POLITECNICO DI TORINO
Repository ISTITUZIONALE

An integro-differential non-local model for cell migration and its efficient numerical solution

Original
An integro-differential non-local model for cell migration and its efficient numerical solution / Colombi, A., Falletta, S., Scianna, M., Scuderi, L.. - In: MATHEMATICS AND COMPUTERS IN SIMULATION. - ISSN 0378-4754. - 180:(2021), pp. 179-204. [10.1016/j.matcom.2020.08.020]

Availability:
This version is available at: 11583/2845082 since: 2021-10-05T16:10:30Z

Publisher:
Elsevier

Published
DOI:10.1016/j.matcom.2020.08.020

Terms of use:

This article is made available under terms and conditions as specified in the corresponding bibliographic description in the repository

Publisher copyright

(Article begins on next page)

26 November 2026

An integro-differential non-local model for cell migration and its efficient numerical solution

Annachiara Colombi[a,*], Silvia Falletta[b], Marco Scianna[b], Letizia Scuderi[b]

[a] *Istituto Nazionale di Alta Matematica "F. Severi" - INdAM Reseach Unit at Politecnico di Torino, Corso Duca degli Abruzzi 24, 10129, Torino, Italy*
[b] *Department of Mathematical Sciences "G. L. Lagrange", Politecnico di Torino, Corso Duca degli Abruzzi 24, 10129, Torino, Italy*

Abstract

Cell migration is fundamental in a wide variety of physiological and pathological phenomena, being exploited in biomedical engineering as well. In this respect, we here present a hybrid non-local integro-differential model where a representative cell, reproduced by a point particle with an orientation, moves on a planar domain upon signals coming from environmental variables. From a numerical point of view, non-locality implies the need to evaluate integral terms which may present non-regular integrand functions because of heterogeneities in the environmental conditions and/or in cell sensing region. Having in mind multicellular applications, we here propose a robust computational method able to handle such non-regularities. The procedure is based on low order Runge-Kutta methods and on an *ad hoc* application of the Gauss-Legendre quadrature rule. The accuracy and efficiency of the resulting computational method is then tested by selected benchmark settings. In this context, the *ad hoc* application of the quadrature rule reveals to be crucial to obtain a high accuracy with a remarkably low number of quadrature nodes with respect to the *standard* Gauss-Legendre quadrature formula, and which thus results in a reduced overall computational cost. Finally, the proposed method is further coupled with the cubic spline interpolation scheme which allows to deal also with possible poor (i.e., point-wise defined) molecular spatial information. The performed simulations (which accounts also for different scenarios) show how the interpolation of the molecular variables affects the efficiency of the overall method and further justify the proposed

*Corresponding author: annachiara.colombi@polito.it

procedure.

Keywords: cell migration, non-local integro-differential model, quadrature rules

1. Introduction

The migratory behavior of cells, considered either as single isolated individuals or as part of a collective systems, is fundamental in a wide range of processes. During embryogenesis, the coordinated movement of cells from different origins is crucial for organogenesis [21, 29]. In adult organisms, cell migration is relevant, for instance, during inflammation or wound healing [21]. Pathologies such as artheriosclerosis and cancer involve several aspects of cell movement as well [40].

Cell motile behavior is determined by a complex set of interconnected mechanisms, which act at different levels, see [23] and references therein. For instance, cells respond to signals coming from the external environment, i.e., from other individuals or from the surrounding tissue, according to their specific biophysical determinants (such as elasticity, sensitivity, adhesiveness). In particular, a critical role in guiding cell migration is played by molecular substances which can be either diffusive or not. The former group includes growth factors that behave as chemoattractants or chemorepellents, thereby establishing a sort of cell directional cue. The latter group includes extracellular matrix proteins (such as collagen, glycoproteins and glycosaminoglycans) which are organized in almost insoluble complex networks that provide microstructural support for moving individuals [44]. In this respect, each cell is able to sense the presence and the distribution of both fixed and diffusive chemicals (and therefore to behave accordingly) in a given region around its actual position, thanks to the activity of proper receptors clustered in selected areas of its plasmamembrane (PM).

Such *extracellular-to-intracellular* signaling mechanisms are here reproduced by a hybrid model in the case of a single representative agent. In particular, the moving individual is identified by a material point, whereas the environmental molecular substances are given by a density (in the case of ECM proteins) or by a concentration (in the case of a diffusive growth factor). Our approach can be therefore defined as *multiscale*, since it spans both the cellular and the subcellular levels (as, among others, in [3, 10, 14, 33]). Entering in more details, cell dynamics are given by a first-order ordinary

differential equation (ODE, see [6, 17, 41] and reference therein), which distinguishes individual speed and direction of locomotion (identified by a polarization vector, as in [37]). The former quantity is assumed to depend on the amount of extracellular proteins, the latter is instead established by the distribution of the chemical factor. In both cases, it is indeed fundamental to quantify the pattern of the molecular substances as *perceived* by the cell in its surrounding (i.e., *non-locally*). To this end, both cell motility and orientation are here determined by proper integrals that involve the local quantity of ECM proteins and of soluble chemical substances, spatially weighted to account for anisotropic aspects, and are evaluated over given cell sensing regions (as, for instance, in [4, 9, 15, 46]).

The numerical evaluation of such integrals may not be trivial. In fact, the integrand functions may loose regularity both because of non-regularities in the distributions of the molecular elements and/or in the profile of the weight functions characterizing cell sensing region. Further, the time discretization of cell motion equations requires to compute these integrals at each time step, thereby strongly affecting the overall computational cost. Moreover, these numerical issues can in principle affect the accuracy of the approximated cell trajectory, and therefore the validity of the modeling approach, decreasing its realism.

In this work, we propose to tackle these problems by combining an explicit classical solver for ODEs with an efficient quadrature formula which is able to handle possible non-regularities of the integrand functions. Among all the possibilities, we consider low order Runge-Kutta schemes (in particular the explicit Euler and the Heun method), for the solution of the ODEs, and a tailored (*ad hoc*) application of the Gauss-Legendre quadrature formula to each subinterval where the integrand function is actually regular (rather than its direct application to the entire domain of integration, as established by the *standard* approach).

The use of the resulting numerical framework is however impossible if the value of molecular variables is unknown at the quadrature nodes (involved in the evaluation of the integral terms). Such a situation may occur, for instance, when the concentrations of the diffusing chemical and the distribution of the ECM proteins are given by experimental measurements taken at specific environmental points, or when they are numerically computed solving proper evolution laws for the environmental factors over a tensor grid that does not match the discretization of the sensing region. To handle also these situations, it is sufficient to couple the proposed numerical approach

with a proper interpolating method. A possible choice that reasonably combines (in terms of accuracy and computational cost) with the Runge-Kutta methods and the tailored Gauss-Legendre quadrature formula is the spline interpolation.

The rest of the paper is then organized as it follows. In Section 2, we will present the modeling framework: in particular, we will introduce and comment the numerical problems that will be first addressed in Section 3, where the proposed computational method will be described in details. In Section 4, the accuracy and the efficacy of our numerical technique will be investigated by dealing with a selected benchmark test. Section 5, will be then devoted to the inclusion of the interpolation method and to a series of simulations involving several patterns of matrix proteins and chemical substance. Conclusive remarks and some hints for the development of the proposed approach, both from a modeling and from a numerical point of view, will be given in Section 6. A discussion of the existence and uniqueness of the solution of the proposed model, and its dependence on the initial condition is finally reported in the Appendix.

2. Mathematical model

Having in mind the application to multicellular situations, the proposed model focuses on the migratory behavior of a single cell represented by a discrete point-wise element and guided by external signals. Given a simplified bidimensional setting, the individual position at time $t \in \mathbb{R}_0^+$ (being $\mathbb{R}_0^+$ the set of not negative real numbers) is identified by the vector $\mathbf{x}(t) \in \mathbb{R}^2$. Its dimensions are instead taken into account by a coefficient R which is set to identify a maximum extension of its plasmamembrane protrusions.

We then differentiate the concepts, and the underlying mechanisms, of cell *polarization* and *motility*, both fundamental for individual dynamics. The former is related to the process of organization of cell cytoskeletal filaments, which are able to align in response to internal (e.g., genetic [20]) and/or external inputs (e.g., diffusible chemicals [24]) and therefore to identify a preferred axis of the individual body, eventually establishing a direction of locomotion. The motility of a cell instead designates the characteristic speed of its movements, which depends on the frequency of retraction/expansion cycles of its plasmamembrane motility structures, such as filopodia and pseudopodia. Such a plasmamembrane ruffling is in turn highly controlled both by intracellular cascades involving specific ions and molecules (e.g., calcium,

Rac1, Rho, see [35, 36, 39] and references therein) and by the presence and the density of extracellular matrix proteins (ECM), such as collagen, elastin and laminin, which are grabbed and used as "climber hand anchors" by the moving individual (through integrin-based adhesion sites, see [22, 26]). Taking all these considerations into account, the velocity of our representative cell can be defined by

$$\frac{d\mathbf{x}(t)}{dt} = v(t)\,\mathbf{w}(t) =: \mathbf{g}_{\mathbf{x}}(t,\,\mathbf{x}(t),\,\mathbf{w}(t)), \tag{1}$$

where the scalar $v \in [0, V_{\max}]$ is its actual speed (being $V_{\max}$ a maximal plausible value), while $\mathbf{w} \in \mathbb{R}^2$, such that $\|\mathbf{w}\| \leq 1$ (being $\|\cdot\|$ the Euclidean norm in $\mathbb{R}^2$), is its polarization vector which, as seen, identifies also the direction of motion.

From Eq. (1), it can be easily seen that the cell cannot move either if its motility v is null nor if it is not polarized (i.e., if $\mathbf{w} = \mathbf{0}$). Further, a partial polarization (i.e., if $\|\mathbf{w}\| < 1$) results in a slow down of cell displacement: if the cytoskeleton is not fully organized, the effective migratory capacity of the individual is in fact downregulated, since it is subjected to inputs coming from different directions that may inhibit each other. A random term may be easily added in Eq. (1) to account for isotropic Brownian crawling, typical of biological individuals. However this effect will be not considered here.

For the sake of simplicity, we hereafter assume that cell polarization processes are regulated only by the activity of an extracellular diffusive substance, which is sensed by specific cell membrane receptors and acts as a directional *chemoattractant* triggering a proper alignment of cell cytoskeletal filaments. On the other hand, the cell intrinsic motility is set to depend only on the density (and the distribution) of a fixed substratum of ECM proteins. In mathematical terms, we have indeed that

$$\mathbf{w}(t) = \mathbf{w}(t;\,c) \qquad \text{and} \qquad v(t) = v(t;\,m), \tag{2}$$

where the functions $c : \mathbb{R}_0^+ \times \mathbb{R}^2 \mapsto \mathbb{R}_0^+$ and $m : \mathbb{R}^2 \mapsto [0, 1]$ define the local amount of the diffusive chemical (for each instant of time t) and the local fraction of the fixed extracellular proteins, respectively. Specifically, as done in [2], it is reasonable to assume that m has a piecewise constant profile since transitions between soft and stiff biological substrates are typically sharp. Therefore, for simplicity, we hereafter deal with piecewise constant distributions m of ECM proteins.

2.1. Cell polarization

Entering in more details, the polarization vector of the representative cell evolves according to

$$\frac{d\mathbf{w}(t)}{dt} = \frac{1}{\tau}\left(\frac{\mathbf{W}(t,\, \mathbf{x}(t),\, \mathbf{w}(t))}{\|\mathbf{W}(t,\, \mathbf{x}(t),\, \mathbf{w}(t))\| + \chi} - \mathbf{w}(t)\right) =: \mathbf{g}_{\mathbf{w}}(t,\, \mathbf{x}(t),\, \mathbf{w}(t)), \quad (3)$$

where $\mathbf{W} \in \mathbb{R}^2$ is the direction along which its cytoskeletal filaments are actually induced to align by the distribution of the external molecular signal (as will be defined by Eq. (4) below). The parameter $\tau > 0$ is instead a sort of persistence time, i.e., the time needed by the cell to re-orient. In Eq. (3), we use a Michaelis-Menten law, with $0 < \chi \ll 1$, to assure that the modulus of the forcing term is smaller than one.

The cell preferred axis $\mathbf{W}$, as well as the direction of motion $\mathbf{w}$, is indeed established by the spatial distribution of the diffusive substance c, as perceived by the individual. In particular, the cell is coherently assumed to be able to sense the extracellular chemotactic stimuli only within a bounded neighborhood, which is reasonably contained in a circular area $\mathcal{S}$ centered at $\mathbf{x}$ with radius R (being, as seen, R the maximum extension of the individual PM protrusions). In this respect, we define:

$$\begin{aligned}
&\mathbf{W}(t,\, \mathbf{x},\, \mathbf{w}) = \\
&= \int_{\mathcal{S}(\mathbf{x})} K\left(\|\mathbf{y}-\mathbf{x}\|,\, \frac{\mathbf{y}-\mathbf{x}}{\|\mathbf{y}-\mathbf{x}\|}\cdot\frac{\mathbf{w}}{\|\mathbf{w}\|}\right) c(t,\mathbf{y})\, \frac{\mathbf{y}-\mathbf{x}}{\|\mathbf{y}-\mathbf{x}\|}\, d\mathbf{y} \\
&= \int_0^R r \int_0^{2\pi} K\left(r,\, \mathbf{n}(\theta)\cdot\frac{\mathbf{w}}{\|\mathbf{w}\|}\right) c\left(t, \mathbf{x} + r\mathbf{n}(\theta)\right) \mathbf{n}(\theta)\, d\theta\, dr,
\end{aligned} \quad (4)$$

where $\mathbf{n}(\theta) = (\cos\theta, \sin\theta)$ for any $\theta \in [0, 2\pi]$, see Fig. 1 (panel A). In Eq. (4), the scalar kernel $K : [0, R]\times[-1, 1] \mapsto [0, 1]$ is a weight function that measures the capacity of the cell to sense the diffusive substance c in a non-local and non-isotropic way, since its support actually identifies the individual *sensing region*. Specifically, we hypothesize that the local amount of chemical receptors of the representative individual: (i) decreases with the distance from its center of mass (identified by the positional vector $\mathbf{x}$) and (ii) is higher towards its front/head (identified by the polarization vector $\mathbf{w}$) than towards its rear/tail. In mathematical terms, such considerations, along with

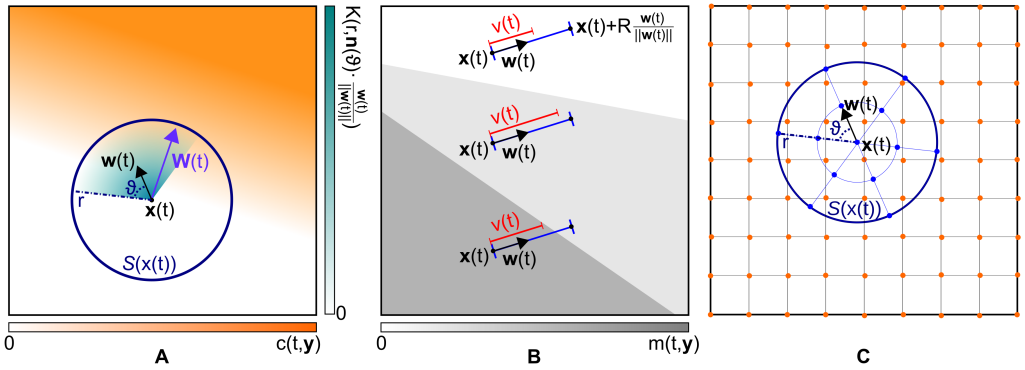

Figure 1: Panel A: Cell polarization process (i.e., the actual polarization vector $\mathbf{w}$ and the preferred axis $\mathbf{W}$) is regulated only by the spatial distribution c of an extracellular diffusive substance (see the orange shadow) within cell sensing region $\mathcal{S}(\mathbf{x}(t))$ (see the blue circle), as perceived by the individual according to the weight function K (see the green shadow), cf. Eqs. (3)-(6). Panel B: Cell intrinsic motility (i.e., $v(t)$, see the red segments) depends only on the spatial distribution $m(t, \mathbf{y})$ of ECM proteins (see the grey shadow) non-locally perceived at its leading front (i.e., over the blue segments joining $\mathbf{x}$ and $\mathbf{x} + R\frac{\mathbf{w}}{||\mathbf{w}||}$), cf. Eqs. (7)-(9). Panel C: Representative case where the distribution of c is defined only over a finite amount of points (i.e., the orange dots) which do not coincide with the quadrature nodes (i.e., the blue dots) required to evaluate $\mathbf{W}$.

an independence assumption of the two variations, suggest to choose

$$K\left(r,\, u\right) = K_r\left(r\right)\, K_\theta\left(u\right), \tag{5}$$

being $K_r : [0, R] \mapsto [0, 1]$ a not increasing law, and $K_\theta : [-1, 1] \mapsto [0, 1]$ an even function, which is not increasing over the interval $[0, 1]$. Eq. (4) can be therefore rewritten as

$$\begin{aligned} &\mathbf{W}(t,\, \mathbf{x},\, \mathbf{w}) = \\ &= \int_0^R r\, K_r\left(r\right) \int_0^{2\pi} K_\theta\left(\mathbf{n}(\theta)\cdot\frac{\mathbf{w}}{\|\mathbf{w}\|}\right)\, c\left(t, \mathbf{x} + r\mathbf{n}(\theta)\right)\, \mathbf{n}(\theta)\, d\theta\, dr, \end{aligned} \tag{6}$$

where the weight function components, K_r and K_θ, and the chemical profile, c, should be suitably defined to result in a Riemann integrable integrand function. In this respect, for the sake of simplicity, we hereafter assume that the integrand functions in Eq. (6) are piecewise regular.

2.2. Cell motility

The effect of the ECM substrate on the behavior of the representative cell is finally included by the following constitutive law for its motility:

$$v(t) = v(t; m) = V_{\max}\, f(t; m), \tag{7}$$

where, as already defined, $V_{\max}$ is a maximal cell speed. In particular, the scalar function $f : \mathbb{R}_0^+ \times [0, 1] \mapsto [0, 1]$ measures the amount of matrix components non-locally (i.e., within its sensing region) perceived by the individual through specific adhesive receptors (i.e., integrins) that are typically clustered at its actual leading front (again identified by the polarization vector $\mathbf{w}$). Among all the possible choices, we define

$$f(t; m) = 4\, M(t,\, \mathbf{x}(t),\, \mathbf{w}(t))\, \left(1 - M(t,\, \mathbf{x}(t),\, \mathbf{w}(t))\right)_+, \tag{8}$$

where $(\,\cdot\,)_+$ stands for the positive part of the quantity within the brackets, and M, which is given by

$$M(t,\, \mathbf{x}(t),\, \mathbf{w}(t)) = \frac{1}{R} \int_0^R m\left(\mathbf{x}(t) + r\, \frac{\mathbf{w}(t)}{\|\mathbf{w}(t)\|}\right) dr, \tag{9}$$

measures the mean ECM ahead up to the maximum protrusion distance R, see Fig. 1 (panel B). Specifically, Eq. (8) gives a bimodal relation between cell motility and the amount of sensed ECM. This assumption is consistent since at low density of matrix elements (i.e., $M \approx 0$), the cell is unable to find sufficient collagen-like sites to hold onto and to use for traction, being therefore unable to significantly displace [22, 26]. At the other extremum, an abundance of ECM proteins (i.e., $M \approx 1$) typically leads to the formation of stable focal adhesions and, hence, low detachment and migration rates, as provided by the experimental literature [19, 25]. Intermediate amounts of ECM densities instead result in optimal attachment-detachment cycles and maximal cell speed [8, 38].

Summarizing, the model in Eqs. (1) and (3), with the closure assumptions in Eqs. (6) and (7)-(9), is a system of non-linear integro-differential equations, which entail the following numerical issues.

(i) The accurate numerical approximation of the integrals $\mathbf{W}$ in Eq. (6) and M in Eq. (9) requires suitable quadrature formulas, i.e., able to

handle possible non-regularities of the integrand functions (which depends on the concentration c of the diffusive chemical, on the weight functions K_r and K_θ, and on the distribution m of the fixed ECM proteins, respectively). As a remark, according to the above assumptions, the integrand functions in Eqs. (6) and (9) (i.e., $r\,K_r\,K_\theta\,c$ and m) are piecewise regular and piecewise constant, respectively, and therefore Riemann integrable.

(ii) The time integration of the proposed model involves at each time step the evaluation of non-linear and integral terms at the r.h.s. of Eqs. (1) and (3). Accounting also for point (i), it is thus important to identify a numerical scheme that minimizes both the number of evaluations of $\mathbf{W}$ and M per instant of time and the total amount of time steps, being therefore characterized by a sufficiently fast convergence and a good degree of accuracy.

(iii) Furthermore, if the distributions c and m are defined, at any $t \in [0, T]$, only over a finite amount of points (e.g., if they are experimental measurements taken at selected places), then the values of the molecular variables at the quadrature nodes (required to evaluate $\mathbf{W}$ and M, see point (i)) may be unknown, see Fig. 1 (panel C). In this case, it is necessary to use a suitable interpolation method to approximate the missing data, preserving the accuracy and efficiency of the overall numerical scheme.

In order to handle all these numerical issues, we will first propose, in Section 3, a numerical approach that takes into account points (i) and (ii), i.e., based on the simplified assumption that c ad m are known in analytical form. The proposed approach will be then applied, in Section 4, to a selected benchmark test, in order to show that it is effective to obtain an accurate solution of the problem with a low computational cost. In Section 5, we will finally add and couple to the proposed method an interpolation scheme useful to address issue (iii). The overall computational approach will be finally applied to selected biological settings.

3. Numerical approach

In this section, we present a numerical procedure to solve the system of non-linear integro-differential equations in Eqs. (1) and (3), that is able to

take into account the numerical issues (i) and (ii) highlighted in the previous section. As a remark, we here assume that the distributions c and m are known, for any instant time $t \in [0, T]$, at any point $\mathbf{y} \in \mathbb{R}^2$ (i.e., they are known in analytical form).

The main idea of the proposed approach is to tackle point (ii) with an explicit solver for ODEs, able to handle the non-linear integral forcing terms (i.e., $\mathbf{g_x}$ in Eq. (1) and $\mathbf{g_w}$ in Eq. (3)), and to address point (i) with an efficient quadrature formula that accounts for regularity properties of the integrand functions of $\mathbf{W}$ in Eq. (6) and of M in Eq. (9).

In particular, to tackle point (ii), for the solution of the ODEs, we discard implicit methods due to the presence of non-linear integral forcing terms, and we consider the first and second-order explicit Runge-Kutta methods, known as Euler and Heun method respectively (see [30]). By a comparison in terms of efficiency, we will show that the Heun method turns out to be a reasonable choice, being it sufficiently accurate with an acceptable computational cost.

On the other hand, for the numerical integration of the piecewise regular functions involved in integrals $\mathbf{W}$ and M (i.e., to deal with point (i)), we propose a tailored application of the Gauss-Legendre formula, consisting of applying the latter in each subregion of the integration domains ($\mathcal{S}$ and $[0, R]$, respectively) where the integrand functions are actually smooth. Having in mind that the accuracy of a Gaussian quadrature rule depends on the smoothness of the integrand function (i.e., the higher the regularity of the integrand function is, the faster the convergence of the quadrature formula is), we expect that such *ad hoc* approach turns out to be more accurate and efficient than a *standard* application of the quadrature formula to the entire domain of integration. In the following, we will refer to the proposed tailored/*ad hoc* strategy by using the acronym "hoc", while the other one, i.e., the *standard* approach, will be denoted by the acronym "std".

3.1. Time discretization with Runge-Kutta methods

In order to apply a Runge-Kutta method, Eqs. (1) and (3) are first rewritten in the following canonical form

$$\begin{cases} \dfrac{d\mathbf{z}(t)}{dt} &= \mathbf{F}(t,\, \mathbf{z}(t)), \quad t \in (0, T]; \\ \mathbf{z}(0) &= \mathbf{z}_0, \end{cases} \tag{10}$$

by setting $\mathbf{z}(t) := (\mathbf{x}(t),\, \mathbf{w}(t))^T$ and $\mathbf{F}(t,\, \mathbf{z}(t)) := (\mathbf{g}_{\mathbf{x}}(t,\, \mathbf{x}(t),\, \mathbf{w}(t)),\, \mathbf{g}_{\mathbf{w}}(t,\, \mathbf{x}(t),\, \mathbf{w}(t)))^T$, where, as seen, $\mathbf{g}_{\mathbf{x}}$ and $\mathbf{g}_{\mathbf{w}}$ denote the right hand sides of Eqs. (1) and (3), respectively. We then introduce a uniform partitioning of the time interval $(0,\, T]$ into N subintervals of equal length $\Delta_t = T/N$, and we denote by $t_n = n\Delta_t$, with $n = 0, \ldots, N$, the time instants associated to the above discretization. The approximated values of $\mathbf{z}(t_n)$, $\mathbf{x}(t_n)$ and $\mathbf{w}(t_n)$ are hereafter denoted by $\mathbf{z}_n$, $\mathbf{x}_n$ and $\mathbf{w}_n$, respectively. We apply the r-stage Runge-Kutta method to Eq. (10)

$$\begin{cases} \mathbf{z}_{n+1} &= \mathbf{z}_n + \Delta_t \sum_{i=1}^{r} b_i \mathbf{K}_i, \quad n = 0, \ldots, N-1; \\ \mathbf{K}_1 &= \mathbf{F}(t_n,\, \mathbf{z}_n); \\ \mathbf{K}_2 &= \mathbf{F}(t_n + c_2\Delta_t,\, \mathbf{z}_n + \Delta_t a_{2,1}\mathbf{K}_1); \\ \vdots & \\ \mathbf{K}_r &= \mathbf{F}(t_n + c_r\Delta_t,\, \mathbf{z}_n + \Delta_t(a_{r,1}\mathbf{K}_1 + a_{r,2}\mathbf{K}_2 + \cdots a_{r,r-1}\mathbf{K}_{r-1})), \end{cases} \tag{11}$$

where the coefficients $b_i, c_i, a_{i,j}$, $1 \le j < i \le r$ are defined by the well known Butcher tableau (see [30]). In particular, in the forthcoming numerical tests we will apply and compare the Euler and Heun schemes, retrieved by choosing in (11) the parameters $r = 1$, $b_1 = 1$ and $r = 2$, $b_1 = b_2 = 1/2$, $c_2 = 1$ and $a_{2,1} = 1$, respectively. It is worth noting that, at each time step, the methods require one and two evaluations of the functions $\mathbf{g}_{\mathbf{x}}$ and $\mathbf{g}_{\mathbf{w}}$, respectively. These involve the computation of the integrals $\mathbf{W}$ in Eq. (6) and M in Eq. (9). This further highlights that the choice of an efficient quadrature formula is a key issue to obtain an accurate approximation of cell trajectory, avoiding excessive overall computational costs.

3.2. Standard *and* ad hoc *application of the Gauss-Legendre quadrature formula to* $\mathbf{W}$ *and* M

To compute the double integral $\mathbf{W}(t_n,\, \mathbf{x}_n,\, \mathbf{w}_n)$ in Eq. (6), we consider a $\nu \times \bar{\nu}$ Gauss-Legendre product quadrature rule, with suitable numbers ν and $\bar{\nu}$ of quadrature nodes. The generic formula reads

$$\int_a^b h_1(r) \int_c^d h_2(r,\theta)\, d\theta\, dr \;\approx\; \sum_{p=1}^{\nu} w_p\, h_1(r_p) \sum_{q=1}^{\bar{\nu}} \bar{w}_q\, h_2(r_p, \theta_q), \tag{12}$$

where w_p and r_p, with $p = 1, \ldots, \nu$, (as well as $\bar{w}_q$ and θ_q, with $q = 1, \ldots, \bar{\nu}$) denote the weights and the nodes of the ν-point (or $\bar{\nu}$-point) Gauss-Legendre formula (see [12]). Following the *standard* approach, the quadrature formula (12) can be directly applied to the double integral in Eq. (6) by setting

$$h_1(r) = r\, K_r(r), \qquad h_2(r,\theta) = K_\theta\left(\mathbf{n}(\theta)\cdot\frac{\mathbf{w}_n}{\|\mathbf{w}_n\|}\right) c\left(t_n, \mathbf{x}_n + r\mathbf{n}(\theta)\right)\, \mathbf{n}(\theta), \tag{13}$$

and $[a,b] = [0,R]$, $[c,d] = [0,2\pi]$. On the other hand, in the case of piecewise regular integrand functions, we propose an *ad hoc* application of the quadrature rule which consists in splitting $[0,R]$ and $[0,2\pi]$ into subintervals of regularity of h_1 and h_2, and in applying the quadrature formula (12) to each one of the resulting double integrals. The following remarks then highlight two properties of this tailored approach.

Remark 3.1. *The total number of quadrature nodes* $\nu_{\text{tot}}^{\text{hoc}}$ *involved in the* ad hoc *approach depends on the number of regularity sub-domains of the integrand functions. Conversely, for the* standard *quadrature approach, the total number of quadrature nodes* $\nu_{\text{tot}}^{\text{std}}$ *is always equal to* $\nu \times \bar{\nu}$ *regardless of the specific choice of the weight functions* K_r *and* K_θ*, and the distribution* c *of the diffusing chemical. The efficiency of the two approaches can be then measured in terms of* ν_{tot}^{*}*, with* $* \in \{\text{hoc}, \text{std}\}$*.*

Due to the well known properties of Gaussian quadrature rules, the proposed tailored procedure results in an increased accuracy of the approximation of $\mathbf{W}(t_n, \mathbf{x}_n, \mathbf{w}_n)$ (w.r.t. the *standard* approach) and in a desirable convenience (or at least equality) in terms of computational cost (i.e., in terms of the total amount of quadrature nodes).

For what concerns the computation of the single integral $M(t_n, \mathbf{x}_n, \mathbf{w}_n)$ in Eq. (9), we can in principle proceed analogously as before. In particular, we can apply a ν-point Gauss-Legendre quadrature rule to each regularity subinterval of the integrand function m. However, since we will consider m piece-wise constant, the corresponding integral M is evaluated analytically.

In the forthcoming sections we will perform numerical simulations where, for the sake of simplicity, we will fix $\bar{\nu} = \nu$, in order to have only one discretization parameter related to the approximation of double integrals.

4. Validation of the proposed numerical approach

The aim of the numerical results presented in this section is to validate the proposed numerical procedure focusing on a selected benchmark test where the chemical profiles c and m are known, for any instant time $t \in [0, T]$, at any point $\mathbf{y} \in \mathbb{R}^2$. In particular, in Section 4.2, we will first give an idea of the computational advantages provided by the *ad hoc* application of the Gauss-Legendre quadrature rule (w.r.t the *standard* approach) to estimate $\mathbf{W}(t_n, \mathbf{x}_n, \mathbf{w}_n)$. Then, in Sections 4.3 and 4.4, we will deal with the numerical resolution of the system of non-linear integro-differential equations in Eqs. (1) and (3). Specifically, we will investigate how the quadrature approach affects both the accuracy of cell trajectory and the overall computational cost. As a remark, the integral $M(t_n, \mathbf{x}_n, \mathbf{w}_n)$ will be always analytically evaluated as specified in the previous section, since the ECM profile m is piecewise constant.

4.1. Simulation details

Benchmark test. Given a bounded numerical domain $\Omega = [0, L] \times [0, L]$, with $L = 700$ μm (that reproduces a Petri dish used in [43]), our benchmark test includes

(i) a horizontally diffusing chemical, which is produced at the left border of Ω, is completely absorbed at its right edge, and decays with a constant rate;

(ii) a piecewise constant distribution of ECM proteins characterized by high values at the left half of Ω, and intermediate densities elsewhere,

see Fig. 2 (panel A). Specifically, denoting by y_1, y_2 the components of $\mathbf{y} \in \Omega$, i.e., $\mathbf{y} = (\mathrm{y}_1, \mathrm{y}_2)$, we consider

$$c(t, \mathbf{y}) = c_0 \, \frac{\exp(\sqrt{1/\delta}\, \mathrm{y}_1) - \exp(\sqrt{1/\delta}\, (2L - \mathrm{y}_1))}{1 - \exp(2L\, \sqrt{1/\delta})}, \tag{14}$$

where c_0 is the amount of diffusive substance located at the left edge of the domain, δ is the diffusion length. Accounting for the measurements reported in [42], we here set $c_0 = 0.0004$ μM and $\delta = 492.6$ μm^2.

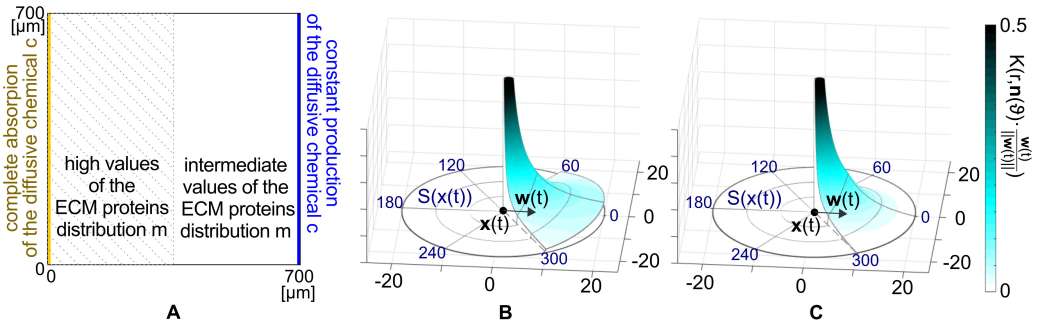

Figure 2: Settings of the benchmark test used to validate the proposed numerical approach. Panel A: Computational domain Ω. Panel B and C: Profiles of the weight function $K = K_r K_\theta$ with K_r as in Eq. (16) and as in Eq. (17), respectively, while K_θ is constantly taken as in Eq. (18) with $\alpha = 1/2$.

The distribution of the fixed ECM substance conversely reads as

$$m\left(\mathbf{y}\right) = \begin{cases} 0.75, & \text{if } \mathrm{y}_1 \leq L/2; \\ 0.5, & \text{otherwise.} \end{cases} \tag{15}$$

The final time T will be specified in the following. We remark that it will be chosen to avoid that the cell (as well as its sensing region) reaches the boundary domain.

Explicit forms for the weight functions. Given the distribution c defined in Eq. (14), the regularity of the integrand function in Eq. (6) depends on the explicit form of the weight functions K_r and K_θ. In this respect, among all the possible options, we consider the following alternative choices for K_r:

$$K_r(r) = \begin{cases} 0, & \text{if } r = 0; \\ \dfrac{1}{r}, & \text{if } r \leq R, \end{cases} \tag{16}$$

and

$$K_r(r) = \begin{cases} 0, & \text{if } r = 0; \\ \dfrac{1}{r}, & r \leq \dfrac{R}{3}; \\ \dfrac{3}{2r}\left(1 - \dfrac{r}{R}\right), & \text{if } \dfrac{R}{3} < r \leq R. \end{cases} \tag{17}$$

On the opposite, K_θ is constantly taken equal to

$$K_\theta(u) = \begin{cases} u^2, & \text{if } u > \alpha; \\ 0, & \text{otherwise,} \end{cases} \tag{18}$$

with $\alpha \in [0, 1)$. We notice that K_r in Eq. (17) vanishes for $r = R$ and involves non-regularities for $r = R/3$, while the one in Eq. (16) does not, see Fig. 2 (panel B and C). Moreover, unless $\alpha = 0$, K_θ is discontinuous.

Model parameter values. The representative cell is able to reach a maximum speed of $V_{\max} = 0.009$ μms^{-1} and to re-orient in a period τ equal to 300 s, according to the observations reported in [31] and [34], respectively. The Michaelis-Menten constant χ in Eq. (3) is conversely arbitrarily fixed to 0.001 for all numerical simulations. The extension of cell PM protrusions R (which defines also where the cell perceives the diffusive chemical and the fixed ECM proteins, respectively) is assumed equal to 20 μm as in [6]. Finally, a reasonable span of cell sensing region is fixed by setting $\alpha = 1/2$ in Eq. (18), corresponding to a sensing angle of 60°.

4.2. Standard *vs.* ad hoc *quadrature approach in computing* $\mathbf{W}$.

In this section, we compare the accuracy of the *standard* and of the *ad hoc* quadrature approaches in computing $\mathbf{W}(t_n, \mathbf{x}_n, \mathbf{w}_n)$ at a given instant of time t_n (i.e., for a fixed $\mathbf{x}_n$ and $\mathbf{w}_n$). To this aim, we denote by $\mathbf{W}^{*,\nu}$ the approximated value of the integral $\mathbf{W}(t_n, \mathbf{x}_n, \mathbf{w}_n)$ obtained by applying the approach $* \in \{\text{std}, \text{hoc}\}$ with ν quadrature nodes per single integral. The accuracy of each procedure is then estimated through the relative error

$$err_{\mathbf{W}}^{*,\nu} = \frac{\|\mathbf{W}^{\text{hoc},64} - \mathbf{W}^{*,\nu}\|}{\|\mathbf{W}^{\text{hoc},64}\|}, \tag{19}$$

where $* \in \{\text{std}, \text{hoc}\}$ and $\mathbf{W}^{\text{hoc},64}$ (i.e., the value obtained applying the *ad hoc* procedure with $\nu = 64$ quadrature nodes) is arbitrary assumed as the reference one. Specifically, to give an idea of the advantages provided by the *ad hoc* strategy in the case of piecewise integrand functions, the above estimate is performed twice by alternatively assuming K_r as in Eq. (16) or in Eq. (17).

Going into details, as defined in Section 3, $\mathbf{W}^{\text{std},\nu}$ is obtained by directly applying the $\nu \times \nu$ Gaussian quadrature rule (i.e., Eq. (12) with $\bar{\nu} = \nu$) to

Eq. (6), regardless of the specific form of K_r and K_θ. Conversely, $\mathbf{W}^{\mathrm{hoc},\nu}$ is here computed by splitting $[0, R]$ and $[0, 2\pi]$ into the subintervals where K_r and K_θ are smooth, respectively, and by applying the $\nu \times \nu$ Gaussian quadrature formula to each one of the resulting double integrals.

In this perspective, the regularity sub-intervals of the proposed weight functions K_r are explicitly given by their definitions, i.e., by Eqs. (16) and (17). Contrarily, the regularity subintervals of K_θ cannot be explicitly evinced from Eq. (18), but their determination requires to solve a trigonometric inequality. In particular, for any $\alpha \in [0, 1)$, the condition $\mathbf{n} \cdot \frac{\mathbf{w}_n}{\|\mathbf{w}_n\|} > \alpha$ identifies the sectors $\mathcal{S}_1$ and $\mathcal{S}_2$ such that $\mathcal{S} = \mathcal{S}_1 \cup \mathcal{S}_2$. In particular, we have

$$\mathcal{S}_1 := \begin{cases} (\theta_{\min},\, \theta_{\max}), & \text{if } \mathrm{w}_{n,1} + \alpha\|\mathbf{w}_n\| > 0; \\ (0,\, \theta_{\min}) \cup (\theta_{\max},\, 2\pi), & \text{if } \mathrm{w}_{n,1} + \alpha\|\mathbf{w}_n\| < 0, \end{cases} \tag{20}$$

where $\theta_{\min} = 2\min_{\epsilon\in\{-1,1\}}\{\operatorname{atan}\zeta_\epsilon\}$, $\theta_{\max} = 2\max_{\epsilon\in\{-1,1\}}\{\operatorname{atan}\zeta_\epsilon\}$ and $\mathbf{w}_n = (\mathrm{w}_{n,1}, \mathrm{w}_{n,2})$, with

$$\zeta_\epsilon = \frac{\mathrm{w}_{n,2} + \epsilon\,\|\mathbf{w}_n\|\sqrt{1-\alpha^2}}{\mathrm{w}_{n,1} + \alpha\|\mathbf{w}_n\|}, \quad \epsilon \in \{-1, 1\}. \tag{21}$$

The case $\mathrm{w}_{n,1} + \alpha\|\mathbf{w}_n\| = 0$ conversely leads to a first order inequality, which can be analogously treated.

Summing up, if K_r reads as in Eq. (16), $\mathbf{W}^{\mathrm{hoc},\nu}$ is obtained by applying the $\nu \times \nu$ Gaussian quadrature rule to each double integral in the following expression

$$\begin{aligned} &\mathbf{W}(t_n, \mathbf{x}_n, \mathbf{w}_n) = \\ &= \int_0^R r K_r(r) \int_{\mathcal{S}_1} \left[K_\theta\left(\mathbf{n}(\theta) \cdot \frac{\mathbf{w}_n}{|\mathbf{w}_n|}\right) c(t_n, \mathbf{x}_n + r\mathbf{n}(\theta))\, \mathbf{n}(\theta)\, d\theta \right] dr, \end{aligned} \tag{22}$$

where $[0, 2\pi]$ reduces to $\mathcal{S}_1$, since K_θ, as defined in Eq. (18), is null over $\mathcal{S}_2$. The total number of quadrature nodes $\nu_{\mathrm{tot}}^{\mathrm{hoc}}$ here results equal to $\nu \times \nu$, as in the *standard* approach (see Remark 3.1). However, in the *ad hoc* approach the quadrature nodes are distributed only over the actual cell sensing region (i.e., the angular sector $\mathcal{S}_1$) rather than over all $\mathcal{S}$.

If K_r reads as in Eq. (17), $\mathbf{W}^{\mathrm{hoc},\nu}$ is instead computed by applying Eq. (12)

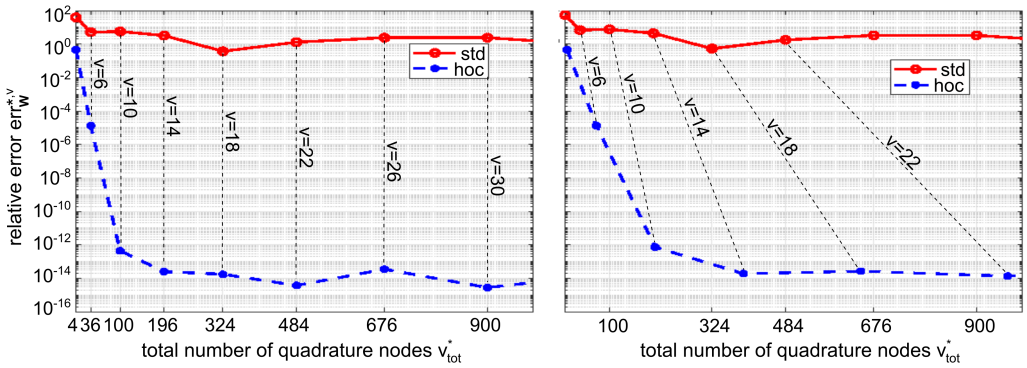

Figure 3: Plots of the relative errors $err^{\mathrm{std},\nu}_{\mathbf{W}}$ (red solid lines) and $err^{\mathrm{hoc},\nu}_{\mathbf{W}}$ (blue dashed lines) defined in Eq. (19) obtained by setting that, at the generical instant time t_n, the representative cell is placed at $\mathbf{x}_n = (L/2, L/2)$ and is oriented in the direction $\mathbf{w}_n = (0, 1)$. Left panel: K_r reads as in Eq. (16) and $\nu^{\mathrm{std}}_{\mathrm{tot}} = \nu^{\mathrm{hoc}}_{\mathrm{tot}} = \nu \times \nu$. Right panel: K_r is given by Eq. (17) and $\nu^{\mathrm{std}}_{\mathrm{tot}} = \nu \times \nu$, while $\nu^{\mathrm{hoc}}_{\mathrm{tot}} = 2\nu \times \nu$.

to each double integral in

$$\mathbf{W}(t_n, \mathbf{x}_n, \mathbf{w}_n) = \\ = \int_{\left[0,\frac{R}{3}\right]\cup\left[\frac{R}{3},R\right]} \left[rK_r(r) \int_{\mathcal{S}_1} K_\theta \left(\mathbf{n}(\theta) \cdot \frac{\mathbf{w}_n}{|\mathbf{w}_n|} \right) c(t_n, \mathbf{x}_n + r\mathbf{n}(\theta))\mathbf{n}(\theta) d\theta dr \right] \tag{23}$$

where, according to the definition of K_θ in Eq. (18), the integral over $\mathcal{S}_2$ is again omitted being null. In this case, we have a total number of quadrature nodes $\nu^{\mathrm{hoc}}_{\mathrm{tot}}$ equal to $2\nu \times \nu$, which all fall within the actual cell sensing region, while the *standard* approach still results in $\nu^{\mathrm{std}}_{\mathrm{tot}} = \nu \times \nu$ nodes distributed over the entire area $\mathcal{S}$.

In Fig. 3, we report the relative errors $err^{*,\nu}_{\mathbf{W}}$ obtained by arbitrarily assuming that, at the generic time instant t_n, the representative cell is placed at $\mathbf{x}_n = (L/2, L/2)$ and is oriented in the direction $\mathbf{w}_n = (0, 1)$ (so that $\mathcal{S}$ is completely embedded in Ω). Specifically, the results in the left panel of Fig. 3 have been obtained dealing with the weight function K_r defined in Eq. (16), while those in the right panel of Fig. 3 refer to the explicit form of K_r provided in Eq. (17). In both cases, the relative errors $err^{*,\nu}_{\mathbf{W}}$ are plotted with respect to increasing total amount of quadrature nodes ν^{*}_{tot}.

As expected, in both cases, the errors produced by the *ad hoc* approach

are much smaller than those obtained with the *standard* one, for any considered value of ν^*_{tot}. In particular, the *standard* approach provides errors in the range $[10^{-1}, 10^2]$ regardless of the number of quadrature nodes. Conversely, the accuracy of the *ad hoc* procedure rapidly improves with the increase of $\nu^{\text{hoc}}_{\text{tot}}$ and almost reaches the machine precision (that in our case is of magnitude order $1.0e-16$) with a reasonably low number of quadrature nodes of about $\nu = 20$. Moreover, an extensive numerical testing (not shown to avoid redundances) has also highlighted that, in our benchmark test, the *ad hoc* procedure with $\nu = 20$ quadrature nodes is able to estimate $\mathbf{W}(t_n, \mathbf{x}_n, \mathbf{w}_n)$ with machine precision, for any choice of $\mathbf{x}_n$ and $\mathbf{w}_n$.

4.3. Standard *vs.* ad hoc *quadrature approach in computing cell trajectory.*

Taking advantage of the considerations presented in previous section, we here turn to investigate the influence of the numerical integration of $\mathbf{W}(t_n, \mathbf{x}_n, \mathbf{w}_n)$ on the approximation of cell trajectory, i.e., on the accuracy of the overall solver for the system of non-linear integro-differential equations in Eqs. (1) and (3), for both Euler and Heun methods. In this perspective, for the sake of simplicity, we here deal only with the explicit form of K_r defined by Eq. (16).

To give an idea of the error introduced in cell trajectory by the *standard* and *ad hoc* quadrature approaches, we here deal with an arbitrarily selected scenario: the representative cell is initially located at $\mathbf{x}(0) = (3L/4, L/4)$ with $\mathbf{w}(0) = (0, 1)$. The period of observation T is set equal to 16 h.

The accuracy of cell trajectory obtained with the two quadrature approaches is then estimated through the relative error

$$err^{*,\nu}_{\mathbf{x}} = \max_{n=0,\ldots,N} \frac{\|\mathbf{x}^{\text{hoc},20}_n - \mathbf{x}^{*,\nu}_n\|}{\|\mathbf{x}^{\text{hoc},20}_n\|}, \tag{24}$$

where $\mathbf{x}^{*,\nu}_n$, with $n = 1, \ldots, N$, denotes the cell trajectory obtained by applying the $* \in \{\text{hoc}, \text{std}\}$ quadrature procedure with ν quadrature nodes per single integral in the computation of $\mathbf{W}(t_n, \mathbf{x}_n, \mathbf{w}_n)$. The value $\mathbf{x}^{\text{hoc},20}_n$ is instead assumed as reference solution according to the considerations presented in the previous section. Specifically, the number of time steps N is here arbitrarily fixed to 100 (both for the reference and for the approximate solutions obtained by the Euler and Heun approaches), which results in a plausible reference trajectory (see the magenta lines in Fig. 4 corresponding to the trajectories obtained by the Heun method).

ν	$\nu_{\text{tot}}^{\text{std}}$	$err_{\mathbf{x}}^{\text{std},\nu}$		ν	$\nu_{\text{tot}}^{\text{hoc}}$	$err_{\mathbf{x}}^{\text{hoc},\nu}$	
		Euler	Heun			Euler	Heun
4	16	$4.84e-01$	$4.84e-01$	4	16	$3.14e-03$	$3.13e-03$
8	64	$3.55e-01$	$3.53e-01$	6	36	$4.19e-06$	$4.14e-06$
16	256	$9.37e-01$	$9.39e-01$	8	64	$1.91e-09$	$1.89e-09$
32	1024	$9.57e-01$	$9.58e-01$	10	100	$3.21e-13$	$3.22e-13$
64	4096	$8.62e-01$	$8.63e-01$	12	144	$6.85e-16$	$1.93e-15$
128	16384	$1.45e-01$	$1.47e-01$	14	196	$--$	$--$
256	65536	$5.38e-02$	$5.77e-02$				
512	262144	$3.37e-03$	$6.87e-03$				
1024	1048576	$2.78e-03$	$5.96e-03$				
2048	4194304	$6.79e-04$	$3.55e-04$				

Table 1: Relative errors $err_{\mathbf{x}}^{*,\nu}$ defined in Eq. (24) and associated to the trajectories (obtained with the Euler and Heun approaches) of a cell initially located at $\mathbf{x}(0) = (3L/4, L/4)$ with $\mathbf{w}(0) = (0, 1)$. The period of observation T is set equal to 16 h, and it is divided into $N = 100$ time steps both for the reference and the approximate solutions. The explicit form of K_r is given in Eq. (16). The symbol "$--$" means that the full accuracy (i.e., double precision) has been reached.

The values of $err_{\mathbf{x}}^{*,\nu}$ reported in Table 1 and the dashed black trajectories plotted in Fig. 4 clearly highlight that, for both ODE solvers, the two quadrature approaches considerably differ one from each other. In particular, the *ad hoc* method quickly converges upon increments in ν (see the values in Table 1) and reaches the machine accuracy with only $\nu = 14$ nodes per integral (i.e., with $\nu_{\text{tot}}^{\text{hoc}} = 196$). On the opposite, the *standard* procedure with similar values of ν (e.g., $\nu = 16$) provides larger relative errors (see Table 1) showing that the obtained solution dramatically differs from the reference trajectory (see Fig. 4, left panel). Moreover, the values of $err_{\mathbf{x}}^{\text{std},\nu}$ reported in Table 1 show that a significative increase in ν does not rapidly improve the accuracy of the approximation. The *standard* approach then turns out to be not competitive with the *ad hoc* one in terms of computational cost, and hence of overall efficiency.

It is worthwhile noting that, in general, a sufficiently accurate (and not excessively expensive) approximation of the cell trajectory is reasonably identified by a relative error approximatively 5.0e-03. In our benchmark test, this is in fact achieved by considering, for both ODE solvers, at least $\nu = 1024$

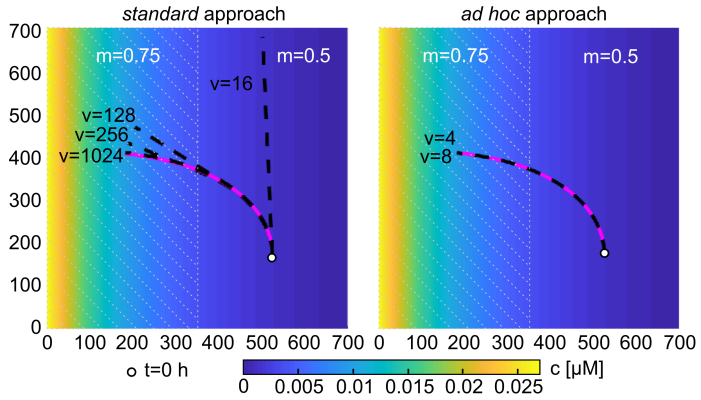

Figure 4: Cell trajectories of a cell initially located at $\mathbf{x}(0) = (3L/4, L/4)$ with $\mathbf{w}(0) = (0, 1)$, obtained by the Heun method, by applying either the *standard* (left panel) or the *ad hoc* (right panel) approach in computing $\mathbf{W}(t_n, \mathbf{x}_n, \mathbf{w}_n)$. The period of observation T is set equal to 16 h, and it is divided into $N = 100$ time steps both for the reference and the approximate solutions. The explicit form of K_r is given in Eq. (16). In each panel, the magenta line denotes the reference trajectory, the dashed black lines are the approximated trajectories obtained with ν quadrature nodes per single integral.

quadrature nodes per single integral if we apply the *standard* approach (see Fig. 4, left panel) and at least $\nu = 4$ quadrature nodes when we deal with the *ad hoc* procedure (see Fig. 4, right panel).

Interestingly, similar results are obtained also assuming K_r as in Eq. (17), thereby justifying the following remark.

Remark 4.1. *According to the considerations presented in Sections 4.2 and 4.3, in the forthcoming examples, we will apply only the* ad hoc *procedure by choosing* $\nu = 20$ *for the reference solution and a proper value of* ν *for the approximate one. This latter will depend on the desired accuracy of the overall method and will be specified within the numerical examples, according to the results of Table 1. Hereafter, we therefore omit the superscripts* ad hoc *and* ν*, in* $\mathbf{x}_n^{\text{hoc},\nu}$*, denoting by* $\overline{\mathbf{x}}_n$ *the reference solution and by* $\mathbf{x}_n$ *the approximate one.*

4.4. Convergence of Runge-Kutta methods with the ad hoc *quadrature approach.*

We here turn to deal with the convergence of the Euler and Heun time marching schemes for the solution of ODEs, and by the *ad hoc* approach for the numerical evaluation of the integrals $\mathbf{W}(t_n, \mathbf{x}_n, \mathbf{w}_n)$ at each time step.

Specifically, we consider the biological scenario introduced previously (with c and m as given in Eqs. (14) and (15), see Fig. 1, and initial condition $\mathbf{x}(0) = (3L/4, L/4)$ and $\mathbf{w}(0) = (0, 1)$), assuming the explicit form for K_r defined in Eq. (17), since it is characterized by non-regularities and, hence, is more interesting from a numerical point of view.

Going into details, we assume as reference trajectory the solution obtained with a very high number of time steps, i.e., $\overline{N} = 102400$, and we consider a sequence of approximated cell trajectories obtained with N time steps. To easily compare the resulting approximated trajectories with the reference one, in all forthcoming simulations, we require that N is a divisor of $\overline{N}$ (i.e., there exists $k \in \mathbb{N}$ such that $\overline{N} = kN$). The convergence of both methods is then evaluated through the relative error

$$err_{\mathbf{x}}^{N} = \max_{n=0,\dots,N} \frac{\|\overline{\mathbf{x}}_{kn} - \mathbf{x}_n\|}{\|\overline{\mathbf{x}}_{kn}\|}, \tag{25}$$

where $\overline{\mathbf{x}}_{kn}$, with $n = 0, \dots, N$, denotes the reference cell trajectory obtained by using the *ad hoc* quadrature formula with $\nu = 20$ and $\overline{N} = 102400$ time steps. The approximated trajectory $\mathbf{x}_n$, with $n = 0, \dots, N$, is instead computed with N time steps and by applying the *ad hoc* quadrature formula with $\nu = 12$, which guarantees the computation of the integrals up to the machine precision (see Table 1). In this way the optimal convergence order of the overall method, for both Euler and Heun solvers, is ensured.

Moreover, we evaluate the Estimated Order of Convergence (EOC), given by

$$EOC = \log_2(err_{\mathbf{x}}^{N}/err_{\mathbf{x}}^{2N}). \tag{26}$$

The values reported in Table 2 show that the accuracy of the approximated cell trajectory improves for increasing values of N, for both solvers, and that the EOC coincides with the theoretical linear order of convergence of the Euler method and quadratic order of the Heun one. This result is a further confirmation of the *ad hoc* procedure efficiency. Furthermore, to compare the two solvers in terms of computing times, in the last column of Table 2 we report the ratio $\mathcal{R}_{\text{CPU}}$ between the CPU time required by the Heun solver over the Euler one. As expected, since the former involves twice the number of integral evaluations per time step, for N large enough the ratio is approximatively 2. However, it is worth noting that, by fairly comparing the two approximate solutions in terms of accuracy, the Heun method reveals to be more efficient than the Euler one. Indeed, for example, if we

	Euler		Heun		
N	$err_{\mathbf{x}}^N$	EOC	$err_{\mathbf{x}}^N$	EOC	$\mathcal{R}_{\text{CPU}}$
100	$2.44e-03$		$2.06e-04$		3.4
		1.0		3.0	
200	$1.26e-03$		$2.62e-05$		2.8
		1.0		2.0	
400	$6.27e-04$		$6.82e-06$		1.6
		1.0		2.0	
800	$3.13e-04$		$1.74e-06$		1.5
		1.0		2.0	
1600	$1.57e-04$		$4.34e-07$		2.2
		1.0		2.0	
3200	$7.83e-05$		$1.09e-07$		2.3
		1.0		2.0	
6400	$3.92e-05$		$2.72e-08$		1.9
		1.0		2.0	
12800	$1.96e-05$		$6.74e-09$		1.9
		1.0		2.0	
25600	$9.79e-06$		$1.60e-09$		1.8

Table 2: Relative errors $err_{\mathbf{x}}^N$, defined in Eq. (25), and estimated orders of convergence EOC, given in Eq. (26), of the Euler and Heun time marching schemes, obtained by using the *ad hoc* procedure ($\nu_{\text{tot}}^{\text{hoc}} = 2\nu \times \nu$), when c and m are given by Eqs. (14) and (15), respectively, and K_r by Eq. (17). The reference solution is obtained by using the *ad hoc* quadrature formula with $\nu = 20$ and $\overline{N} = 102400$ time steps. The approximate solutions are obtained by using the *ad hoc* quadrature formula with $\nu = 8$ and N time steps.

fix the accuracy to the magnitude order $1.0e-04$ (retrieved by Euler with $N = 800$ and by Heun with $N = 100$), we get $\mathcal{R}_{\text{CPU}} \approx 1.7e-01$; by fixing the accuracy order to $1.0e-05$ (retrieved by Euler with $N = 12800$ and by Heun with $N = 200$), we get $\mathcal{R}_{\text{CPU}} \approx 3.2e-02$. In percentage terms, Heun saves the 83% of the computing time in the first case and 97% in the second one.

Finally we highlight that, by testing smaller choices of N (i.e., larger time steps Δ_t), the smallest value, which guarantees a realistic approximated trajectory $\mathbf{x}_n$, is $N = 94$, for both Euler and Heun solvers. As an example, in Fig. 5 we report cell trajectories obtained only by the Heun method with

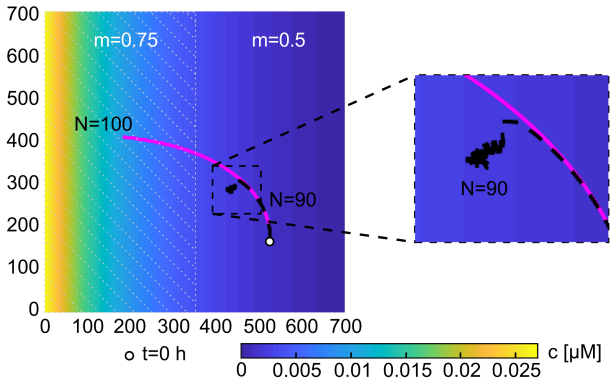

Figure 5: Trajectory of a cell initially located at $\mathbf{x}(0) = (3L/4, L/4)$ with $\mathbf{w}(0) = (0, 1)$. The period of observation T is set equal to 16 h. The explicit form of K_r is given in Eq. (17). Cell trajectory obtained by the Heun method with $N = 100$ (magenta line) recovers the reference solution obtained with $\overline{N} = 102400$. Cell trajectory computed with $N = 90$ (dashed black line) conversely, after initial time steps, diverges from the reference solution.

$N = 100$ (magenta line) and with $N = 90$ (dashed black line). In the former case, the approximated cell trajectory (magenta line) recovers the reference solution while, in the latter case, the numerical method fails. This is a clear example of how a not sufficiently accurate method can lead to implausible biological conclusions. In this respect, in the absence of a computationally more expensive reference solution, the trajectory obtained with $N = 90$ time steps could be in fact wrongly interpreted as the representation of cell random wandering, rather than as the result of numerical errors.

According to the above comparison in terms of accuracy and efficiency, in the forthcoming numerical examples we will apply only the Heun method, being this choice a good compromise among accuracy and computational cost.

5. Extension of the numerical approach for experimental data

This section is devoted to the numerical issue (iii) presented in Section 2. In this respect, we recall that if the distributions of the diffusive chemical and/or of the fixed ECM proteins are constantly known only over a given set of points that, in general, do not coincide with the quadrature nodes needed to compute $\mathbf{W}(t_n, \mathbf{x}_n, \mathbf{w}_n)$ and $M(t_n, \mathbf{x}_n, \mathbf{w}_n)$ (see Fig. 1, panel C), then the approach proposed in Section 3 has to be coupled with a proper interpolation

method to retrieve the unknown data.

Obviously, in this context, the accuracy of the resulting overall numerical method may be strongly affected by the choice of the interpolation technique. On the other hand, such an accurate interpolation scheme should be also not too expensive, since it has to be coupled with the Heun method and the *ad hoc* quadrature approach. In this respect, accounting for the considerations reported in Section 4, a reasonable choice is to interpolate the unknown nodal values of the molecular variables c and m with a bi-dimensional and mono-dimensional cubic spline, respectively, using not-a-knot conditions (see [13]).

In the rest of this section, we will first give an idea of how experimental data (treated with spline interpolation) affect the accuracy of the *ad hoc* quadrature approach and of the convergence of the overall numerical method. To this aim, we will deal with a modified version of the benchmark test analyzed in Section 4. Successively, taking advantage of these results, we will further test the efficacy of the proposed numerical method, by applying it to solve Eqs. (1) and (3) in selected scenarios.

For the sake of simplicity, the interpolation scheme is here used to approximate only the concentration of the diffusive chemical c: since the distribution of the ECM proteins m is always given in analytical form, $M(t_n, \mathbf{x}_n, \mathbf{w}_n)$ is computed as specified in Section 3. By considering the domain Ω and the model parameter values introduced in Section 4.1, in all forthcoming numerical simulations, we assume that, at any instant time t, the diffusive chemical is known only at the points $\mathbf{y}_{ij} \in \Omega$ of a $\Lambda \times \Lambda$ tensor grid with the length step $\Delta_{\mathbf{y}} = L/\Lambda$ in both directions. In particular, we introduce

$$c_\Lambda(t) = \left\{ c_{ij}(t), \text{ with } i,j = 1,\ldots,\Lambda \,:\, c_{ij}(t) = c(t, \mathbf{y}_{ij}) \right\}, \tag{27}$$

where c is a given continuous profile, which will be used to compute the corresponding reference solution.

5.1. Errors introduced by experimental data treated with spline interpolation

To highlight how the spline interpolation of c_Λ affects the accuracy of the *ad hoc* approximation of $\mathbf{W}(t_n, \mathbf{x}_n, \mathbf{w}_n)$ and the convergence of the overall method, we here deal with a modified version of the benchmark test analyzed in Section 4. Specifically, we here assume that, at any instant time $t \in [0, T]$, the concentration of the diffusing chemical is point-wise defined as in Eq. (27),

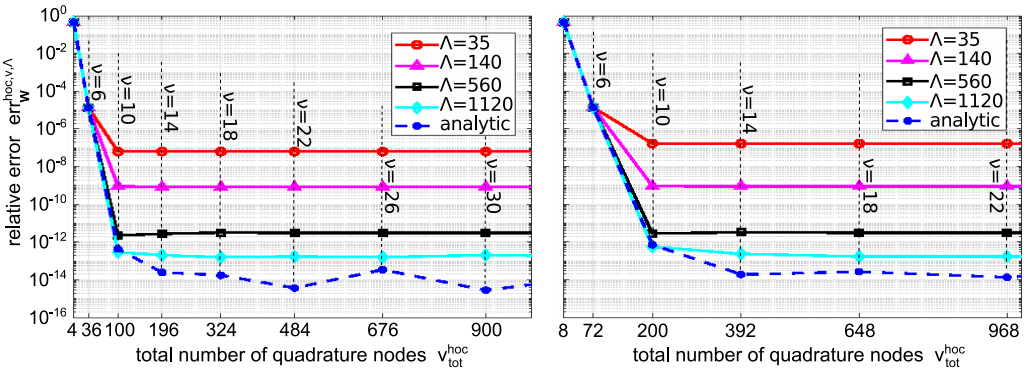

Figure 6: Relative errors $err_{\mathbf{W}}^{\mathrm{hoc},\nu,\Lambda}$ (solid lines) defined in Eq. (28) associated to piecewise distribution c_Λ defined in Eq. (27) with c as in Eq. (14). For the reader's convenience, we here report also the error $err_{\mathbf{W}}^{\mathrm{hoc},\nu}$ (dashed blue line) already plotted in Fig. 3, and we remark that $\Delta_{\mathbf{y}} = L/\Lambda$. Left panel: K_r reads as in Eq. (16) with $\nu_{\mathrm{tot}}^{\mathrm{hoc}} = \nu \times \nu$. Right panel: K_r is given by Eq. (17) with $\nu_{\mathrm{tot}}^{\mathrm{hoc}} = 2\nu \times \nu$.

with c given in Eq. (14). The distribution of the ECM proteins m is still defined in Eq. (15).

Experimental data and the ad hoc *approach in computing* $\mathbf{W}$. The accuracy of the *ad hoc* approximation of the cell preferred direction, i.e., of $\mathbf{W}(t_n, \mathbf{x}_n, \mathbf{w}_n)$, when the diffusive chemical is point-wise defined, is here estimated through the relative error

$$err_{\mathbf{W}}^{\mathrm{hoc},\nu,\Lambda} = \frac{\|\mathbf{W}^{\mathrm{hoc},64} - \mathbf{W}^{\mathrm{hoc},\nu,\Lambda}\|}{\|\mathbf{W}^{\mathrm{hoc},64}\|}. \tag{28}$$

In Eq. (28), $\mathbf{W}^{\mathrm{hoc},\nu,\Lambda}$ is computed by replacing c in $\mathbf{W}(t_n, \mathbf{x}_n, \mathbf{w}_n)$ with the cubic spline interpolation of c_Λ, and then by applying the *ad hoc* quadrature approach with ν nodes for each single integral. $\mathbf{W}^{\mathrm{hoc},64}$ is instead the reference solution already introduced in Eq. (19) (i.e., it is obtained by using the analytical expression of c given in Eq. (14)).

In Fig. 6, we then report the relative errors obtained for different values of Λ (and of the relative length step $\Delta_{\mathbf{y}}$), assuming that, at a generical instant time t_n, the representative cell is placed at $\mathbf{x}_n = (L/2, L/2)$ and is oriented in the direction $\mathbf{w}_n = (0, 1)$ (as already done in Section 4.2). Specifically, the results in the left panel have been obtained by dealing with the weight function K_r defined in Eq. (16), while those in the right panel refer to the

explicit form of K_r provided in Eq. (17). In both cases, it emerges that, for any fixed choice of ν, the error $err_{\mathbf{W}}^{\mathrm{hoc},\nu,\Lambda}$ reasonably decreases as Λ increases (i.e., as the length step $\Delta_{\mathbf{y}}$ decreases). Indeed, it is well known that the smaller the length step is, the smaller the interpolation error is. On the other hand, for any considered value of Λ, an increment in the total amount of the quadrature nodes $\nu_{\mathrm{tot}}^{\mathrm{hoc}}$ results in an initial drop of $err_{\mathbf{W}}^{\mathrm{hoc},\nu,\Lambda}$, until it reaches the maximum possible accuracy allowed by the interpolation method which, as known, actually depends on the length step $\Delta_{\mathbf{y}}$. Interestingly, regardless of the value of Λ, the maximal accuracy is here reached with $\nu \leq 10$, i.e., as the error introduced by the *ad hoc* quadrature formula almost converges to the machine precision (i.e., $err_{\mathbf{W}}^{\mathrm{hoc},\nu} < 10^{-12}$, see Fig. 3 and Table 1).

Experimental data and the convergence of the overall method. We here turn to investigate how the use of experimental data c_Λ, retrieved with spline interpolation, influences the accuracy of cell trajectory. To this aim, we introduce the relative error

$$err_{\mathbf{x}}^{N,\Lambda} = \max_{n=0,\dots,N} \frac{\|\overline{\mathbf{x}}_{kn} - \mathbf{x}_n^\Lambda\|}{\|\overline{\mathbf{x}}_{kn}\|}, \tag{29}$$

where $\mathbf{x}_n^\Lambda$, with $n = 0, \dots, N$, denotes the approximated cell trajectory obtained by applying, at any time step, the *ad hoc* procedure with $\nu = 8$ nodes (see Remark 4.1), and by replacing the molecular variable c with the cubic spline interpolating function of c_Λ. Conversely, $\overline{\mathbf{x}}_{kn}$, with $n = 0, \dots, N$, denotes the reference solution obtained by dividing the period of observation into $\overline{N} = kN = 3200$ time steps, and by using, in computing cell desired direction, the *ad hoc* quadrature formula with $\nu = 20$ (see Remark 4.1) and the analytic expression of c in Eq. (14).

Referring to the numerical simulations performed in Section 4.4, we here consider again a cell initially located at $\mathbf{x}(0) = (3L/4, L/4)$ with $\mathbf{w}(0) = (0, 1)$, whose capacity to radially sense the diffusive chemical K_r reads as in Eq. (17). The period of observation T is set equal to 16 h and divided into $\overline{N} = 3200$ time steps (so that the reference solution $\overline{\mathbf{x}}_{kn}$ is actually the same used in Section 4.4).

In Table 3, we thus report the relative errors obtained with different values of both discretization parameters Λ and N (and relative $\Delta_{\mathbf{y}}$ and Δ_t, respectively). Again, we also compute the estimated order of convergence

$$EOC^\Lambda = \log_2(err_{\mathbf{x}}^{N,\Lambda}/err_{\mathbf{x}}^{2N,\Lambda}). \tag{30}$$

N	$err_{\mathbf{x}}^{N,10}$	EOC^{10}	$err_{\mathbf{x}}^{N,25}$	EOC^{25}	$err_{\mathbf{x}}^{N,35}$	EOC^{35}
100	$2.24e-04$		$2.06e-04$		$2.07e-04$	
		2.8		3.0		3.0
200	$3.20e-05$		$2.62e-05$		$2.62e-05$	
		0.6		2.0		2.0
400	$2.12e-05$		$6.75e-06$		$6.72e-06$	
		0.1		2.0		2.0
800	$1.92e-05$		$1.70e-06$		$1.64e-06$	
		0.1		1.7		2.2
1600	$1.86e-05$		$5.23e-07$		$3.60e-07$	

Table 3: Relative error $err_{\mathbf{x}}^{N,\Lambda}$ defined in Eq. (29) and estimated order of convergence EOC^{Λ} given in Eq. (30). The cell is initially located at $\mathbf{x}(0) = (3L/4, L/4)$ with $\mathbf{w}(0) = (0, 1)$, with K_r defined as in Eq. (17). The period of observation T is set equal to 16 h, while the concentration of the diffusive chemical is point-wise defined by Eq. (27) with c as in Eq. (14). ECM proteins distribution m is finally given analytically as in Eq. (15).

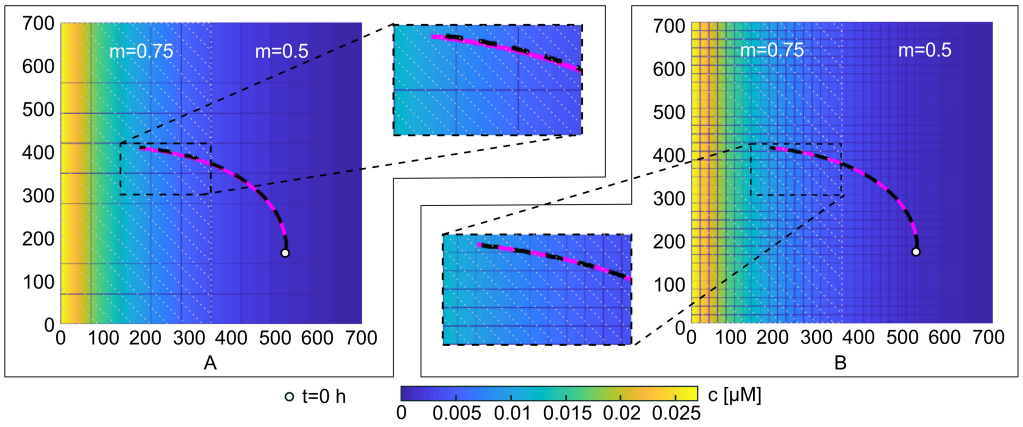

Figure 7: Role of the spatial discretization parameter Λ (and $\Delta_{\mathbf{y}}$) on the approximated cell trajectory. In all panels: the reference solution (magenta line) is computed by applying the *ad hoc* approach ($\nu = 20$), with $\overline{N} = 3200$ and c given by Eq. (14). The approximated cell trajectories (dashed black line) are obtained by the *ad hoc* approach ($\nu = 8$), with $N = 100$ and by replacing c with the cubic spline interpolation of c_{Λ}. Panel A: $\Lambda = 10$ (i.e., $\Delta_{\mathbf{y}} = 70$ μm). Panel B: $\Lambda = 35$ (i.e., $\Delta_{\mathbf{y}} = 20$ μm).

As expected, to obtain the theoretical quadratic order of convergence of the Heun method when the spline interpolation of c_{Λ} is considered, it

is necessary to use a sufficiently refined underlying tensor mesh, i.e., we need a high enough amount of length steps Λ (i.e., sufficiently small $\Delta_{\mathbf{y}}$). Interestingly, among the considered cases, the choice $\Lambda = 35$ (i.e., $\Delta_{\mathbf{y}} = 20$ μm) returns the same errors of the analytic case (reported in Table 2). However, it is worthwhile noting that the quadratic order of convergence is not mandatory to obtain a realistic cell trajectory. In fact, as shown in Fig. 7, to have a good approximation of cell dynamics, it is sufficient to choose $\Delta_{\mathbf{y}} = 70$ μm, which corresponds to the very coarse discretization parameter $\Lambda = 10$ (see Fig. 7, panel A), while $\Lambda = 35$ (i.e., $\Delta_{\mathbf{y}} = 20$ μm) returns a very accurate trajectory (see Fig. 7, panel B).

Interestingly, the values reported in Table 3 further indicate that the stability of the global method depends only on the choice of the temporal discretization (i.e., of N and Δ_t), according to the considerations reported in the previous paragraph. Conversely, the accuracy of the solution is affected also by the tensor mesh refinement (i.e., by the values of Λ and $\Delta_{\mathbf{y}}$), namely by the interpolation error.

5.2. Application of the overall method to selected scenarios

Taking advantage of the results presented in previous sections, we here finally apply the overall numerical method to three distinct scenarios, where the concentration of the diffusive chemical is defined only over a tensor grid with $\Lambda = 70$ (i.e., with $\Delta_{\mathbf{y}} = 10$ μm, so that the interpolation error does not affect the global accuracy), while the distribution of the ECM proteins is instead always known in analytic form. First, we fix c_Λ as defined in Section 5.1, and we alternatively consider two piecewise constant distributions of m that implement an abundance of ECM proteins respectively within a vertical stripe at the middle of the domain Ω (Case 1); and within a small square at the center of Ω (Case 2). Finally, in Case 3, we introduce a proper explicit form of c_Λ describing the equilibrium distribution of a diffusive chemical produced at the center of the domain and completely absorbed at the domain boundary; while the ECM is either assumed uniformly distributed over all the domain or as defined in Eq. (15).

In all forthcoming examples, cell capacity to non-locally sense the diffusive substance c is here assumed to radially vary according to the explicit form of K_r defined in Eq. (16). Referring again to Remark 4.1, the approximate solutions $\mathbf{x}_n$, with $n = 0, \ldots, N$, are computed by applying, at each step of the Heun scheme, the *ad hoc* procedure with $\nu = 8$ quadrature nodes

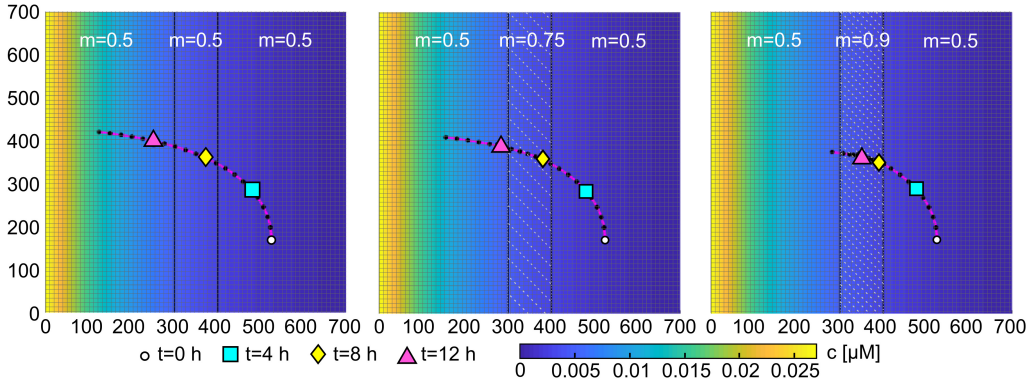

Figure 8: Case 1. Cell trajectory associated to c defined in Eq. (14) (reference solution, magenta line) and the relative point-wise distribution c_Λ in Eq. (27) with $\Lambda = 70$ (approximated solution, dashed black line). m is given by Eq. (31) with $\ell = 0.5$ (left panel), $\ell = 0.75$ (central panel) and $\ell = 0.9$ (right panel). The circle, square, diamond and triangle symbols indicate the cell position at the time instants $t = 0, 4, 8, 12$ h, respectively.

per single integral in $\mathbf{W}(t_n, \mathbf{x}_n, \mathbf{w}_n)$ where c is replaced by the cubic spline interpolation of c_Λ. The reference solutions $\overline{\mathbf{x}}_{kn}$, with $n = 0, \ldots, N$ and $kn = \overline{N}$, are obtained by using the *ad hoc* procedure with $\nu = 20$ quadrature nodes per single integral in $\mathbf{W}(t_n, \mathbf{x}_n, \mathbf{w}_n)$, where c is used in its analytical form. The number of time steps is defined below, according to the considered final instant time of the specific test: however, we constantly choose $\bar{N} = N$ for the reference solution.

Case 1. In this first case, the diffusive chemical is assumed point-wise defined by Eq. (27) with c as in Eq. (14). The distribution of ECM protein conversely reads

$$m(\mathbf{y}) = \begin{cases} \ell, & \text{if } \mathbf{y} \in [L/2 - \epsilon, L/2 + \epsilon] \times [0, L]; \\ 0.5, & \text{otherwise,} \end{cases} \tag{31}$$

where ϵ is fixed equal to 50 μm, while ℓ is alternatively set equal to 0.5, 0.75 and 0.9. The former value actually results into a uniform distribution of the ECM substance over Ω, while the others implement the presence at the middle of the domain of a vertical stripe of abundant ECM proteins. In all cases, we deal with a cell initially located at $\mathbf{x}(0) = (3L/4, L/4)$ and oriented along the direction $\mathbf{w}(0) = (1, 0)$. In Fig. 8, we plot cell trajectories obtained by setting $T = 16$ h and by dividing the period of observation into $N = 100$ time steps.

Interestingly, regardless of the value of ℓ, the approximated solutions re-

cover the reference one: this is a further justification of the proposed method and of the chosen discretization parameters, i.e., ν, Δ_t and $\Delta_{\mathbf{y}}$.

Comparing the three numerical outcomes, we can also notice that variations in the distribution of ECM proteins affect the length of cell trajectory rather than its shape. In particular, cell positions at $t = 4, 8, 12$ h (indicated in Fig. 8 by square, diamond and triangle symbols, respectively) highlight that a uniform distribution of the ECM proteins (i.e., $\ell = 0.5$, see Fig. 8, left panel) does not affect cell speed (according to Eqs. (7)-(9)). Conversely, cell net displacement decreases as the amount of ECM proteins at the central stripe increases: the abundance of fixed matrix elements in fact results in the formation of cell stable focal adhesions, thereby slowing down the individual motion. After the slowing down in the central area, the migrating cell then recovers its initial speed as soon as it reaches the left part of the domain. We finally remark that, assuming $\ell = 0.25$ or $\ell = 0.1$ in Eq. (31) would result in the asymmetric bimodal cell behavior (i.e., low speed, then accelerated speed, and finally decelerated movement), as obtained with $\ell = 0.75$ or $\ell = 0.9$, respectively.

Case 2. The diffusive chemical is here again assumed point-wise defined by Eq. (27) with c as in (14), while the ECM proteins are distributed according to

$$m\left(\mathbf{y}\right) = \begin{cases} 0.9, & \text{if } \mathbf{y} \in [L/2-\epsilon, L/2+\epsilon] \times [L/2-\epsilon, L/2+\epsilon]; \\ 0.5, & \text{otherwise,} \end{cases} \tag{32}$$

with $\epsilon = 50$ μm. In Fig. 9, we plot five cell trajectories obtained in independent realizations, performed by starting from distinct choices of the initial conditions, i.e.,

$$\begin{aligned} &\mathbf{x}^{(1)}(0) = (150, 100), \quad \text{and} \quad \mathbf{w}^{(1)}(0) = \left(\tfrac{1}{\sqrt{2}}, \tfrac{1}{\sqrt{2}}\right); \\ &\mathbf{x}^{(2)}(0) = (400, 100) \quad \text{and} \quad \mathbf{w}^{(2)}(0) = \mathbf{w}^{(1)}(0); \\ &\mathbf{x}^{(3)}(0) = (600, 100) \quad \text{and} \quad \mathbf{w}^{(3)}(0) = \mathbf{w}^{(1)}(0); \\ &\mathbf{x}^{(4)}(0) = (650, 380) \quad \text{and} \quad \mathbf{w}^{(4)}(0) = (-1, 0); \\ &\mathbf{x}^{(5)}(0) = (650, 500) \quad \text{and} \quad \mathbf{w}^{(5)}(0) = \mathbf{w}^{(4)}(0), \end{aligned} \tag{33}$$

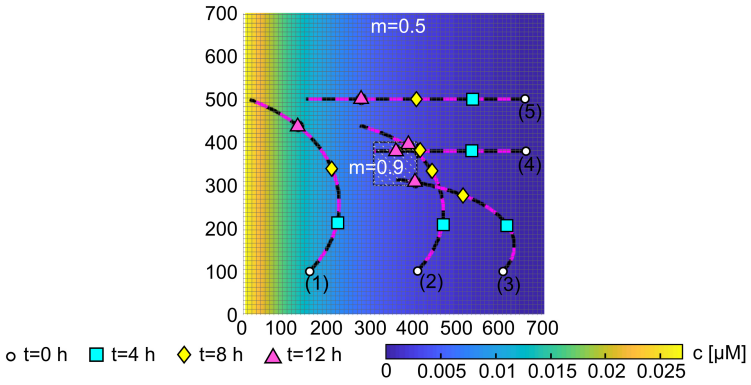

Figure 9: Case 2. Cell trajectories associated to c as defined in Eq. (14) (reference solution, magenta line) and the relative point-wise distribution c_Λ introduced in Eq. (27) with $\Lambda = 70$. m is given by Eq. (32). The circle, square, diamond and triangle symbols indicate the i-cell position for $i = 1, \dots, 5$ and at the time instants $t = 0, 4, 8, 12$ h, respectively.

where the superscript "$(\cdot)$" denotes the specific realization. In all simulations, the period of observation is $T = 16$ h, by setting $N = 100$ as in Case 1.

We first notice that, regardless of the specific initial condition, there is no visual difference between the reference solution (magenta lines) and the approximated one (dashed black lines), i.e., all approximated solutions coincide with the respective reference one (thereby further highlighting the accuracy of the numerical approach). Going into details, we have that in tests (3), (4) and (5) the representative cell crosses the central square, while in test (1) and (2) it does not. Due to the absence of random effects, such a distinction actually depends on the selected initial condition. Specifically, cell positions at $t = 4, 8, 12$ h (indicated in Fig. 9 by square, diamond and triangle symbols, respectively) highlight that the trajectories obtained in realizations (1) and (2) have the same length, while in other cases cell migration results in shorter paths. In these latter cases, the representative individual indeed slows down as soon as its sensing region (i.e., the segment which goes from $\mathbf{x}_n$ to $\mathbf{x}_n + R\frac{\mathbf{w}_n}{\|\mathbf{w}_n\|}$) intersects the central square. Consistently with Case 1, an abundance of ECM proteins in fact results in an increased formation of cell stable focal adhesions that anchor the individual to the substrate, therefore decreasing its migratory potential.

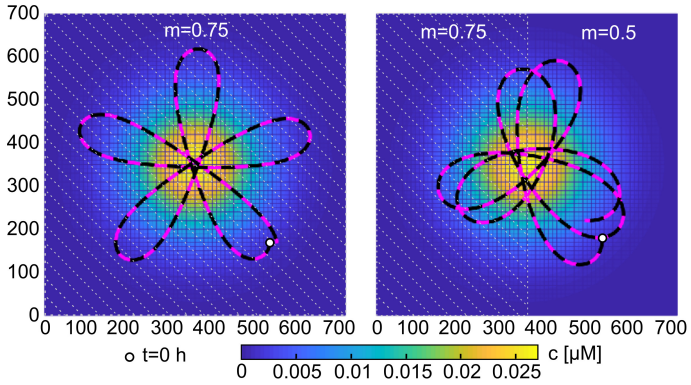

Figure 10: Case 3. Cell trajectories associated to c defined in Eq. (34) (reference solution, magenta line) and the relative point-wise distribution c_Λ with $\Lambda = 70$ (approximated solution, dashed black line). Left panel: $m(\mathbf{x}) = 0.75$ for any $\mathbf{x} \in \Omega$. Right panel: m is given by Eq. (15).

Case 3. In this last test, the point-wise distribution c_Λ reads as in Eq. (27), with c given by

$$c(t, \mathbf{y}) = c_0 \exp\left(\frac{-20\left((\mathrm{y}_1 - L/2)^2 - (\mathrm{y}_2 - L/2)^2\right)}{L^2}\right). \tag{34}$$

The distribution of ECM is instead alternatively assumed uniform over all the domain, i.e., $m(\mathbf{y}) = 0.75$ for any $\mathbf{y} \in \Omega$, or piecewise constant according to Eq. (15). In both cases, the cell is initially placed at $\mathbf{x}(0) = (3L/4, L/4)$ and oriented in the direction $\mathbf{w}(0) = (-1, 0)$. The period of observation T is finally set equal to 126 h and it is divided into $N = 1000$ time steps.

As expected, the trajectories reported in Fig. 10 show that, in both cases the approximated solutions (dashed black lines) recover the reference ones (magenta lines). More in details, regardless of the explicit expression of m, cell moves around the central region of the domain where there is the higher concentration of the diffusive chemical, i.e., the maximal chemotactic stimulus. In particular, it does not stop where c reaches the maximum concentration, because of the presence of a persistence time (a sort of biological inertia), see Eq. (3). This means that the representative cell overcomes the center of the domain and takes a given time before reorienting toward the source of the molecular substance. Interestingly, in the presence of a uniform distribution for m (see Fig. 10, left panel), the maximum distance from the

center reached by the individual is always the same. Conversely, in the case of a piecewise constant distribution m (see Fig. 10, right panel), we have a different behavior according to the actual amount of ECM: in particular, in the left half of Ω, the cell, being slowed by the abundance of extracellular proteins, moves away from the center less than it does in the right half of the domain.

6. Conclusions

Due to the increasingly recognized importance of cell migration processes in physio-pathological phenomena (both during embryogenesis and in adult life) and its exploitation for therapy and for tissue engineering, an increasing number of models have been developed.

The fact that cells determine both their polarization and speed by non-locally sensing their environment implies that several models proposed in the literature are characterized by the presence of integral terms. This feature may considerably increase the computational cost, especially when modeling collective phenomena that involve hundreds of cells or more. To overcome such difficulties our analysis suggests to combine the Heun method with tailored Gauss-Legendre formulas, possibly associated with interpolation techniques in the case of information on the environmental cues that are either poor or non-matching from the point of view of the numerical discretization.

The efficacy of the proposed method (also in terms of computational cost) has been then tested with several simulations, reproducing cell trajectories upon different micro-environmental conditions. In this respect, we remark that, among several well known methods with similar features, we chose the Heun method as it represents a right compromise among accuracy and computational cost. Moreover, the choices concerning the quadrature nodes and the polynomial interpolation degree are related to the accuracy of the Heun method. However, in principle, a different ODE method can be considered as well, taking into account that the accuracy of the quadrature and interpolation methods might be accordingly returned. For example, if a higher (lower) order ODEs solver is chosen, to preserve the global accuracy and efficiency, it might be necessary to consider a higher (lower) number of quadrature nodes, and/or a higher (lower) piece-wise interpolant polynomial degree for the approximation of the diffusive chemical.

It is finally worth to remark that the proposed *ad hoc* Gauss-Legendre quadrature formula is able to handle discontinuous distributions of the en-

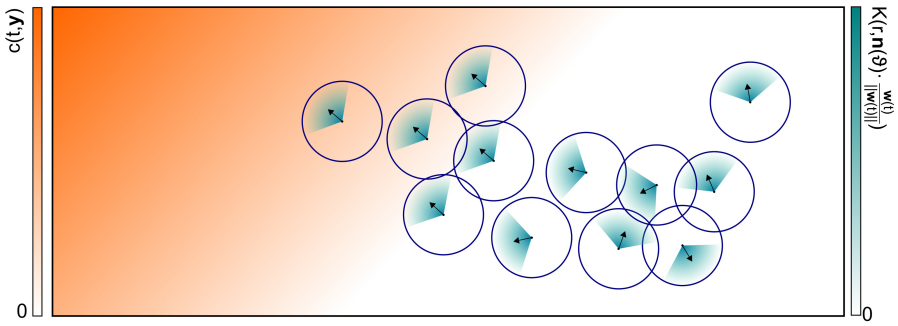

Figure 11: The extension of the proposed model to collective migration phenomena implies, for each time step, at least the evaluation of the integral terms W and M for each individual and eventually the computation of a non local cell-cell interaction velocity term.

vironmental cues (arising, for instance, when diffusive-based evolution laws for micro-environmental molecular variables are included), provided that the resulting integrand function is Riemann-integrable and that possible singularities are detectable.

Possible model and computational improvements and developments. From a pure modeling perspective, it would be interesting to extend the model to achieve a more realistic description of the distribution of fixed matrix proteins. In fact, it has been widely shown that the efficiency of cell migration is highly affected by the orientation and the spacing of ECM components. For instance, several experimental models have demonstrated the cell preference to migrate along aligned matrix fibers within 3D environments, such as fibroblasts in collagen [16] or neuronal cells in fibrin substrates [18]. Similarly, *in vivo* intravital imaging studies of carcinoma cells in the mammary fat pad have pointed out the preferential chemotactic movement of invasive malignant cells along thick bundles of collagen [11], while, in the lymph node paracortex, the aligned microarchitecture of fibronectin elements significantly influences the migratory behavior of T-cells [5].

Finally, the migratory behavior of cells is also affected by their proteolytic activity that impacts on the distribution of the matrix components [45]. This would amount to add a proper remodeling law for the fixed ECM elements as well as a reaction-diffusion equation for the low-diffusing cell matrix metalloproteinases.

From a numerical point of view, it would be interesting to analyze the extension of the presented computational procedure to cellular ensembles to

describe, in particular, collective migration phenomena (see, among others, [6, 7, 15, 32]). Referring to Fig. 11, the evolution of a system of N non-interacting particles can be obtained by solving Eqs. (1) and (3) for each individual. The overall computational cost of the proposed approach therefore obviously scales with N (unless parallel computing techniques are used). However, the dynamics of cellular ensembles are strongly affected by mutual interactions. On one hand, cells tend in fact to move away from other individuals, when they are excessively close, in order to preserve the minimal space they need to live. On the other hand, cells are also able to form adhesive bounds with surrounding agents to maintain the cohesion of an aggregate. A multiagent extension of our method therefore calls for the introduction in Eqs. (1) and (3) of proper interaction terms, which may have non-local characteristics, that have to be evaluated at each time step. For instance, each individual may be affected by the presence of all other agents falling within a suitable sensing regions. This may in turn amount to set a interaction velocity term which is given by the sum of pairwise contributions. From a numerical point of view, in the case of N cells, the direct evaluation of such sums would require a computational complexity of order $O(N^2)$. In this respect, computational strategies able to increase the numerical efficacy would be convenient (e.g., the so called fast summation methods, see for instance [1, 27, 28]).

Another non-trivial issue arising when cell dynamics is coupled with the evolution laws of some molecular quantity is that their dynamics are typically characterized by different spatio-temporal scales. From a numerical point of view, in these cases it is then crucial to choose suitable computational strategies able to both preserve a good degree of accuracy and avoid excessive costs. As an example, it may be convenient to use different tensor grids (and specific interpolation schemes) to deal with c and m. Analogously, differences in temporal scales may be tackled by choosing distinct time steps or time discretization schemes (e.g., implicit or adaptive schemes) for cell dynamics and for the molecular variables, and by properly coupling them. It is however clear that the proper numerical approach strongly depends on the explicit form of the mathematical model, as well as on the required accuracy.

Appendix A. Analytical results

This section is devoted to the discussion of the existence and uniqueness of the solution of the proposed model, and its dependence on the initial

condition. In this perspective, we first introduce two preliminary results and details on the notation used in this Section.

Preliminary results and notation

Proposition 1. *If the initial polarization vector* $\mathbf{w}(0)$ *is such that* $\|\mathbf{w}(0)\| =: w_0 \in (0,1]$*, Eq.* (3) *ensures that cell polarization* $\mathbf{w}(t)$ *has* $\|\mathbf{w}(t)\| \leq 1$ *for any* $t > 0$*.*

Proof. The inner product of Eq. (3) by $\mathbf{w}$ gives

$$\begin{aligned} \mathbf{w}(t) \cdot \frac{d\mathbf{w}(t)}{dt} &= \frac{1}{\tau}\left(\frac{\mathbf{W}(t,\, \mathbf{x}(t),\, \mathbf{w}(t)) \cdot \mathbf{w}(t)}{\|\mathbf{W}(t,\, \mathbf{x}(t),\, \mathbf{w}(t))\| + \chi} - \|\mathbf{w}(t)\|^2\right) \\ &\leq \frac{1}{\tau}\left(\frac{\|\mathbf{W}(t,\, \mathbf{x}(t),\, \mathbf{w}(t))\|}{\|\mathbf{W}(t,\, \mathbf{x}(t),\, \mathbf{w}(t))\| + \chi} - \|\mathbf{w}(t)\|\right)\|\mathbf{w}(t)\|. \end{aligned} \tag{A.1}$$

Since $\mathbf{w}(t) \cdot \frac{d\mathbf{w}(t)}{dt} = \|\mathbf{w}(t)\|\frac{d\|\mathbf{w}(t)\|}{dt}$, Eq. (A.1) becomes

$$\frac{d\|\mathbf{w}(t)\|}{dt} \leq \frac{1}{\tau}\left(\frac{\|\mathbf{W}(t,\, \mathbf{x}(t),\, \mathbf{w}(t))\|}{\|\mathbf{W}(t,\, \mathbf{x}(t),\, \mathbf{w}(t))\| + \chi} - \|\mathbf{w}(t)\|\right) \leq \frac{1}{\tau}\left(1 - \|\mathbf{w}(t)\|\right), \tag{A.2}$$

which gives $\|\mathbf{w}(t)\| \leq 1 - (1 - w_0)\exp(-t/\tau)$, and thus implies the thesis. □

Proposition 2. *For any* $\mathbf{w}_1, \mathbf{w}_2 \in \mathbb{R}^2$ *such that* $\|\mathbf{w}_1\|, \|\mathbf{w}_2\| \geq w_0$*, with* $w_0 \in (0,1]$*, there exists a constant* $C_{w_0} > 0$ *such that*

$$\left\| \frac{\mathbf{w}_1}{\|\mathbf{w}_1\|} - \frac{\mathbf{w}_2}{\|\mathbf{w}_2\|} \right\| \leq C_{w_0}\|\mathbf{w}_1 - \mathbf{w}_2\|. \tag{A.3}$$

Proof. Let us assume that $\|\mathbf{w}_1\| < \|\mathbf{w}_2\|$,

$$\left\| \frac{\mathbf{w}_1}{\|\mathbf{w}_1\|} - \frac{\mathbf{w}_2}{\|\mathbf{w}_2\|} \right\| = \frac{1}{\|\mathbf{w}_1\|}\left\| \mathbf{w}_1 - \frac{\|\mathbf{w}_1\|}{\|\mathbf{w}_2\|}\mathbf{w}_2 \right\| \leq \frac{\|\mathbf{w}_1 - \mathbf{w}_2\|}{\|\mathbf{w}_1\|} \leq \frac{\|\mathbf{w}_1 - \mathbf{w}_2\|}{w_0}. \tag{A.4}$$

Analogous considerations hold when $\|\mathbf{w}_2\| < \|\mathbf{w}_1\|$. □

Notation. Taking in account the above preliminary results, we denote by $\mathcal{I} := [0, T]$ the time interval such that, given an initial condition $\mathbf{w}(0)$ with $\|\mathbf{w}(0)\| = w_0 \in (0,1]$, it results $\|\mathbf{w}(t)\| \geq w_0$ for any $t \in \mathcal{I}$. In this respect, we introduce the set $\mathcal{D} := \{\mathbf{w} \in \mathbb{R}^2 \text{ such that } \|\mathbf{w}\| \in [w_0, 1]\} \subset \mathbb{R}^2$, i.e., the

set of possible cell polarization vectors. Moreover, in the rest of this section, cell speed $v(t)$ will by denoted by $v(t, \mathbf{x}, \mathbf{w})$, in order to highlight that it depends on the actual position $\mathbf{x}(t)$ and on the polarization vector $\mathbf{w}(t)$ of the cell, see Eqs. (7)-(9).

Main results

Let us introduce a set of assumptions that ensures existence and uniqueness of the solution of the system in Eq. (10) (and specified in Eqs. (1) and (3)), as well as its continuous dependence on the initial condition. This is stated by the following Theorem and relative proof.

Assumption 1. *$m : \mathbb{R}^2 \mapsto [0, 1]$ in Eq.(9) is Lipschitz continuous on $\mathbb{R}^2$.*

Assumption 2. *$c : \mathbb{R}_0^+ \times \mathbb{R}^2 \mapsto \mathbb{R}_0^+$ in Eq.(6) is bounded and Lipschitz continuous on $\mathbb{R}^2$.*

Assumption 3. *$K_\theta : [-1, 1] \mapsto [0, 1]$ in Eq.(6) has compact support over $(-\alpha, \alpha)$, with $\alpha \in [0, 1)$, and is Lipschitz continuous on its support.*

Assumption 4. *$\mathbf{z}_0 := (\mathbf{x}(0), \mathbf{w}(0))^T$ in Eq. (10) is such that $\|\mathbf{w}(0)\| = w_0 \in (0, 1]$.*

Remark 1. *We further remark that the kernel $K_r : [0, R] \mapsto [0, 1]$ in Eq.(6) is bounded on $[0, R]$.*

Theorem 1. *If the system in Eq. (10), and specified in Eqs. (1) and (3), satisfies Assumptions 1-4, the following statements hold true*

1. *$\mathbf{F} : \mathbb{R}_0^+ \times \mathbb{R}^2 \times \mathcal{D} \mapsto \mathbb{R}^2 \times \mathcal{D}$ is Lipschitz continuous on $\mathbb{R}^2 \times \mathcal{D}$ for any $t \in \mathcal{I}$;*
2. *(existence and uniqueness) Eq. (10) admits unique solution on $\mathcal{I}$;*
3. *(dependence on the initial condition) let $\boldsymbol{Z}_1(t) = (\mathbf{x}_1(t), \mathbf{w}_1(t))^T : \mathcal{I} \mapsto \mathbb{R}^2 \times \mathcal{D}$ and $\boldsymbol{Z}_2(t) = (\mathbf{x}_2(t), \mathbf{w}_2(t))^T : \mathcal{I} \mapsto \mathbb{R}^2 \times \mathcal{D}$ denote two cell trajectory-polarization pairs obtained when the initial condition is set equal to $\mathbf{z}_{0,1} := (\mathbf{x}_1(0), \mathbf{w}_1(0))^T \in \mathbb{R}^2 \times \mathcal{D}$ and $\mathbf{z}_{0,2} := (\mathbf{x}_2(0), \mathbf{w}_2(0))^T \in \mathbb{R}^2 \times \mathcal{D}$, respectively, it then results that*

$$\|\boldsymbol{Z}_1(t) - \boldsymbol{Z}_2(t)\|_{2\times 2} \leq \exp(\mathrm{Lip}(\mathbf{F})\, t)\, \|\mathbf{z}_{0,1} - \mathbf{z}_{0,2}\|_{2\times 2} \tag{A.5}$$

for any $t \in \mathcal{I}$, being $\|\cdot\|_{2\times 2}$ the norm on $\mathbb{R}^2 \times \mathbb{R}^2$ defined by $\|\mathbf{z}\|_{2\times 2} = \|\mathbf{x}\| + \|\mathbf{w}\|$ for any $\mathbf{z} = (\mathbf{x}, \mathbf{w})^T$.

Proof. By recalling that classical Cauchy-Lipschitz Theorem and the Grönwall's Lemma ensure that the Lipschitz continuity of $\mathbf{F}$ implies Statements 2 and 3, we have to demonstrate Statement 1 only.

Step 1. Lipschitz continuity of $\mathbf{F}$. Let us first remark that $\mathbf{F} : \mathbb{R}_0^+ \times \mathbb{R}^2 \times \mathcal{D} \mapsto \mathbb{R}^2 \times \mathcal{D}$ is Lipschitz continuous on $\mathbb{R}^2 \times \mathcal{D}$ for any $t \in \mathcal{I}$, if its components $\mathbf{g_x} : \mathbb{R}_0^+ \times \mathbb{R}^2 \times \mathcal{D} \mapsto \mathbb{R}^2$ and $\mathbf{g_w} : \mathbb{R}_0^+ \times \mathbb{R}^2 \times \mathcal{D} \mapsto \mathbb{R}^2$ (defined in Eq. (1) and (3), respectively) are both Lipschitz continuous on $\mathbb{R}^2 \times \mathcal{D}$ for any $t \in \mathcal{I}$. We therefore focus on each component, separately.

Step 2. Lipschitz continuity of $\boldsymbol{g_x}$. Let us fix $t \in \mathcal{I}$, for any $\mathbf{x}_1, \mathbf{x}_2 \in \mathbb{R}^2$ and $\mathbf{w}_1, \mathbf{w}_2 \in \mathcal{D}$, we have that

$$\begin{aligned} \|\mathbf{g_x}(t, \mathbf{x}_1, \mathbf{w}_1) - \mathbf{g_x}(t, \mathbf{x}_2, \mathbf{w}_2)\| &= \|v(t, \mathbf{x}_1, \mathbf{w}_1)\mathbf{w}_1 - v(t, \mathbf{x}_2, \mathbf{w}_2)\mathbf{w}_2\| \\ &\leq |v(t, \mathbf{x}_1, \mathbf{w}_1)| \, \|\mathbf{w}_1 - \mathbf{w}_2\| + |v(t, \mathbf{x}_1, \mathbf{w}_1) - v(t, \mathbf{x}_2, \mathbf{w}_2)| \|\mathbf{w}_2\|. \end{aligned} \tag{A.6}$$

Recalling Eq. (7)-(9) and that $m : \mathbb{R}^2 \mapsto [0, 1]$, it results that $M(t, \mathbf{x}, \mathbf{w}) \in [0, 1]$ and, in turn, that $v(t, \mathbf{x}, \mathbf{w}) \in [0, V_{\max}]$ for any $(t, \mathbf{x}, \mathbf{w}) \in \mathcal{I} \times \mathbb{R}^2 \times \mathcal{D}$. On the other hand, the Lipschitz continuity of m (see Assumption 1) implies the Lipschitz continuity of $M(t, \mathbf{x}, \mathbf{w})$ on $\mathbb{R}^2 \times \mathcal{D}$ for any $t \in \mathcal{I}$. Indeed, we have

$$\begin{aligned} &|M(t, \mathbf{x}_1, \mathbf{w}_1) - M(t, \mathbf{x}_2, \mathbf{w}_2)| \\ &\quad \leq \frac{1}{R} \int_0^R \left| m\left(\mathbf{x}_1 + r\frac{\mathbf{w}_1}{\|\mathbf{w}_1\|}\right) - m\left(\mathbf{x}_2 + r\frac{\mathbf{w}_2}{\|\mathbf{w}_2\|}\right) \right| dr \\ &\quad \leq \frac{\mathrm{Lip}(m)}{R} \int_0^R \left(\|\mathbf{x}_1 - \mathbf{x}_2\| + r \left\| \frac{\mathbf{w}_1}{\|\mathbf{w}_1\|} - \frac{\mathbf{w}_2}{\|\mathbf{w}_2\|} \right\| \right) dr \\ &\quad \leq \mathrm{Lip}(m) \max\left\{1, \frac{R}{2}\right\} \left(\|\mathbf{x}_1 - \mathbf{x}_2\| + \left\| \frac{\mathbf{w}_1}{\|\mathbf{w}_1\|} - \frac{\mathbf{w}_2}{\|\mathbf{w}_2\|} \right\| \right) \\ &\quad \leq \underbrace{\mathrm{Lip}(m) \max\left\{1, \frac{R}{2} C_{w_0}\right\}}_{=: \mathrm{Lip}(M)} \Big(\|\mathbf{x}_1 - \mathbf{x}_2\| + \|\mathbf{w}_1 - \mathbf{w}_2\| \Big), \end{aligned} \tag{A.7}$$

where the last inequality follows from Proposition 2 and the definition of $\mathcal{I}$. It is then easy to verify that also $v(t, \mathbf{x}, \mathbf{w})$ results Lipschitz continuous on $\mathbb{R}^2 \times \mathcal{D}$ for any $t \in \mathcal{I}$:

$$\begin{aligned} |v(t, \mathbf{x}_1, \mathbf{w}_1) - v(t, \mathbf{x}_2, \mathbf{w}_2)| &\leq 8V_{\max} |M(t, \mathbf{x}_1, \mathbf{w}_1) - M(t, \mathbf{x}_2, \mathbf{w}_2)| \\ &\leq \underbrace{8V_{\max}\mathrm{Lip}(M)}_{=:\,\mathrm{Lip}(v)} \Big(\|\mathbf{x}_1 - \mathbf{x}_2\| + \|\mathbf{w}_1 - \mathbf{w}_2\| \Big). \end{aligned} \tag{A.8}$$

Eq. (A.6) therefore reads

$$\begin{aligned} &\|\mathbf{g}_{\mathbf{x}}(t, \mathbf{x}_1, \mathbf{w}_1) - \mathbf{g}_{\mathbf{x}}(t, \mathbf{x}_2, \mathbf{w}_2)\| \\ &\quad \leq \mathrm{Lip}(v) \max\{1, V_{\max}\} \Big(\|\mathbf{x}_1 - \mathbf{x}_2\| + \|\mathbf{w}_1 - \mathbf{w}_2\| \Big). \end{aligned} \tag{A.9}$$

Step 3. Lipschitz continuity of $\boldsymbol{g}_{\mathbf{w}}$. Let us fix again $t \in \mathcal{I}$, for any $\mathbf{x}_1, \mathbf{x}_2 \in \mathbb{R}^2$ and $\mathbf{w}_1, \mathbf{w}_2 \in \mathcal{D}$, it results

$$\begin{aligned} &\|\mathbf{g}_{\mathbf{w}}(t, \mathbf{x}_1, \mathbf{w}_1) - \mathbf{g}_{\mathbf{w}}(t, \mathbf{x}_2, \mathbf{w}_2)\| \\ &\quad \leq \frac{1}{\tau} \left(\left\| \frac{\mathbf{W}(t, \mathbf{x}_1, \mathbf{w}_1)}{\|\mathbf{W}(t, \mathbf{x}_1, \mathbf{w}_1)\| + \chi} - \frac{\mathbf{W}(t, \mathbf{x}_2, \mathbf{w}_2)}{\|\mathbf{W}(t, \mathbf{x}_2, \mathbf{w}_2)\| + \chi} \right\| + \|\mathbf{w}_1 - \mathbf{w}_2\| \right). \end{aligned} \tag{A.10}$$

Since $\chi > 0$, we have that $(\|\mathbf{W}(t, \mathbf{x}, \mathbf{w})\| + \chi)^{-1} \leq \chi^{-1}$ and therefore

$$\begin{aligned} &\left\| \frac{\mathbf{W}(t, \mathbf{x}_1, \mathbf{w}_1)}{\|\mathbf{W}(t, \mathbf{x}_1, \mathbf{w}_1)\| + \chi} - \frac{\mathbf{W}(t, \mathbf{x}_2, \mathbf{w}_2)}{\|\mathbf{W}(t, \mathbf{x}_2, \mathbf{w}_2)\| + \chi} \right\| \\ &\quad \leq \frac{2\|\mathbf{W}(t, \mathbf{x}_2, \mathbf{w}_2)\| + \chi}{\chi^2} \, \|\mathbf{W}(t, \mathbf{x}_1, \mathbf{w}_1) - \mathbf{W}(t, \mathbf{x}_2, \mathbf{w}_2)\|. \end{aligned} \tag{A.11}$$

It is then straightforward to verify that the boundedness of c, K_θ, and K_r imply the boundedness of $\mathbf{W}$. Specifically, it results $\|\mathbf{W}(t, \mathbf{x}, \mathbf{w})\| \leq c_0 \pi R^2 =: C_{\mathbf{W}}$ for any $(t, \mathbf{x}, \mathbf{w}) \in \mathcal{I} \times \mathbb{R}^2 \times \mathcal{D}$. Concerning instead the last norm in Eq. (A.11), we now show that the Lipschitz continuity of $\mathbf{W}$ follows from the boundedness and Lipschitz continuity of c and K_θ (stated in Assumptions 2 and 3), and the boundedness of K_r (see

Remark 1). Specifically, we start by setting

$$\begin{aligned}&\|\mathbf{W}(t,\mathbf{x}_1,\mathbf{w}_1)-\mathbf{W}(t,\mathbf{x}_2,\mathbf{w}_2)\|\\&\quad\le\big\|\mathbf{W}(t,\mathbf{x}_1,\mathbf{w}_1)-\mathbf{W}(t,\mathbf{x}_2,\mathbf{w}_1)\big\|+\big\|\mathbf{W}(t,\mathbf{x}_2,\mathbf{w}_1)-\mathbf{W}(t,\mathbf{x}_2,\mathbf{w}_2)\big\|,\end{aligned}\tag{A.12}$$

and recalling the definition of $\mathbf{W}(t,\mathbf{x},\mathbf{w})$ given in Eq. (6). On one hand, dealing with the first term in Eq. (A.12), i.e.,

$$\begin{aligned}&\big\|\mathbf{W}(t,\mathbf{x}_1,\mathbf{w}_1)-\mathbf{W}(t,\mathbf{x}_2,\mathbf{w}_1)\big\|\\&\quad=\Big\|\int_{\mathcal{S}(\mathbf{x}_1)}K_r(\|\mathbf{y}-\mathbf{x}_1\|)\,K_\theta\left(\frac{\mathbf{y}-\mathbf{x}_1}{\|\mathbf{y}-\mathbf{x}_1\|}\cdot\frac{\mathbf{w}_1}{\|\mathbf{w}_1\|}\right)c(t,\mathbf{y})\frac{\mathbf{y}-\mathbf{x}_1}{\|\mathbf{y}-\mathbf{x}_1\|}d\mathbf{y}\\&\qquad-\int_{\mathcal{S}(\mathbf{x}_2)}K_r(\|\mathbf{y}-\mathbf{x}_2\|)\,K_\theta\left(\frac{\mathbf{y}-\mathbf{x}_2}{\|\mathbf{y}-\mathbf{x}_2\|}\cdot\frac{\mathbf{w}_1}{\|\mathbf{w}_1\|}\right)c(t,\mathbf{y})\frac{\mathbf{y}-\mathbf{x}_2}{\|\mathbf{y}-\mathbf{x}_2\|}d\mathbf{y}\Big\|,\end{aligned}\tag{A.13}$$

we take the advantage that the direction of polarization $\mathbf{w}_1$ is fixed. Proper changes of variables (i.e., $\mathbf{y}=\mathbf{x}_1+\boldsymbol{\xi}$ in the first integral and $\mathbf{y}=\mathbf{x}_2+\boldsymbol{\xi}$ in the second one) in fact allow us to rewrite both integrals on the circular area $\mathcal{S}(\mathbf{0})$ centered at the origin and with radius R, and hence to use the boundedness of kernels K_r and K_θ, and the Lipschitz continuity of c. Specifically, we have

$$\begin{aligned}\text{(A.13)}&=\Big\|\int_{\mathcal{S}(\mathbf{0})}K_r(\|\boldsymbol{\xi}\|)\,K_\theta\left(\frac{\boldsymbol{\xi}}{\|\boldsymbol{\xi}\|}\cdot\frac{\mathbf{w}_1}{\|\mathbf{w}_1\|}\right)\Big(c(t,\mathbf{x}_1+\boldsymbol{\xi})\\&\qquad-c(t,\mathbf{x}_2+\boldsymbol{\xi})\Big)\frac{\boldsymbol{\xi}}{\|\boldsymbol{\xi}\|}d\boldsymbol{\xi}\Big\|\\&\le\int_{\mathcal{S}(\mathbf{0})}\Big|K_r(\|\boldsymbol{\xi}\|)\Big|\,\Big|K_\theta\left(\frac{\boldsymbol{\xi}}{\|\boldsymbol{\xi}\|}\cdot\frac{\mathbf{w}_1}{\|\mathbf{w}_1\|}\right)\Big|\,\Big|c(t,\mathbf{x}_1+\boldsymbol{\xi})\\&\qquad-c(t,\mathbf{x}_2+\boldsymbol{\xi})\Big|\,d\boldsymbol{\xi}\\&\le\underbrace{R^2\operatorname{acos}\alpha\operatorname{Lip}(c)}_{=:\operatorname{Lip}(\mathbf{W}_\mathbf{x})}\|\mathbf{x}_1-\mathbf{x}_2\|.\end{aligned}\tag{A.14}$$

On the other hand, the second term in Eq. (A.12), where cell position $\mathbf{x}_1$ is fixed, reads as follows

$$
\begin{aligned}
&\left\|\mathbf{W}(t,\mathbf{x}_1,\mathbf{w}_1)-\mathbf{W}(t,\mathbf{x}_1,\mathbf{w}_2)\right\| \\
&\quad= \Big\| \int_{\mathcal{S}(\mathbf{x}_1)} K_r(\|\mathbf{y}-\mathbf{x}_1\|)\, K_\theta\left(\frac{\mathbf{y}-\mathbf{x}_1}{\|\mathbf{y}-\mathbf{x}_1\|}\cdot\frac{\mathbf{w}_1}{\|\mathbf{w}_1\|}\right) c(t,\mathbf{y})\frac{\mathbf{y}-\mathbf{x}_1}{\|\mathbf{y}-\mathbf{x}_1\|}d\mathbf{y} \\
&\qquad - \int_{\mathcal{S}(\mathbf{x}_1)} K_r(\|\mathbf{y}-\mathbf{x}_1\|)\, K_\theta\left(\frac{\mathbf{y}-\mathbf{x}_1}{\|\mathbf{y}-\mathbf{x}_1\|}\cdot\frac{\mathbf{w}_2}{\|\mathbf{w}_2\|}\right) c(t,\mathbf{y})\frac{\mathbf{y}-\mathbf{x}_1}{\|\mathbf{y}-\mathbf{x}_1\|}d\mathbf{y}\Big\| \\
&\quad\le \int_{\mathcal{S}(\mathbf{x}_1)} \Big|K_r(\|\mathbf{y}-\mathbf{x}_1\|)\Big| \Big|K_\theta\left(\frac{\mathbf{y}-\mathbf{x}_1}{\|\mathbf{y}-\mathbf{x}_1\|}\cdot\frac{\mathbf{w}_1}{\|\mathbf{w}_1\|}\right) \\
&\qquad - K_\theta\left(\frac{\mathbf{y}-\mathbf{x}_1}{\|\mathbf{y}-\mathbf{x}_1\|}\cdot\frac{\mathbf{w}_2}{\|\mathbf{w}_2\|}\right)\Big| \Big|c(t,\mathbf{y})\Big|\, d\mathbf{y},
\end{aligned}
\tag{A.15}
$$

and thus the boundedness of c and K_r imply that

$$
\begin{aligned}
\text{(A.15)} \le\ c_0 \int_{\mathcal{S}(\mathbf{x}_1)} &\Big|K_\theta\left(\frac{\mathbf{y}-\mathbf{x}_1}{\|\mathbf{y}-\mathbf{x}_1\|}\cdot\frac{\mathbf{w}_1}{\|\mathbf{w}_1\|}\right) \\
&- K_\theta\left(\frac{\mathbf{y}-\mathbf{x}_1}{\|\mathbf{y}-\mathbf{x}_1\|}\cdot\frac{\mathbf{w}_2}{\|\mathbf{w}_2\|}\right)\Big| d\mathbf{y}.
\end{aligned}
\tag{A.16}
$$

Recalling that K_θ is Lipschitz continuous over its support (see Assumption 3), we denote by $U_{\mathbf{x},\mathbf{w}}$ the support of the integrand function in Eq. (6), and define three subsets of $\mathcal{S}(\mathbf{x}_1)$: $\mathcal{A}_{1,2} := U_{\mathbf{x}_1,\mathbf{w}_1} \cap U_{\mathbf{x}_1,\mathbf{w}_2}$, (which results not empty if the angle between $\mathbf{w}_1$ and $\mathbf{w}_2$ is sufficiently small), $\mathcal{A}_1 := U_{\mathbf{x}_1,\mathbf{w}_1} \setminus \mathcal{A}_{1,2}$ and $\mathcal{A}_2 := U_{\mathbf{x}_1,\mathbf{w}_2} \setminus \mathcal{A}_{1,2}$ (see the sketch in Fig. A.12). Eq. (A.16) then becomes

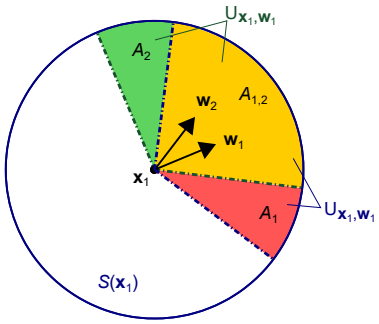

Figure A.12: Representation of the circular sectors $U_{\mathbf{x}_1,\mathbf{w}_1}$ and $U_{\mathbf{x}_1,\mathbf{w}_2}$, when $\mathbf{w}_1$ and $\mathbf{w}_2$ are sufficiently close to have not empty intersection $\mathcal{A}_{1,2} := U_{\mathbf{x}_1,\mathbf{w}_1} \cap U_{\mathbf{x}_1,\mathbf{w}_2}$ (yellow area). The other circular sectors are $\mathcal{A}_1 := U_{\mathbf{x}_1,\mathbf{w}_1} \setminus \mathcal{A}_{1,2}$ (orange area) and $\mathcal{A}_2 := U_{\mathbf{x}_1,\mathbf{w}_2} \setminus \mathcal{A}_{1,2}$ (green area).

$$
\begin{aligned}
(\text{A.16}) \leq c_0 \Big(& \int_{\mathcal{A}_1} \left| K_\theta \left(\frac{\mathbf{y}-\mathbf{x}_1}{\|\mathbf{y}-\mathbf{x}_1\|} \cdot \frac{\mathbf{w}_1}{\|\mathbf{w}_1\|} \right) \right| d\mathbf{y} \\
& + \int_{\mathcal{A}_{1,2}} \left| K_\theta \left(\frac{\mathbf{y}-\mathbf{x}_1}{\|\mathbf{y}-\mathbf{x}_1\|} \cdot \frac{\mathbf{w}_1}{\|\mathbf{w}_1\|} \right) - K_\theta \left(\frac{\mathbf{y}-\mathbf{x}_1}{\|\mathbf{y}-\mathbf{x}_1\|} \cdot \frac{\mathbf{w}_2}{\|\mathbf{w}_2\|} \right) \right| d\mathbf{y} \\
& + \int_{\mathcal{A}_2} \left| K_\theta \left(\frac{\mathbf{y}-\mathbf{x}_1}{\|\mathbf{y}-\mathbf{x}_1\|} \cdot \frac{\mathbf{w}_2}{\|\mathbf{w}_2\|} \right) \right| d\mathbf{y} \Big),
\end{aligned}
\tag{A.17}
$$

and the Lipschitz continuity of K_θ on $\mathcal{A}_{1,2}$, and its boundedness on $\mathcal{A}_1$

and $\mathcal{A}_2$ give

$$
\begin{aligned}
(\text{A.17}) \leq\ & c_0\Big(\mathrm{Lip}(K_\theta)\int_{\mathcal{A}_{1,2}}\Big|\frac{\mathbf{y}-\mathbf{x}_1}{\|\mathbf{y}-\mathbf{x}_1\|}\cdot\Big(\frac{\mathbf{w}_1}{\|\mathbf{w}_1\|}-\frac{\mathbf{w}_2}{\|\mathbf{w}_2\|}\Big)\Big|d\mathbf{y}+2\int_{\mathcal{A}_1}d\mathbf{y}\Big)\\
\leq\ & c_0\Big(\mathrm{Lip}(K_\theta)\int_{\mathcal{A}_{1,2}}\Big\|\frac{\mathbf{y}-\mathbf{x}_1}{\|\mathbf{y}-\mathbf{x}_1\|}\Big\|\Big\|\frac{\mathbf{w}_1}{\|\mathbf{w}_1\|}-\frac{\mathbf{w}_2}{\|\mathbf{w}_2\|}\Big\|d\mathbf{y}\\
&\qquad +2\,R^2\Big\|\frac{\mathbf{w}_1}{\|\mathbf{w}_1\|}-\frac{\mathbf{w}_2}{\|\mathbf{w}_2\|}\Big\|\Big)\\
=\ & c_0\,R^2\,(\mathrm{Lip}(K_\theta)\mathrm{acos}\alpha+2)\Big\|\frac{\mathbf{w}_1}{\|\mathbf{w}_1\|}-\frac{\mathbf{w}_2}{\|\mathbf{w}_2\|}\Big\|\\
\leq\ & \underbrace{c_0\,R^2\,(\mathrm{Lip}(K_\theta)\mathrm{acos}\alpha+2)\,C_{w_0}}_{\mathrm{Lip}(\mathbf{W}_\mathbf{w})}\|\mathbf{w}_1-\mathbf{w}_2\|.
\end{aligned}
\tag{A.18}
$$

Specifically, in Eq. (A.18), the area of the circular sector $\mathcal{A}_1$ is bounded by $R^2\|\mathbf{w}_1/\|\mathbf{w}_1\|-\mathbf{w}_2/\|\mathbf{w}_2\|\|$, and the last inequality follows from Proposition 2 and the definition of $\mathcal{I}$. Summing up, Eqs. (A.12)-(A.18) give

$$
\begin{aligned}
&\|\mathbf{W}(t,\mathbf{x}_1,\mathbf{w}_1)-\mathbf{W}(t,\mathbf{x}_2,\mathbf{w}_2)\|\\
&\quad\leq\underbrace{\max\{\mathrm{Lip}(\mathbf{W}_\mathbf{x}),\mathrm{Lip}(\mathbf{W}_\mathbf{w})\}}_{\mathrm{Lip}(\mathbf{W})}\Big(\|\mathbf{x}_1-\mathbf{x}_2\|+\|\mathbf{w}_1-\mathbf{w}_2\|\Big),
\end{aligned}
\tag{A.19}
$$

and Eq. (A.10) thus reads

$$
\begin{aligned}
&\|\mathbf{g}_\mathbf{w}(t,\mathbf{x}_1,\mathbf{w}_1)-\mathbf{g}_\mathbf{w}(t,\mathbf{x}_2,\mathbf{w}_2)\|\\
&\quad\leq\frac{1}{\tau}\left(\frac{(2\,C_\mathbf{W}+\chi)\mathrm{Lip}(\mathbf{W})}{\chi^2}+1\right)\Big(\|\mathbf{x}_1-\mathbf{x}_2\|+\|\mathbf{w}_1-\mathbf{w}_2\|\Big).
\end{aligned}
\tag{A.20}
$$

Eqs. (A.9) and (A.20) then imply the Lipschitz continuity of $\mathbf{F}$ on $\mathbb{R}^2\times\mathcal{D}$ for any $t\in\mathcal{I}$ and, as already said, Statements 2 and 3. □

Acknowledgments. The research that lead to the present work has been supported by GNCS-INdAM 2019 research program: "Metodi di approssimazione locale con applicazioni all'analisi isogeometrica e alle equazioni integrali di contorno" and by MIUR grant "Dipartimenti di Eccellenza 2018-2022" (Project no. E11G18000350001). AC has been also supported by a Research Grant from the Istituto Nazionale di Alta Matematica "F. Severi" (INdAM). The authors extend warm thanks to Prof. Luigi Preziosi for many stimulating and fruitful discussions.

References

[1] Albi, G., Pareschi, L., 2012. Binary interaction algorithms for the simulation of flocking and swarming dynamics. Multiscale Modeling & Simulation 11, 1–29. doi:10.1137/120868748.

[2] Allena, R., Scianna, M., Preziosi, L., 2016. A cellular potts model of single cell migration in presence of durotaxis. Mathematical Biosciences 275, 57–70. doi:10.1016/j.mbs.2016.02.011.

[3] Anderson, A.R., Weaver, A.M., Cummings, P.T., Quaranta, V., 2006. Tumor morphology and phenotypic evolution driven by selective pressure from the microenvironment. Cell 127, 905–915. doi:10.1016/j.cell.2006.09.042.

[4] Armstrong, N.J., Painter, K.J., Sherratt, J.A., 2006. A continuum approach to modelling cell-cell adhesion. Journal of Theoretical Biology 243, 98–113. doi:10.1016/j.jtbi.2006.05.030.

[5] Bajénoff, M., Egen, J.G., Koo, L.Y., Laugier, J.P., Brau, F., Glaichenhaus, N., Germain, R.N., 2006. Stromal cell networks regulate lymphocyte entry, migration, and territoriality in lymph nodes. Immunity 25, 989–1001. doi:10.1016/j.immuni.2006.10.011.

[6] Carrillo, J.A., Colombi, A., Scianna, M., 2018. Adhesion and volume constraints via nonlocal interactions determine cell organisation and migration profiles. Journal of Theoretical Biology 445, 75–91. doi:10.1016/j.jtbi.2018.02.022.

[7] Carrillo, J.A., Fornasier, M., Toscani, G., Vecil, F., 2010. Particle, kinetic, and hydrodynamic models of swarming, in: Naldi, G., Pareschi, L., Toscani, G. (Eds.), Mathematical Modeling of Collective Behavior in Socio-Economic and Life Sciences. Birkhäuser Boston. Modeling and Simulation in Science, Engineering and Technology, pp. 297–336. doi:10.1007/978-0-8176-4946-3_12.

[8] Cavalcanti-Adam, E.A., Volberg, T., Micoulet, A., Kessler, H., Geiger, B., Spatz, J.P., 2007. Cell spreading and focal adhesion dynamics are regulated by spacing of integrin ligands. Biophysical Journal 92, 2964–2974. doi:10.1529/biophysj.106.089730.

[9] Chaplain, M., Lachowicz, M., Szymanska, Z., Wrzosek, D., 2011. Mathematical modelling of cancer invasion: the importance of cell-cell adhesion and cell-matrix adhesion. hematical Models and Methods in Applied Sciences 21, 719–743. doi:10.1142/S0218202511005192.

[10] Colombi, A., Scianna, M., Painter, K.J., Preziosi, L., 2019. Modelling chase-and-run migraiton in heterogeneous populations. Journal of Mathemathical Biology doi:10.1007/s00285-019-01421-9. (Accepted/In press).

[11] Condeelis, J., Segall, J.E., 2003. Intravital imaging of cell movement in tumours. Nature Reviews Cancer 3, 921–930. doi:10.1038/nrc1231.

[12] Davis, P.J., Rabinowitz, P., 1984. Methods of numerical integration. Academic Press. doi:10.1016/C2013-0-10566-1.

[13] De Boor, C., 1978. A practical guide to splines, in: Applied Mathematical Sciences. doi:10.1007/978-1-4612-6333-3.

[14] Di Costanzo, E., Menci, M., Messina, E., Natalini, R., Vecchio, A., 2020. A hybrid model of collective motion of discrete particles under alignment and continuum chemotaxis. Discrete & Continuous Dynamical Systems - B 25, 443–472. doi:10.3934/dcdsb.2019189.

[15] Di Costanzo, E., Natalini, R., Preziosi, L., 2015. A hybrid mathematical model for self-organizing cell migration in the zebrafish lateral line. Journal of Mathematical Biology 71, 171–214. doi:10.1007/s00285-014-0812-9.

[16] Dickinson, R.B., Guido, S., Tranquillo, R.T., 1994. Biased cell migration of fibroblasts exhibiting contact guidance in oriented collagen gels. Annals of Biomedical Engineering 22, 342–356. doi:10.1007/BF02368241.

[17] Drasdo, D., 2003. On selected individual-based approaches to the dynamics of multicellular systems, in: Alt, W., Chaplain, M., Griebel, M., Lenz, J. (Eds.), Polymer and Cell Dynamics. Birkhäuser, Basel. Mathematics and Biosciences in Interaction, pp. 169–203. doi:10.1007/978-3-0348-8043-5_15.

[18] Dubey, N., Letourneau, P.C., Tranquillo, R.T., 2001. Neuronal contact guidance in magnetically aligned fibrin gels: effect of variation in gel mechano-structural properties. Biomaterials 22, 1065–1075. doi:10.1016/S0142-9612(00)00341-0.

[19] Engler, A., Bacakova, L., Newman, C., Hategan, A., Griffin, M., Discher, D., 2004. Substrate compliance versus ligand density in cell on gel responses. Biophysical Journal 86, 617–628. doi:10.1016/S0006-3495(04)74140-5.

[20] Etienne-Manneville, S., 2004. Cdc42–the centre of polarity. Journal of Cell Science 117, 1291–1300. doi:10.1242/jcs.01115.

[21] Friedl, P., Gilmour, D., 2009. Collective cell migration in morphogenesis, regeneration and cancer. Nature Reviews Molecular Cell Biology 10, 445–457. doi:10.1038/nrm2720.

[22] Friedl, P., Wolf, K., 2003. Tumour-cell invasion and migration: Diversity and escape mechanisms. Nature Reviews Cancer 3, 362–374. doi:10.1038/nrc1075.

[23] Friedl, P., Wolf, K., 2010. Plasticity of cell migration: a multiscale tuning model. Journal of Cell Biology 188, 11–19. doi:10.1083/jcb.200909003.

[24] Funamoto, S., Meili, R., Lee, S., Parry, L., Firtel, R.A., 2002. Spatial and temporal regulation of 3-phosphoinositides by pi 3-kinase and pten mediates chemotaxis. Cell 109, 611–623. doi:10.1016/S0092-8674(02)00755-9.

[25] Gaudet, C., Marganski, W., Kim, S., Brown, C.T., Gunderia, V., Dembo, M., Wong, J., 2003. Influence of type i collagen surface density on fibroblast spreading, motility, and contractility. Biophysical Journal 85, 3329–3335. doi:10.1016/S0006-3495(03)74752-3.

[26] Goodman, S.L., Risse, G., Vondermark, K., 1989. The e8 subfragment of laminin promotes locomotion of myoblasts over extracellular matrix. The Journal of Cell Biology 109, 799–809. doi:10.1083/jcb.109.2.799.

[27] Greengard, L., . Fast algorithms for classical physics. Science .

[28] Greengard, L., Rokhlin, V., 1987. A fast algorithm for particle simulations. Journal of Computational Physics 73, 325–348. doi:10.1016/0021-9991(87)90140-9.

[29] Kurosaka, S., Kashina, A., 2008. Cell biology of embryonic migration. Birth Defects Res C Embryo Today 84, 102–122. doi:10.1002/bdrc.20125.

[30] Lambert, J.D., 1991. Numerical methods for ordinary differential systems: the initial value problem. John Wiley & Sons, New York.

[31] Lo, C., Wang, H., Dembo, M., Wang, Y., 2000. Cell movement is guided by the rigidity of the substrate. Biophysical Journal 79, 144–152. doi:10.1016/S0006-3495(00)76279-5.

[32] Mackey, A., Kolokolnikov, T., Bertozzi, A.L., 2014. Two-species particle aggregation and stability of co-dimension one solutions. Discrete & Continuous Dynamics Systems - B 19, 1411–1436. doi:10.3934/dcdsb.2014.19.1411.

[33] Macklin, P., Edgerton, M.E., Thompson, A.M., Cristini, V., 2012. Patient-calibrated agent-based modelling of ductal carcinoma in situ (dcis): from microscopic measurements to macroscopic predictions of clinical progression. Journal of Theoretical Biology 301, 122–140. doi:10.1016/j.jtbi.2012.02.002.

[34] McCann, C.P., Kriebel, P.W., Parent, C.A., Losert, W., 2010. Cell speed, persistence and information transmission during signal relay and collective migration. Journal of Cell Science 123, 1724–1731. doi:10.1242/jcs.060137.

[35] Munaron, L., Fiorio Pla, A., 2009. Endothelial calcium machinery and angiogenesis: understanding physiology to interfere with pathology. Current Medical Chemistry 16, 4691–4703. doi:10.2174/092986709789878210.

[36] Nobes, C.D., Hall, A., 1999. Rho gtpases control polarity, protrusion, and adhesion during cell movement. Journal of Cell Biology 144, 1235–1244. doi:10.1083/jcb.144.6.1235.

[37] Painter, J.K., 2009. Modelling cell migration strategies in the extracellular matrix. Journal of Mathematical Biology 58, 511–543. doi:10.1007/s00285-008-0217-8.

[38] Palecek, S.P., Loftus, J.C., Ginsberg, M.H., Lauffenburger, D.A., Horwitz, A.F., 1997. Integrin-ligand binding properties govern cell migration speed through cell-substratum adhesiveness. Nature 385, 537–540. doi:10.1038/385537a0.

[39] Ridley, A.J., Paterson, H. F. Johnston, C.L., Diekmann, D., Hall, A., 1992. The small gtp-binding protein rac regulates growth factor-induced membrane ruffling. Cell 70, 401–410. doi:10.1016/0092-8674(92)90164-8.

[40] Sahai, E., 2007. Illuminating the metastatic process. Nature Reviews Cancer 7, 737–749. doi:10.1038/nrc2229.

[41] Scianna, M., Preziosi, L., 2012. Multiscale developments of the cellular potts model. SIAM Journal on Multiscale Modeling and Simulation 10, 342–382. doi:10.1137/100812951.

[42] Szabó, A., Melchionda, M., Nastasi, G., Woods, M.L., Campo, S., Perris, R., Mayor, R., 2016. In vivo confinement promotes collective migration of neural crest cells. Journal of Cell Biology 213, 543–555. doi:10.1083/jcb.201602083.

[43] Theveneau, E., Steventon, B., Scarpa, E., Garcia, S., Trepat, X., Streit, A., Mayor, R., 2013. Chase-and-run between adjacent cell populations promotes directional collective migration. Nature Cell Biology 15, 763–772. doi:10.1038/ncb2772.

[44] Wolf, K., Alexander, S., Schacht, V., Coussens, L.M., Von Andrian, U.H., Van Rheenen, J., Deryugina, E., Friedl, P., 2009. Collagen-based

cell migration models in vitro and in vivo. Seminars in Cell & Developmental Biology 20, 931–941. doi:10.1016/j.semcdb.2009.08.005.

[45] Wolf, K., Wu, Y.I., Liu, Y., Geiger, J., Tam, E., Overall, C., Stack, M.S., P., F., 2007. Multi-step pericellular proteolysis controls the transition from individual to collective cancer cell invasion. Nature Cell Biology 9, 893–904. doi:10.1038/ncb1616.

[46] Zhigun, A., Surulescu, C., Painter, K., Chen, L., 2019. Mathematical models for cell migration: a nonlocal perspective. Philosophical Transactions of the Royal Society B (Accepted/In press).