CONVERGENCE ANALYSIS OF FINITE VOLUME SCHEME FOR NONLINEAR TENSOR ANISOTROPIC DIFFUSION IN IMAGE PROCESSING

OLGA DRBLÍKOVÁ AND KAROL MIKULA *

Abstract. In this article we design the semi-implicit finite volume scheme for coherence enhancing diffusion in image processing and prove its convergence to the weak solution of the problem. The finite volume methods are natural tools for image processing applications since they use piecewise constant representation of approximate solutions similarly to the structure of digital images. They have been successfully applied in image processing, e.g., for solving the Perona-Malik equation or curvature driven level set equations, where the nonlinearities are represented by a scalar function dependent on solution gradient. Design of suitable finite volume schemes for tensor diffusion is a nontrivial task, here we present first such scheme with convergence proof for the practical nonlinear model used in coherence enhancing image smoothing. We provide basic information about this type of nonlinear diffusion including a construction of its diffusion tensor, and we derive a semi-implicit finite volume scheme for this nonlinear model with the help of co-volume mesh. This method is well-known as diamond-cell method owing to the choice of co-volume as a diamond-shaped polygon. Further, we prove a convergence of a discrete solution given by our scheme to the weak solution of the problem. The proof is based on Kolmogorov's compactness theorem and a bounding of a gradient in tangential direction by using a gradient in normal direction. Finally computational results illustrated in figures are discussed.

Key words. image processing, nonlinear tensor diffusion, numerical solution, semi-implicit scheme, diamond-cell finite volume method, convergence.

1. Introduction. Nonlinear diffusion models are widely used nowadays in many practical tasks of image processing. In this paper we deal with the numerical solution of the model of tensor nonlinear anisotropic diffusion introduced by Weickert (see [23], [24] and [22]) in the following form

$$\frac{\partial u}{\partial t} - \nabla \cdot (D \nabla u) = 0, \qquad \text{in } Q_T \equiv I \times \Omega, \tag{1.1}$$

$$u(x,0) = u_0(x), \quad \text{in } \Omega, \tag{1.2}$$

$$(D \nabla u) \cdot \mathbf{n} = 0, \qquad \text{on } I \times \partial\Omega, \tag{1.3}$$

where D is a matrix depending on the eigenvalues and eigenvectors of the so-called (regularized) structure tensor, $u_0 \in L^2(\Omega)$ and $\mathbf{n}$ is the outer normal unit vector to $\partial\Omega$. Such model is useful in any situation, where strong smoothing is desirable in a preferred direction and a low smoothing is expected in the perpendicular direction, e.g. for images with interrupted coherence of structures. To that goal the matrix

$$J_0(\nabla u_{\tilde{t}}) = \nabla u_{\tilde{t}} \otimes \nabla u_{\tilde{t}} = \nabla u_{\tilde{t}} \nabla u_{\tilde{t}}^{\,T}, \tag{1.4}$$

where

$$u_{\tilde{t}}(x,t) = (G_{\tilde{t}} * u(\cdot,t))(x), \qquad (\tilde{t} > 0). \tag{1.5}$$

is used. The matrix J_0 is symmetric and positive semidefinite and its eigenvectors are parallel and orthogonal to $\nabla u_{\tilde{t}}$, respectively. We can average J_0 by applying another convolution with Gaussian G_ρ and define

$$J_\rho(\nabla u_{\tilde{t}}) = G_\rho * (\nabla u_{\tilde{t}} \otimes \nabla u_{\tilde{t}}), \qquad (\rho > 0). \tag{1.6}$$

*Department of Mathematics, Slovak University of Technology, Radlinského 11, 813 68 Bratislava, Slovakia (e-mail: drblikov@math.sk, mikula@math.sk)

In computer vision the matrix $J_\rho = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$ is known as a structure tensor or interest operator or second moment matrix (see [9]). It is again symmetric and positive semidefinite and its eigenvalues are given by

$$\mu_{1,2} = \frac{1}{2}\left(a + c \pm \sqrt{(a-c)^2 + 4b^2}\right), \qquad \mu_1 \geq \mu_2. \tag{1.7}$$

Since the eigenvalues integrate the variation of the grey values within a neighbourhood of size $O(\rho)$, they describe the average contrast in the eigendirections v and w.

With the help of the eigenvalues of J_ρ we can obtain useful information on the coherence. The expression $(\mu_1 - \mu_2)^2$ is large for anisotropic structures and tends to zero for isotropic structures, constant areas are characterized by $\mu_1 = \mu_2 = 0$, straight edges by $\mu_1 \gg \mu_2 = 0$ and corners by $\mu_1 \geq \mu_2 \gg 0$.

The corresponding orthogonal set of eigenvectors (v, w) to eigenvalues (μ_1, μ_2) is given by

$$\begin{array}{lll} v = (v_1, v_2), & w = (w_1, w_2), & \\ v_1 = 2b, & v_2 = c - a + \sqrt{(a-c)^2 + 4b^2}, & \\ w \perp v, & w_1 = -v_2, & w_2 = v_1. \end{array} \tag{1.8}$$

The orientation of the eigenvector w, which corresponds to the smaller eigenvalue μ_2 is called coherence orientation. This orientation has the lowest fluctuations.

One can use the above mentioned structure tensor information into a construction of specific nonlinear diffusion filter [23, 24, 22]. The idea of the tensor nonlinear diffusion filtering is as follows. We get a processed version $u(x,t)$ of an original image $u_0(x)$ with a scale parameter $t \geq 0$ as the solution of mathematical model (1.1)-(1.3), where matrix D depends on solution u, satisfies smoothness, symmetry and uniform positive definiteness properties, and steers a filtering process such that diffusion is strong along the coherence direction w and increases with the coherence $(\mu_1 - \mu_2)^2$. To that goal D must possess the same eigenvectors v and w as the structure tensor $J_\rho(\nabla u_{\tilde{t}})$ and we choose the eigenvalues of D as

$$\begin{aligned} &\kappa_1 = \alpha, \quad \alpha \in (0,1),\ \alpha \ll 1, \\ &\kappa_2 = \begin{cases} \alpha, & \text{if } \mu_1 = \mu_2, \\ \alpha + (1-\alpha)\exp\left(\frac{-C}{(\mu_1-\mu_2)^2}\right),\ C > 0 & \text{else.} \end{cases} \end{aligned} \tag{1.9}$$

The matrix D then has following form

$$D = ABA^{-1}, \tag{1.10}$$

where $A = \begin{pmatrix} v_1 & -v_2 \\ v_2 & v_1 \end{pmatrix}$ and $B = \begin{pmatrix} \kappa_1 & 0 \\ 0 & \kappa_2 \end{pmatrix}$. The exponential function is used in (1.9) because it ensures that the smoothness of the structure tensor carries over to the diffusion tensor and that κ_2 does not exceed 1. The positive parameter α guarantees that the process never stops. Even if $(\mu_1 - \mu_2)^2$ tends to zero so the structure becomes isotropic, there still remains some small linear diffusion with diffusivity $\alpha > 0$. Such α is a regularization parameter, which keeps the diffusion tensor uniformly positive definite. C has a role of a threshold parameter. If $(\mu_1 - \mu_2)^2 \gg C$ then $\kappa_2 \approx 1$, and, in opposite if $(\mu_1 - \mu_2)^2 \ll C$ then $\kappa_2 \approx \alpha$. Due to the convolutions in (1.5) and (1.6), the elements of matrix D are C^∞ functions. Such model is a nontrivial extension of

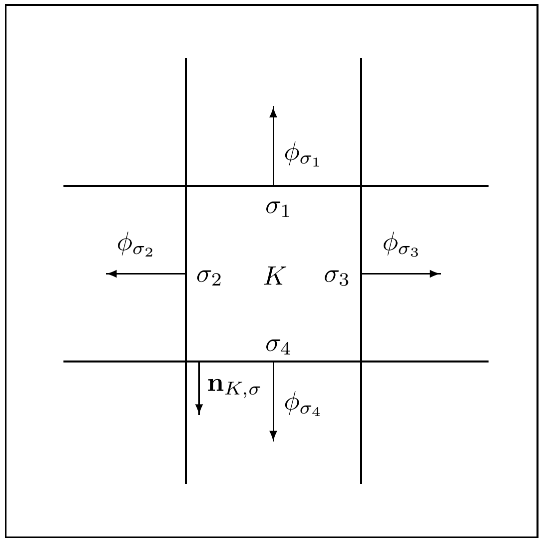

FIG. 2.1. *A detail of finite volume mesh - a finite volume K, its boundaries σ_i, $i = 1, 2, 3, 4$ and the fluxes outward to a finite volume K.*

the regularized Perona-Malik equation [17, 1, 15] and, as well as further PDEs employing tensor diffusion, it is used in many practical image processing applications, see e.g. [23, 24, 22, 6, 13, 19, 18]. In section 5 of this paper we also illustrate its usefulness by smoothing and segmenting the cell membrane images obtained by a confocal microscope. We show that after application of the nonlinear tensor anisotropic diffusion using our numerical scheme the coherent structures are attenuated. If such improved edge information is used in the so-called subjective surface segmentation method [20, 16, 2] the cell boundaries are correctly segmented.

There are only few purely finite volume methods designed and studied from numerical analysis point of view for solving tensor diffusion problems, see e.g. [3, 4] devoted to discretization of the elliptic operators. On the other hand, finite volume schemes for nonlinear parabolic problems as arising in image analysis are natural since they use piecewise constant representation of approximate solutions similarly to the structure of digital images. Finite and complementary volume schemes have been used successively in image processing for solving the Perona-Malik equation and its generalizations (see e.g. [15, 11, 12, 10, 7, 21]) and for solving the generalized curvature driven level set equations (see e.g. [8, 16, 2]) where the nonlinearities are represented by a scalar function dependent on solution gradient. Here we present the first finite volume scheme with convergence proof for the highly nonlinear anisotropic tensor diffusion model arising in coherence enhancing image smoothing.

The next section is devoted to derivation of our numerical scheme, in section 3 we study existence and uniqueness of discrete solutions, section 4 contains our convergence proof and finally, in section 5, we discuss numerical experiments.

2. Finite volume scheme for nonlinear tensor anisotropic diffusion. The aim of this section is to derive our computational method. Let the image be represented by $n_1 \times n_2$ pixels (finite volumes) such that it looks like a mesh with n_1 rows and n_2 columns. Let $\Omega = (0, n_1 h) \times (0, n_2 h)$, h is a pixel size and let the image $u(x)$ be given by a bounded mapping $u : \Omega \to R$. The filtering process is considered in a time interval $I = [0, T]$. Let $0 = t_0 \leq t_1 \leq \cdots \leq t_{N_{max}} = T$ denote the time discretization with $t_n = t_{n-1} + k$, where k is a length of discrete time step. In our scheme we will look for u^n an approximation of solution at time t_n, for every $n = 1, ..., N_{max}$. As usual in finite volume methods, we integrate equation (1.1) over finite volume K, then provide a semi-implicit time discretization and use a divergence theorem to get

$$\frac{u_K^n - u_K^{n-1}}{k} m(K) - \sum_{\sigma \in \mathcal{E}_K \cap \mathcal{E}_{int}} \int_\sigma (D^{n-1} \nabla u^n) \cdot \mathbf{n}_{K,\sigma} ds = 0, \tag{2.1}$$

where u_K^n, $K \in \mathcal{T}_h$, represents the mean value of u^n on K. $\mathcal{T}_h$ is an admissible finite volume mesh (see [4]) and further quantities and notations are described as follows: $m(K)$ is the measure of the finite volume K with boundary ∂K, $\sigma_{KL} = K \cap L = K|L$ is an edge of the finite volume K, where $L \in \mathcal{T}_h$ is an adjacent finite volume to K such that $m(K \cap L) \neq 0$. Due to simplifying notation, we use σ instead of σ_{KL} at several places if no confusion can appear. $\mathcal{E}_K$ is set of edges such that $\partial K = \bigcup_{\sigma \in \mathcal{E}_K} \sigma$ and $\mathcal{E} = \bigcup_{K \in \mathcal{T}_h} \mathcal{E}_K$. The set of boundary edges is denoted by $\mathcal{E}_{ext}$, that is $\mathcal{E}_{ext} = \mathcal{E} \cap \partial\Omega$ and let $\mathcal{E}_{int} = \mathcal{E} \setminus \mathcal{E}_{ext}$. Υ is the set of pairs of adjacent finite volumes, defined by $\Upsilon = \{(K, L) \in \mathcal{T}_h^2,\ K \neq L,\ m(K|L) \neq 0\}$ and $\mathbf{n}_{K,\sigma}$ is the normal unit vector to σ outward to K.

Let us define our discrete numerical solution by

$$u_{h,k}(x, t) = \sum_{n=0}^{N_{max}} \sum_{K \in \mathcal{T}_h} u_K^n \chi\{x \in K\} \chi\{t_{n-1} < t \leq t_n\}, \tag{2.2}$$

where the function $\chi(A)$ is defined as

$$\chi_{\{A\}} = \begin{cases} 1, & \text{if A is true,} \\ 0, & \text{elsewhere.} \end{cases} \tag{2.3}$$

The extension of the function (2.2) outside Ω is given first by its periodic mirror reflection in $\Omega_{\tilde{t}}$, where $\tilde{t}$ is the width of the smoothing kernel,

$$\Omega_{\tilde{t}} = \Omega \cup B_{\tilde{t}}(x), \quad x \in \partial\Omega, \tag{2.4}$$

$B_{\tilde{t}}(x)$ is a ball centered at x with radius $\tilde{t}$, and then we extend this periodic mirror reflection by 0 outside $\Omega_{\tilde{t}}$ and denote it by $\tilde{u}_{h,k}$.

In our scheme we will start computation by defining initial values

$$u_K^0 = \frac{1}{m(K)} \int_K u_0(x) dx,\ K \in \mathcal{T}_h \tag{2.5}$$

and let $u_{h,k}^n(x) = \sum\limits_{K \in \mathcal{T}_h} u_K^n \chi\{x \in K\}$ denote a finite volume approximation at the n-th time step. In order to get the scheme we write (2.1) in the form $\frac{u_K^n - u_K^{n-1}}{k} - \frac{1}{m(K)} \sum\limits_{\sigma \in \mathcal{E}_K \cap \mathcal{E}_{int}} \phi_\sigma^n(u_{h,k}^n) m(\sigma) = 0$, where $m(\sigma)$ is the measure of edge σ and $\phi_\sigma^n(u_{h,k}^n)$ denotes an approximation of the exact averaged flux $\frac{1}{m(\sigma)} \int_\sigma (D^{n-1} \nabla u^n) \cdot\ \mathbf{n}_{K,\sigma} ds$ for any K and $\sigma \in \mathcal{E}_K$.

We construct $\phi_\sigma^n(u_{h,k}^n)$ with the help of a co-volume mesh (see e.g. [3]). The co-volume χ_σ associated to σ is constructed around each edge by joining endpoints of this edge and midpoints of finite volumes which are common to this edge, see Fig.2.2.

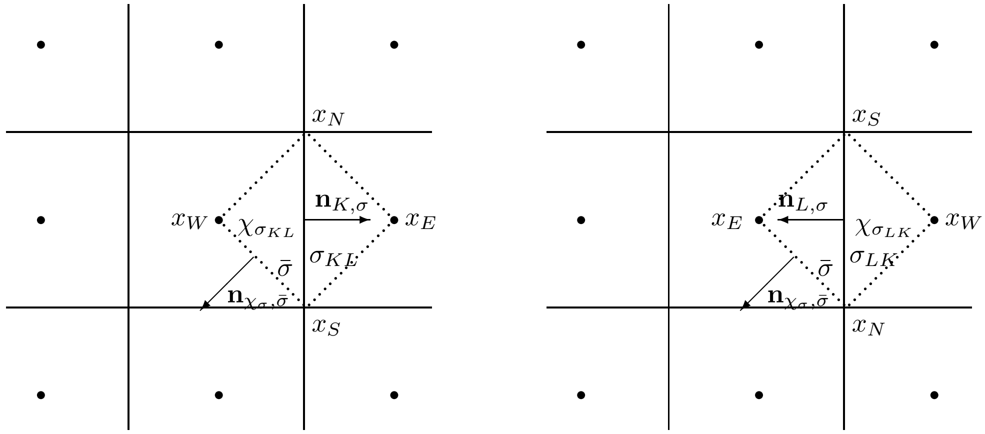

FIG. 2.2. *A detail of a mesh. The co-volumes associated to edges $\sigma = \sigma_{KL}$ (left) and $\sigma = \sigma_{LK}$ (right).*

We denote the endpoints of an edge $\bar\sigma \subset \partial\chi_\sigma$ by $N_1(\bar\sigma)$ and $N_2(\bar\sigma)$ and let $\mathbf{n}_{\chi_\sigma,\bar\sigma}$ be the normal unit vector to $\bar\sigma$ outward to χ_σ. In order to approximate diffusion flux, using divergence theorem, we first derive an approximation of the averaged gradient on χ_σ, namely $\frac{1}{m(\chi_\sigma)}\int\limits_{\chi_\sigma} \nabla u^n dx = \frac{1}{m(\chi_\sigma)}\int\limits_{\partial\chi_\sigma} u^n \mathbf{n}_{\chi_\sigma,\bar\sigma} ds$ and then we approximate it by $p^n_\sigma(u) = \frac{1}{m(\chi_\sigma)} \sum\limits_{\bar\sigma\in\partial\chi_\sigma} \frac{1}{2}\left(u^n_{N_1(\bar\sigma)} + u^n_{N_2(\bar\sigma)}\right) m(\bar\sigma)\mathbf{n}_{\chi_\sigma,\bar\sigma}$. Let the values at x_E and x_W be taken as u_E and u_W, and let the values u_S and u_N at the vertices x_N and x_S be computed as the arithmetic mean of u_K, where K are finite volumes which are common to this vertex.

Since our mesh is uniform and squared, we can use the following relations: $m(\chi_\sigma) = \frac{h^2}{2}$, $m(\bar\sigma) = \frac{\sqrt{2}}{2}h$ and after a short calculation we are ready to write

$$p^n_\sigma(u) = \frac{u^n_E - u^n_W}{h}\mathbf{n}_{K,\sigma} + \frac{u^n_N - u^n_S}{h}\mathbf{t}_{K,\sigma}, \tag{2.6}$$

where $\mathbf{t}_{K,\sigma}$ is a unit vector parallel to σ such that $(x_N - x_S)\cdot\mathbf{t}_{K,\sigma} > 0$. Although such u^n_N, u^n_W, u^n_E and u^n_S correspond to particular edge σ, and so we should denote them by $u^n_{N_\sigma}$, $u^n_{W_\sigma}$, $u^n_{E_\sigma}$ and $u^n_{S_\sigma}$ in (2.6), we will use the above simplified notations. Replacing the exact gradient ∇u^n by the numerical gradient $p^n_\sigma(u)$ in approximation of $\phi^n_\sigma(u^n_{h,k})$ we get the numerical flux in the form

$$\phi^n_\sigma(u^n_{h,k}) = (D_\sigma p^n_\sigma(u))\cdot\mathbf{n}_{K,\sigma}, \tag{2.7}$$

where $D_\sigma = D^{n-1}_\sigma = \begin{pmatrix} \bar\lambda_\sigma & \bar\beta_\sigma \\ \bar\beta_\sigma & \bar\nu_\sigma \end{pmatrix}$ is an approximation of the mean value of matrix D along σ evaluated at the previous time step. To that goal we take $u^{n-1}_{h,k}$ for constructing the structure and diffusion tensor and evaluate them at x_{KL}, where x_{KL} is a point of $\sigma_{KL} = K|L$ intersecting the segment $x_K x_L$. From implementation point of view, the structure and then diffusion tensor evaluation can be done in two ways. Either we can replace gradients of u appearing in structure tensor by their numerical approximation $p^n_\sigma(u)$ and then smooth them by weighted average (convolution), or we can evaluate $\nabla G_{\tilde{t}} * u^{n-1}_{h,k}$ using weights given by $\nabla G_{\tilde{t}}$ applied to discrete piecewise constant values of $u^{n-1}_{h,k}$ as convolution realization. In the latter way we do not introduce additional approximation into the scheme, and, in the part devoted to convergence analysis we use the latter approach, although both are realizable computationally.

It is important to note that in (2.7) we always consider the matrix D_σ written in the basis $(\mathbf{n}_{K,\sigma}, \mathbf{t}_{K,\sigma})$, cf. [3]. Although it may look artificial, it will simplify further considerations. In practice it means that, cf. Fig. 2.1, if the matrix D is given in standard basis on edge σ by $\begin{pmatrix} \lambda_\sigma & \beta_\sigma \\ \beta_\sigma & \nu_\sigma \end{pmatrix}$ then $D_\sigma = \begin{pmatrix} \lambda_\sigma & \beta_\sigma \\ \beta_\sigma & \nu_\sigma \end{pmatrix}$, i.e. $\bar{\lambda}_\sigma = \lambda_\sigma$, $\bar{\beta}_\sigma = \beta_\sigma$, $\bar{\nu}_\sigma = \nu_\sigma$ for the two edges $\sigma = \sigma_2$ and σ_3. On the other hand, $D_\sigma = \begin{pmatrix} \nu_\sigma & -\beta_\sigma \\ -\beta_\sigma & \lambda_\sigma \end{pmatrix}$, i.e. $\bar{\lambda}_\sigma = \nu_\sigma$, $\bar{\beta}_\sigma = -\beta_\sigma$, $\bar{\nu}_\sigma = \lambda_\sigma$ for other two edges $\sigma = \sigma_1$ and σ_4. Using such matrix representation the definition (2.7) can be written in this compact form

$$(2.8)\quad \phi^n_\sigma(u^n_{h,k}) = \left[\begin{pmatrix} \bar{\lambda}_\sigma & \bar{\beta}_\sigma \\ \bar{\beta}_\sigma & \bar{\nu}_\sigma \end{pmatrix} \begin{pmatrix} \frac{u^n_E - u^n_W}{h} \\ \frac{u^n_N - u^n_S}{h} \end{pmatrix}\right] \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \bar{\lambda}_\sigma \frac{u^n_E - u^n_W}{h} + \bar{\beta}_\sigma \frac{u^n_N - u^n_S}{h},$$

since in the basis $(\mathbf{n}_{K,\sigma}, \mathbf{t}_{K,\sigma})$ the formula (2.6) can be written for each edge as

$$(2.9)\qquad p^n_\sigma(u) = \begin{pmatrix} \frac{u^n_E - u^n_W}{h} \\ \frac{u^n_N - u^n_S}{h} \end{pmatrix}$$

and $\mathbf{n}_{K,\sigma}$ is equal to $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ in the basis $(\mathbf{n}_{K,\sigma}, \mathbf{t}_{K,\sigma})$ for each edge σ. Because of the convolutions in (1.5) and (1.6), the elements of matrix D_σ are C^∞ functions.

Finally, let us summarize our **semi-implicit finite volume scheme**:

$$(2.10)\qquad \frac{u^n_K - u^{n-1}_K}{k} - \frac{1}{m(K)} \sum_{\sigma \in \mathcal{E}_K \cap \mathcal{E}_{int}} \phi^n_\sigma(u^n_{h,k}) m(\sigma) = 0,$$

where

$$(2.11)\qquad \phi^n_\sigma(u^n_{h,k}) = \bar{\lambda}_\sigma \frac{u^n_E - u^n_W}{h} + \bar{\beta}_\sigma \frac{u^n_N - u^n_S}{h}.$$

3. Existence and uniqueness of the solution to discrete scheme. In order to fulfill main goal of this section, to prove existence and uniqueness of u^n_K, $K \in \mathcal{T}_h$, we estimate the expressions $u^n_N - u^n_S$ by means of $u^n_E - u^n_W$ for all edges σ. To that goal we use mainly results of [3] in our situation. Let us note that, due to simplification of notation, we do not use upper index n in the sequel and at some places we relate u_E and u_W to particular edge σ using u_{E_σ}, u_{W_σ}, etc. In the sequel we denote by C_i constants which may depend on the properties of diffusion tensor.

Definition 3.1. *Let P_σ be the set of all edges δ perpendicular to σ (see Fig. 3.1 for two illustrative situations when $\sigma = \sigma_{WE}$ and $\sigma = \sigma_{EW}$), which have common vertex with σ and fulfill the following conditions: $x_{E_\delta} - x_{W_\delta} > 0$ if $x_{N_\sigma} - x_{S_\sigma} > 0$ and $x_{E_\delta} - x_{W_\delta} < 0$ if $x_{N_\sigma} - x_{S_\sigma} < 0$. Let us note that $x_{W_\sigma} = x^1_{W_\delta} = x^3_{E_\delta}$, for $\sigma = \sigma_{WE}$, $x_{E_\sigma} = x^2_{W_\delta} = x^4_{E_\delta}$, for $\sigma = \sigma_{WE}$, $x_{W_\sigma} = x^2_{E_\delta} = x^4_{W_\delta}$, for $\sigma = \sigma_{EW}$ and $x_{E_\sigma} = x^1_{E_\delta} = x^3_{W_\delta}$, for $\sigma = \sigma_{EW}$.*

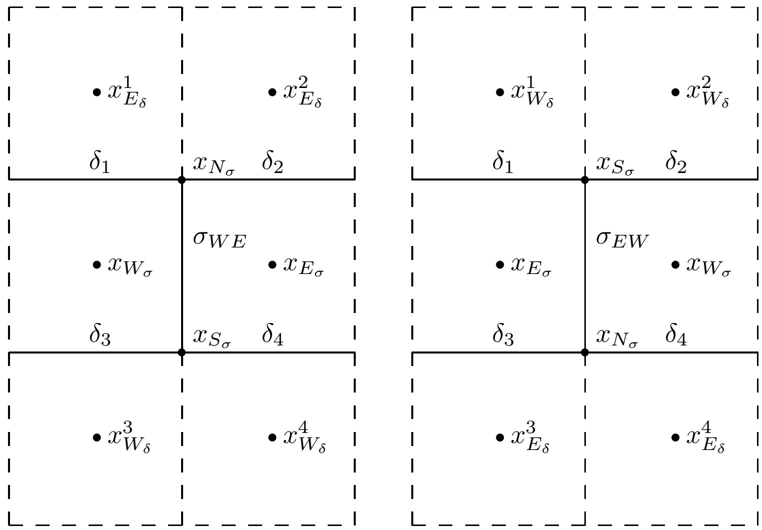

FIG. 3.1. *Left: an edge σ_{WE} and edges δ_1, δ_2, δ_3, $\delta_4 \in P_{\sigma_{WE}}$. Right: an edge σ_{EW} and edges δ_1, δ_2, δ_3, $\delta_4 \in P_{\sigma_{EW}}$.*

Using definitions given in the previous section we can write

$$u_N - u_S = \frac{1}{4}\left[u_E^1 - u_W^1) + (u_E^3 - u_W^3) + (u_E^2 - u_W^2) + (u_E^4 - u_W^4)\right], \tag{3.1}$$

where $u_E^1 = u_{E_{\delta_1}}$ and $u_W^1 = u_{W_{\delta_1}}$ correspond to edge δ_1 and similarly u_E^2, u_W^2, u_E^3, u_W^3, u_E^4 and u_W^4 correspond to edges δ_2, δ_3 and δ_4. Applying the inequality $(a+b)^2 \le 2a^2 + 2b^2$ we have

$$(u_{N_\sigma} - u_{S_\sigma})^2 \le \sum_{\delta \in P_\sigma \cap \mathcal{E}_{int}} \frac{1}{4}(u_{E_\delta} - u_{W_\delta})^2. \tag{3.2}$$

Multiplying (3.2) by $\left(\frac{\bar\beta_\sigma}{\bar\lambda_\sigma}\right)^2 \frac{\bar\lambda_\sigma}{h^2}$ and summing for all $\sigma \in \mathcal{E}_{int}$ (by σ we mean σ_{WE}) we obtain

$$\sum_{\sigma\in\mathcal{E}_{int}} \left(\frac{\bar\beta_\sigma}{\bar\lambda_\sigma}\right)^2 \left(\frac{u_{N_\sigma} - u_{S_\sigma}}{h}\right)^2 \bar\lambda_\sigma \le \sum_{\sigma\in\mathcal{E}_{int}} \left(\frac{\bar\beta_\sigma}{\bar\lambda_\sigma}\right)^2 \sum_{\delta\in P_\sigma\cap\mathcal{E}_{int}} \frac{1}{4}\left(\frac{u_{E_\delta} - u_{W_\delta}}{h}\right)^2 \bar\lambda_\sigma. \tag{3.3}$$

Then we swap the two sums on the right hand side of (3.3) to get

$$\sum_{\sigma\in\mathcal{E}_{int}} \left(\frac{\bar\beta_\sigma}{\bar\lambda_\sigma}\right)^2 \left(\frac{u_{N_\sigma} - u_{S_\sigma}}{h}\right)^2 \bar\lambda_\sigma \le \sum_{\delta\in\mathcal{E}_{int}} \gamma_\delta \left(\frac{u_{E_\delta} - u_{W_\delta}}{h}\right)^2 \bar\lambda_\delta \tag{3.4}$$

where

$$\gamma_\delta = \sum_{\sigma\in P_\delta\cap\mathcal{E}_{int}} \frac{1}{4}\left(\frac{\bar\beta_\sigma}{\bar\lambda_\sigma}\right)^2 \frac{\bar\lambda_\sigma}{\bar\lambda_\delta}. \tag{3.5}$$

Let us consider the matrix $\begin{pmatrix} \bar\lambda_{\sigma^\perp} & \bar\beta_{\sigma^\perp} \\ \bar\beta_{\sigma^\perp} & \bar\nu_{\sigma^\perp} \end{pmatrix}$, which is the matrix D written in the basis $(\mathbf{t}_{K,\sigma}, -\mathbf{n}_{K,\sigma})$ on edge σ. Due to smoothness of D we get

$$\bar\lambda_\sigma = \bar\nu_{\sigma^\perp} = \bar\nu_\delta(1 + O(h)) = \bar\lambda_{\delta^\perp}(1 + O(h)), \quad \delta \in P_\sigma, \tag{3.6}$$

$$\bar{\beta}_\sigma = -\bar{\beta}_{\sigma^\perp} = -\bar{\beta}_\delta(1+O(h)) = \bar{\beta}_{\delta^\perp}(1+O(h)), \quad \delta \in P_\sigma, \tag{3.7}$$

$$\bar{\nu}_\sigma = \bar{\lambda}_{\sigma^\perp} = \bar{\lambda}_\delta(1+O(h)) = \bar{\nu}_{\delta^\perp}(1+O(h)), \quad \delta \in P_\sigma. \tag{3.8}$$

Applying (3.6)-(3.8) in (3.5) and using $\frac{1}{1-Ch} \leq 1 + (C + 2C^2 h)h$ for h sufficiently small ($h \leq \frac{1}{2C}$) we have

$$\gamma_\delta \leq \sum_{\sigma \in P_\delta \cap \mathcal{E}_{int}} \frac{1}{4} \left(\frac{\bar{\beta}_{\delta^\perp}}{\bar{\lambda}_{\delta^\perp}} \right)^2 \frac{\bar{\lambda}_{\delta^\perp}}{\bar{\lambda}_\delta} (1+O(h)) = \left(\frac{\bar{\beta}_{\delta^\perp}}{\bar{\lambda}_{\delta^\perp}} \right)^2 \frac{\bar{\lambda}_{\delta^\perp}}{\bar{\lambda}_\delta} (1+O(h)) \, .$$

Using the positive definiteness of the diffusion tensor written in a standard basis as $\begin{pmatrix} \lambda_\delta & \beta_\delta \\ \beta_\delta & \nu_\delta \end{pmatrix}$ we obtain for its determinant

$$\lambda_\delta \nu_\delta - \beta_\delta^2 > 0. \tag{3.9}$$

Now, we have two possibilities for γ_δ. Let δ be an arbitrary edge in the mesh parallel to σ_3 (see Fig. 2.1). Then $\gamma_\delta \leq \left(\frac{-\beta_\delta}{\nu_\delta} \right)^2 \frac{\nu_\delta}{\lambda_\delta}(1+O(h)) = \frac{(\beta_\delta)^2}{\lambda_\delta \nu_\delta}(1+O(h)) < 1$ for h sufficiently small due to (3.9). Similarly, if δ is any edge oriented perpendicularly to σ_3 we have $\gamma_\delta \leq \left(\frac{\beta_\delta}{\lambda_\delta} \right)^2 \frac{\lambda_\delta}{\nu_\delta}(1+O(h)) = \frac{(\beta_\delta)^2}{\lambda_\delta \nu_\delta}(1+O(h)) < 1$ for h sufficiently small. Thus, due to the fact that $\lambda_\sigma \geq C > 0$ and $\nu_\sigma \geq C > 0$, we obtain $0 \leq \gamma_\delta < 1$ for h sufficiently small. Since this condition is fulfilled for each edge δ we can rewrite (3.4) as

$$\sum_{\sigma \in \mathcal{E}_{int}} \left(\frac{\bar{\beta}_\sigma}{\bar{\lambda}_\sigma} \right)^2 \left(\frac{u_N - u_S}{h} \right)^2 \bar{\lambda}_\sigma \leq \gamma \sum_{\sigma \in \mathcal{E}_{int}} \left(\frac{u_E - u_W}{h} \right)^2 \bar{\lambda}_\sigma, \tag{3.10}$$

where $0 \leq \gamma < 1, \quad \gamma = \max_{\sigma \in \mathcal{E}} \gamma_\sigma$.

Let us now introduce the space of piecewise constant functions associated to our mesh and discrete H^1 norm for this space. This discrete norm will be used to obtain some estimates on the approximate solution given by the finite volume scheme.

Definition 3.2. *Let Ω be an open bounded polygonal subset of R^2. Let $\mathcal{T}_h$ be an admissible finite volume mesh (see [4]). We define $\mathcal{P}_0(\mathcal{T}_h)$ as the set of functions from Ω to R which are constant over each finite volume K of the mesh $\mathcal{T}_h$.*

Definition 3.3. *Let Ω be an open bounded polygonal subset of R^2. For $u \in \mathcal{P}_0(\mathcal{T}_h)$ we define*

$$|u^n_{h,k}|_{1,\mathcal{T}_h} = \left(\sum_{(K,L) \in \Upsilon} \frac{(u_L - u_K)^2}{d_{K,L}} m(\sigma) \right)^{\frac{1}{2}}, \tag{3.11}$$

where $d_{K,L}$ is the Euclidean distance between x_K and x_L.

Remark that (3.11) can be rewritten for our uniform mesh into the following form

$$|u^n_{h,k}|_{1,\mathcal{T}_h} = \left(2 \sum_{\sigma \in \mathcal{E}_{int}} \left(\frac{u_E - u_W}{h} \right)^2 m(\chi_\sigma) \right)^{\frac{1}{2}}, \tag{3.12}$$

where $\sigma = \sigma_{WE}$. Let us define a discrete operator $\mathcal{L}_h$ by

$$\mathcal{L}_h(u^n_{h,k}) = u^n_K m(K) - k \sum_{\sigma \in \mathcal{E}_K \cap \mathcal{E}_{int}} \phi^n_\sigma(u^n_{h,k}) m(\sigma).$$

Then solution $u_{h,k}^n \in \mathcal{P}_0(\mathcal{T}_h)$ of our scheme at time t_n is given by

$$\mathcal{L}_h(u_{h,k}^n) = f_{h,k}(u_{h,k}^{n-1}), \tag{3.13}$$

where $f_{h,k}(u_{h,k}^{n-1}) = u_K^{n-1} m(K)$, $K \in \mathcal{T}_h$, and u_K^{n-1} is value of the piecewise constant function $u_{h,k}^{n-1}$ in K. This equality is a linear system of N equations with N unknowns u_K^n, $K \in \mathcal{T}_h$, $N = card(\mathcal{T}_h)$.

Multiplying $\mathcal{L}_h(u_{h,k})$ by u_K^n, summing over K and splitting into a part A and B leads to

$$\sum_{K \in \mathcal{T}_h} \mathcal{L}_h(u_{h,k}^n) u_K^n = A + B, \tag{3.14}$$

with

$$A = \sum_{K \in \mathcal{T}_h} (u_K^n)^2 m(K) = ||u_{h,k}^n||^2_{L^2(\Omega)} \tag{3.15}$$

and

$$B = k \sum_{K \in \mathcal{T}_h} u_K^n \sum_{\sigma \in \mathcal{E}_K \cap \mathcal{E}_{int}} - \phi_\sigma^n(u_{h,k}^n) m(\sigma).$$

The above expression can be written in the following form

$$\begin{aligned} B &= k \sum_{W \in \mathcal{T}_h} u_W^n \sum_{\sigma \in \mathcal{E}_W \cap \mathcal{E}_{int}} -\phi_\sigma^n(u_{h,k}^n) m(\sigma) \\ &= \frac{k}{2} \sum_{\sigma \in \mathcal{E}_{int}} \phi_\sigma^n(u_{h,k}^n) \frac{u_E - u_W}{h} 2m(\chi_\sigma) = Q(u_{h,k}^n) \end{aligned} \tag{3.16}$$

owing to property $\phi_\sigma^n(u_{h,k}^n) = \phi_{\sigma_{WE}}^n(u_{h,k}^n) = -\phi_{\sigma_{EW}}^n(u_{h,k}^n)$. Since $\phi_\sigma^n(u_{h,k}^n) = 0$ for $\sigma \in \mathcal{E}_{ext}$ we can extend the sum in (3.16) and write

$$Q(u_{h,k}^n) = \frac{k}{2} \sum_{\sigma \in \mathcal{E}} (D_\sigma p_\sigma^\star) \cdot p_\sigma 2m(\chi_\sigma) = k(D_h p_h^\star, p_h)_{L^2(\Omega)}.$$

where $p_\sigma^\star = \frac{u_E - u_W}{h} \mathbf{n}_{W,\sigma}$ for $\sigma = \sigma_{WE}$ is the normal component of the gradient and $D_h, p_h, p_h^\star$ are piecewise constant functions with values extended from σ to χ_σ. Further, we use the following inequality

$$(D_h p_h^\star, p_h)_{L^2(\Omega)} \geq (D_h p_h^\star, p_h^\star)_{L^2(\Omega)} - |(D_h p_h^\star, p_h - p_h^\star)_{L^2(\Omega)}|. \tag{3.17}$$

It is clear that $(D_h p_h^\star, p_h^\star)_{L^2(\Omega)} = \sum_{\sigma \in \mathcal{E}} \bar{\lambda}_\sigma \left(\frac{u_E - u_W}{h}\right)^2 m(\chi_\sigma)$, due to fact that $u_E - u_W = 0$ for $\sigma \in \mathcal{E}_{ext}$ thanks to reflexion of $u_{h,k}$ in $\Omega_{\tilde{t}}$ (see page 4). Applying Young's inequality in the second term on the right hand side of (3.17) leads to

$$\begin{aligned} |(D_h p_h^\star, p_h - p_h^\star)_{L^2(\Omega)}| &= \left| \sum_{\sigma \in \mathcal{E}} \bar{\beta}_\sigma \frac{u_E - u_W}{h} \frac{u_N - u_S}{h} m(\chi_\sigma) \right| \\ &\leq \sum_{\sigma \in \mathcal{E}_{int}} \frac{1}{2} \left[\left(\frac{u_E - u_W}{h} \right)^2 + \left(\frac{\bar{\beta}_\sigma}{\bar{\lambda}_\sigma} \right)^2 \left(\frac{u_N - u_S}{h} \right)^2 \right] \bar{\lambda}_\sigma m(\chi_\sigma), \end{aligned} \tag{3.18}$$

since $\phi_\sigma^n(u_{h,k}^n) = 0$ for $\sigma \in \mathcal{E}_{ext}$. Using inequalities (3.10) we get

$$|(D_h p_h^\star, p_h - p_h^\star)_{L^2(\Omega)}| \le \frac{1+\gamma}{2} \sum_{\sigma\in\mathcal{E}_{int}} \bar{\lambda}_\sigma \left(\frac{u_E - u_W}{h}\right)^2 m(\chi_\sigma) =$$

$$= \frac{1+\gamma}{2} (D_h p_h^\star, p_h^\star)_{L^2(\Omega)}. \tag{3.19}$$

Using (3.12), it in turn implies

$$Q(u_{h,k}^n) \ge \left(1 - \frac{1+\gamma}{2}\right) k (D p_h^\star, p_h^\star)_{L^2(\Omega)} \ge \bar{\lambda}_{\min} \frac{1-\gamma}{2} \frac{k}{2} |u_{h,k}^n|_{1,\mathcal{T}_h}^2, \tag{3.20}$$

where $\bar{\lambda}_{\min} = \inf_{\sigma\in\mathcal{E}} \bar{\lambda}_\sigma \ge C > 0$. Applying (3.15), (3.16) and (3.20) in (3.14) we get for h sufficiently small and any $u_{h,k}^n \in \mathcal{P}_0(\mathcal{T}_h)$ that

$$\sum_{K\in\mathcal{T}_h} \mathcal{L}_h(u_{h,k}^n) u_K^n \ge \alpha \left(|u_{h,k}^n|_{1,\mathcal{T}_h}^2 + ||u_{h,k}^n||_{L^2(\Omega)}^2 \right)$$

with $\alpha = \min\left(\bar{\lambda}_{\min}(1-\gamma)\frac{k}{4}, 1\right)$.

THEOREM 3.4. *For h sufficiently small, there exists unique solution $u_{h,k}^n$ given by the scheme (2.10)-(2.11) at any discrete time step t_n.*

Proof. Assume that u_K, $K \in \mathcal{T}_h$ satisfy the linear system (3.13) and let the right hand side of (3.13) be equal to 0. Then

$$\alpha \left(|u_{h,k}^n|_{1,\mathcal{T}_h}^2 + ||u_{h,k}^n||_{L^2(\Omega)}^2 \right) \le \sum_{K\in\mathcal{T}_h} \mathcal{L}_h(u_{h,k}^n) u_K^n = \sum_{K\in\mathcal{T}_h} f_{h,k}(u_{h,k}^{n-1}) u_K^n = 0. \tag{3.21}$$

Due to relation (3.21) and strict positivity of α we know that $u_K^n = 0$, $\forall K \in \mathcal{T}_h$. It means that kernel of the linear transformation represented by the matrix of the system (3.13) contains only $\bar{0}$ vector, which implies that the matrix is regular. And thus it implies that there exists unique solution for any right hand side. □

4. Convergence of the scheme to the weak solution.

DEFINITION 4.1. *Weak solution of the problem (1.1)-(1.3) is a function $u \in L^2(0,T;H^1(\Omega))$ which satisfies the identity*

$$\int_0^T\!\!\int_\Omega u \frac{\partial \varphi}{\partial t}(x,t)\, dxdt + \int_\Omega u_0(x)\, \varphi(x,0)\, dx - \int_0^T\!\!\int_\Omega (D\nabla u)\cdot \nabla\varphi dxdt = 0, \ \forall \varphi \in \Psi \tag{4.1}$$

where $\Psi = \left\{ \varphi \in C^{2,1}\left(\overline{\Omega}\times[0,T]\right), (D\nabla\varphi)\cdot \overrightarrow{n} = 0 \text{ on } \partial\Omega\times(0,T), \varphi(.,T) = 0 \right\}$.

REMARK 1. *Existence and uniqueness of the weak solution and extremum principle for the model (1.1)-(1.3) are given in [24]. The proofs are based on theory built in [1].*

In the proof of convergence we will use strategy based on application of the Kolmogorov's compactness criterion in L^2 which gives relative compactness of the approximate solutions given by the scheme refining the space and time discretization step. Using relative compactness we can choose convergent subsequence which in the limit gives the weak solution. In order to use Kolmogorov's compactness criterion we shall prove following four lemmata.

LEMMA 4.2. **(Uniform boundedness)** *There exists a positive constant C such that*

$$\left\| u_{h,k} \right\|_{L^2(Q_T)} \leq C. \tag{4.2}$$

LEMMA 4.3. **(Time translate estimate)** *For any $s \in (0,T)$ there exists a positive constant C such that*

$$\int\limits_{\Omega \times (0,T-s)} \left(u_{h,k}\left(x,t+s\right) - u_{h,k}\left(x,t\right) \right)^2 dxdt \leq Cs. \tag{4.3}$$

LEMMA 4.4. **(Space translate estimate I)** *There exists a positive constant C such that*

$$\int\limits_{\Omega_\xi \times (0,T)} \left(u_{h,k}\left(x+\xi,t\right) - u_{h,k}\left(x,t\right) \right)^2 dxdt \leq C \left|\xi\right| \left(\left|\xi\right| + 2h\right) \tag{4.4}$$

for any vector $\xi \in R^d$, where $\Omega_\xi = \{x \in \Omega, [x, x+\xi] \in \Omega\}$.

LEMMA 4.5. **(Space translate estimate II)** *There exists a positive constant C such that*

$$\int\limits_{\Omega \times (0,T)} \left(u_{h,k}\left(x+\xi,t\right) - u_{h,k}\left(x,t\right) \right)^2 dxdt \leq C \left|\xi\right|, \tag{4.5}$$

for any vector $\xi \in R^d$.

To prove (4.2)-(4.5) we will use following a-priori estimates.

LEMMA 4.6. *The scheme (2.10)-(2.11) leads to the following estimates. For h sufficiently small, there exists a positive constant C which does not depend on h, k such that*

$$\max_{0 \leq n \leq N_{\max}} \sum_{K \in \mathcal{T}_h} \left(u_K^n\right)^2 m(K) \leq C, \tag{4.6}$$

$$\sum_{n=1}^{N_{\max}} k \sum_{(K,L) \in \Upsilon} \frac{\left(u_K^n - u_L^n\right)^2}{d_{K,L}} m\left(\sigma\right) \leq C, \tag{4.7}$$

$$\sum_{n=1}^{N_{\max}} \sum_{K \in \mathcal{T}_h} \left(u_K^n - u_K^{n-1}\right)^2 m\left(K\right) \leq C. \tag{4.8}$$

Proof. We multiply (2.10) by u_K^n, sum it over $K \in \mathcal{T}_h$, over $n = 1, ..., m < N_{max}$, and use the property $(a-b)a = \frac{1}{2}a^2 - \frac{1}{2}b^2 + \frac{1}{2}(a-b)^2$ to obtain

$$\begin{aligned} &\frac{1}{2} \sum_{K \in \mathcal{T}_h} (u_K^m)^2 m(K) + \frac{1}{2} \sum_{n=1}^{m} \sum_{K \in \mathcal{T}_h} (u_K^n - u_K^{n-1})^2 m(K) \\ &- \sum_{n=1}^{m} k \sum_{K \in \mathcal{T}_h} u_K^n \sum_{\sigma \in \mathcal{E}_K \cap \mathcal{E}_{int}} \phi_\sigma^n(u_{h,k}^n) m(\sigma) = \frac{1}{2} \sum_{K \in \mathcal{T}_h} (u_K^0)^2 m(K). \end{aligned} \tag{4.9}$$

Then using (3.16) and (3.20) we have

$$\begin{aligned}\frac{1}{2}\sum_{K\in\mathcal{T}_h}(u_K^m)^2 m(K)+\frac{1}{2}\sum_{n=1}^{m}\sum_{K\in\mathcal{T}_h}(u_K^n-u_K^{n-1})^2 m(K)\\ +\bar{\alpha}\sum_{n=1}^{m}k|u_{h,k}^n|_{1,\mathcal{T}_h}^2\le\frac{1}{2}\sum_{K\in\mathcal{T}_h}(u_K^0)^2 m(K)\end{aligned}\tag{4.10}$$

with positive constant $\bar{\alpha}=\bar{\lambda}_{\min}\frac{1-\gamma}{4}$. Since $u_0\in L^2(\Omega)$, the right hand side is bounded by a positive constant C. Using the first term of (4.10) we get a-priori estimate (4.6) and from the second term of (4.10) we get a-priori estimate (4.8). From the strict positiveness of $\bar{\alpha}$ in the third term of (4.10) and from definition (3.11) we obtain a-priori estimate (4.7). □

Proof. (of Lemma 4.2) It follows from the first $L^2(\Omega)$ - a priori estimate (4.6). □

Proof. (of Lemma 4.3) First, for fixed $s\in(0,T)$, we define function

$$f(t)=\int_\Omega\left(u_{h,k}(x,t+s)-u_{h,k}(x,t)\right)^2dx.$$

Using the fact that $u_{h,k}$ is a piecewise constant function, we get

$$f(t)=\sum_{K\in\mathcal{T}_h}\left(u_K^{n_{t+s}}-u_K^{n_t}\right)^2 m(K),\tag{4.11}$$

with $n_t=\lceil\frac{t}{k}\rceil$ and $n_{t+s}=\lceil\frac{t+s}{k}\rceil$, where $\lceil\cdot\rceil$ means the upper integer part of positive real number. We rearrange (4.11) to obtain

$$f(t)=\sum_{K\in\mathcal{T}_h}\left(u_K^{n_{t+s}}-u_K^{n_t}\right)\sum_{t\le(n+1)k<t+s}\left(u_K^n-u_K^{n-1}\right)m(K).\tag{4.12}$$

Using the scheme (2.10)-(2.11) in (4.12) (replacing K by W) we get

$$f(t)=\sum_{t\le(n+1)k<t+s}k\sum_{W\in\mathcal{T}_h}\left(\left(u_W^{n_{t+s}}-u_W^{n_t}\right)\sum_{\sigma\in\mathcal{E}_W\cap\mathcal{E}_{int}}\bar{\lambda}_\sigma(u_E^n-u_W^n)+\bar{\beta}_\sigma(u_N^n-u_S^n)\right),\tag{4.13}$$

and due to conservativity of numerical fluxes (antisymmetry of term $\bar{\lambda}_\sigma(u_E^n-u_W^n)+\bar{\beta}_\sigma(u_N^n-u_S^n)$) we have

$$\begin{aligned}f(t)=\sum_{t\le(n+1)k<t+s}\frac{k}{2}\sum_{\sigma\in\mathcal{E}_{int}}\left(u_W^{n_{t+s}}-u_W^{n_t}-u_E^{n_{t+s}}+u_E^{n_t}\right)\\ \left(\bar{\lambda}_\sigma(u_E^n-u_W^n)+\bar{\beta}_\sigma(u_N^n-u_S^n)\right).\end{aligned}\tag{4.14}$$

Using Young's inequality leads to the relation

$$\begin{aligned}f(t)\le\sum_{t\le(n+1)k<t+s}\frac{k}{4}\sum_{\sigma\in\mathcal{E}_{int}}\bar{\lambda}_\sigma(u_W^{n_{t+s}}-u_W^{n_t}-u_E^{n_{t+s}}+u_E^{n_t})^2+\\ +\sum_{t\le(n+1)k<t+s}\frac{k}{4}\sum_{\sigma\in\mathcal{E}_{int}}\bar{\lambda}_\sigma\left((u_E^n-u_W^n)+\frac{\bar{\beta}_\sigma}{\bar{\lambda}_\sigma}(u_N^n-u_S^n)\right)^2\end{aligned}\tag{4.15}$$

where the right hand side can be further estimated and we get

$$f(t) \le f_1(t) + f_2(t) + f_3(t) + f_4(t), \tag{4.16}$$

$$f_1(t) = \sum_{t \le (n+1)k < t+s} \frac{k}{2} \sum_{\sigma \in \mathcal{E}_{int}} \bar{\lambda}_\sigma \left(u_E^{n_t} - u_W^{n_t}\right)^2, \tag{4.17}$$

$$f_2(t) = \sum_{t \le (n+1)k < t+s} \frac{k}{2} \sum_{\sigma \in \mathcal{E}_{int}} \bar{\lambda}_\sigma \left(u_E^{n_{t+s}} - u_W^{n_{t+s}}\right)^2, \tag{4.18}$$

$$f_3(t) = \sum_{t \le (n+1)k < t+s} \frac{k}{2} \sum_{\sigma \in \mathcal{E}_{int}} \bar{\lambda}_\sigma (u_E^n - u_W^n)^2, \tag{4.19}$$

$$f_4(t) = \sum_{t \le (n+1)k < t+s} \frac{k}{2} \sum_{\sigma \in \mathcal{E}_{int}} \bar{\lambda}_\sigma \left(\frac{\bar{\beta}_\sigma}{\bar{\lambda}_\sigma}\right)^2 (u_N^n - u_S^n)^2. \tag{4.20}$$

Next we integrate (4.16) in time interval $(0, T-s)$, and replacing $\sum\limits_{\sigma \in \mathcal{E}_{int}}$ by $\sum\limits_{(K,L) \in \Upsilon}$ (the edge $\sigma \in \mathcal{E}_K$ is an intersection of K and its adjacent finite volume L) we get an estimate of the first integral term

$$\int_0^{T-s} f_1(t)dt = \int_0^{T-s} \frac{k}{2} \sum_{(K,L) \in \Upsilon} \bar{\lambda}_\sigma \left(u_L^{n_t} - u_K^{n_t}\right)^2 \sum_{n \in \mathbf{N}} \chi_{\{t \le (n+1)k < t+s\}} dt\,. \tag{4.21}$$

We substitute the integral over $(0, T-s)$ by the sum of time step intervals and use the property $\chi_{\{t \le (n+1)k < t+s\}} = \chi_{\{(n+1)k - s < t \le (n+1)k\}}$ to obtain

$$\int_0^{T-s} f_1(t)dt \le \sum_{n_t=0}^{N_{\max}-1} \frac{k}{2} \sum_{(K,L) \in \Upsilon} \bar{\lambda}_\sigma \left(u_L^{n_t} - u_K^{n_t}\right)^2 \int_{n_t k}^{(n_t+1)k} \sum_{n \in \mathbf{N^0}} \chi_{\{(n+1)k - s < t \le (n+1)k\}} dt. \tag{4.22}$$

Since $\int\limits_{n_t k}^{(n_t+1)k} \sum\limits_{n \in \mathbf{N^0}} \chi_{\{(n+1)k - s < t \le (n+1)k\}} dt = s$, and $m(\sigma) = d_{K,L}$ for our uniform mesh, the relation (4.22) leads to

$$\int_0^{T-s} f_1(t)dt \le s \sum_{n_t=0}^{N_{\max}} \frac{k}{2} \sum_{(K,L) \in \Upsilon} \frac{m(\sigma)}{d_{K,L}} \bar{\lambda}_\sigma \left(u_L^{n_t} - u_K^{n_t}\right)^2. \tag{4.23}$$

Next step is to prove the following relation

$$0 < C_1 \le \bar{\lambda}_\sigma \le C_2 < \infty, \text{ for all } \sigma \in \mathcal{E}. \tag{4.24}$$

Let K be any fixed finite volume. Since at any time step the matrix $D_\sigma = \begin{pmatrix} \bar{\lambda}_\sigma & \bar{\beta}_\sigma \\ \bar{\beta}_\sigma & \bar{\nu}_\sigma \end{pmatrix}$ is uniformly (strictly) positive definite, $\bar{\lambda}_\sigma \ge C_3 > 0$ and $\bar{\nu}_\sigma \ge C_4 > 0$ for all σ. The structure tensor evaluated numerically at point x_{KL} is given by

$$J_\rho \left(\nabla u_{h,k\ \tilde{t}}^{n-1}\right)(x_{KL}) = G_\rho * \begin{pmatrix} A & B \\ B & C \end{pmatrix}, \tag{4.25}$$

where

$$A=\left(\left(\frac{\partial G_{\tilde{t}}}{\partial x}*\tilde{u}_{h,k}^{n-1}\right)(x_{KL})\right)^2,\quad C=\left(\left(\frac{\partial G_{\tilde{t}}}{\partial y}*\tilde{u}_{h,k}^{n-1}\right)(x_{KL})\right)^2, \tag{4.26}$$

$$B=\left(\frac{\partial G_{\tilde{t}}}{\partial x}*\tilde{u}_{h,k}^{n-1}\right)(x_{KL})\left(\frac{\partial G_{\tilde{t}}}{\partial y}*\tilde{u}_{h,k}^{n-1}\right)(x_{KL}). \tag{4.27}$$

Using Young's inequality, a-priori estimate (4.6) and definition of extension $\tilde{u}_{h,k}^n$ (see (2.4)) we subsequently get for $i=1,2$ ($x_1=x, x_2=y$)

$$\begin{aligned}\left|\left(\frac{\partial}{\partial x_i}G_{\tilde{t}}*\tilde{u}_{h,k}^n\right)(x_{KL})\right| &\le \int\limits_{R^d}\left|\frac{\partial}{\partial x_i}G_{\tilde{t}}(x_{KL}-\xi)\tilde{u}_{h,k}^n(\xi)\right|d\xi \le \\ &\le \frac{1}{2}\int\limits_{R^d}\left|\frac{\partial}{\partial x_i}G_{\tilde{t}}(x_{KL}-\xi)\right|^2 d\xi+\frac{1}{2}\int\limits_{R^d}|\tilde{u}_{h,k}^n(\xi)|^2 d\xi \le C_{\tilde{t}}+ \\ &+C_5\int\limits_{\Omega_{\tilde{t}}}|\tilde{u}_{h,k}^n(\xi)|^2 d\xi \le C_{\tilde{t}}+C_5\sum_{K\in\mathcal{T}_h}(u_K^n)^2 m(K)\le C_6.\end{aligned} \tag{4.28}$$

Inspecting relations (1.6)-(1.10) we may observe that if the elements of matrix J_ρ are finite then also the elements of matrix D_σ are finite which give (4.24). Applying (4.24) and (4.7) in (4.23) we get $\int\limits_0^{T-s} f_1(t)dt \le Cs$. Using similar approach as in [15] and relation (3.10) all further integrals can be estimated in the same way which end the proof. ◻

Proof. (of Lemma 4.4) Let us define $\xi_{K,L}=\frac{\xi}{|\xi|}.n_{K,\sigma}$ for all $(K,L)\in\Upsilon$ and let for all $x\in\Omega_\xi$

$$E(x,K,L)=\begin{cases}1 & \text{if } [x,x+\xi] \text{ intersects } \sigma=\sigma_{KL},\ K \text{ and } L;\ \text{and } \xi_{K,L}>0\\ 0 & \text{otherwise.}\end{cases}$$

For any $t\in(0,T)$ there exists $n\in N$ which satisfies $(n-1)k<t\le nk$. Then for almost all $x\in\Omega_\xi$ we can see that

$$u_{h,k}(x+\xi,t)-u_{h,k}(x,t)=u_{K(x+\xi)}^n-u_{K(x)}^n=\sum_{(K,L)\in\Upsilon}E(x,K,L)\left(u_L^n-u_K^n\right),$$

where $K(x)$ denotes the volume $K\in\mathcal{T}_h$, where $x\in K$. Using these notations we get the proof in a similar lines as in [15] - proof of Lemma 3.2. ◻

Proof. (of Lemma 4.5) In this proof, for simplicity, let us consider that $\Omega_{\tilde{t}}=\Omega$, i.e., we extend $u_{h,k}$ outside Ω by 0. The results which are obtained bellow can be straightforwardly adjusted to situation with reflexion to $\Omega_{\tilde{t}}$, the derivation is just technically more complicated, for details we refer to [10]. Let us define the set

$$\mathcal{E}_{ext}=\{\varpi,\text{ such that there exists } K\in\mathcal{T}_h,\ \varpi\subset\partial K\cap\partial\Omega\}$$

and let $u_\varpi:=u_K$ where $K\in\mathcal{T}_h,\ \varpi\subset\partial K\cap\partial\Omega$. Since now for $x\in\Omega-\Omega_\xi$ the point $x+\xi$ can be outside Ω we see that

$$u_{h,k}(x+\xi,t)-u_{h,k}(x,t)=$$

$$\sum_{(K,L)\in\Upsilon} E\left(x,K,L\right)\left(u_L^n - u_K^n\right) - \sum_{\varpi\in\mathcal{E}_{ext}} \chi\left(\left[x,x+\xi\right]\cap\varpi\right) u_\varpi^n. \tag{4.29}$$

Using the Cauchy-Schwarz and Young inequalities we obtain

$$\left(u_{h,k}\left(x+\xi,t\right) - u_{h,k}\left(x,t\right)\right)^2 \leq$$

$$\leq 2\left(\sum_{(K,L)\in\Upsilon} E\left(x,p,q\right)\xi_{K,L} d_{K,L}\right)\left(\sum_{(K,L)\in\Upsilon} E\left(x,K,L\right)\frac{\left(u_L^n - u_K^n\right)^2}{\xi_{K,L} d_{K,L}}\right) \tag{4.30}$$

$$+ 2\sum_{\varpi\in\mathcal{E}_{ext}} \chi\left(\left[x,x+\xi\right]\cap\varpi\right)\left(u_\varpi^n\right)^2,$$

$$\int\limits_{\Omega\times(0,T)} \left(u_{h,k}\left(x+\xi,t\right) - u_{h,k}\left(x,t\right)\right)^2 dxdt \leq$$

$$\left(2h + |\xi|\right)|\xi| C + 2\sum_{n=0}^{N_{\max}} k \int\limits_{\Omega} \sum_{\varpi\in\mathcal{E}_{ext}} \chi\left(\left[x,x+\xi\right]\cap\varpi\right)\left(u_\varpi^n\right)^2 dxdt, \tag{4.31}$$

which can be written as

$$\int\limits_{\Omega\times(0,T)} \left(u_{h,k}\left(x+\xi,t\right) - u_{h,k}\left(x,t\right)\right)^2 dxdt \leq$$

$$\left(2h + |\xi|\right)|\xi| C + 2|\xi| \sum_{n=0}^{N_{\max}} k \sum_{\varpi\in\mathcal{E}_{ext}} \left(u_\varpi^n\right)^2 m\left(\varpi\right). \tag{4.32}$$

For the last term in 4.32 we use the trace inequality given in [4].

Lemma 4.7. *Let Ω be an open bounded polygonal connected subset of R^d. Let $\overline{\gamma}\left(u_{h,k}\right)$ be defined by $\overline{\gamma}\left(u_{h,k}\right) = u_\varpi$ a.e. for the $(d-1)-$Lebesgue measure on $\varpi \in \mathcal{E}_{ext}$. Then there exists positive C depending only on Ω, such that*

$$\left\|\overline{\gamma}\left(u_{h,k}^n\right)\right\|_{L^2(\partial\Omega)} \leq C\left(\left|u_{h,k}^n\right|_{1,\mathcal{T}_h} + \left\|u_{h,k}^n\right\|_{L^2(\Omega)}\right).$$

Using the trace operator $\overline{\gamma}\left(u_{h,k}\right) = u_\varpi$ we can write

$$\int\limits_{\Omega\times(0,T)} \left(u_{h,k}\left(x+\xi,t\right) - u_{h,k}\left(x,t\right)\right)^2 dxdt \leq$$

$$\left(2h + |\xi|\right)|\xi| C + 2|\xi| \sum_{n=0}^{N_{\max}} k\left\|\overline{\gamma}\left(u_{h,k}^n\right)\right\|_{L^2(\partial\Omega)}^2 \tag{4.33}$$

and applying the trace inequality implies that

$$\int\limits_{\Omega\times(0,T)} \left(u_{h,k}\left(x+\xi,t\right) - u_{h,k}\left(x,t\right)\right)^2 dxdt \leq$$

$$\leq \left(2h + |\xi|\right)|\xi| C + 4C|\xi| \sum_{n=0}^{N_{\max}} k\left(\left|u_{h,k}^n\right|_{1,\mathcal{T}_h}^2 + \left\|u_{h,k}^n\right\|_{L^2(\Omega)}^2\right). \tag{4.34}$$

Then using a-priori estimates (4.6) and (4.7) ends the proof. □

Lemmas 4.2, 4.5, 4.3 guarantee that sequence $u_{h,k}$ is relatively compact in $L^2(Q_T)$ which imply following convergence result.

LEMMA 4.8. *There exists $u \in L^2(Q_T)$ such that $u_{h,k} \to u$ in $L^2(Q_T)$ as $h, k \to 0$ in the sense of subsequences.*

For the sake of simplicity, we denote the subsequence converging to u again by $u_{h,k}$, and we are going to prove that its limit u is the weak solution of (1.1)-(1.3) in the sense of Definition 4.1.

To that goal, let $\varphi \in \Psi$ be given and multiply the scheme (2.10) by $\varphi(x_K, t_n)$. Then we sum it over all $K \in \mathcal{T}_h$ and for $n = 1, ..., N_{\max}$ to get

$$\sum_{n=1}^{N_{\max}} k \sum_{K\in\mathcal{T}_h} \frac{\left(u_K^n - u_K^{n-1}\right)}{k} \varphi\left(x_K, t_{n-1}\right) m\left(K\right) = \sum_{n=1}^{N_{\max}} k \sum_{K\in\mathcal{T}_h} \varphi\left(x_K, t_{n-1}\right) \sum_{\sigma\in\mathcal{E}_K\cap\mathcal{E}_{int}} \phi_{\sigma_{KL}}^n(u_{h,k}^n) m(\sigma). \tag{4.35}$$

In order to have a structure similar to the weak solution identity (4.1), we rearrange (4.35) using a discrete integration by parts and gathering the sums over $K \in \mathcal{T}_h$ and $\sigma \in \mathcal{E}_K \cap \mathcal{E}_{int}$, and we get

$$\sum_{n=1}^{N_{\max}} k \sum_{K\in\mathcal{T}_h} u_K^n \frac{\varphi\left(x_K, t_n\right) - \varphi\left(x_K, t_{n-1}\right)}{k} m\left(K\right) + \sum_{K\in\mathcal{T}_h} u_K^0 \varphi\left(x_K, 0\right) m\left(K\right) - \frac{1}{2} \sum_{n=1}^{N_{\max}} k \sum_{\sigma\in\mathcal{E}_{int}} \phi_\sigma^n(u_{h,k}^n) m(\sigma) \left(\varphi\left(x_L, t_{n-1}\right) - \varphi\left(x_K, t_{n-1}\right)\right) = 0. \tag{4.36}$$

In the same way as in [15] we can prove that

$$\sum_{n=1}^{N_{\max}} k \sum_{K\in\mathcal{T}_h} u_K^n \frac{\varphi\left(x_K, t_n\right) - \varphi\left(x_K, t_{n-1}\right)}{k} m\left(K\right) \to \int_0^T \int_\Omega u \frac{\partial\varphi}{\partial t}\left(x, t\right) dxdt\,, \tag{4.37}$$

$$\sum_{K\in\mathcal{T}_h} u_K^0 \varphi\left(x_K, 0\right) m\left(K\right) \to \int_\Omega u_0\left(x\right) \varphi\left(x, 0\right) dx \tag{4.38}$$

as $h, k \to 0$ for all $\varphi \in \Psi$. The main point in proving convergence of the scheme is to get that

$$\frac{1}{2} \sum_{n=1}^{N_{\max}} k \sum_{\sigma\in\mathcal{E}_{int}} \phi_\sigma^n(u_{h,k}^n) m(\sigma) \left(\varphi\left(x_L, t_{n-1}\right) - \varphi\left(x_K, t_{n-1}\right)\right) \to \to \int_0^T \int_\Omega -\nabla\cdot\left(D\nabla\varphi\right) u dxdt. \tag{4.39}$$

as $h, k \to 0$ for all $\varphi \in \Psi$. Using then the space translate estimate (4.4) we know (see e.g. [4] or [14]) that the limit function u is in the space $L^2((0,T), H^1)$, so we can use

Green's theorem and applying the boundary conditions we have

$$\int_0^T \int_\Omega -\nabla \cdot (D\nabla\varphi)\, u dx dt = \int_0^T \int_\Omega (D\nabla\varphi) \cdot \nabla u dx dt.$$

Proving (4.39) thus leads to overall convergence of the scheme to the weak solution in the sense of (4.1). To deal with (4.39) we rewrite it and then split into the sum of five terms

$$\begin{aligned} &\frac{1}{2}\sum_{n=1}^{N_{\max}} k \sum_{(W,E)\in\Upsilon} \phi_\sigma^n(u_{h,k}^n) m(\sigma) \left(\varphi\left(x_E, t_{n-1}\right) - \varphi\left(x_W, t_{n-1}\right)\right) + \\ &\qquad + \int_0^T \int_\Omega \nabla \cdot \left(D\nabla\varphi(x,t)\right) u(x,t) dx dt = \sum_{i=1}^{5} T_i, \end{aligned} \tag{4.40}$$

where

$$T_1 = \frac{1}{2}\sum_{n=1}^{N_{\max}} k \sum_{(W,E)\in\Upsilon} \left[(D_\sigma p_\sigma(u)) \cdot \mathbf{n}_{W,\sigma}(\varphi_E^{n-1} - \varphi_W^{n-1}) - (u_E^n - u_W^n)(D_\sigma p_\sigma(\varphi^{n-1})) \cdot \mathbf{n}_{W,\sigma}\right] m(\sigma),$$

$$T_2 = \frac{1}{2}\sum_{n=1}^{N_{\max}} k \sum_{(W,E)\in\Upsilon} (u_E^n - u_W^n) m(\sigma) \left[(D_\sigma p_\sigma(\varphi^{n-1})) \cdot \mathbf{n}_{W,\sigma} - (D_\sigma \nabla\varphi(x_{WE}, t_{n-1})) \cdot \mathbf{n}_{W,\sigma}\right],$$

$$T_3 = \frac{1}{2}\sum_{n=1}^{N_{\max}} \sum_{(W,E)\in\Upsilon} (u_E^n - u_W^n) \left(m(\sigma) k (D_\sigma \nabla\varphi(x_{WE}, t_{n-1})) \cdot \mathbf{n}_{W,\sigma} - \int_{t_{n-1}}^{t_n} \int_\sigma (D_\sigma \nabla\varphi(s,t)) \cdot \mathbf{n}_{W,\sigma} ds dt \right),$$

$$T_4 = \frac{1}{2}\sum_{n=1}^{N_{\max}} \sum_{(W,E)\in\Upsilon} (u_E^n - u_W^n) \int_{t_{n-1}}^{t_n} \int_\sigma ((D_\sigma - D)\nabla\varphi(s,t)) \cdot \mathbf{n}_{W,\sigma} ds dt,$$

$$T_5 = \int_0^T \int_\Omega \nabla \cdot \left(D\nabla\varphi(x,t)\right) \left(u(x,t) - u_{h,k}(x,t)\right) dx dt,$$

where $\varphi_W^{n-1} = \varphi(x_W, t_{n-1})$, $\varphi_E^{n-1} = \varphi(x_E, t_{n-1})$ and $\varphi^{n-1} = \varphi(x, t_{n-1})$. Since

$$\begin{aligned} &\frac{1}{2}\sum_{n=1}^{N_{\max}} \sum_{(W,E)\in\Upsilon} (u_E^n - u_W^n) \int_{t_{n-1}}^{t_n} \int_\sigma (D\nabla\varphi(s,t)) \cdot \mathbf{n}_{W,\sigma} ds dt \\ &= -\sum_{n=1}^{N_{\max}} \sum_{W\in\mathcal{T}_h} u_W^n \int_{t_{n-1}}^{t_n} \sum_{\sigma\in\mathcal{E}_W} \int_\sigma (D\nabla\varphi(s,t)) \cdot \mathbf{n}_{W,\sigma} ds dt \\ &= -\sum_{n=1}^{N_{\max}} \sum_{W\in\mathcal{T}_h} u_W^n \int_{t_{n-1}}^{t_n} \int_W \nabla \cdot (D\nabla\varphi(x,t)) dx dt = -\int_0^T \int_\Omega \nabla \cdot (D\nabla\varphi(x,t)) u_{h,k}(x,t) dx dt \end{aligned}$$

one can see correspondence of terms in T_4 and T_5.

First, let us deal with T_1 and rewrite it using (2.7)-(2.8) into the form

$$T_1 = \frac{1}{2}\sum_{n=1}^{N_{\max}} k \sum_{(W,E)\in\Upsilon} \big([\bar\lambda_\sigma(u_E^n - u_W^n) + \bar\beta_\sigma(u_N^n - u_S^n)](\varphi_E^{n-1} - \varphi_W^{n-1}) \\ - (u_E^n - u_W^n)[\bar\lambda_\sigma(\varphi_E^{n-1} - \varphi_W^{n-1}) + \bar\beta_\sigma(\varphi_N^{n-1} - \varphi_S^{n-1})]\big)\,,$$

which can be easily simplified to

$$T_1 = \frac{1}{2}\sum_{n=1}^{N_{\max}} k \sum_{(W,E)\in\Upsilon} \left[\bar\beta_\sigma(u_N^n - u_S^n)(\varphi_E^{n-1} - \varphi_W^{n-1}) - \bar\beta_\sigma(u_E^n - u_W^n)(\varphi_N^{n-1} - \varphi_S^{n-1})\right]. \tag{4.41}$$

Applying (3.1) for u and similarly for φ we get

$$T_1 = \frac{1}{2}\sum_{n=1}^{N_{\max}} k \sum_{(W,E)\in\Upsilon}\sum_{i=1}^{4}\left[\frac{\bar\beta_\sigma}{4}(\varphi_{E_\sigma} - \varphi_{W_\sigma})(u_{E_{\delta_i}} - u_{W_{\delta_i}}) \right. \\ \left. -\frac{\bar\beta_\sigma}{4}(u_{E_\sigma} - u_{W_\sigma})(\varphi_{E_{\delta_i}} - \varphi_{W_{\delta_i}})(1+O(h))\right], \tag{4.42}$$

where we omit time indexes due to simplification; for graphical explanation of notations see Fig. 3.1 and Fig. 4.1. For each term with positive sign in (4.42) one can

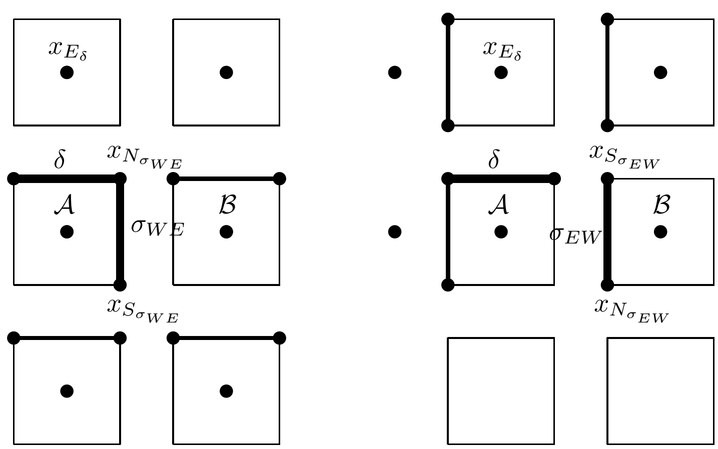

FIG. 4.1. *Left: the edge $\sigma = \sigma_{WE}$ and one $\delta \in P_\sigma$. Right: the same edge δ and one corresponding $\sigma = \sigma_{EW} \in P_\delta$. The thickest lines represent one particular couple $T_{\sigma\delta}$ in (4.42), which vanishes up to $O(h)$ term. $\mathcal{A} = x_{W_{\sigma_{WE}}} = x_{E_{\sigma_{EW}}} = x_{W_\delta}$, $\mathcal{B} = x_{E_{\sigma_{WE}}} = x_{W_{\sigma_{EW}}}$.*

find a corresponding term in the group of terms with negative signs ($\bar\beta_\sigma$ corresponds to some $\bar\beta_\delta$). We denote these couples by $T_{\sigma\delta}$. E.g., for $\sigma = \sigma_{WE}$ and δ as plotted in Fig. 4.1 left we can write the couple as follows

$$T_{\sigma_{WE}\delta} = \frac{\bar\beta_{\sigma_{WE}}}{4}(\varphi^{n-1}_{E_{\sigma_{WE}}} - \varphi^{n-1}_{W_{\sigma_{WE}}})(u^n_{E_\delta} - u^n_{W_\delta}) -$$

$$\frac{\bar{\beta}_\delta}{4}(\varphi^{n-1}_{E_{\sigma_{EW}}} - \varphi^{n-1}_{W_{\sigma_{EW}}})(u^n_{E_\delta} - u^n_{W_\delta})(1+O(h)) =$$
$$\frac{\bar{\beta}_{\sigma_{WE}}}{4}(\varphi^{n-1}_{E_{\sigma_{WE}}} - \varphi^{n-1}_{W_{\sigma_{WE}}})(u^n_{E_\delta} - u^n_{W_\delta}) - \frac{\bar{\beta}_\delta}{4}(\varphi^{n-1}_{W_{\sigma_{WE}}} - \varphi^{n-1}_{E_{\sigma_{WE}}})(u^n_{E_\delta} - u^n_{W_\delta})(1+O(h))$$

because $\bar{\beta}_{\sigma_{WE}} = \bar{\beta}_{\sigma_{EW}}$, $\varphi^{n-1}_{E_{\sigma_{WE}}} = \varphi^{n-1}_{W_{\sigma_{EW}}}$ and $\varphi^{n-1}_{W_{\sigma_{WE}}} = \varphi^{n-1}_{E_{\sigma_{EW}}}$, see Fig. 4.1 right. Using previous expression for every $\sigma = \sigma_{WE}$ yields

$$T_{\sigma\delta} = \left[\frac{\bar{\beta}_\sigma}{4} + \left(-\frac{\bar{\beta}_\sigma}{4}(1+O(h))\right)\right](\varphi^{n-1}_{E_\sigma} - \varphi^{n-1}_{W_\sigma})(u^n_{E_\delta} - u^n_{W_\delta})$$
$$= \frac{\bar{\beta}_\sigma}{4}O(h)(\varphi^{n-1}_{E_\sigma} - \varphi^{n-1}_{W_\sigma})(u^n_{E_\delta} - u^n_{W_\delta}). \tag{4.43}$$

Then T_1 can be estimated as follows

$$|T_1| \le C_1 h \left| \sum_{n=1}^{N_{\max}} k \sum_{\sigma\in\mathcal{E}_{int}} \sum_{\delta\in P_\sigma\cap\mathcal{E}_{int}} (\varphi^{n-1}_{E_\sigma} - \varphi^{n-1}_{W_\sigma})(u^n_{E_\delta} - u^n_{W_\delta}) \right|$$

with positive constant C_1 due to fact that $\bar{\beta}_\sigma$ for each edge of mesh is finite (see page 14). By the Cauchy-Schwarz inequality we have

$$|T_1| \le C_2 h \left(\sum_{n=1}^{N_{\max}} k \sum_{\sigma\in\mathcal{E}_{int}} \sum_{\delta\in P_\sigma\cap\mathcal{E}_{int}} (\varphi^{n-1}_{E_\sigma} - \varphi^{n-1}_{W_\sigma})^2 \right)^{\frac{1}{2}}$$
$$\left(\sum_{n=1}^{N_{\max}} k \sum_{\sigma\in\mathcal{E}_{int}} \sum_{\delta\in P_\sigma\cap\mathcal{E}_{int}} (u^n_{E_\delta} - u^n_{W_\delta})^2 \right)^{\frac{1}{2}}. \tag{4.44}$$

It comes from regularity of φ that there exists a positive constant C_3 such that $(\varphi^{n-1}_{E_\sigma} - \varphi^{n-1}_{W_\sigma})^2 \le C_3 h^2$. Thanks to geometrical arguments, we know that

$$\sum_{\sigma\in\mathcal{E}_{int}} d_{WE} m(\sigma) \le C_4 |\Omega| \tag{4.45}$$

which straightforwardly gives for our uniform square mesh $\sum\limits_{\sigma\in\mathcal{E}_{int}} h^2 \le C_4|\Omega|$. The above mentioned facts lead to

$$\left(\sum_{n=1}^{N_{\max}} k \sum_{\sigma\in\mathcal{E}_{int}} \sum_{\delta\in P_\sigma\cap\mathcal{E}_{int}} (\varphi^{n-1}_{E_\sigma} - \varphi^{n-1}_{W_\sigma})^2 \right) \le C_5 T |\Omega|,$$

which together with a priori estimate (4.7) gives that $|T_1| \le C_6 h$, which implies

$$|T_1| \to 0 \text{ as } h, k \to 0. \tag{4.46}$$

Term T_2 can be written as $T_2 = \frac{1}{2}\sum\limits_{n=1}^{N_{\max}} k \sum\limits_{(W,E)\in\Upsilon} (u^n_E - u^n_W) m(\sigma) T_{2WE}$ where

$$T_{2WE} = \bar{\lambda}_\sigma \frac{\varphi^{n-1}_E - \varphi^{n-1}_W}{h} + \bar{\beta}_\sigma \frac{\varphi^{n-1}_N - \varphi^{n-1}_S}{h} - \bar{\lambda}_\sigma \left(\varphi^{n-1}_{WE}\right)_x - \bar{\beta}_\sigma \left(\varphi^{n-1}_{WE}\right)_y$$

and $\begin{pmatrix} (\varphi_{WE}^{n-1})_x \\ (\varphi_{WE}^{n-1})_y \end{pmatrix} = \nabla\varphi(x_{WE}, t_{n-1})$ in the basis $(\mathbf{n}_{W,\sigma}, \mathbf{t}_{W,\sigma})$. Since $\varphi \in C^{2,1}\left(\overline{\Omega} \times [0,T]\right)$, there exist positive constants C_7 and C_8 such that

$$\left|\frac{\varphi_E^{n-1} - \varphi_W^{n-1}}{h} - \left(\varphi_{WE}^{n-1}\right)_x\right| \leq C_7 h, \quad \left|\frac{\varphi_N^{n-1} - \varphi_S^{n-1}}{h} - \left(\varphi_{WE}^{n-1}\right)_y\right| \leq C_8 h.$$

From there and the property that elements of D_σ are finite (see page 14) we have $|T_{2WE}| \leq C_9 h$ with a positive constant C_9, which implies

$$|T_2| \leq C_{10} h \left| \sum_{n=1}^{N_{\max}} k \sum_{(W,E)\in\Upsilon} (u_E^n - u_W^n) m(\sigma) \right|. \tag{4.47}$$

Then we multiply the right hand side of (4.47) by $\frac{\sqrt{d_{WE}}}{\sqrt{d_{WE}}}$ and apply Cauchy-Schwarz inequality to obtain

$$|T_2| \leq C_{10} h \left(\sum_{n=1}^{N_{\max}} k \sum_{(W,E)\in\Upsilon} (u_E^n - u_W^n)^2 \frac{m(\sigma)}{d_{WE}} \right)^{\frac{1}{2}} \left(\sum_{n=1}^{N_{\max}} k \sum_{(W,E)\in\Upsilon} m(\sigma) d_{WE} \right)^{\frac{1}{2}}.$$

A-priori estimate (4.7) together with (4.45) gives $|T_2| \leq C_{10}(C_{11}C_4|\Omega|T)^{\frac{1}{2}} h$ and finally

$$|T_2| \to 0 \text{ as } h, k \to 0. \tag{4.48}$$

We consider the third term in the form $T_3 = \dfrac{1}{2}\displaystyle\sum_{n=1}^{N_{\max}} \sum_{(W,E)\in\Upsilon} (u_E^n - u_W^n) T_{3WE}$, where

$$T_{3WE} = m(\sigma) k (D_\sigma \nabla\varphi(x_{WE}, t_{n-1})) \cdot \mathbf{n}_{W,\sigma} - \int_{t_{n-1}}^{t_n} \int_\sigma (D_\sigma \nabla\varphi(s,t)) \cdot \mathbf{n}_{W,\sigma} ds dt.$$

Due to smoothness of φ, the mean value theorem and finiteness of D_σ we have $|T_3| \leq C_{12}(h+k) \left| \displaystyle\sum_{n=1}^{N_{\max}} k \sum_{(W,E)\in\Upsilon} (u_E^n - u_W^n) m(\sigma) \right|$ and similarly as above using (4.7) and (4.45) we get

$$|T_3| \to 0 \text{ as } h, k \to 0. \tag{4.49}$$

The basic ingredients of the proof of convergence $|T_4| \to 0$ as $h, k \to 0$ are given by the properties of functions λ, β, ν as Lipschitz functions of the partial derivatives of solution (smoothed by spatial convolutions) and from the convergence of $u_{h,k}$ to u in $L_2(Q_T)$. Term T_4 is written in detailed way as follows

$$T_4 = \frac{1}{2} \sum_{n=1}^{N_{\max}} \sum_{(W,E)\in\Upsilon} (u_E^n - u_W^n) \int_{t_{n-1}}^{t_n} \int_\sigma T_\sigma \, ds dt \tag{4.50}$$

where

$$T_\sigma = (D_\sigma(u_{h,k})(x_{KL}, t_{n-1}) - D(u)(s,t))\nabla\varphi(s,t)\cdot \mathbf{n}_{W,\sigma},\ s\in\sigma,\ t\in(t_{n-1},t_n).$$

Let us note that in the previous definition the construction of D_σ and D, respectively, is the same (the index σ is used only because the tensor D_σ is evaluated using numerical solution on edge σ, such notation was introduced and used in previous sections), but the arguments are different. In the first case it is the numerical solution $u_{h,k}$, which is used in evaluation of diffusion tensor at point $x_{KL}\in\sigma$ and at time t_{n-1}. In the second case the argument is given by the limit function u and the tensor is evaluated at any point $s\in\sigma$ and $t\in(t_{n-1},t_n)$. So, we can write

$$T_\sigma = \left(\bar\lambda(u_{h,k})(x_{KL},t_{n-1}) - \bar\lambda(u)(s,t)\right)\bar\varphi_x + \left(\bar\beta(u_{h,k})(x_{KL},t_{n-1}) - \bar\beta(u)(s,t)\right)\bar\varphi_y, \tag{4.51}$$

where $\bar\varphi_x$ and $\bar\varphi_y$ are elements of $\nabla\varphi(s,t)$ in the basis $(\mathbf{n}_{K,\sigma},\mathbf{t}_{K,\sigma})$. I.e. $\begin{pmatrix}\bar\varphi_x\\ \bar\varphi_y\end{pmatrix}$ is equal to $\begin{pmatrix}\varphi_x\\ \varphi_y\end{pmatrix}$ on σ_3, $\begin{pmatrix}-\varphi_x\\ -\varphi_y\end{pmatrix}$ on σ_2, $\begin{pmatrix}\varphi_y\\ -\varphi_x\end{pmatrix}$ on σ_1 and $\begin{pmatrix}-\varphi_y\\ \varphi_x\end{pmatrix}$ on σ_4, cf. Fig. 2.1. In order to get bounds of the term T_σ. we use properties of functions λ, β, ν, because $\bar\lambda$ and $\bar\beta$ may be equal to one of them depending on the local basic in which both matrices D_σ and D are written. From the diffusion tensor construction it follows that we can write it in the following form

$$T_\sigma = \left(\bar\lambda(a^{n-1}_{h,k}, b^{n-1}_{h,k}, c^{n-1}_{h,k}) - \bar\lambda(a,b,c)\right)\bar\varphi_x + \left(\bar\beta(a^{n-1}_{h,k}, b^{n-1}_{h,k}, c^{n-1}_{h,k}) - \bar\beta(a,b,c)\right)\bar\varphi_y, \tag{4.52}$$

where $\bar\lambda$ and $\bar\beta$ is equal to one of the functions λ, β, ν depending on three arguments as follows

$$\begin{aligned}
\lambda(a,b,c) &= \frac{\kappa_1 v_1^2 + \kappa_2 v_2^2}{v_1^2+v_2^2} = \\
&= \begin{cases} \alpha, \quad \text{if } \mu_1=\mu_2 \text{ (i.e. if} \sqrt{4b^2+(a-c)^2}=0), \\ \alpha + (1-\alpha)\left(\frac{1}{2} + \frac{c-a}{2\sqrt{4b^2+(a-c)^2}}\right) e^{-\frac{1}{4b^2+(a-c)^2}}, \quad \text{else.}\end{cases} \\
\beta(a,b,c) &= \frac{v_1 v_2(\kappa_1-\kappa_2)}{v_1^2+v_2^2} = \\
&= \begin{cases} 0, \quad \text{if } \mu_1=\mu_2 \text{ (i.e. if} \sqrt{4b^2+(a-c)^2}=0), \\ \frac{(\alpha-1)b}{\sqrt{4b^2+(a-c)^2}} e^{-\frac{1}{4b^2+(a-c)^2}}, \quad \text{else.}\end{cases} \\
\nu(a,b,c) &= \frac{\kappa_2 v_1^2 + \kappa_1 v_2^2}{v_1^2+v_2^2} = \\
&= \begin{cases} \alpha, \quad \text{if } \mu_1=\mu_2 \text{ (i.e. if} \sqrt{4b^2+(a-c)^2}=0), \\ \alpha + (1-\alpha)\left(\frac{1}{2} - \frac{c-a}{2\sqrt{4b^2+(a-c)^2}}\right) e^{-\frac{1}{4b^2+(a-c)^2}}, \quad \text{else.}\end{cases}
\end{aligned}$$

If we denote by $F = 4b^2+(a-c)^2$ the values of λ, β, ν for $F=0$ are defined as limit values as $F\to 0$ and thus the functions are continuous. In (4.52) we have

$$a = \left(G_\rho * \left(\frac{\partial G_{\tilde t}}{\partial x} * u\right)^2\right)(s,t),$$

$$b = \left(G_\rho * \left[\left(\frac{\partial G_{\tilde t}}{\partial x} * u\right)\left(\frac{\partial G_{\tilde t}}{\partial y} * u\right)\right]\right)(s,t),$$

$$c = \left(G_\rho * \left(\frac{\partial G_{\tilde t}}{\partial y} * u\right)^2\right)(s,t),$$

$$a_{h,k}^{n-1} = \left(G_\rho * \left(\frac{\partial G_{\tilde t}}{\partial x} * u_{h,k}\right)^2\right)(x_{KL}, t_{n-1}),$$

$$b_{h,k}^{n-1} = \left(G_\rho * \left[\left(\frac{\partial G_{\tilde t}}{\partial x} * u_{h,k}\right)\left(\frac{\partial G_{\tilde t}}{\partial y} * u_{h,k}\right)\right]\right)(x_{KL}, t_{n-1}),$$

$$c_{h,k}^{n-1} = \left(G_\rho * \left(\frac{\partial G_{\tilde t}}{\partial y} * u_{h,k}\right)^2\right)(x_{KL}, t_{n-1}).$$

First of all, from such form of λ, β, ν one can simply see that they are uniformly bounded. We can write

$$\lambda(a,b,c) = \frac{\kappa_1 v_1^2 + \kappa_2 v_2^2}{v_1^2 + v_2^2} = \alpha + (1-\alpha)\left(\frac{1}{2} + \frac{c-a}{2\sqrt{F}}\right)e^{-\frac{1}{F}}.$$

Since $|a-c| \le \sqrt{4b^2 + (a-c)^2} = \sqrt{F}$ and $\alpha \in (0,1)$, we get

$$|\lambda| \le \alpha + (1-\alpha)\left(\frac{1}{2} + \frac{\sqrt{F}}{2\sqrt{F}}\right)e^{-\frac{1}{F}} \le 1.$$

Similarly it is for β and ν.

From the structure of λ, β, ν we can see that their partial derivatives, with respect to a, b, c of any order will contain the term $e^{-\frac{1}{F}}$ and some rational polynomial which can be estimated by the powers of F and which together give uniform bounds on derivatives. In the convergence proof it will be sufficient to have Lipschitz continuity of λ, β, ν, so we show that their first partial derivatives are uniformly bounded. First we have

$$\frac{\partial \lambda}{\partial a} = (1-\alpha)(a-c)\frac{e^{-\frac{1}{F}}}{F^2} + (\alpha-1)(8b^4 + 2b^2(a-c)^2 + (a-c)^2)\frac{e^{-\frac{1}{F}}}{F^{\frac{5}{2}}}.$$

Since $|\alpha| < 1$, $|1-\alpha| < 1$ and $|a-c| \le F^{\frac{1}{2}}$ we have

$$\begin{aligned}\left|\frac{\partial \lambda}{\partial a}\right| &\le |a-c|\frac{e^{-\frac{1}{F}}}{F^2} + ((4b^2 + (a-c)^2)^2 + (a-c)^2 + 4b^2)\frac{e^{-\frac{1}{F}}}{F^{\frac{5}{2}}} \\ &\le \frac{Fe^{-\frac{1}{F}}}{F^{\frac{5}{2}}} + \frac{(F^2+F)e^{-\frac{1}{F}}}{F^{\frac{5}{2}}} = \frac{F+2}{F^{\frac{3}{2}}}e^{-\frac{1}{F}} \\ &= h_1(F) \le \max_{F \ge 0} h_1(F) = h_1(F_M),\end{aligned}$$

where $F_M = 2(\sqrt{2}-1)$, so

$$\left|\frac{\partial \lambda}{\partial a}\right| \le C_1 \le 1.2\,. \tag{4.53}$$

Then we have

$$\frac{\partial\lambda}{\partial b} = (1-\alpha)\frac{4be^{-\frac{1}{F}}}{F^2} + (1-\alpha)2b(a-c)(4b^2+(a-c)^2-2)\frac{e^{-\frac{1}{F}}}{F^{\frac{5}{2}}}.$$

Since $2|b| \le \sqrt{4b^2+(a-c)^2} = F^{\frac{1}{2}}$ we have

$$\left|\frac{\partial\lambda}{\partial b}\right| \le \frac{2F^{\frac{1}{2}}e^{-\frac{1}{F}}}{F^2} + \frac{F^{\frac{1}{2}}F^{\frac{1}{2}}(F+2)e^{-\frac{1}{F}}}{F^{\frac{5}{2}}} = \frac{(F+4)e^{-\frac{1}{F}}}{F^{\frac{3}{2}}}$$
$$= h_2(F) \le \max_{F\ge 0} h_2(F) = h_2(F_M),$$

where $F_M = \sqrt{33}-5$, so

$$\left|\frac{\partial\lambda}{\partial b}\right| \le C_2 \le 2\,. \tag{4.54}$$

Since

$$\frac{\partial\lambda}{\partial c} = (\alpha-1)(a-c)\frac{e^{-\frac{1}{F}}}{F^2} + (1-\alpha)(8b^4+2b^2(a-c)^2+(a-c)^2)\frac{e^{-\frac{1}{F}}}{F^{\frac{5}{2}}} = -\frac{\partial\lambda}{\partial a}$$

we get

$$\left|\frac{\partial\lambda}{\partial c}\right| = \left|\frac{\partial\lambda}{\partial a}\right| \le C_1 \le 1.2\,. \tag{4.55}$$

For the function ν we get

$$\frac{\partial\nu}{\partial a} = (1-\alpha)(a-c)\frac{e^{-\frac{1}{F}}}{F^2} + (1-\alpha)(8b^4+2b^2(a-c)^2+(a-c)^2)\frac{e^{-\frac{1}{F}}}{F^{\frac{5}{2}}}$$

so we get the same estimate as in (4.53) and (4.55) also for

$$\left|\frac{\partial\nu}{\partial a}\right| \le C_1 \le 1.2\,. \tag{4.56}$$

Then

$$\frac{\partial\nu}{\partial b} = (1-\alpha)\frac{4be^{-\frac{1}{F}}}{F^2} + (\alpha-1)2b(a-c)(4b^2+(a-c)^2-2)\frac{e^{-\frac{1}{F}}}{F^{\frac{5}{2}}},$$

so we get in the same way as in (4.54) the estimate for

$$\left|\frac{\partial\nu}{\partial b}\right| \le C_2 \le 2\,. \tag{4.57}$$

Since $\frac{\partial\nu}{\partial c} = -\frac{\partial\nu}{\partial a}$ we have again

$$\left|\frac{\partial\nu}{\partial c}\right| \le C_1 \le 1.2\,. \tag{4.58}$$

For the function β we have

$$\frac{\partial\beta}{\partial a} = (1-\alpha)b(a-c)(4b^2+(a-c)^2-2)\frac{e^{-\frac{1}{F}}}{F^{\frac{5}{2}}},$$

thus

$$\left|\frac{\partial\beta}{\partial a}\right| \leq \frac{F^{\frac{1}{2}}F^{\frac{1}{2}}(F+2)e^{-\frac{1}{F}}}{F^{\frac{5}{2}}} = \frac{(F+2)e^{-\frac{1}{F}}}{F^{\frac{3}{2}}} \leq \max_{F\geq 0} h_1(F) = h_1(F_M)$$

and therefore

$$\left|\frac{\partial\beta}{\partial a}\right| \leq C_1 \leq 1.2\,. \tag{4.59}$$

We also have

$$\frac{\partial\beta}{\partial b} = (\alpha-1)(8b^2+4b^2(a-c)^2+(a-c)^4)\frac{e^{-\frac{1}{F}}}{F^{\frac{5}{2}}},$$

so

$$\begin{aligned}\left|\frac{\partial\beta}{\partial b}\right| &\leq (2(4b^2+(a-c)^2)+(4b^2+(a-c)^2)^2)\frac{e^{-\frac{1}{F}}}{F^{\frac{5}{2}}} = \frac{2F+F^2}{F^{\frac{5}{2}}}e^{-\frac{1}{F}} \\ &= \frac{F+2}{F^{\frac{3}{2}}}e^{-\frac{1}{F}} \leq \max_{F\geq 0} h_1(F) = h_1(F_M)\end{aligned}$$

and from it

$$\left|\frac{\partial\beta}{\partial b}\right| \leq C_1 \leq 1.2\,. \tag{4.60}$$

Using $\frac{\partial\beta}{\partial c} = -\frac{\partial\beta}{\partial a}$ we have the estimate

$$\left|\frac{\partial\beta}{\partial c}\right| \leq C_1 \leq 1.2\,. \tag{4.61}$$

Just as an illustration we also show boundedness of $\frac{\partial^2\lambda}{\partial a^2}$ (all other second partial derivatives can be treated similarly):

$$\begin{aligned}\frac{\partial^2\lambda}{\partial a^2} &= (1-\alpha)(16b^4-8b^2(a-c)^2-3(a-c)^4+2(a-c)^2)\frac{e^{-\frac{1}{F}}}{F^4} \\ &+ (1-\alpha)(96b^6(a-c)+6b^2(a-c)^5+3(a-c)^5-2(a-c)^3 \\ &+ 48b^4(a-c)^3-48b^4(a-c))\frac{e^{-\frac{1}{F}}}{F^{\frac{9}{2}}}\end{aligned}$$

Applying $|a-c|^2 \leq F$, $|a-c| \leq F^{\frac{1}{2}}$, $4|b|^2 \leq F$ and $2|b| \leq F^{\frac{1}{2}}$ we get

$$\begin{aligned}\left|\frac{\partial^2\lambda}{\partial a^2}\right| &\leq ((4b^2+(a-c)^2)^2+3(a-c)^4+2(a-c)^2)\frac{e^{-\frac{1}{F}}}{F^4} \\ &+ \left(\frac{3}{2}(4b^2)^3|a-c|+6b^2|a-c|^5+3|a-c|^5+2|a-c|^3\right. \\ &+ \left.48b^4|a-c|^3+48b^4|a-c|\right)\frac{e^{-\frac{1}{F}}}{F^{\frac{9}{2}}} \\ &\leq \frac{F^2+3F^2+2F}{F^4}e^{-\frac{1}{F}} +\end{aligned}$$

$$+\frac{\frac{3}{2}F^3F^{\frac{1}{2}}+\frac{3}{2}FF^{\frac{5}{2}}+3F^{\frac{5}{2}}+2F^{\frac{3}{2}}+3F^2F^{\frac{3}{2}}+3F^2F^{\frac{1}{2}}}{F^{\frac{9}{2}}}e^{-\frac{1}{F}}$$
$$=\frac{4F+2}{F^3}e^{-\frac{1}{F}}+\frac{\frac{3}{2}F^2+\frac{3}{2}F^2+3F+2+3F^2+3F}{F^3}e^{-\frac{1}{F}}$$
$$=\frac{6F^2+10F+4}{F^3}e^{-\frac{1}{F}}\leq\max_{F\geq 0}h_3(F)=h_3(F_M)\leq C\leq 12,$$

where $F_M=\frac{\sqrt{13}-1}{6}$.

The term T_σ in (4.52) contains differences of either λ, β or ν evaluated in different arguments. Using their Lipschitz continuity those differences can be estimated by the differences of arguments. We will do it only for λ, all other situations are treated similarly. So we have

$$|\lambda(a_{h,k}^{n-1},b_{h,k}^{n-1},c_{h,k}^{n-1})-\lambda(a,b,c)|\leq L_\lambda\sqrt{(a_{h,k}^{n-1}-a)^2+(b_{h,k}^{n-1}-b)^2+(c_{h,k}^{n-1}-c)^2}$$
$$\leq L_\lambda(|a_{h,k}^{n-1}-a|+|b_{h,k}^{n-1}-b|+|c_{h,k}^{n-1}-c|), \tag{4.62}$$

where L_λ is Lipschitz constant of function λ. Since all terms in the absolute values on the right hand side of (4.62) can be estimated similarly, we do it in details just for the first one $|a_{h,k}^{n-1}-a|$ (a slight difference is only when treating $|b_{h,k}^{n-1}-b|$, we will mention it later in the text). We can use the following splitting and get

$$|a_{h,k}^{n-1}-a|\leq\left|\left(G_\rho*\left(\frac{\partial G_{\tilde t}}{\partial x}*u_{h,k}\right)^2\right)(x_{KL},t_{n-1})-\left(G_\rho*\left(\frac{\partial G_{\tilde t}}{\partial x}*u_{h,k}\right)^2\right)(s,t_{n-1})\right|$$
$$+\left|\left(G_\rho*\left(\frac{\partial G_{\tilde t}}{\partial x}*u_{h,k}\right)^2\right)(s,t_{n-1})-\left(G_\rho*\left(\frac{\partial G_{\tilde t}}{\partial x}*u_{h,k}\right)^2\right)(s,t)\right|$$
$$+\left|\left(G_\rho*\left(\frac{\partial G_{\tilde t}}{\partial x}*u_{h,k}\right)^2\right)(s,t)-\left(G_\rho*\left(\frac{\partial G_{\tilde t}}{\partial x}*u\right)^2\right)(s,t)\right|$$
$$=A_1+A_2+A_3.$$

Then subsequently

$$A_1\leq\left|\int_{R^N}G_\rho(x_{KL}-\xi)\left(\int_{R^N}\frac{\partial G_{\tilde t}}{\partial x}(\xi-\eta)u_{h,k}(\eta,t_{n-1})d\eta\right)^2d\xi-\right.$$
$$\left.-\int_{R^N}G_\rho(s-\xi)\left(\int_{R^N}\frac{\partial G_{\tilde t}}{\partial x}(\xi-\eta)u_{h,k}(\eta,t_{n-1})d\eta\right)^2d\xi\right|\leq$$
$$\leq\int_{R^N}|G_\rho(x_{KL}-\xi)-G_\rho(s-\xi)|\left(\int_{R^N}\frac{\partial G_{\tilde t}}{\partial x}(\xi-\eta)u_{h,k}(\eta,t_{n-1})d\eta\right)^2d\xi\leq Ch,$$

because of the fact that $|x_{KL}-s|\leq h$, C^∞ smoothness of G_ρ and because

$$\left(\int_{R^N}\frac{\partial G_{\tilde t}}{\partial x}(\xi-\eta)u_{h,k}(\eta,t_{n-1})d\eta\right)^2\leq C\int_{R^N}\left(\frac{\partial G_{\tilde t}}{\partial x}(\xi-\eta)\right)^2d\eta\int_\Omega u_{h,k}^2(\eta,t_{n-1})d\eta\leq$$
$$\leq\ C||u_{h,k}(t_{n-1})||_{L_2(\Omega)}\leq C$$

holds for any $\xi \in R^N$, using Cauchy-Schwartz inequality, C^∞ smoothness of $G_{\tilde{t}}$, extension by 0 of $u_{h,k}$ outside a neighbourhood of Ω, and due to a-priori estimate (4.6).

For the second part we have

$$A_2 = \int_{R^N} G_\rho(s-\xi)\left[\left(\int_{R^N} \frac{\partial G_{\tilde{t}}}{\partial x}(\xi-\eta)u_{h,k}(\eta,t_{n-1})d\eta\right)^2 - \left(\int_{R^N} \frac{\partial G_{\tilde{t}}}{\partial x}(\xi-\eta)u_{h,k}(\eta,t)d\eta\right)^2\right]d\xi = \int_{R^N} G_\rho(s-\xi)B_1 d\xi,$$

and, for the term B_1, using the relation $|p^2-q^2| = |p+q||p-q|$, we get for any $\xi \in R^N$, $t \in (t_{n-1}, t_n)$ that

$$|B_1| = \left|\int_{R^N} \frac{\partial G_{\tilde{t}}}{\partial x}(\xi-\eta)(u_{h,k}(\eta,t_{n-1}) + u_{h,k}(\eta,t))d\eta\right| \cdot \left|\int_{R^N} \frac{\partial G_{\tilde{t}}}{\partial x}(\xi-\eta)(u_{h,k}(\eta,t_{n-1}) - u_{h,k}(\eta,t))d\eta\right| \leq$$

$$\leq C\left(||u_{h,k}(\eta,t_{n-1})||_{L_2(\Omega)} + ||u_{h,k}(\eta,t_n)||_{L_2(\Omega)}\right) ||u_{h,k}(t-k) - u_{h,k}(t)||_{L_2(\Omega)} \leq$$

$$\leq C||u_{h,k}(t-k) - u_{h,k}(t)||_{L_2(\Omega)}, \quad (4.63)$$

where Cauchy-Schwartz inequality, a-priori estimate (4.6) and piecewise constant in time definition of $u_{h,k}$ was used. Then also

$$|A_2| \leq C||u_{h,k}(t-k) - u_{h,k}(t)||_{L_2(\Omega)}$$

because $\int_{R^N} G_\rho(s-\xi)d\xi = 1$ for any s.

For the third term we have

$$A_3 = \left|\int_{R^N} G_\rho(s-\xi)\left[\left(\int_{R^N} \frac{\partial G_{\tilde{t}}}{\partial x}(\xi-\eta)u_{h,k}(\eta,t)d\eta\right)^2 d\xi - \left(\int_{R^N} \frac{\partial G_{\tilde{t}}}{\partial x}(\xi-\eta)u(\eta,t)d\eta\right)^2 d\xi\right]\right| = \left|\int_{R^N} G_\rho(s-\xi)B_2 d\xi\right|$$

and again due to a-priori estimate (4.6) and Cauchy-Schwartz inequality we get

$$|B_2| = \left|\int_{R^N} \frac{\partial G_{\tilde{t}}}{\partial x}(\xi-\eta)(u_{h,k}(\eta,t) + u(\eta,t))d\eta\right| \left|\int_{R^N} \frac{\partial G_{\tilde{t}}}{\partial x}(\xi-\eta)(u_{h,k}(\eta,t) - u(\eta,t))d\eta\right|$$

$$\leq C||u_{h,k}(t) - u(t)||_{L_2(\Omega)} + C||u(t)||_{L_2(\Omega)}||u_{h,k}(t) - u(t)||_{L_2(\Omega)}.$$

Let us note that for $|b_{h,k}^{n-1} - b|$ we can use the same approach as above, but in the terms which would correspond to A_2 and A_3, we would use $|p_1 q_1 - p_2 q_2| \le |p_1(q_1 - q_2)| + |(p_1 - p_2)q_2|$ in order to get the same estimates as above. For the term $|c_{h,k}^{n-1} - c|$ we can use completely same approach as above. Putting together all previous estimates and because of smoothness of φ_x and φ_y we have that

$$|T_4| \le C \sum_{n=1}^{N_{\max}} \sum_{(W,E)\in\Upsilon} |u_E^n - u_W^n| \int_{t_{n-1}}^{t_n} \int_{\sigma} (h + ||u_{h,k}(t-k) - u_{h,k}(t)||_{L_2(\Omega)}$$
$$+||u_{h,k}(t) - u(t)||_{L_2(\Omega)} + ||u(t)||_{L_2(\Omega)} ||u(t) - u_{h,k}(t)||_{L_2(\Omega)}) ds dt.$$

The most important observation now is that, the terms inside the double integral do not depend on s. So it will be sufficient to use several times Cauchy-Schwarz inequality and the same trick as in treating the term T_2 to get convergence of T_4 to 0. First term is simple and is estimated exactly as T_2, i.e.

$$C \sum_{n=1}^{N_{\max}} \sum_{(W,E)\in\Upsilon} |u_E^n - u_W^n| \int_{t_{n-1}}^{t_n} \int_{\sigma} h \, ds dt \le Ch \sum_{n=1}^{N_{\max}} k \sum_{(W,E)\in\Upsilon} |u_E^n - u_W^n| m(\sigma)$$
$$\le Ch \sum_{n=1}^{N_{\max}} \sqrt{k} \left(\sum_{(W,E)\in\Upsilon} \frac{(u_E^n - u_W^n)^2}{d_{KL}} m(\sigma) \right)^{\frac{1}{2}} \sqrt{k} \left(\sum_{(W,E)\in\Upsilon} m(\sigma) d_{KL} \right)^{\frac{1}{2}}$$
$$\le C|\Omega|^{\frac{1}{2}} h \left(\sum_{n=1}^{N_{\max}} k \sum_{(W,E)\in\Upsilon} \frac{(u_E^n - u_W^n)^2}{d_{KL}} m(\sigma) \right)^{\frac{1}{2}} \left(\sum_{n=1}^{N_{\max}} k \right)^{\frac{1}{2}} \le C|\Omega|^{\frac{1}{2}} T^{\frac{1}{2}} h$$

due to a-priori estimate (4.7). For the second term we have

$$C \sum_{n=1}^{N_{\max}} \int_{t_{n-1}}^{t_n} \sum_{(W,E)\in\Upsilon} |u_E^n - u_W^n| m(\sigma) ||u_{h,k}(t-k) - u_{h,k}(t)||_{L_2(\Omega)} dt$$
$$\le C|\Omega|^{\frac{1}{2}} \sum_{n=1}^{N_{\max}} \int_{t_{n-1}}^{t_n} \left(\sum_{(W,E)\in\Upsilon} \frac{(u_E^n - u_W^n)^2}{d_{KL}} m(\sigma) \right)^{\frac{1}{2}} ||u_{h,k}(t-k) - u_{h,k}(t)||_{L_2(\Omega)} dt$$
$$\le C|\Omega|^{\frac{1}{2}} \sum_{n=1}^{N_{\max}} \left(\int_{t_{n-1}}^{t_n} \sum_{(W,E)\in\Upsilon} \frac{(u_E^n - u_W^n)^2}{d_{KL}} m(\sigma) dt \right)^{\frac{1}{2}} \left(\int_{t_{n-1}}^{t_n} ||u_{h,k}(t-k) - u_{h,k}(t)||_{L_2(\Omega)}^2 dt \right)^{\frac{1}{2}}$$
$$\le C|\Omega|^{\frac{1}{2}} \left(\sum_{n=1}^{N_{\max}} k \sum_{(W,E)\in\Upsilon} \frac{(u_E^n - u_W^n)^2}{d_{KL}} m(\sigma) \right)^{\frac{1}{2}} \left(\sum_{n=1}^{N_{\max}} \int_{t_{n-1}}^{t_n} ||u_{h,k}(t-k) - u_{h,k}(t)||_{L_2(\Omega)}^2 dt \right)^{\frac{1}{2}}$$
$$\le C|\Omega|^{\frac{1}{2}} \left(\int_0^T \int_{\Omega} (u_{h,k}(x,t-k) - u_{h,k}(x,t))^2 dx dt \right)^{\frac{1}{2}} \le C|\Omega|^{\frac{1}{2}} k^{\frac{1}{2}},$$

because of (4.7) and the time translate estimate (4.3). The third term is treated similarly, and we get

$$C\sum_{n=1}^{N_{\max}}\sum_{(W,E)\in\Upsilon}|u_E^n-u_W^n|\int\limits_{t_{n-1}}^{t_n}\int\limits_{\sigma}||u_{h,k}(t)-u(t)||_{L_2(\Omega)}dsdt$$

$$\le C|\Omega|^{\frac{1}{2}}\left(\int\limits_0^T\int\limits_\Omega (u_{h,k}(x,t)-u(x,t))^2dxdt\right)^{\frac{1}{2}}\le C|\Omega|^{\frac{1}{2}}||u_{h,k}-u||_{L_2(Q_T)}.$$

For the fourth term we get similarly as above, just using once more Cauchy-Schwartz inequality, that

$$C\sum_{n=1}^{N_{\max}}\sum_{(W,E)\in\Upsilon}|u_E^n-u_W^n|\int\limits_{t_{n-1}}^{t_n}\int\limits_{\sigma}||u(t)||_{L_2(\Omega)}||u(t)-u_{h,k}(t)||_{L_2(\Omega)}dsdt$$

$$\le C|\Omega|^{\frac{1}{2}}\left(\int\limits_0^T ||u(t)||_{L_2(\Omega)}\,||u_{h,k}(t)-u(t)||_{L_2(\Omega)}dt\right)^{\frac{1}{2}}$$

$$\le C|\Omega|^{\frac{1}{2}}\left(\int\limits_0^T\int\limits_\Omega u^2dxdt\right)^{\frac{1}{4}}\left(\int\limits_0^T\int\limits_\Omega (u_{h,k}(x,t)-u(x,t))^2dxdt\right)^{\frac{1}{4}}$$

$$\le C|\Omega|^{\frac{1}{2}}||u_{h,k}-u||^{\frac{1}{2}}_{L_2(Q_T)}$$

because $u\in L_2(Q_T)$ and thus its norm is bounded by a constant. So finally we have

$$|T_4|\le Ch+Ck^{\frac{1}{2}}+C||u_{h,k}-u||_{L_2(Q_T)}+C||u_{h,k}-u||^{\frac{1}{2}}_{L_2(Q_T)}$$

which means that

$$|T_4|\to 0 \text{ as } h,k\to 0. \tag{4.64}$$

The last term is given by $T_5=\int\limits_0^T\int\limits_\Omega \nabla\cdot(D\nabla\varphi(x,t))\,(u(x,t)-u_{h,k}(x,t))dxdt$. We use the property that $D\in C^\infty(R^{2\times 2},R^{2\times 2})$ (see [24], page 115) to state that $\nabla\cdot(D\nabla\varphi(x,t))\in L^2(Q_T)$. Then using strong convergence of $u_{h,k}$ to u one can see that

$$|T_5|\to 0 \text{ as } h,k\to 0. \tag{4.65}$$

Now, we can state following convergence result.

THEOREM 4.9. *The sequence $u_{h,k}$ converges strongly in $L^2(Q_T)$ to the unique weak solution u of (1.1)-(1.3) as $h,k\to 0$.*

Proof. The result comes from (4.46), (4.48), (4.49), (4.64) and (4.65) and the fact that the limit u of $u_{h,k}$ is in space $L^2(0,T;H^1(\Omega))$. Due to uniqueness of the weak solution, which can be found in [24] not only subsequence but also the sequence $u_{h,k}$ itself converges to u. ▯

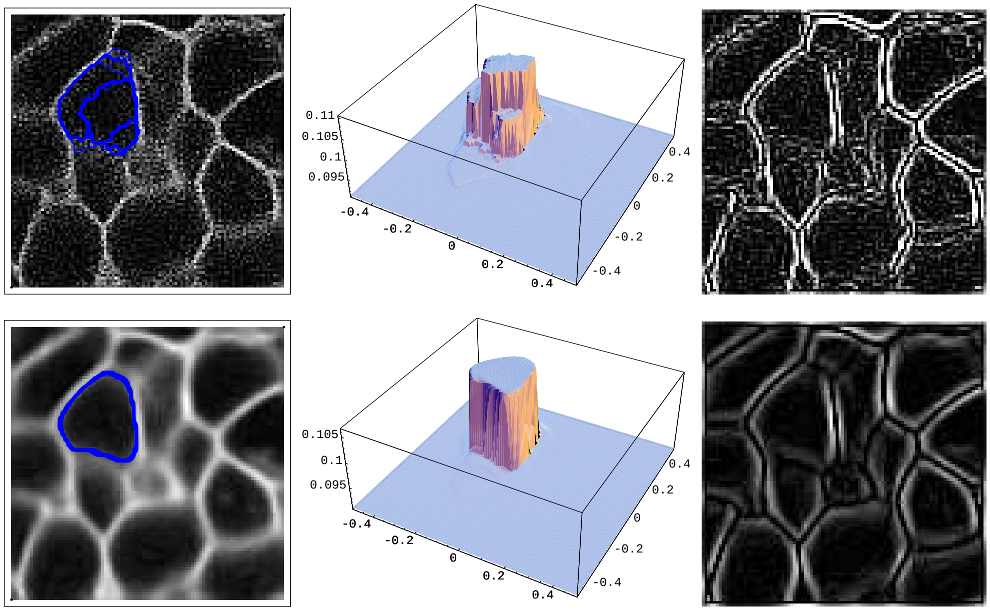

Fig. 5.1. *The cell segmentation and edge detection for the original image and the image filtered by* 4 *diffusion steps. Top row: the original image* (100×100 *pixels detail) with the isolines of the final state of the segmentation function (left), the graph of the final state of the segmentation function (middle), the edge detection for the original image (right). Bottom row: the filtered image with the isolines of the final state of the segmentation function (left), the graph of the final state of the segmentation function (middle), the edge detection for the filtered image (right).*

5. Numerical experiments. In this section we present results of several computational examples using real 2D images coming from multiphoton laser scanning microscopy. They represent the membranes and nuclei of cells in the early stages of zebrafish embryogenesis. Especially the images of membranes are well suited for processing by this type of diffusion which is documented by comparing the edge detection and cell segmentation results before and after filtering in Figures 5.1-5.4. In the experiments we use the spatial step $h = 0.01$, time step $k = 0.0001$, $C = 1$, $\alpha = 0.001$, $\tilde{t} = 0.00001$, $\rho = 0.002$. The arising linear systems are solved using Gauss-Seidel iterations. The satisfactory results were obtained after few filtering steps, so the denoising method is really fast. In the presented experiments we do not observe any stability problems which is a usual drawback of explicit schemes, cf. [25].

The nonlinear tensor anisotropic diffusion smoothes out the noise and improves significantly the connectivity of the coherent structures. Although the filtered image seems to be more blurred compared to the original one, cf. Fig. 5.1 left top and bottom, the enhancement of the structure connectivity and improvement of the quality of edge detection, cf. cf. Fig. 5.1 right top and bottom, enable us to get much more precise results of segmentation algorithms based on image intensity gradient information like the subjective surface method [20, 16, 2]. In the subjective surface method, the initial segmentation function in the form of peak centered in approximate centroid of the segmented object is created. Then the initial function is evolved by nonlinear PDE, it forms a shock profile during the evolution and the segmented object is detected as an isoline of the final (numerically steady) state of the segmentation function, for details we refer to [20, 16]. Since many spurious noisy structures can be seen

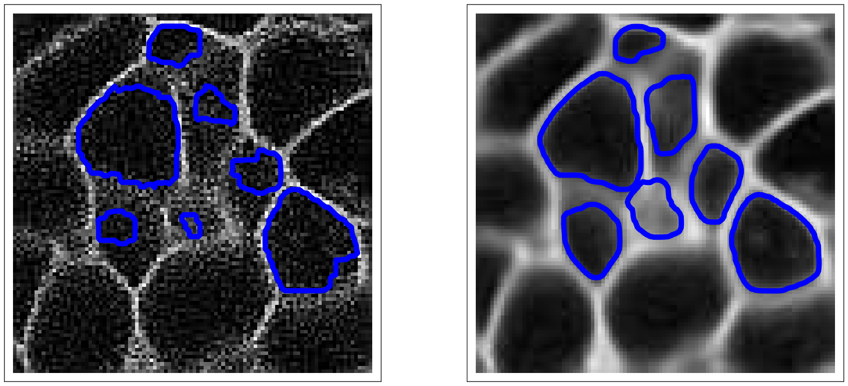

FIG. 5.2. *The results of subjective surface cell segmentation when using unfiltered and filtered images, respectively.*

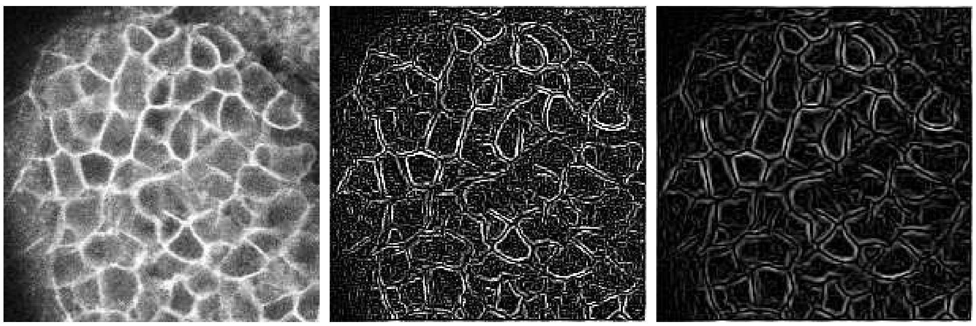

FIG. 5.3. *The image of the cell membranes (*200×200 *pixels, left), its edge detection (middle), and the edge detection for the image filtered by* 2 *diffusion steps where the strong improvement of structure coherence can be seen (right).*

in the original image, which is expressed in highly noisy edge detection result, the segmentation algorithm can hardly find the correct cell boundary using the noisy data. It is difficult to chose proper isoline when several shocks are formed in the irregular steady state which is created due to a noise in the image, cf. Fig. 5.1 left and middle top. On the other hand, using few steps of the nonlinear tensor anisotropic diffusion, all level lines are accumulated along the cell boundary (just one shock is created in the final state of segmentation function), cf. Fig. 5.1 left and middle bottom, and the cell can be segmented precisely. Now, it is easy to choose isoline for the cell boundary representation, we take the average of minimal and maximal values of the final segmentation function, and, in Fig. 5.2 we show the segmentation results for several cells visualizing it both for unfiltered and filtered images. We use the same parameters of the subjective surface segmentation method in both cases and one can see much more precise segmentation results after filtering.

In figures 5.3-5.4 we show two other real images, originals (on the left) and edge detection results for originals (middle) and after few steps of filtering (on the right). Again, one can clearly see coherence enhancement and edge detection improvement.

In the last experiment we test experimental order of convergence (EOC) of our method. In the theoretical part we prove its convergence, the rigorous error estimates will be an objective of a further study, cf. [7]. Here, we consider function $u(x, y, t) =$

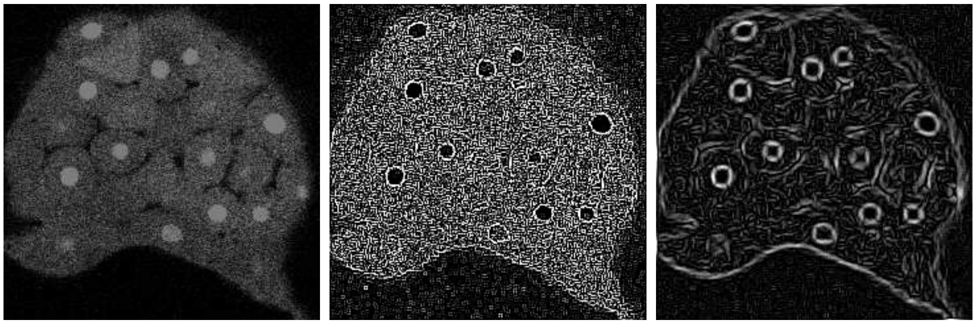

FIG. 5.4. *The image of the cell membranes and nuclei (240×240 pixels, left), its edge detection (middle), and the edge detection for the image filtered by* 5 *diffusion steps.*

TABLE 5.1
Error in $L_2(I, L_2(\Omega))$-norm and EOC comparing numerical and exact solution.

n	h	k	error	EOC
10	0.2	0.04	$1.809572 \cdot 10^{-4}$	
20	0.1	0.01	$0.3835138 \cdot 10^{-4}$	2.2383
40	0.05	0.0025	$0.09159927 \cdot 10^{-4}$	2.06587
80	0.025	0.000625	$0.02238713 \cdot 10^{-4}$	2.03267
160	0.0125	0.00015625	$0.00495121 \cdot 10^{-4}$	2.17682

$t\cos(\pi x)\cos(\pi y)$ which fulfills the boundary conditions in the domain $\Omega = [-1, 1]^2$ and in time interval $I = [0, 1]$. Putting this function into the model equation (1.1), without convolutions, because we do not need to smooth neither the function nor the structure tensor, we get the nonzero right hand side and we modify the scheme accordingly. We take $C = 1$ and $\alpha = 0.001$ so the diffusion matrix D has eigenvalues between α and 1 and the process is strongly anisotropic. Then we take subsequently refined grids with $M = n^2$ finite volumes, $n = 10, 20, 40, 80, 160$, the time step $k = h^2$ and we measure errors in $L_2(I, L_2(\Omega))$-norm, which is natural for testing the schemes for solving parabolic problems. In Table 5.1 we report the errors for different grid sizes and we observe that EOC of our numerical scheme is equal to 2.

6. Acknowledgement. The authors thanks to Dr. Nadine Peyrieras from CNRS-DEPSN, Paris, coordinator of the European project Embryomics, for providing data used in our numerical experiments. We thank also to Prof. Robert Eymard for giving us a hint how to treat T_1 term in the convergence proof. This work was supported by EC projects Embryomics and BioEmergences of the 6th Framework programme and by the projects VEGA 1/3321/06 and APVV-RPEU-0004-06.

REFERENCES

[1] Catté, F., Lions, P. L., Morel, J. M., Coll, T.: Image selective smoothing and edge detection by nonlinear diffusion, SIAM J. Numer. Anal. 129, 182–193, 1992.

[2] Corsaro, S., Mikula, K., Sarti, A., Sgallari, F.: Semi-implicit co-volume method in 3D image segmentation, SIAM J. Sci. Comput., Vol. 28, No. 6, 2248–2265, 2006.

[3] Coudiere, Y., Vila, J. P., Villedieu, P.: Convergence rate of a finite volume scheme for a two-dimensional convection-diffusion problem. M2AN Math. Model. Numer. Anal., 33 , 493–516, 1999.

[4] Eymard, R., Gallouët, T., Herbin, R.: Finite Volume Methods, in : Handbook for Numerical Analysis, Vol. 7 (Ph.Ciarlet, J.L.Lions, eds.), Elsevier, 2000.

[5] Eymard, R., Gallouët, T., Herbin, R.: A cell-centred finite-volume approximation for anisotropic diffusion operators on unstructured meshes in any space dimension. IMA J. Numer. Anal., Vol. 26 (2), 326–353, 2006.

[6] Gilboa, G., Sochen, N., Zeevi, Y.Y.: Forward-and-backward diffusion processes for adaptive image enhancement and denoising. IEEE Trans. Image Processing 11 (7), 689–703, 2002.

[7] Handlovičová, A., Krivá, Z.: Error estimates for finite volume scheme for Perona-Malik equation, Acta Math. Univ. Comenianae, Vol. 74, No. 1, 79–94, 2005.

[8] Handlovičová, A., Mikula, K., Sgallari, F.: Semi-implicit complementary volume scheme for solving level set like equations in image processing and curve evolution. Numer. Math. Vol. 93, No. 4, 675–695, 2003.

[9] Jähne, B.: Spatio-temporal image processing, Lecture Notes in Comp. Science, Vol. 751, Springer, Berlin, 1993.

[10] Krivá, Z.: Explicit finite volume scheme for the Perona-Malik equation, Computational Methods in Applied Mathematics, Vol. 5, No. 2, 170–200, 2005.

[11] Krivá, Z., Mikula, K.: An adaptive finite volume method in processing of color images, Proceedings of ALGORITMY 2000, Conference on Scientific Computing, Podbanské, 174–187, 2000.

[12] Krivá, Z., Mikula, K.: An adaptive finite volume scheme for solving nonlinear diffusion equations arising in image processing, Journal for Visual Communication and Image Representation, Vol. 13, No.1/2, 22–35, 2002.

[13] Lange, C., Polthier, K.: Anisotropic smoothing of point sets. Com. Geom. Design 22 (7), 680–692, 2005.

[14] Mikula, K.: Image processing with partial differential equations, A. Bourlioux and M.J. Gander (eds.), Modern Methods in Scientific Computing and Applications, 283–321, 2002.

[15] Mikula, K., Ramarosy, N.: Semi-implicit finite volume scheme for solving nonlinear diffusion equations in image processing, Numer. Math. 89 (3), 561–590, 2001.

[16] Mikula, K., Sarti, A., Sgallari, F.: Semi-implicit co-volume level set method in medical image segmentation, in Handbook of Biomedical Image Analysis: Segmentation and Registration Models (J.Suri et al., Eds.), Springer, New York, 583–626, 2005.

[17] Perona, P, Malik, J.: Scale space and edge detection using anisotropic diffusion, Proc. IEEE Computer Society Workshop on Computer Vision, 1987.

[18] Preusser, T., Rumpf, M.: An adaptive finite element method for large scale image processing. Journal of Visual Comm. and Image Repres., 11: 183–195, 2000.

[19] Preusser, T., Rumpf, M.: A level set method for anisotropic geometric diffusion in 3D image processing. SIAM J. Appl. Math 62 (5), 1772–1793, 2002.

[20] Sarti, A., Malladi, R., Sethian, J.A.: Subjective Surfaces: A Method for Completing Missing Boundaries, Proceedings of the National Academy of Sciences of the United States of America, Vol. 12, No. 97 (2000) 6258–6263, 2000.

[21] Sarti, A., Mikula, K., Sgallari, F.: Nonlinear multiscale analysis of 3D echocardiographic sequences, IEEE Transactions on Medical Imaging, Vol. 18, No. 6, 453–467, 1999.

[22] Weickert, J.: Anisotropic Diffusion in Computer Vision, Teubner-Stuttgart, 1998.

[23] Weickert, J.: Coherence-enhancing diffusion of colour images. Image and Vision Computing 17, 201–212, 1999.

[24] Weickert, J.: Coherence-enhancing diffusion filtering, Int. J. Comput. Vision, Vol. 31, 111–127, 1999.

[25] Weickert, J., Scharr, H.: A Scheme for Coherence-Enhancing Diffusion Filtering with Optimized Rotation Invariance. Journal of Visual Comm. and Image Repres., 13 (1-2): 103–118, 2002.

[26] Ye, X.: A new discontinuous finite volume method for elliptic problems. SIAM J. Numer. Anal. 42 (3), 1062–1072, 2004.