = 0, \tag{4.3a}$$

$$\beta^-(\tau,0)\big(u(\tau,0+) - \tilde{u}(\tau)\big) = 0. \tag{4.3b}$$

Thus, we have

$$\begin{aligned}
0 &\ge g(0+) - g_0, \quad \text{by (4.1c)},\\
&= g_0^* - g_0, \quad \text{by (4.1a), (4.2), and (4.3a)},\\
&= \int_0^T \big[U(u(\tau,0+)-c) - v(\tau)\big]\,\beta^-(\tau,0)\,\omega(\tau)\,d\tau, \quad \text{by (4.1b), and (4.2)},\\
&= \int_0^T \big[U(\tilde{u}(\tau)-c) - v(\tau)\big]\,\beta^-(\tau,0)\,\omega(\tau)\,d\tau, \quad \text{by (4.3b)},\\
&= \int_0^T \big[U(\int_0^{u^*}(\lambda-c)\,d\gamma_\tau(\lambda)) - v(\tau)\big]\,\beta^-(\tau,0)\,\omega(\tau)\,d\tau\\
&\ge \int_0^T \big[\int_0^{u^*} U(\lambda-c)\,d\gamma_\tau(\lambda) - v(\tau)\big]\,\beta^-(\tau,0)\,\omega(\tau)\,d\tau, \text{ by Jensen's inequality},\\
&= 0, \quad \text{by the definition of } v.
\end{aligned}$$

This proves that $g_0^* = g_0$, and ends the proof for the case $\varphi \in C_0^1([0,T]\times[0,1))$.

A similar argument shows that the result is also true for $\varphi \in C^1([0,T]\times(0,1])$. This completes the proof. □

4c. A discrete entropy inequality. In this section we obtain a discrete entropy inequality, which is crucial for obtaining an upper bound for $\Theta(u_{h'}, c; \beta_{h'}; \varphi)$.

PROPOSITION 4.5 (A discrete entropy inequality). *Suppose that the* CFL *condition* (2.9) *is satisfied. Then for every* $c \in \mathbf{R}$,

$$U(u_i^{n+1}-c) - U(u_i^n - c) + \frac{\Delta\tau^n}{\Delta x_i}(G_{i+\frac12}^n - G_{i-\frac12}^n) - ((\beta_h)_x)_i^n V(u_i^{n+1}, c)\Delta\tau^n \le 0,$$

where

$$G_{i+\frac12}^n = \beta_{i+\frac12}^{n+} U(u_i^n - c) + \beta_{i+\frac12}^{n-} U(u_{i+1}^n - c),$$
$$V(u_i^{n+1}, c) = U(u_i^{n+1} - c) - u_i^{n+1} U'(u_i^{n+1} - c).$$

Proof. By the definition of the scheme (2.6), and by the fact that β_h^n is piecewise linear, we have, for $c \in \mathbf{R}$,

$$\begin{aligned}
u_i^{n+1} - c = {} & \{-\frac{\Delta\tau^n}{\Delta x_i}\beta_{i+\frac12}^{n-}\}(u_{i+1}^n - c) + \{1 - \frac{\Delta\tau^n}{\Delta x_i}(\beta_{i+\frac12}^{n+} - \beta_{i-\frac12}^{n-})\}(u_i^n - c)\\
& + \{\frac{\Delta\tau^n}{\Delta x_i}\beta_{i-\frac12}^{n+}\}(u_{i-1}^n - c) - ((\beta_h)_x)_i^n \Delta\tau^n c.
\end{aligned}$$

Since, by the CFL condition (2.9), the terms between braces are nonnegative, we obtain, after multiplying by $\mathrm{sgn}(u_i^{n+1} - c)$,

$$|u_i^{n+1} - c| \le |u_i^n - c| - \frac{\Delta\tau^n}{\Delta x_i}(F_{i+\frac12}^n - F_{i-\frac12}^n) + ((\beta_h)_x)_i^n W(u_i^{n+1}, c)\,\Delta\tau^n,$$

where

$$F^n_{i+\frac{1}{2}} = \beta^{n+}_{i+\frac{1}{2}}\,|u^n_i - c| + \beta^{n-}_{i+\frac{1}{2}}\,|u^n_{i+1} - c|,$$
$$W(u^{n+1}_i, c) = -\mathrm{sgn}(u^{n+1}_i - c)\,c.$$

This proves the result in the case in which $U(\omega) = |\omega|$. To prove the result for a general even entropy function U, with Lipschitz second-order derivative having support in $[-\alpha, \alpha]$ and such that $U'(0) = 0$, we use the following representation of U:

$$U(u) = \int_{-\alpha}^{\alpha} j(s)\{|u - s| - |s|\}ds,$$

where $j = \frac{1}{2}U''$. We replace c by $c + s$ in the above inequality, multiply by $j(s) \geq 0$ (since U is convex), and integrate over $[-\alpha, \alpha]$. The result follows from the above representation result and from the fact that

$$V(u^{n+1}_i, c) = \int_{-\alpha}^{\alpha} j(s)\{-\mathrm{sgn}(u^{n+1}_i - (c+s))(c+s) - |s|\}ds. \quad \square$$

In §4d we shall need a result concerning our estimate of the Lipschitz constant of $u \mapsto V(u, c)$.

LEMMA 4.6. *We have*

$$|V(u_1, c) - V(u_2, c)| \leq M \max\{|u_1|, |u_2|\}|u_1 - u_2|,$$

where $M = \sup_{u \in \mathbf{R}} |U''(u)|$.

Proof. By hypothesis, $U(0) = 0$, and since U is even and $C^1, U'(0) = 0$. Hence, it is easy to see that

$$\begin{aligned} V(u, c) &= U(u - c) - uU'(u - c) \\ &= -\int_0^{u-c} (c + s)U''(s)ds. \end{aligned}$$

Hence,

$$V(u_1, c) - V(u_2, c) = -\int_{u_2 - c}^{u_1 - c} (c + s)U''(s)ds,$$

and the result follows. $\square$

4d. An upper bound for $\Theta(u_{h'}, c; \beta_{h'}; \varphi)$. To obtain an upper bound for $\Theta(u_{h'}, c; \beta_{h'}; \varphi)$, we decompose $\Theta(u_{h'}, c; \beta_{h'}; \varphi)$ as the sum $\Theta_{\mathrm{ent}}(u_{h'}, c; \beta_{h'}; \varphi) + \Theta_{\mathrm{comp}}(u_{h'}, c; \beta_{h'}; \varphi)$; this is done in Lemma 4.7. The upper bound for $\Theta_{\mathrm{ent}}(u_{h'}, c; \beta_{h'}; \varphi)$ is obtained by using the discrete entropy inequality of Proposition 4.5 and the upper bound for $\Theta_{\mathrm{comp}}(u_{h'}, c; \beta_{h'}; \varphi)$ by using the compactness properties obtained in §3; this is done in Lemma 4.8.

LEMMA 4.7 (Decomposition of $\Theta(u_{h'}, c; \beta_{h'}; \varphi)$). *We have*

$$\Theta(u_{h'}, c; \beta_{h'}; \varphi) = \Theta_{\mathrm{ent}}(u_{h'}, c; \beta_{h'}; \varphi) + \Theta_{\mathrm{comp}}(u_{h'}, c; \beta_{h'}; \varphi),$$

where

$$\begin{aligned} \Theta_{\mathrm{ent}}(u_{h'}, c; \beta_{h'}; \varphi) = \sum_{n=0}^{n_T-1} \sum_{i=1}^{n_x} &\{U(u^{n+1}_i - c) - U(u^n_i - c) \\ &+ \frac{\Delta\tau^n}{\Delta x_i}(G^n_{i+\frac{1}{2}} - G^n_{i-\frac{1}{2}}) - ((\beta_{h'})_x)^n_i V(u^{n+1}_i, c)\Delta\tau^n\}\varphi^{n+1}_i \Delta x_i, \end{aligned}$$

and

$$\begin{aligned} \Theta_{\mathrm{comp}}(u_{h'}, c; \beta_{h'}; \varphi) = &\sum_{n=0}^{n_T-1} \sum_{i=1}^{n_x} \{[U(u^n_{i+1} - c) - U(u^n_i - c)](-\beta^{n-}_{i+\frac{1}{2}})[\varphi^{n+1}_i - \varphi^{n+\frac{1}{2}}_{i+\frac{1}{2}}] \\ &+ [U(u^n_{i-1} - c) - U(u^n_i - c)](\beta^{n+}_{i-\frac{1}{2}})[\varphi^{n+1}_i - \varphi^{n+\frac{1}{2}}_{i-\frac{1}{2}}]\}\Delta\tau^n \\ &- \sum_{n=0}^{n_T-1} \sum_{i=1}^{n_x} ((\beta_{h'})_x)^n_i\, u^n_i\, U'(u^n_i - c)\,[\varphi^{n+1}_i - \varphi^{n+\frac{1}{2}}_i]\Delta x_i \Delta\tau^n \\ &+ \sum_{n=0}^{n_T-1} \sum_{i=1}^{n_x} ((\beta_{h'})_x)^n_i\,(V(u^{n+1}_i, c) - V(u^n_i, c))\,\varphi^{n+1}_i \Delta x_i \Delta\tau^n, \end{aligned}$$

where

$$\begin{aligned} \varphi^{n+1}_i &= \frac{1}{\Delta x_i}\int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} \varphi(\tau^{n+1}, x)\,dx, \\ \varphi^{n+\frac{1}{2}}_{i+\frac{1}{2}} &= \frac{1}{\Delta\tau^n}\int_{\tau^n}^{\tau^{n+1}} \varphi(\tau, x_{i+\frac{1}{2}})\,d\tau, \\ \varphi^{n+\frac{1}{2}}_i &= \frac{1}{\Delta\tau^n}\int_{\tau^n}^{\tau^{n+1}} \frac{1}{\Delta x_i}\int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} \varphi(\tau, x)\,dx\,d\tau. \end{aligned}$$

Proof. From the definition of the form $\Theta(u_h, c; \beta_h; \varphi)$, (2.11b), and the fact that $u_{h'}$ and $(\beta_{h'})_x$ are piecewise constant functions, we get

$$\Theta(u_{h'}, c; \beta_{h'}; \varphi) = \Psi_t(u_{h'}, c; \beta_{h'}; \varphi) + \Psi_x(u_{h'}, c; \beta_{h'}; \varphi) + \Psi(u_{h'}, c; \beta_{h'}; \varphi),$$

where

$$\Psi_t(u_{h'},c;\beta_{h'};\varphi) = -\sum_{n=0}^{n_T-1}\sum_{i=1}^{n_x} U(u_i^n - c)(\varphi_i^{n+1} - \varphi_i^n)\Delta x_i$$
$$+ \sum_{i=1}^{n_x} U(u_i^{n_T} - c)\varphi_i^{n_T}\Delta x_i - \sum_{i=1}^{n_x} U(u_i^0 - c)\varphi_i^0 \Delta x_i,$$
$$\Psi_x(u_{h'},c;\beta_{h'};\varphi) = -\sum_{n=0}^{n_T-1}\sum_{i=1}^{n_x} U(u_i^n - c)(\beta_{i+\frac{1}{2}}^n \varphi_{i+\frac{1}{2}}^{n+\frac{1}{2}} - \beta_{i-\frac{1}{2}}^n \varphi_{i-\frac{1}{2}}^{n+\frac{1}{2}})\,\Delta\tau^n$$
$$+ \sum_{n=0}^{n_T-1} G_{n_x+\frac{1}{2}}^n\, \varphi_{n_x+\frac{1}{2}}^{n+\frac{1}{2}}\,\Delta\tau^n - \sum_{n=0}^{n_T-1} G_{\frac{1}{2}}^n\, \varphi_{\frac{1}{2}}^{n+\frac{1}{2}}\,\Delta\tau^n,$$
$$\Psi(u_{h'},c;\beta_{h'};\varphi) = \sum_{n=0}^{n_T-1}\sum_{i=1}^{n_x}((\beta_{h'})_x)_i^n\, u_i^n\, U'(u_i^n - c)\,\varphi_i^{n+\frac{1}{2}}\,\Delta x_i\,\Delta\tau^n.$$

Consider Ψ_t. After a simple integration by parts, we get

$$\Psi_t(u_{h'},c;\beta_{h'};\varphi) = \sum_{n=0}^{n_T-1}\sum_{i=1}^{n_x}\big(U(u_i^{n+1} - c) - U(u_i^n - c)\big)\,\varphi_i^{n+1}\Delta x_i,$$

and by the definition of $\Theta_{\text{ent}}(u_{h'},c;\beta_{h'};\varphi)$,

$$\Psi_t(u_{h'},c;\beta_{h'};\varphi) = \Theta_{\text{ent}}(u_{h'},c;\beta_{h'};\varphi)$$
$$- \sum_{n=0}^{n_T-1}\sum_{i=1}^{n_x}(G_{i+\frac{1}{2}}^n - G_{i-\frac{1}{2}}^n)\,\varphi_i^{n+1}\,\Delta\tau^n$$
$$+ \sum_{n=0}^{n_T-1}\sum_{i=1}^{n_x}((\beta_{h'})_x)_i^n\, V(u_i^{n+1},c)\,\varphi_i^{n+1}\,\Delta\tau^n\Delta x_i.$$

Since β_h is piecewise linear in space, $((\beta_{h'})_x)_i^n = (\beta_{i+1/2}^n - \beta_{i-1/2}^n)/\Delta x_i$, and so

$$\Psi_t(u_{h'},c;\beta_{h'};\varphi) = \Theta_{\text{ent}}(u_{h'},c;\beta_{h'};\varphi)$$
$$- \sum_{n=0}^{n_T-1}\sum_{i=1}^{n_x}\{[G_{i+\frac{1}{2}}^n - U(u_i^n - c)\,\beta_{i+\frac{1}{2}}^n]\,\varphi_i^{n+1}$$
$$- [G_{i-\frac{1}{2}}^n - U(u_i^n - c)\,\beta_{i-\frac{1}{2}}^n]\,\varphi_i^{n+1}\}\,\Delta\tau^n$$
$$- \sum_{n=0}^{n_T-1}\sum_{i=1}^{n_x}((\beta_{h'})_x)_i^n\, U(u_i^n - c)\,\varphi_i^{n+1}\,\Delta\tau^n\Delta x_i$$
$$+ \sum_{n=0}^{n_T-1}\sum_{i=1}^{n_x}((\beta_{h'})_x)_i^n\, V(u_i^{n+1},c)\,\varphi_i^{n+1}\,\Delta\tau^n\Delta x_i.$$

Since $G_{i+1/2}^n = U(u_{i+1}^n - c)\,\beta_{i+1/2}^- + U(u_i^n - c)\,\beta_{i+1/2}^+$, by (2.11c), and $U(u_i^n - c) = V(u_i^n, c) + u_i^n U'(u_i^n - c)$, by (2.11d), we get

$$\Psi_t(u_{h'},c;\beta_{h'};\varphi) = \Theta_{\text{ent}}(u_{h'},c;\beta_{h'};\varphi)$$
$$+ \sum_{n=0}^{n_T-1}\sum_{i=1}^{n_x}\{[U(u_{i+1}^n - c) - U(u_i^n - c)](-\beta_{i+\frac{1}{2}}^{n-})\,\varphi_i^{n+1}$$
$$+ [U(u_{i-1}^n - c) - U(u_i^n - c)](\beta_{i-\frac{1}{2}}^{n+})\,\varphi_i^{n+1}\}\Delta\tau^n$$
$$- \sum_{n=0}^{n_T-1}\sum_{i=1}^{n_x}((\beta_{h'})_x)_i^n\, u_i^n\, U'(u_i^n - c)\,\varphi_i^{n+1}\Delta x_i\Delta\tau^n$$
$$+ \sum_{n=0}^{n_T-1}\sum_{i=1}^{n_x}((\beta_{h'})_x)_i^n(V(u_i^{n+1},c) - V(u_i^n,c))\,\varphi_i^{n+1}\Delta x_i\Delta\tau^n.$$

By the definition of $\Psi(u_{h'},c;\beta_{h'};\varphi)$, we have

$$\Psi_t(u_{h'},c;\beta_{h'};\varphi) = \Theta_{\text{ent}}(u_{h'},c;\beta_{h'};\varphi) - \Psi(u_{h'},c;\beta_{h'};\varphi)$$
$$+ \sum_{n=0}^{n_T-1}\sum_{i=1}^{n_x}\{[U(u_{i+1}^n - c) - U(u_i^n - c)]\,(-\beta_{i+\frac{1}{2}}^{n-})\,\varphi_i^{n+1}$$
$$+ [U(u_{i-1}^n - c) - U(u_i^n - c)]\,(\beta_{i+\frac{1}{2}}^{n+})\,\varphi_i^{n+1}\}\,\Delta\tau^n$$
$$- \sum_{n=0}^{n_T-1}\sum_{i=1}^{n_x}((\beta_{h'})_x)_i^n\, u_i^n\, U'(u_i^n - c)\,[\varphi_i^{n+1} - \varphi_i^{n+\frac{1}{2}}]\,\Delta x_i\Delta\tau^n$$
$$+ \sum_{n=0}^{n_T-1}\sum_{i=1}^{n_x}((\beta_{h'})_x)_i^n\,(V(u_i^{n+1},c) - V(u_i^n,c))\,\varphi_i^{n+1}\,\Delta x_i\Delta\tau^n.$$

Finally, by the definition of $\Theta_{\text{comp}}(u_{h'},c;\beta_{h'};\varphi)$, we have

$$\Psi_t(u_{h'},c;\beta_{h'};\varphi) = \Theta_{\text{ent}}(u_{h'},c;\beta_{h'};\varphi) + \Theta_{\text{comp}}(u_{h'},c;\beta_{h'};\varphi) - \Psi(u_{h'},c;\beta_{h'};\varphi)$$
$$+ \sum_{n=0}^{n_T-1}\sum_{i=1}^{n_x}\{[U(u_{i+1}^n - c) - U(u_i^n - c)]\,(-\beta_{i+\frac{1}{2}}^{n-})\,\varphi_{i+\frac{1}{2}}^{n+\frac{1}{2}}$$
$$+ [U(u_{i-1}^n - c) - U(u_i^n - c)]\,(\beta_{i-\frac{1}{2}}^{n+})\,\varphi_{i-\frac{1}{2}}^{n+\frac{1}{2}}\}\Delta\tau^n$$
$$= \Theta_{\text{ent}}(u_{h'},c;\beta_{h'};\varphi) + \Theta_{\text{comp}}(u_{h'},c;\beta_{h'};\varphi) - \Psi(u_{h'},c;\beta_{h'};\varphi)$$
$$- \sum_{n=0}^{n_T-1}\sum_{i=1}^{n_x}\{[G_{i+\frac{1}{2}}^n - U(u_i^n - c)\,\beta_{i+\frac{1}{2}}^n]\,\varphi_{i+\frac{1}{2}}^{n+\frac{1}{2}}$$
$$- [G_{i-\frac{1}{2}}^n - U(u_i^n - c)\,\beta_{i-\frac{1}{2}}^n]\,\varphi_{i-\frac{1}{2}}^{n+\frac{1}{2}}\}\Delta\tau^n$$

$$
\begin{aligned}
&= \Theta_{\text{ent}}(u_{h'}, c; \beta_{h'}; \varphi) + \Theta_{\text{comp}}(u_{h'}, c; \beta_{h'}; \varphi) - \Psi(u_{h'}, c; \beta_{h'}; \varphi) \\
&+ \sum_{n=0}^{n_T-1} \sum_{i=1}^{n_x} U(u_i^n - c)(\beta_{i+\frac{1}{2}}^n \varphi_{i+\frac{1}{2}}^{n+\frac{1}{2}} - \beta_{i-\frac{1}{2}}^n \varphi_{i-\frac{1}{2}}^{n+\frac{1}{2}}) \Delta\tau^n \\
&- \sum_{n=0}^{n_T-1} \sum_{i=1}^{n_x} (G_{i+\frac{1}{2}}^n \varphi_{i+\frac{1}{2}}^{n+\frac{1}{2}} - G_{i-\frac{1}{2}}^n \varphi_{i-\frac{1}{2}}^{n+\frac{1}{2}}) \Delta\tau^n \\
&= \Theta_{\text{ent}}(u_{h'}, c; \beta_{h'}; \varphi) + \Theta_{\text{comp}}(u_{h'}, c; \beta_{h'}; \varphi) \\
&- \Psi(u_{h'}, c; \beta_{h'}; \varphi) - \Psi_x(u_{h'}, c; \beta_{h'}; \varphi),
\end{aligned}
$$

by the definition of $\Psi_x(u_{h'}, c; \beta_{h'}; \varphi)$. This completes the proof. □

We are now ready to obtain the upper bounds we were looking for.

LEMMA 4.8 (Upper bounds for $\Theta_{\text{ent}}(u_{h'}, c; \beta_{h'}; \varphi)$ and $\Theta_{\text{comp}}(u_{h'}, c; \beta_{h'}; \varphi)$). *Suppose that the* CFL *condition* (2.9) *is satisfied. Then*

$$
\begin{aligned}
\Theta_{\text{ent}}(u_{h'}, c; \beta_{h'}; \varphi) &\le 0, \\
\Theta_{\text{comp}}(u_{h'}, c; \beta_{h'}; \varphi) &\le L\,\hat{C}_2\,(\Delta x\,\|\varphi_x\|_{L^\infty(0,T;L^\infty(0,1))} + \Delta\tau\,\|\varphi_\tau\|_{L^\infty(0,T;L^\infty(0,1))}) \\
&\quad + M\,\hat{C}_3\,\Delta\tau\|\varphi\|_{L^\infty((0,T)\times(0,1))},
\end{aligned}
$$

where $L = \sup_{u\in\mathbf{R}} |U'(u)|$, $M = \sup_{u\in\mathbf{R}} |U''(u)|$, *and*

$$
\begin{aligned}
\hat{C}_2 &= \{(\phi_1^* + \frac{1}{2}\max\{1, u^* - 1\})C_1 + \frac{1}{2}u^* \max\{1, u^* - 1\}\}\,T, \\
\hat{C}_3 &= u^* \max\{1, u^* - 1\}\,T\,C_6.
\end{aligned}
$$

The constant C_6 is given in Proposition 3.14.

Proof. The first inequality follows immediately from Lemma 4.7 and the discrete entropy inequality of Proposition 4.5.

To estimate $\Theta_{\text{comp}}(u_{h'}, c; \beta_{h'}; \varphi)$, we proceed as follows. By Lemmas 4.7 and 4.6,

(4.4)

$$
\begin{aligned}
\Theta_{\text{comp}}(u_{h'}, c; \beta_{h'}; \varphi) &\le L\,\|\beta_{h'}\|_{L^\infty(0,T;L^\infty(0,1))} \sum_{n=0}^{n_T-1} \sum_{i=1}^{n_x} \{|u_{i+1}^n - u_i^n|\,|\varphi_i^{n+1} - \varphi_{i+\frac{1}{2}}^{n+\frac{1}{2}}| + \\
&\qquad |u_i^n - u_{i-1}^n|\,|\varphi_i^{n+1} - \varphi_{i-\frac{1}{2}}^{n+\frac{1}{2}}|\}\Delta\tau^n \\
&+ L\,\|(\beta_{h'})_x\|_{L^\infty(0,T;L^\infty(0,1))} \sum_{n=0}^{n_T-1} \sum_{i=1}^{n_x} u_i^n\,|\varphi_i^{n+1} - \varphi_i^{n+\frac{1}{2}}|\Delta x_i \Delta\tau^n \\
&+ M\,\|(\beta_{h'})_x\|_{L^\infty(0,T;L^\infty(0,1))}\,u^* \sum_{n=0}^{n_T-1} \sum_{i=1}^{n_x} |u_i^{n+1} - u_i^n|\,|\varphi_i^{n+1}\Delta x_i \Delta\tau^n,
\end{aligned}
$$

where $L = \sup_{u\in\mathbf{R}} |U'(u)|$ and $M = \sup_{u\in\mathbf{R}} |U''(u)|$.

Since (see Lemma 4.7)

$$
\begin{aligned}
|\varphi_i^{n+1} - \varphi_{i+\frac{1}{2}}^{n+\frac{1}{2}}| &\le \frac{1}{2}\Delta x_i\,\|\varphi_x\|_{L^\infty(0,T;L^\infty(0,1))} + \frac{1}{2}\Delta\tau^n\,\|\varphi_\tau\|_{L^\infty(0,T;L^\infty(0,1))}, \\
|\varphi_i^{n+1} - \varphi_i^{n+\frac{1}{2}}| &\le \frac{1}{2}\Delta\tau^n\,\|\varphi_\tau\|_{L^\infty(0,T;L^\infty(0,1))},
\end{aligned}
$$

we obtain

$$
\begin{aligned}
&\Theta_{\text{comp}}(u_{h'}, c; \beta_{h'}; \varphi) \le \\
&\quad L\,\|\beta_{h'}\|_{L^\infty(0,T;L^\infty(0,1))}\,\|u_{h'}\|_{L^\infty(0,T;\mathrm{BV}(0,1))}\,T \\
&\qquad \cdot (\Delta x\,\|\varphi_x\|_{L^\infty(0,T;L^\infty(0,1))} + \Delta\tau\,\|\varphi_\tau\|_{L^\infty(0,T;L^\infty(0,1))}) \\
&\quad + \frac{1}{2}L\,\|(\beta_{h'})_x\|_{L^\infty(0,T;L^\infty(0,1))}\,\|u_{h'}\|_{L^\infty(0,T;L^1(0,1))}\,T\,\Delta\tau\,\|\varphi_\tau\|_{L^\infty(0,T;L^\infty(0,1))} \\
&\quad + M\|(\beta_{h'})_x\|_{L^\infty(0,T;L^\infty(0,1))}\,u^*\,T \sup_{0\le n\le n_T-1} \|u_{h'}^{n+1} - u_{h'}^n\|_{L^1(0,1)}\,\|\varphi\|_{L^\infty((0,T)\times(0,1))},
\end{aligned}
$$

and the result follows by using Theorem 2.1 and Proposition 3.14. □

We end this section with the following result.

COROLLARY 4.9 (An upper bound for $E^{\epsilon,\epsilon_0}(u_{h'}, v; \beta_{h'})$). *Suppose that the* CFL *condition* (2.9) *is satisfied. Then*

$$
E^{\epsilon,\epsilon_0}(u_{h'}, v; \beta_{h'}) \le L\,C_2\,(\Delta x/\epsilon + \Delta\tau/\epsilon_0) + M\,C_3\,\Delta\tau,
$$

where $L = \sup_{u\in\mathbf{R}} |U'(u)|$, $M = \sup_{u\in\mathbf{R}} |U''(u)|$, $C_2 = C_0\,\hat{C}_2$, $C_3 = C_0\,\hat{C}_3$, *and* C_0 *depends solely on the function* w *used to define the form* E^{ϵ,ϵ_0}.

Proof. In the definition of E^{ϵ,ϵ_0}, (2.11), the function φ is a function of four variables:

$$
\varphi(\tau, x; \tau', x') = \frac{1}{\epsilon_0}\,w(\frac{\tau - \tau'}{\epsilon_0})\;\frac{1}{\epsilon}\,w(\frac{x - x'}{\epsilon}),
$$

two of which are 'frozen' in the form Θ; see (2.11). Taking this into consideration, we can easily obtain that

$$
\begin{aligned}
\int_0^T \int_0^1 |\varphi_i^{n+1} - \varphi_{i+\frac{1}{2}}^{n+\frac{1}{2}}| &\le \frac{1}{2}C_0\Delta x_i/\epsilon + \frac{1}{2}C_0\Delta\tau^n/\epsilon_0, \\
\int_0^T \int_0^1 |\varphi_i^{n+1} - \varphi_i^{n+\frac{1}{2}}| &\le \frac{1}{2}C_0\Delta\tau^n/\epsilon_0, \\
\int_0^T \int_0^1 \varphi_i^{n+1} &\le 1,
\end{aligned}
$$

where C_0 depends solely on the function w.

The result follows easily from the above inequalities, (4.4), (2.11a) and the proof of Lemma 4.8. □

4e. Proof of Theorems 2.3 and 2.4. First, let us consider Theorem 2.3. By Corollary 4.3, and Lemmas 4.4 and 4.8, there is a subsequence $\{(u_{h'}, \beta_{h'}, \phi_{h'})\}_{h'>0}$ converging (in the topology indicated by Theorem 2.3) to a limit (u, β, ϕ) satisfying (2.3) a.e. in $(0, T)$ such that for $c \in \mathbf{R}$ and $\varphi \in \mathcal{C}^1([0,T] \times [0,1])$,

$$\Theta(u, c; \beta; \varphi) = \lim_{h' \to 0} \Theta(u_{h'}, c; \beta_{h'}; \varphi) \leq 0.$$

This implies that (u, β, ϕ) is the unique solution of (2.2), (2.3). As a consequence, the whole sequence $\{(u_h, \beta_h, \phi_h)\}_{h>0}$ converges to (u, β, ϕ), and thus Theorem 2.3 is proved.

The first inequality of Theorem 2.4 follows from the above inequality and (2.11). The second inequality follows from Corollary 4.9.