

Semi-implicit DGM Applied to a Model of Flocking

Andrea Živčáková and Václav Kučera

Abstract We present the numerical solution of a hydrodynamics model of flocking using a suitable modified semi-implicit discontinuous Galerkin method. The investigated model describing the dynamics of flocks of birds or other individual entities forming herds or swarms was introduced by Fornasier et al. (*Physica D* 240(1):21–31, 2011). The main idea of this model comes from the well known Cucker-Smale model. The resulting equations consist of the Euler equations for compressible flow with an additional non-local non-linear source term.

The model is discretized by the semi-implicit discontinuous Galerkin method for the compressible Euler equations of Feistauer and Kučera (*J Comput Phys* 224(1):208–221, 2007). We show that with a suitable treatment of the source term we can use this method for models like the model of flocking and find a numerical solution very efficiently.

1 Continuous Problem

In the paper [4], a hydrodynamic limit of a modification of the famous Cucker-Smale model is derived. The equations describe, using macroscopic quantities, the dynamics of flocks of birds or other self-organizing entities. The equations are highly nonlinear and nonlocal and are therefore extremely expensive to treat numerically, in [4] a first simple simulation was performed using the finite volume method. In this paper, we discretize the model more efficiently using the discontinuous Galerkin method.

A. Živčáková (✉) • V. Kučera

Faculty of Mathematics and Physics, Department of Numerical Mathematics, Charles University in Prague, Sokolovská 83, 186 75 Praha 8, Czech Republic
e-mail: zivcakova@karlin.mff.cuni.cz; kucera@karlin.mff.cuni.cz

Let $\Omega = (0, 1) \subset \mathbb{R}$ and for $0 < L < +\infty$, we set $Q_L := \Omega \times (0, L)$. We treat the following problem: Find $\rho, u, E : Q_L \rightarrow \mathbb{R}$ such that

$$\begin{aligned} \frac{\partial \rho}{\partial t} + \operatorname{div}(\rho u) &= 0, \\ \frac{\partial(\rho u)}{\partial t} + \operatorname{div}(\rho u^2 + p) &= \lambda \mathcal{A}(\rho, u), \\ \frac{\partial E}{\partial t} + \operatorname{div}(u(E + p)) &= \lambda \mathcal{B}(\rho, u, T), \end{aligned} \tag{1}$$

where ρ denotes the density, u velocity, E energy, T temperature and p pressure. The right-hand side functions \mathcal{A} and \mathcal{B} are given by

$$\begin{aligned} \mathcal{A}(\rho, u)(x, t) &= \int_{\mathbb{R}} b(|x - y|) (u(y, t) - u(x, t)) \rho(x, t) \rho(y, t) dy, \\ \mathcal{B}(\rho, u, T)(x, t) &= \int_{\mathbb{R}} b(|x - y|) \rho(x, t) (\rho(y, t) u(x, t) u(y, t) - 2E(y, t)) dy, \end{aligned}$$

where

$$b(|x - y|) = \frac{K}{(\lambda + |x - y|^2)^{\beta+1}}$$

for $K, \lambda > 0$ and $\beta \geq 0$ given constants. The relations between E, p, T are

$$E = \rho \left(\frac{3}{2} T + \frac{1}{2} u^2 \right), \quad p = \rho T.$$

By omitting the right-hand side terms \mathcal{A}, \mathcal{B} from (1), we obtain the compressible Euler equations for a 1D monoatomic gas. In this light, we rewrite system (1) as a system of conservation laws with right-hand side source terms:

$$\frac{\partial \mathbf{w}}{\partial t} + \frac{\mathbf{f}(\mathbf{w})}{\partial x} = \mathbf{g}(\mathbf{w}) \quad \text{in } Q_L, \tag{2}$$

where

$$\begin{aligned} \mathbf{w} &= (\rho, \rho u, E)^\top \in \mathbb{R}^3, \\ \mathbf{f}(\mathbf{w}) &= (f_1(\mathbf{w}), f_2(\mathbf{w}), f_3(\mathbf{w}))^\top = (\rho u, \rho u^2 + p, (E + p)u)^\top, \\ \mathbf{g}(\mathbf{w}) &= \lambda (0, \mathcal{A}(\mathbf{w}), \mathcal{B}(\mathbf{w}))^\top. \end{aligned} \tag{3}$$

The vector-valued function \mathbf{w} is called the *state vector* and the function \mathbf{f} is the so-called *Euler* or *inviscid flux*. In (3), we write the right-hand side terms \mathcal{A}, \mathcal{B} as functions of the state vector \mathbf{w} , although in (1), they are written in terms of the

nonconservative variables. Expressing \mathcal{A}, \mathcal{B} in \mathbf{w} in a suitable way is a key ingredient in our scheme and will be described in Sect. 2.3.1.

The resulting system is equipped with the initial condition

$$\mathbf{w}(x, 0) = \mathbf{w}^0(x), \quad x \in \Omega,$$

and periodic boundary conditions, for simplicity.

The Euler flux is a homogeneous function, which implies

$$\mathbf{f}(\mathbf{w}) = \mathbb{A}(\mathbf{w})\mathbf{w}, \quad (4)$$

where $\mathbb{A} = \frac{D\mathbf{f}}{D\mathbf{w}}$. Furthermore, the Jacobi matrix of the Euler flux is *diagonally hyperbolic*. In 1D this means the matrix

$$\mathbb{P}(\mathbf{w}, n) := \mathbb{A}(\mathbf{w})n$$

is diagonalizable with real eigenvalues, where $n = \pm 1$. I.e. there exists a matrix $\mathbb{T}(\mathbf{w}, n) \in \mathbb{R}^{3,3}$ and a diagonal matrix $\mathbb{D}(\mathbf{w}, n) \in \mathbb{R}^{3,3}$ with eigenvalues $\lambda_1, \lambda_2, \lambda_3 \in \mathbb{R}$ such that

$$\mathbb{P}(\mathbf{w}, n) = \mathbb{T}\mathbb{D}\mathbb{T}^{-1}, \quad \text{where } \mathbb{D}(\mathbf{w}, n) = \text{diag}(\lambda_1, \lambda_2, \lambda_3). \quad (5)$$

2 Discretization

We shall use the multidimensional notation for $\Omega \subset \mathbb{R}^d$, although in our computations we have $d = 1$. Let \mathcal{T}_h be triangulation of Ω and \mathcal{F}_h the system of all faces (nodes in 1D) of \mathcal{T}_h . For each $\Gamma \in \mathcal{F}_h$ we choose a unit normal $n_\Gamma = \pm 1$, which, for $\Gamma \subset \partial\Omega$, has the same orientation as the outer normal to Ω . For each interior face $\Gamma \in \mathcal{F}_h$ there exist two neighbours $K_\Gamma^{(L)}, K_\Gamma^{(R)} \in \mathcal{T}_h$ such that n_Γ is the outer normal to $K_\Gamma^{(L)}$. For v piecewise defined on \mathcal{T}_h and $\Gamma \in \mathcal{F}_h$ we introduce $v|_\Gamma^{(L)}$ is the trace of $v|_{K_\Gamma^{(L)}}$ on Γ , $v|_\Gamma^{(R)}$ is the trace of $v|_{K_\Gamma^{(R)}}$ on Γ and $[v]_\Gamma = v|_\Gamma^{(L)} - v|_\Gamma^{(R)}$ is the *jump* of v . On $\partial\Omega$, we define $v|_\Gamma^{(L)}, v|_\Gamma^{(R)}$ using periodic boundary conditions. If $[\cdot]_\Gamma, v|_\Gamma^{(L)}, v|_\Gamma^{(R)}$ appear in an integral over $\Gamma \in \mathcal{F}_h$, we omit the subscript Γ .

Let $p \in \mathbb{N}$ and let $P^p(K)$ be the space of polynomials on $K \in \mathcal{T}_h$ of degree $\leq p$. The approximate solution will be sought in the space of discontinuous piecewise polynomial functions

$$\mathcal{S}_h := [S_h]^3, \quad \text{where } S_h = \{v; v|_K \in P^p(K), \forall K \in \mathcal{T}_h\}.$$

2.1 Discontinuous Galerkin Space Semidiscretization

To derive the discrete problem, we assume that \mathbf{w} is a classical solution of problem (2). We multiply (2) by a test function $\varphi_h \in S_h$, integrate over $K \in \mathcal{T}_h$ and apply Green's theorem in the convective terms. Summing over all $K \in \mathcal{T}_h$ and rearranging, we obtain

$$\int_{\Omega} \frac{\partial \mathbf{w}}{\partial t} \cdot \varphi \, dx + \int_{\mathcal{F}_h} \mathbf{f}(\mathbf{w}) \mathbf{n} \cdot [\varphi] \, dS - \sum_{K \in \mathcal{T}_h} \int_K \mathbf{f}(\mathbf{w}) \cdot \frac{\partial \varphi}{\partial \mathbf{x}} \, dx = \int_{\Omega} \mathbf{g}(\mathbf{w}) \cdot \varphi \, dx.$$

The discrete approximation of \mathbf{w} will be sought in S_h , we need to give proper meaning to the boundary integral term. Similarly as in the finite volume method, we approximate the physical flux $\mathbf{f}(\mathbf{w}) \mathbf{n}$ through an edge $\Gamma \in \mathcal{F}_h$ by a so-called *numerical flux* $\mathbf{H}(\mathbf{w}^{(L)}, \mathbf{w}^{(R)}, \mathbf{n})$

$$\int_{\mathcal{F}_h} \mathbf{f}(\mathbf{w}) \mathbf{n} \cdot [\varphi] \, dS \approx \int_{\mathcal{F}_h} \mathbf{H}(\mathbf{w}^{(L)}, \mathbf{w}^{(R)}, \mathbf{n}) \cdot [\varphi] \, dS. \quad (6)$$

The specific choice of \mathbf{H} will be discussed in Sect. 2.2.

For $\mathbf{w}, \varphi \in H^1(\Omega, \mathcal{T}_h)$, we can define the following forms.

Convective form:

$$b_h(\mathbf{w}, \varphi) = \int_{\mathcal{F}_h} \mathbf{H}(\mathbf{w}^{(L)}, \mathbf{w}^{(R)}, \mathbf{n}) \cdot [\varphi] \, dS - \sum_{K \in \mathcal{T}_h} \int_K \mathbf{f}(\mathbf{w}) \cdot \frac{\partial \varphi}{\partial \mathbf{x}} \, dx,$$

right-hand side source term form:

$$l_h(\mathbf{w}, \varphi) = - \int_{\Omega} \mathbf{g}(\mathbf{w}) \cdot \varphi \, dx.$$

Finally, we introduce the space semi-discrete problem: We seek $\mathbf{w}_h \in C^1([0, T]; S_h)$ such that for all $\varphi_h \in S_h$ and for all $t \in (0, T)$

$$\frac{d}{dt}(\mathbf{w}_h(t), \varphi_h) + b_h(\mathbf{w}_h(t), \varphi_h) + l_h(\mathbf{w}_h(t), \varphi_h) = 0. \quad (7)$$

2.2 Numerical Flux

The choice of the numerical flux is a very important question in the finite volume and DG schemes. As such, it has been extensively studied from theoretical and practical points of view and many different constructions exist. Here, we will use the *Vijayasundaram* numerical flux, cf. [5], which is suitable for our semi-implicit time

discretization. This numerical flux is based on the flux vector splitting concept, and can be viewed as an extension of the upwind numerical flux to nonlinear systems of equations. We use the diagonal hyperbolicity (5) and define the *positive* and *negative* parts of matrix \mathbb{P} :

$$\mathbb{P}^\pm(\mathbf{w}, n) = \mathbb{T}(\mathbf{w}, n)\mathbb{D}^\pm(\mathbf{w}, n)\mathbb{T}^{-1}(\mathbf{w}, n), \quad \mathbb{D}^\pm(\mathbf{w}, n) = \text{diag}(\lambda_1^\pm, \lambda_2^\pm, \lambda_3^\pm), \tag{8}$$

where $\lambda^+ = \max\{0, \lambda\}$, $\lambda^- = \min\{0, \lambda\}$. Then $\mathbb{P}(\mathbf{w}, n) = \mathbb{P}^+(\mathbf{w}, n) + \mathbb{P}^-(\mathbf{w}, n)$ and we can define the Vijayasundaram numerical flux as

$$\mathbf{H}_{\text{VS}}(\mathbf{w}_L, \mathbf{w}_R, n) = \mathbb{P}^+\left(\frac{\mathbf{w}_L + \mathbf{w}_R}{2}, n\right)\mathbf{w}_L + \mathbb{P}^-\left(\frac{\mathbf{w}_L + \mathbf{w}_R}{2}, n\right)\mathbf{w}_R. \tag{9}$$

Explicit formulas for \mathbb{P} , \mathbb{T} , \mathbb{T}^{-1} and \mathbb{D} can be found e.g. in [3].

2.3 Time Discretization

After choosing some basis of the space \mathcal{S}_h , Eq. (7) represents a system of nonlinear ordinary differential equations, which must be discretized with respect to time. Due to severe time step restrictions, we want to avoid using an explicit scheme. However an implicit time discretization is also very expensive due to its nonlinearity. Therefore we choose the semi-implicit scheme of [2] as a basis and apply it to our problem.

Let $0 = t_0 < t_1 < t_2 < \dots$ be a partition of time interval $[0, T]$ and define $\tau_k = t_{k+1} - t_k$. We approximate $\mathbf{w}_h^k \approx \mathbf{w}_h(t_k)$, where $\mathbf{w}_h^k \in \mathcal{S}_h$. We use a first order backward difference approximation for the time derivative. The resulting scheme reads

$$\left(\frac{\mathbf{w}_h^{k+1} - \mathbf{w}_h^k}{\tau_k}, \boldsymbol{\varphi}_h\right) + b_h(\mathbf{w}_h^{k+1}, \boldsymbol{\varphi}_h) + l_h(\mathbf{w}_h^{k+1}, \boldsymbol{\varphi}_h) = 0, \quad \forall \boldsymbol{\varphi}_h \in \mathcal{S}_h, \tag{10}$$

for all $k = 0, 1, \dots$. Equation (10) is nonlinear with respect to the unknown \mathbf{w}_h^{k+1} , therefore we linearize the scheme.

In the convective form, we linearized the interior terms using the homogeneity (4) as $f(\mathbf{w}_h^{k+1}) \approx \mathbb{A}(\mathbf{w}_h^k)\mathbf{w}_h^{k+1}$. In the boundary terms, we use the Vijayasundaram numerical flux (9) and linearize by taking the matrices \mathbb{P}^+ and \mathbb{P}^- at t_k . Thus we get the linearized convective form

$$\begin{aligned} \tilde{b}_h(\mathbf{w}_h^k, \mathbf{w}_h^{k+1}, \boldsymbol{\varphi}_h) &= - \sum_{K \in \mathcal{T}_h} \int_K \mathbb{A}(\mathbf{w}_h^k)\mathbf{w}_h^{k+1} \cdot \frac{\partial \boldsymbol{\varphi}_h}{\partial x} \, dx \\ &+ \int_{\mathcal{F}_h} \left(\mathbb{P}^+(\langle \mathbf{w}_h^k \rangle, n)\mathbf{w}_h^{k+1, (L)} + \mathbb{P}^-(\langle \mathbf{w}_h^k \rangle, n)\mathbf{w}_h^{k+1, (R)} \right) \cdot [\boldsymbol{\varphi}_h] \, dS. \end{aligned}$$

As for the source terms, they also need to be linearized to obtain the approximation $l_h(\mathbf{w}_h^{k+1}, \boldsymbol{\varphi}_h) \approx \tilde{l}_h(\mathbf{w}_h^k, \mathbf{w}_h^{k+1}, \boldsymbol{\varphi}_h)$. The specific form of this linearization will be derived in the following. Collecting all these considerations, we obtain the semi-implicit DG scheme:

We seek $\mathbf{w}_h^k \in \mathcal{S}_h, k = 0, 1, \dots$, such that for all $\boldsymbol{\varphi}_h \in \mathcal{S}_h$

$$\left(\frac{\mathbf{w}_h^{k+1} - \mathbf{w}_h^k}{\tau_k}, \boldsymbol{\varphi}_h \right) + \tilde{b}_h(\mathbf{w}_h^k, \mathbf{w}_h^{k+1}, \boldsymbol{\varphi}_h) + \tilde{l}_h(\mathbf{w}_h^k, \mathbf{w}_h^{k+1}, \boldsymbol{\varphi}_h) = 0. \quad (11)$$

Equation (11) represents a linear equation for the unknown \mathbf{w}_h^{k+1} . If we choose a basis of the space \mathcal{S}_h consisting of functions whose support is exactly one element, we can rewrite Eq. (11) as a system of linear algebraic equations for the coefficients of \mathbf{w}_h^{k+1} in the chosen basis. If $\tilde{l}_h \equiv 0$ (i.e. we solve the Euler equations), by grouping together basis functions with a common supporting element, the structure of the system matrix is block-tridiagonal with lower-left and upper-right corner blocks corresponding to the periodic boundary conditions. Such systems can be efficiently solved e.g. by a direct solver, in our case UMFPACK, [1].

2.3.1 Linearization of the Source Terms l_h

First, it is necessary to rewrite the right-hand side integrals \mathcal{A}, \mathcal{B} in terms of \mathbf{w} . For \mathcal{A} , we obtain

$$\begin{aligned} \mathcal{A} &= \int_{\mathbb{R}} b(|x-y|) \left(\underbrace{\rho(x,t)}_{w_1(x,t)} \underbrace{\rho(y,t)u(y,t)}_{w_2(y,t)} - \underbrace{\rho(y,t)}_{w_1(y,t)} \underbrace{\rho(x,t)u(x,t)}_{w_2(x,t)} \right) dy \\ &= \int_{\mathbb{R}} b(|x-y|) \left(w_1(x,t)w_2(y,t) - w_1(y,t)w_2(x,t) \right) dy \\ &= \int_{\mathbb{R}} b(|x-y|) \mathbf{w}(x,t) \cdot (w_2(y,t), -w_1(y,t), 0) dy. \end{aligned}$$

Similarly, we can write \mathcal{B} as

$$\begin{aligned} \mathcal{B} &= \int_{\mathbb{R}} b(|x-y|) \left(\underbrace{\rho(x,t)u(x,t)}_{w_2(x,t)} \underbrace{\rho(y,t)u(y,t)}_{w_2(y,t)} - 2 \underbrace{\rho(x,t)}_{w_1(x,t)} \underbrace{E(y,t)}_{w_3(y,t)} \right) dy \\ &= \int_{\mathbb{R}} b(|x-y|) \left(w_2(x,t)w_2(y,t) - 2w_1(x,t)w_3(y,t) \right) dy \\ &= \int_{\mathbb{R}} b(|x-y|) \mathbf{w}(x,t) \cdot (-2w_3(y,t), w_2(y,t), 0) dy. \end{aligned}$$

Therefore, we can rewrite the vector $\mathbf{g}(\mathbf{w})$ as

$$\mathbf{g}(\mathbf{w})(x, t) = \lambda \int_{\mathbb{R}} b(|x - y|) \mathbb{U}_2(\mathbf{w}(y, t)) \mathbf{w}(x, t) dy, \quad (12)$$

where $\mathbb{U}_2(\mathbf{w}) \in \mathbb{R}^{3 \times 3}$ is the matrix

$$\mathbb{U}_2(\mathbf{w}) = \begin{pmatrix} 0 & 0 & 0 \\ w_2 & -w_1 & 0 \\ -2w_3 & w_2 & 0 \end{pmatrix}.$$

If we approximate $\mathbf{w}(x, t) \approx \mathbf{w}_h^{k+1}(x)$ and $\mathbf{w}(y, t) \approx \mathbf{w}_h^k(y)$, we obtain the linearized form

$$\tilde{l}_h(\mathbf{w}_h^k, \mathbf{w}_h^{k+1}, \boldsymbol{\varphi}_h) = \int_{\mathbb{R}} \left(\int_{\mathbb{R}} b(|x - y|) \mathbb{U}_2(\mathbf{w}_h^k(y)) dy \right) \mathbf{w}_h^{k+1}(x) \cdot \boldsymbol{\varphi}_h(x) dx. \quad (13)$$

For a basis for S_h formed by functions whose support is only one element, adding (13) does not change the structure of the system matrix, since it contributes only to the block-diagonal. This is important, since other expressions than (12) are possible, however they lead to a full system matrix.

The computation of (13) is very time consuming due to the nonlocal nature. Even if the basis functions of S_h are local, in order to evaluate \tilde{l}_h , we must compute the inner integral $\int_{\mathbb{R}} b(|x - y|) \mathbb{U}_2(\mathbf{w}_h^k(y)) dy$, which is expensive due to the slow decay of the function $b(|x - y|)$. We note that in our implementation, we do not compute this integral over the whole of \mathbb{R} , but only over one periodically taken copy of Ω centered at point x .

3 Numerical Experiment

In this numerical experiment, we set the initial density to have a Gaussian distribution $\rho(x) = \exp(-10(x - 0.5)^2)$. The temperature is taken constant, $T = 10$, and the velocity is given by $u(x) = -\sin(2\pi x)$. We used 400 piecewise quadratic elements. We observed the formation of a sharp peak in the density, as seen in Fig. 1. Due to the discontinuities in the solution, artificial diffusion was added, as described in [2]. Furthermore, in large regions of Ω , a state close to *vacuum* occurs, i.e. $\rho \approx 0$, $T \approx 0$. In fact, the minimum density and temperature over Ω seems to decay exponentially, cf. Fig. 2. To avoid this complication, at each time step, the \mathbf{w}_h^k was *postprocessed* to avoid the vacuum state. Specifically, if $\rho < \varepsilon$, then set $\rho := \varepsilon$ and recompute the energy, so that $T > \varepsilon$, where $\varepsilon := 10^{-5}$ in our case. A uniform time step $\tau = 10^{-3}$ was chosen as a balance between discretization error in time and computational efficiency.

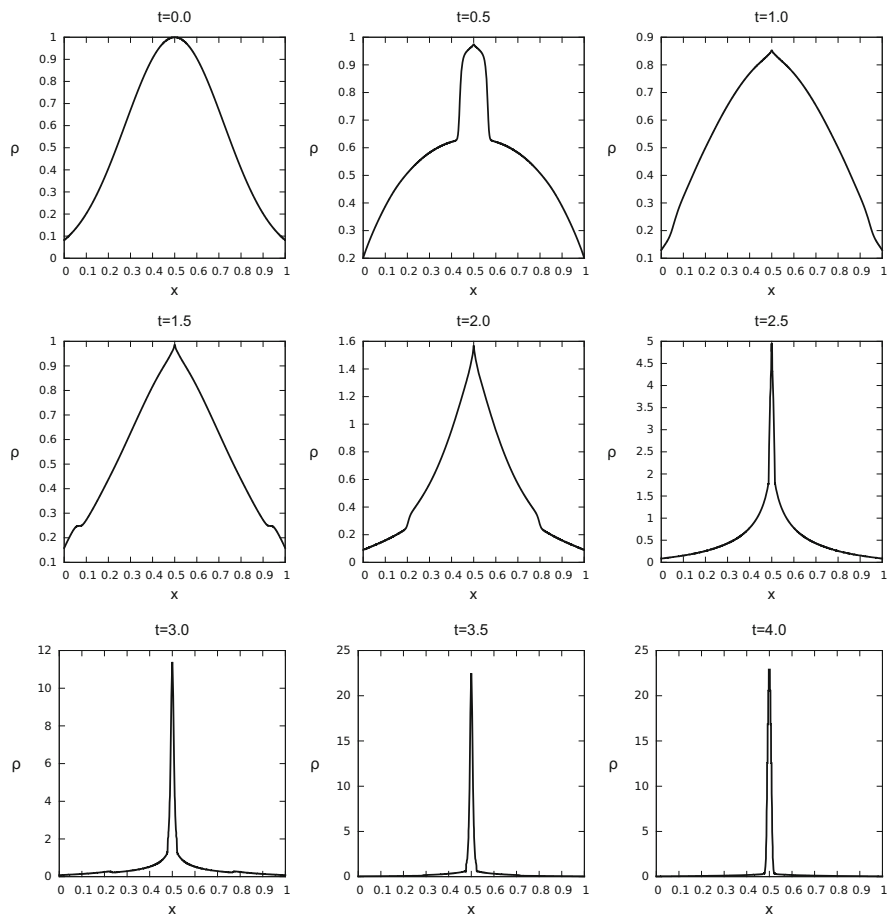


Fig. 1 Time evolution of the density distribution

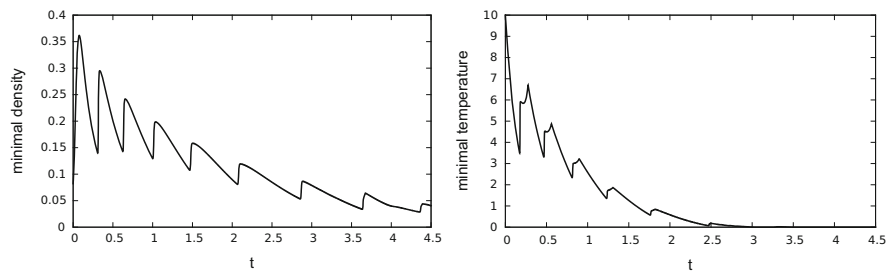


Fig. 2 Time evolution of minimal density and temperature

4 Conclusion

We have presented an efficient numerical method for the solution of a complicated nonlinear and nonlocal version of the compressible Euler equations describing the dynamics of flocks of birds, cf. [4]. To avoid severe time step restrictions and consequently the need to evaluate the expensive nonlocal terms too many times, a semi-implicit discontinuous Galerkin scheme is applied. A suitable treatment of the nonlocal terms is given, which leads to sparse matrices. Shock capturing and postprocessing of vacuum must be added to obtain a stable scheme.

Acknowledgements The research of V. Kučera is supported by the Grant No. P201/13/00522S of the Czech Science Foundation. He is currently a Fulbright visiting scholar at Brown University, Providence, RI, USA, supported by the J. William Fulbright Commission in the Czech Republic. The research of A. Živčáková is supported by the Charles University in Prague, project GA UK No. 758214.

References

1. T.A. Davis, I.S. Duff, A combined unifrontal/multifrontal method for unsymmetric sparse matrices. *ACM Trans. Math. Softw.* **25**, 1–19 (1999)
2. M. Feistauer, V. Kučera, On a robust discontinuous Galerkin technique for the solution of compressible flow. *J. Comput. Phys.* **224**, 208–221 (2007)
3. M. Feistauer, J. Felcman, I. Straškraba, *Mathematical and Computational Methods for Compressible Flow* (Clarendon Press, Oxford, 2003)
4. M. Fornasier, J. Haškovec, G. Toscani, Fluid dynamic description of flocking via Povzner-Boltzmann equation. *Physica D* **240**(1), 21–31 (2011)
5. G. Vijayasundaram, Transonic flow simulation using upstream centered scheme of Godunov type in finite elements. *J. Comput. Phys.* **63**, 416–433 (1986)