Research article

Discrete attachment to a cellulolytic biofilm modeled by an Itô stochastic differential equation

Yousef Rohanizadegan[1], Stefanie Sonner[2] and Hermann J. Eberl[1,*]

[1] Department of Mathematics and Statistics, University of Guelph, 50 Stone Rd E, Guelph ON, N1G 2W1, Canada

[2] Radboud University, Department of Mathematics, Postbus 9010, 6500 GL Nijmegen, The Netherlands

* **Correspondence:** Email: heberl@uoguelph.ca; Tel: +15198244120 ext. 52622.

Abstract: We propose a mathematical framework for introducing random attachment of bacterial cells in a deterministic continuum model of cellulosic biofilms. The underlying growth model is a highly nonlinear coupled PDE-ODE system. It is regularised and discretised in space. Attachment is described then via an auxiliary stochastic process that induces impulses in the biomass equation. The resulting system is an Itô stochastic differential equation. Unlike the more direct approach of modeling attachment by additive noise, the proposed model preserves non-negativity of solutions. Our numerical simulations are able to reproduce characteristic features of cellulolytic biofilms with cell attachment from the aqueous phase. Grid refinement studies show convergence for the expected values of spatially integrated biomass density and carbon concentration. We also examine the sensitivity of the model with respect to the parameters that control random attachment.

Keywords: attachment; cellulolytic biofilm; cellulosic biofilm; mathematical model; nonlinear diffusion; numerical simulation; stochastic differential equation

1. Introduction

Bacterial biofilms are microbial depositions on, usually immersed, abiotic or biotic surfaces. The attachment of cells from the surrounding aqueous phase to the substratum, or to already existing colonies, plays a prominent role in the biofilm life cycle and for the initiation of biofilm colony formation [1]. Despite its importance, cell attachment is normally not accounted for in mathematical models of biofilm growth that have been studied in the literature for more than three decades. Many of these models are deterministic continuum models, whereas cell attachment on the biofilm scale is typically considered a discrete stochastic process. Therefore incorporating attachment in existing

models is not straightforward.

The traditional one-dimensional Wanner-Gujer biofilm model, which is frequently used in engineering applications to study biofilm performance, is a coupled hyperbolic-elliptic free boundary value problem with nonlocal effects. Underlying this model is the assumption that a biofilm covers the substratum homogeneously and that biofilm processes are governed only by substrate gradients normal to the surface area. The few studies that include attachment processes usually assume a constant attachment rate, i.e., describe attachment as a continuous, deterministic process [2–5]. The same assumption is also made in a multiscale model of biofilm formation in porous media, in which exchange of biomass between the actual biofilm and the aqueous phase is considered [6]. This model concept, due to its one-dimensional nature, does not allow to track spatial heterogeneities in biofilm structure or how they may arise from sparsely distributed attachment events.

To overcome the limitations of one-dimensional models for the study of biofilm structure-function-dynamics relationships, several multi-dimensional partial differential equation biofilm models have been proposed. This includes models that can be considered direct extensions of the Wanner-Gujer model [7, 8], or models that are based on different basic assumptions, such as [9–11] which describe biofilms as fluids, or the density-dependent diffusion-reaction model [12] which takes the view of biofilms as spatially structured populations. Whereas biofilms have been characterised as both, the model [12] is the only one that has been derived from both perspectives, see [13]. Also in such models attachment may be accounted for as a deterministic continuous process, e.g., employing a constant attachment rate as in [14] in a fluids model, or via chemotaxis as for example in a semilinear diffusion-reaction equation that does not explicitly distinguish between bacteria in the biofilm and the surrounding aqueous phase [15]. The random and discrete localised aspects of attachment in these partial differential equation models are, if at all, usually only included in the form of randomly generated initial data that prescribe location, and sometimes size, of colonies at the onset of the simulation, e.g., in [7, 13, 16].

To overcome this limitation and to be able to account for discrete stochastic attachment in deterministic continuum biofilm models is the objective of the present study. More specifically, we will be considering the extension of a two-dimensional model for cellulolytic biofilms. Cellulolytic biofilms play an important role in the production of cellulosic ethanol, a biofuel that can be implemented in the currently existing transportation infrastructure.

The production of biofuels from food products, such as sugar, starch, or fats and oils, are not as sustainable as their production from inedible resources, *e.g.*, lignocellulose*, due to the ever-increasing demand for food and feedstock [17]. In the US, the annual production of cellulosic ethanol from lignocellulosic resources such as woody biomass, switchgrass, and crop waste was estimated at 53 billion gallons in 2016 [18]. The commercial production of cellulosic ethanol in large quantities is deemed achievable for lignocellulosic biomass. However, several technical difficulties arose with utilising this type of feedstock: biomass recalcitrance, which is associated with the resistance of cell walls to enzymatic deconstruction; the heterogeneity of the substrate, which arises from the complex structure of lignocellulosic feedstock [19]. As a result, the biomass needs to be treated and prepared before ethanol can be extracted. In consolidated bioprocessing (CBP) preparation and production stages, namely enzyme production, polysaccharide hydrolysis and sugar fermentation, are combined in one single operation [19, 20]. CBP requires microorganisms that are

---

*Lignocellulose is composed of carbohydrate polymers (cellulose, hemicellulose), and an aromatic polymer (lignin).

highly efficient at producing cellulolytic enzymes and at fermenting resultant sugars; at the same time they should produce small amounts of other byproducts and be able to live in high levels of ethanol [20].

Strains of bacteria used in CBP, such as *Clostridium thermocellum* and *Caldicellulosiruptor obsidiansis*, consume a cellulosic substratum and form cellulolytic biofilm colonies, which are distributed spatially and evolve temporally [21–23]. Such cellulolytic biofilms, unlike the more traditional biofilms one encounters for example in wastewater engineering, tend to produce only very little amounts of extracellular polymeric substance (EPS). They attach directly to the fibrous substratum and form monolayer biofilms [21, 23]. Whereas traditional biofilms form colonies that grow into the aqueous phase and receive their nutrients from there via diffusion, cellulolytic biofilms do not form such structures and instead consume and degrade the cellulosic substratum that supports them. This leads to the development of crater-like structures in the substratum that are sometimes also referred to as inverse colonies [23]. Cellulolytic biofilm formation and growth progresses through a number of stages (see Figure 1):

1. Attachment to the cellulosic fibres,
2. Cell growth and division,
3. Inverted colony formation,
4. Formation of crater-like depressions due to substrate consumption and substratum degradation,
5. Radial growth of the depressions and their eventual coalescence,
6. Formation of uniform-thickness biofilm.

Due to cell detachment from cellulolytic biofilms, there are free cells in the aqueous phase, the existence of which has been corroborated by experimental observations [21, 23]. Among possible reasons for the existence of such free floating cells in general are mechanical detachment as a result of shear forces exerted by the aqueous phase on the biofilm [24], unavailability of space for new cells, biological detachment induced by quorum sensing or starvation [25–27], *etc*. Cells from the aqueous phase can form new colonies upon attachment to random locations on the substratum, as depicted in Figure 1. The different stages of spatio-temporal evolution of biofilm formed by *C. obsidiansis* on cellulose chads were captured by a confocal laser scanning microscope in [23] and are illustrated in Figure 1.

The possibility of adhesion and formation of a thriving bacterial colony at a certain position on the substratum can be affected by several aspects, including [28]:

- Environment: temperature, bacterial concentration, chemical treatment, etc.,
- Bacterial characteristics: hydrophobicity, surface charge, appendages, etc.,
- Substratum properties: chemical composition, roughness, physical configuration, etc.

Consequently, because of the sheer number of parameters required to characterize the attachment and adhesion of cells, the mathematical formulation of these phenomena on the basis of physical principles is very complex. One possible way to simplify the mathematical modeling of the attachment phenomenon is to think of it as an emergent random occurrence where the effect of all the deterministic underlying mechanisms are lumped together and treated as a single stochastic process. We treat them as a stochastic driving force that causes a single attachment event to take place arbitrarily at any point in space and time. As soon as such attachment of biomass occurs, a new colony can start developing.
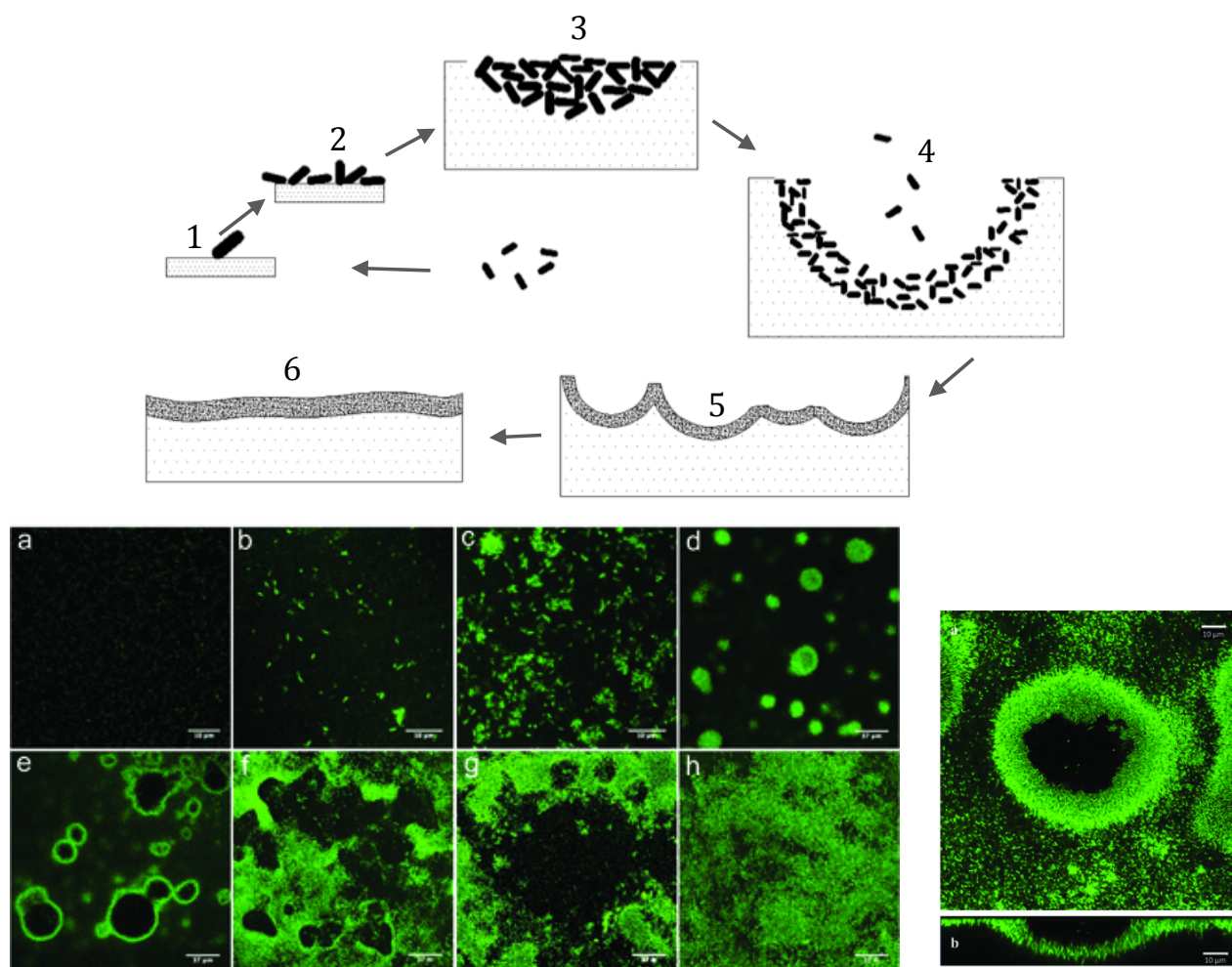
Figure 1. Top panel: Schematic diagram showing the six stages of cellulolytic biofilm formation: (1) initial cell attachment, (2) cell growth and division, (3) inverted colony formation, (4) emergence of crater-like depression, (5) growth and merging of craters, (6) uniform biofilm formation. **Bottom:** Confocal Laser Scanning Micrographs of cellulosic biofilms; cells are displayed in green. **Left panel:** Formation of cell clusters and craters of *C. obsidiansis* on a cellulose chad with attachment from the aqueous phase, after a) 0h, b) 8h, c) 16h, d) 24h, e) 44h, f) 48h, g) 56h and h) 68h (top view). Radius of craters $\approx 40\mu$m. **Right panel:** Crater like depression formed by *C. thermocellum.* a) top view, b) cross-sectional view. Length scale bar = 10μm. All images are from [23], where also more details about the microscopy and experimental setup are provided.

The mathematical modeling of stochastic cell attachment and the possible effects it may have on the growth and temporal evolution of biofilms is the main focus of our study. Specifically, we are concerned with the question how such discrete stochastic initial attachment can be included in the deterministic continuum model for cellulosic biofilm growth that has been introduced previously

in [29]. The overall goal is to introduce a straightforward mathematical framework for modeling the random cell attachment and to perform numerical simulations by which the model can be tested and computationally analyzed. Of paramount importance is to cast this model in a setting for which a theoretical framework exists, and numerical methods for which it is known that they represent the stochasticity of the underlying model correctly and reliably. Therefore, we choose to develop a model in the context of Itô stochastic differential equations [30, 31]. We suggest that, barring computational costs, the modeling framework that we aim to establish here could potentially also be applied to more traditional biofilms and incorporated in other continuum models of biofilm structure formation.

Our work builds on and aims to overcome limitations of an earlier stochastic attachment model that was used in an *ad-hoc* manner by modifying a deterministic model of cellulolytic biofilms [29]. Here, the deterministic part deals with the formation and temporal evolution of cellulolytic biofilms and is described by a continuum model, namely a highly nonlinear degenerate coupled system of a partial differential equation (PDE) and an ordinary differential equation (ODE). The biomass production and mechanics is governed by the PDE and the uptake of the substrate is governed by the ODE. The *ad-hoc* inclusion of stochastic cell attachment in that study was basically implemented by the addition of an impulse function to the biomass PDE: if a certain quantity (Q) is greater than a uniformly distributed U[0,1]-random number, the impulse function becomes non-zero and a fixed amount of biomass is added to the PDE; otherwise the value of the impulse function is zero and no biomass is added. The quantity, Q, encodes that an attachment site is less probable to be occupied if it already accommodates some biomass (M) and if it is deficient in nutrient substrate (C), *i.e.*, $Q \propto (1 - M)C$. By formulating the attachment phenomenon this way, the proposed continuum model in [29] has reproduced the qualitative behavior of cellulolytic biofilms with cell attachment. In Figure 2 we show a snapshot of that model from [29], for comparison with the experimental observations shown in Figure 1. Here, as in the experiments in [32] the substratum was initially inoculated with a single colony and it was investigated how random attachment deforms the circular structures that form. Although this approach is successful in generating qualitative results that resemble the experimental observations, due to the form of the stochastic component, standard results from stochastic differential equations do not apply and there is no guarantee that the stochasticity observed in the numerical simulations is a true representation of the stochasticity in the underlying model and not amplified or smeared by numerical artifacts.

Our goal is to introduce a modeling framework that can be supported by established stochastic theories, and for which numerical algorithms are known that faithfully represent the stochasticity of the underlying equation.

One possible path to modeling random attachment of cells would be to formalize it as random noise. For instance, one could add white noise to the PDE that describes the bacterial density, which turns it into a stochastic PDE (SPDE). Whereas there is some theory established for semi-linear stochastic diffusion-reaction equations [33] this seems not to be the case for degenerate diffusion-reaction equations. Moreover, further undesirable side effects will be associated with this direct approach to incorporate stochasticity in the model: it could lead to negative values of the solutions - especially at locations where no biomass is attached - and to an uncontrollable rate and magnitude of cell attachment. These are due to the nature of white noise, or more specifically additive white noise, which can assume both positive and negative values with arbitrary magnitudes. Instead, one could consider noise terms with state-dependent coefficients, e.g., multiplicative noise, which ensures that solutions remain non-negative. For instance, for systems of semi-linear parabolic SPDEs
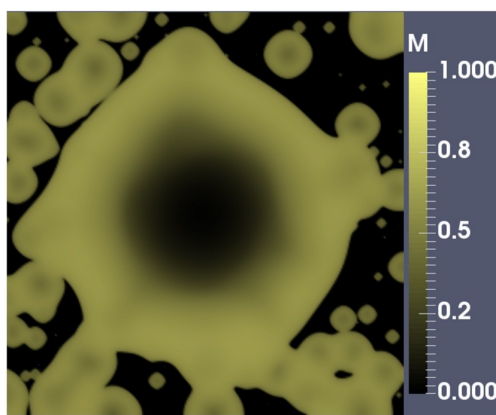
Figure 2. An example of simulated PDE-ODE system from [29] with an initial central inoculation of cells and random attachment of cells. Here *M* stands for the biomass density ($0 \leq M \leq 1$). Note the formation of inverse colonies especially in the center, *cf.* Figure 1.

it has been shown that if the deterministic and stochastic components satisfy certain conditions, the non-negativity of solutions is guaranteed [34]; or in case of systems of (ordinary) SDEs, even necessary and sufficient conditions for the non-negativity and/or boundedness of solutions are known [35]. Another possible remedy would be to directly modify the numerical schemes that estimate the solution by including reflecting or absorbing barriers: reflecting barriers send negative values to their corresponding positive counterparts and absorbing barriers swap negative values with zeros, cf. e.g., [36–38]. However, the relation between the numerical simulations and the mathematical model are unclear and it is therefore questionable whether the simulations are valid tests of the actual model's behavior.

Furthermore, the uncontrollability of attachment makes the simple addition of white noise problematic. All positive values of the noise term correspond to attachment events for which the amount of the attached bacteria takes arbitrary values. This behavior has been corroborated by our exploratory numerical simulations (data not shown) that have demonstrated that the high frequency of attachment instances and the randomly varying amount of attaching bacteria in each instance do not qualitatively replicate the experimental observations, e.g., the prominent features of cellulolytic biofilms shown in Figure 1. This problem follows from the mathematical properties of the added white noise. Since the noise is primarily realized by sampling the normal distribution in the numerical simulations of SPDEs and SDEs, the high frequency of attachment events and their arbitrary magnitude are natural outcomes.

The mentioned issues arise if white noise is simply added to the PDE-ODE model. Therefore, the starting point in our approach to model the random attachment of cells is to conceive an attachment mechanism that is driven by a local stochastic property of the substratum and the environment. We call this local stochastic property the *attachment factor* (AF). It is assumed to be a continuous stochastic process that can be roughly interpreted as a particular local background noise in the system. We also assume that the AF is driven by a local Itô SDE whose stochastic part is proportional to the local values of state variables. This SDE's deterministic part is assumed to be zero so that there are no contributions to the temporal evolution of the AF from any deterministic sources.

Random cell attachment is a discrete event. Our approach treats these random events as impulsive phenomena, similar in spirit to the *ad hoc* approach in [29] but technically refined. We consider an impulse function (the AF operates as the independent variable of this function) that provides the means for introducing the stochastic cell attachment to arbitrary locations on the substratum and also allows controlling different aspects of the attachment. Most importantly, the issue with negative solutions can be readily addressed since we have the freedom to choose an impulse function with suitable properties: a function that always outputs positive values will not add negative biomass especially to locations where there is no biomass. Furthermore, the effective domain of the chosen impulse function can be adjusted by its characterizing parameters. For instance, one may manipulate the frequency of attachment by careful exploitation of the function parameters that affect its effective domain. Also, the amount of attaching bacteria in a single event can be controlled by the coefficient of the impulse function.

In order to obtain numerical solutions for the new model that includes random attachment, the governing biomass PDE is discretized spatially by a uniform grid in the same manner as in [29]. The discretization converts the PDE into a finite system of lattice ODEs where each ODE governs the biomass dynamics locally in a specific grid cell. The stochasticity in each cell is simulated via the corresponding locally-defined Itô SDE, which governs the AF. The conversion of the PDE to a system of ODEs has the added benefit of allowing us to use established SDE solvers for each cell. The alternative would have been to add the stochastic terms to the underlying PDE and then solve the stochastic PDE. Since the deterministic PDE is highly nonlinear with a porous medium equation like degeneracy as biomass vanishes and a super-diffusion singularity where it attains its threshold value, no numerical method is known to guarantee that the stochasticity would be correctly rendered. For a first proof of concept we have used the simplest SDE solver, the Euler-Maruyama method for the numerical integration of the resulting model. We have made this choice in order to capture the effect of stochasticity in the model.

Finally, we point out that in addition to deterministic continuum biofilm models, also biofilm models are known which themselves are discrete and stochastic, such as individual based models that track individual cells, or cellular automata models that are based on them. To include a discrete stochastic attachment process in those would be straightforward. The focus of our work is to include attachment in deterministic biofilm models. A comparison with such innately discrete stochastic models is beyond our scope.

2. The underlying spatially explicit continuum model

2.1. Model assumptions

Our model is based on the spatially explicit model of cellulolytic biofilms in [29], which is a modification of the previous prototype biofilm growth model in [12]. In case of cellulolytic biofilms the bacteria hydrolyse their substratum instead of receiving nutrients from the surrounding aqueous phase. The basic model assumptions are:

1. The local capacity per unit area/volume of the substratum for accommodating bacterial biomass is limited by a finite quantity.
2. The only consumed substrate for biomass production is carbon. Since it is part of the chemical composition of the substratum, it is immobile, *i.e.*, it does not diffuse.

3. The production, spatial distribution and movement of biomass only depends on the local availability of space and nutrients. Biomass is produced if nutrients are available and the biofilm does not expand into neighboring positions if space is locally available to allocate new biomass.

4. Carbon is consumed and converted into new cells. Substrate uptake and production of biomass are driven by standard saturation kinetics: if substrate becomes locally limited its uptake by cells is proportional to both local biomass and available substrate. If it is abundant, substrate degradation is proportional to the local biomass only, close to the maximum uptake rate.

5. Cell loss in the biofilm occurs either due to cell death or due to detachment of viable cells. In our model formulation, both these mechanisms are subsumed in a single cell loss process. The amount of lost biomass is assumed to be proportional to the local biomass density. Once carbon substrate is locally degraded growth ceases and cell loss becomes the dominating process.

6. In the original model in [29] there is no attachment of cells from the aqueous phase.

We will cast our model here in a two-dimensional setting. More specifically, we will consider the limiting case of cellulosic biofilms that spread across a very thin cellulosic substratum, the depth of which is negligible compared to the length scales of the colonies. This is one of the two setups that were considered in [29]. Inoculation of the substratum in this case is possible everywhere in the domain. Furthermore, in [29] it was shown that the same 2D model can also be applied to simulate the penetration of substrata by cellulosic biofilms. In that case the domain of the model is a vertical slice through the substratum and some parts of the boundary of the domain constitute the interface with the external aqueous phase that is not explicitly considered. Inoculation, in this case is restricted to that part of the boundary. We will comment on this setup from the perspective of the current study briefly in Section 6.

2.2. Governing equations

The model comprises a PDE that governs the spatiotemporal evolution of the biomass density, M, and this PDE is coupled to an ODE that describes the evolution of the carbon concentration, C:

$$\partial_t M = \nabla \cdot (D(M)\nabla M) + F(C)M, \qquad\qquad (t, x) \in (0, \infty) \times \Omega, \qquad (2.1)$$

$$\partial_t C = -G(C)M, \qquad\qquad (t, x) \in (0, \infty) \times \Omega. \qquad (2.2)$$

Here $M = M(t, x)$, $C = C(t, x)$, $D(M)$ is the biomass diffusion coefficient, $F(C)$ the net biomass growth rate and $G(C)$ the uptake rate of the carbon substrate. The spatial domain is $\Omega \subset \mathbb{R}^d$, $d = 1, 2, 3$. Due to high computational costs, the model is only simulated for $d = 1$ and $d = 2$. In the two-dimensional case the domain is assumed to be rectangular. The actual biofilm consists of the subregion

$$\Omega_0(t) = \{x \in \Omega \,|\, M(t, x) > 0\}.$$

Following [12], the density dependent biomass diffusion coefficient $D(M)$ is given by

$$D(M) = \delta \frac{M^\alpha}{(M_\infty - M)^\beta}, \qquad \alpha \geq 1, \quad \beta \geq 1, \qquad (2.3)$$

where $\delta > 0$ is the motility coefficient and $M_\infty > 0$ the maximum biomass density.

The diffusion coefficient exhibits two distinct limiting behaviors: (i) if the local biomass density approaches its maximum value, M_∞, additionally produced biomass moves to neighboring positions

on the substratum (in accordance with model assumption 1); (ii) if sufficient space is available, newly produced biomass is accommodated locally (in accordance with model assumption 3). These behaviors emerge from the underlying nonlinear modes of diffusion. For $0 \le M \ll M_\infty$, the diffusion coefficient can be approximated by a degenerate power law, as known from the porous medium equation, $D(M) \approx M^\alpha$, which vanishes as M approaches zero. In fact, the degenerate power law creates an expanding biofilm colony with a sharp distinct front: if the initial data has compact support, then the corresponding solution has compact support [39]. The exponent α determines how much biomass can be accommodated locally before a noticeable expansion to neighboring regions takes place. When $0 \ll M < M_\infty$, the biomass density approaches the singularity and super diffusion becomes dominant, $D(M) \approx (M_\infty - M)^{-\beta}$, which prevents the biomass density to attain its maximum value M_∞ [39]. Here, the exponent β determines the strength of this effect. We emphasize that both nonlinear diffusion effects are needed; on the one hand, to ensure an expansion with finite speed (due to the porous medium nonlinearity) and on the other hand, for the local volume filling effect (due to the super diffusion singularity). Neither one alone can cause both of these effects [29].

The net biomass growth rate is composed of the true biomass growth rate and the cell loss rate. By model assumption 2 carbon is the sole growth limiting substrate, and the true growth rate is given by a Monod reaction function [46] (model assumption 4). Cell loss, by model assumption 5, occurs through two different mechanisms: cell death and cell detachment to the aqueous phase. Both mechanisms are combined in a constant, called the cell loss rate. Hence, the net growth rate is given by

$$F(C) = \mu \frac{C}{\kappa + C} - \lambda, \tag{2.4}$$

where $\mu > 0$ is the maximum specific growth rate, $\lambda > 0$ the cell loss rate and $\kappa > 0$ is the half saturation concentration.

The uptake rate of the carbon substrate is modeled by standard saturation kinetics and given by a Monod function (model assumption 4),

$$G(C) = \Upsilon \frac{C}{\kappa + C}, \tag{2.5}$$

where $\Upsilon > 0$ is the maximum consumption rate. The uptake rate G approaches a constant value when substrate is locally abundant. When and where the substrate is limited, the uptake rate is proportional to the local carbon concentration.

The substrate uptake rate in (2.5) is the only driving factor in (2.2). There are no transport mechanisms since carbon is assumed to be immobilized in the substratum (by model assumption 2). This distinguishes the current model from the original modeling framework in [12].

The biomass is confined to the domain Ω, and homogeneous Neumann boundary conditions are imposed,

$$\partial_n M(t, x) = 0, \qquad (t, x) \in [0, \infty) \times \partial\Omega, \tag{2.6}$$

where ∂_n denotes the outward unit normal derivative and $\partial\Omega$ the boundary of Ω. The initial condition for the biomass density is

$$M(0, x) = M_0(x), \qquad x \in \Omega, \tag{2.7}$$

where $M_0 > 0$ in a few small patches inside Ω and $M_0 \equiv 0$ in the rest of Ω.

The initial condition for the carbon concentration is

$$C(0, x) = C_0(x), \qquad x \in \Omega, \tag{2.8}$$

where we assume that $C_0 > 0$ in Ω. Although C_0 can depend on the spatial position, in the simulations we set C_0 to its maximal value C_∞ everywhere in Ω, which is in agreement with experiments conducted on homogeneous paper chads as the cellulosic substrata [21, 23, 40].

For diffusion-reaction equations with the specific diffusion coefficient (2.3) only few analytical results are available in the literature, hence such models were mainly studied in numerical experiments. To prepare it for computer simulations, the model is cast into the non-dimensional form by re-scaling x, t, M and C as follows

$$\tilde{x}_i = \frac{x_i}{L} \quad \text{for} \quad i \leq d, \quad \tilde{t} = \mu t, \quad \widetilde{M} = \frac{M}{M_\infty}, \quad \widetilde{C} = \frac{C}{C_\infty},$$

where L is a characteristic length of Ω; for instance, in the two-dimensional case it can be chosen as the length of the longest side of a rectangular domain. Moreover, $1/\mu$ is the characteristic time for biomass growth and the time scale of the dimensionless problem. The dimensionless parameters are given by

$$\tilde{\kappa} = \frac{\kappa}{C_\infty}, \quad \widetilde{\Upsilon} = \frac{M_\infty \Upsilon}{C_\infty \mu}, \quad \tilde{\lambda} = \frac{\lambda}{\mu}, \quad \tilde{\delta} = \frac{\delta M_\infty^{(\alpha-\beta)}}{\mu L^2}.$$

This leads to the model in non-dimensional form

$$\partial_{\tilde{t}} \widetilde{M} = \widetilde{\nabla} \cdot (\widetilde{D}(\widetilde{M}) \widetilde{\nabla} \widetilde{M}) + \widetilde{F}(\widetilde{C}) \widetilde{M}, \tag{2.9}$$

$$\partial_{\tilde{t}} \widetilde{C} = -\widetilde{G}(\widetilde{C}) \widetilde{M}, \tag{2.10}$$

where

$$\widetilde{D}(\widetilde{M}) = \tilde{\delta} \frac{\widetilde{M}^\alpha}{(1 - \widetilde{M})^\beta}, \quad \widetilde{F}(\widetilde{C}) = \frac{\widetilde{C}}{\tilde{\kappa} + \widetilde{C}} - \tilde{\lambda}, \quad \widetilde{G}(\widetilde{C}) = \widetilde{\Upsilon} \frac{\widetilde{C}}{\tilde{\kappa} + \widetilde{C}},$$

and the corresponding initial data for a homogeneous cellulosic substratum are

$$\widetilde{M}(0, \tilde{x}) = \widetilde{M}_0(\tilde{x}), \quad \widetilde{C}(0, \tilde{x}) = 1. \tag{2.11}$$

To simplify notations, from now on we will omit the tilde.

2.3. Spatial discretization

For simulation purposes, following the Method of Lines approach, the PDE-ODE system (2.9)–(2.10) is converted into a system of ODEs by spatial discretization. We use for this a standard finite volume method as introduced in [42].

We place an $N \times M$ uniform grid over the rectangular domain $\Omega := [0, L] \times [0, W]$. Equation (2.9) is integrated over each grid cell $v_{i,j}$, and using the divergence theorem we obtain

$$\frac{d}{dt} \int_{v_{i,j}} M = \int_{\partial v_{i,j}} J_n dS + \int_{v_{i,j}} F(C)M dV, \quad i = 1, ..., N, \quad j = 1, ..., M, \tag{2.12}$$

where the area integral involving the term $\nabla \cdot (D(M)\nabla M)$ is converted into the surface integral for the outward normal flux through the boundary $\partial v_{i,j}$ of the grid cell. The biomass density and carbon concentration are both evaluated at the centers of the grid cells, i.e., in terms of the grid indices and grid lengths we have

$$M_{i,j}(t) \ := \ M\left(t, x_i, y_j\right) \approx M\left(t, \left(i - \frac{1}{2}\right)\Delta x, \left(j - \frac{1}{2}\right)\Delta x\right), \tag{2.13}$$

$$C_{i,j}(t) \ := \ C\left(t, x_i, y_j\right) \approx C\left(t, \left(i - \frac{1}{2}\right)\Delta x, \left(j - \frac{1}{2}\right)\Delta x\right), \tag{2.14}$$

where the sides of each grid cell are equal to $\Delta x = \frac{L}{N} = \frac{W}{M}$. The area integrals are evaluated by the mid-point rule. The flux integrals are evaluated for each cell edge separately by the mid-point rule as well. Since the fluxes are computed at the grid cell edges, we need to evaluate the diffusion coefficient in these positions. To this end, we follow the approach in [39] where the flux across a shared interface of two grid cells is approximated by the arithmetic average of the diffusion at the centers of those cells. The derivative of M across each cell edge is approximated by the central finite difference formula. By applying this discretization to Eq (2.9) we end up with the following ODE for the biomass density

$$\frac{d}{dt}M_{i,j} = \frac{1}{\Delta x}\left(J_{i+\frac{1}{2},j} + J_{i-\frac{1}{2},j} + J_{i,j+\frac{1}{2}} + J_{i,j-\frac{1}{2}}\right) + F_{i,j}M_{i,j}, \tag{2.15}$$

where $F_{i,j} := F(C_{i,j}) = \frac{C_{i,j}}{\kappa + C_{i,j}} - \lambda$. With the boundary conditions (2.6), the flux terms in (2.15) are obtained as

$$J_{i+\frac{1}{2},j} = \begin{cases} \frac{1}{2\Delta x}\left(D\left(M_{i+1,j}\right) + D\left(M_{i,j}\right)\right)\left(M_{i+1,j} - M_{i,j}\right), & \text{for } i < N, \\ 0, & \text{for } i = N, \end{cases} \tag{2.16a}$$

$$J_{i-\frac{1}{2},j} = \begin{cases} 0, & \text{for } i = 1, \\ \frac{1}{2\Delta x}\left(D\left(M_{i,j}\right) + D\left(M_{i-1,j}\right)\right)\left(M_{i-1,j} - M_{i,j}\right), & \text{for } i > 1, \end{cases} \tag{2.16b}$$

$$J_{i,j+\frac{1}{2}} = \begin{cases} \frac{1}{2\Delta x}\left(D\left(M_{i,j+1}\right) + D\left(M_{i,j}\right)\right)\left(M_{i,j+1} - M_{i,j}\right), & \text{for } j < M, \\ 0, & \text{for } j = M, \end{cases} \tag{2.16c}$$

$$J_{i,j-\frac{1}{2}} = \begin{cases} 0, & \text{for } j = 1, \\ \frac{1}{2\Delta x}\left(D\left(M_{i,j}\right) + D\left(M_{i,j-1}\right)\right)\left(M_{i,j-1} - M_{i,j}\right), & \text{for } j > 1. \end{cases} \tag{2.16d}$$

Applying the same discretization to (2.10) we obtain

$$\frac{d}{dt}C_{i,j} = -G_{i,j}C_{i,j}, \tag{2.17}$$

with $G_{i,j} := G(C_{i,j}) = \Upsilon\frac{C_{i,j}}{\kappa + C_{i,j}}$. There are no fluxes across cell edges since the substrate is immobile.

For computational purposes, the lexicographical grid ordering is introduced,

$$\pi : \{1, ..., N\} \times \{1, ..., M\} \to \{1, ..., NM\}, \quad (i, j) \mapsto p = (i - 1)M + j, \tag{2.18}$$

and we end up with two NM-component vectors for the biomass density and carbon concentration: $\mathbf{M} = (M_1, ..., M_{NM})^T$ and $\mathbf{C} = (C_1, ..., C_{NM})^T$ with $M_p := M_{\pi(i,j)} = M_{i,j}$ and $C_p := C_{\pi(i,j)} = C_{i,j}$. Hence, spatial discretization, converts the PDE-ODE system into the following system of $2NM$ ODEs,

$$\frac{d\mathbf{M}}{dt} = \mathfrak{D}\,(\mathbf{M})\,\mathbf{M} + \mathfrak{F}\,(\mathbf{C})\,\mathbf{M},$$

$$\frac{d\mathbf{C}}{dt} = -\mathfrak{G}\,(\mathbf{C})\,\mathbf{M}. \tag{2.19}$$

Here $\mathfrak{D}\,(\mathbf{M})$, $\mathfrak{F}\,(\mathbf{C})$ and $\mathfrak{G}\,(\mathbf{C})$ are $NM \times NM$ matrices. The entries of $\mathfrak{D}\,(\mathbf{M})$ are calculated using the first term in (2.15) and the corresponding definitions in (2.16), after the biomass density indices are mapped to p using (2.18). The entries of $\mathfrak{F}\,(\mathbf{C})$ and $\mathfrak{G}\,(\mathbf{C})$ are evaluated by using the growth rate and uptake rate formulae for each C_p, respectively.

Remark 1. *The matrix $\mathfrak{D}\,(\mathbf{M})$ is symmetric and weakly diagonally dominant with non-positive diagonal entries and non-negative off-diagonal entries. The matrices $\mathfrak{F}\,(\mathbf{C})$ and $\mathfrak{G}\,(\mathbf{C})$ are diagonal. If $C_p = 0$ the pth diagonal entry of $\mathfrak{G}\,(\mathbf{C})$ becomes zero while the same entry in $\mathfrak{F}\,(\mathbf{C})$ becomes negative with a magnitude equal to the cell loss rate.*

3. A Spatially explicit model with random cell attachment

As mentioned in the introduction, we aim at formulating a new approach using SDEs to model random attachment of cells from the aqueous phase to the substratum. Hence, we discard model assumption 6 in Section 2.1 and allow for attachment to occur.

3.1. Model assumptions for attachment

Attachment refers to the process by which cells from the aqueous phase adhere to the substratum or join an already existing biofilm colony. In many controlled experiments, no bacteria are present in the inflow, so that the attaching cells are cells that have previously detached, but have not yet been washed out. This is for example the case for the experiments described in [32, 40], which formed the basis for the development of the model in [29], which we extend here. The close proportionality relationship between growth activity in the reactor and biomass in the effluent reported in [32] suggests that availability of cells in the aqueous phase is not a limiting factor for attachment, and that detachment is proportional to the number of active cells in the system. Therefore, we make here the simplifying assumption that attachment and detachment can be modeled independently. This allows us to assume an unlimited reservoir of cells available for attachment in the aqueous phase, and thus that the occurrence of an attachment event is independent of other detachment events and of detachment activity. To describe detachment as a continuous phenomenon subsumed in a constant cell loss rate λ, as introduced in [29], is consistent with the above proportionality between cell yield and active biomass and the description of cell growth activity as a continuous process. Hence, we keep this assumption for the cell yield process.

3.2. Mathematical model for stochastic cell attachment

We add an impulse function to the biomass ODE (2.15) in order to devise a mechanism for picking up stochastic attachment signals received as inputs and converting these signals into attached biomass

at some arbitrary position on the substratum. These input signals, which we call the attachment factors (AFs), are denoted by ϕ. We aim to choose an impulse function, $f(\phi)$ with properties that allow to control different aspects of attachment and appropriate characteristics for numerical implementation. For instance, the normal distribution possesses most of the required properties, however, its maximum value depends on its width. This can cause that the biomass density exceeds its upper bound which is not admissible. Therefore, choosing an impulse function with a maximum equal to the theoretical upper bound for the biomass density is required.

We construct an impulse function from the Fermi-Dirac distribution mostly used in physics, which suits our modeling framework well. The range of this function is $[0, 1]$ and its shape can be adjusted by tweaking its characterizing parameters a, b and σ,

$$f(\phi) = \frac{1}{e^{\frac{\phi-b}{\sigma}} + 1} - \frac{1}{e^{\frac{\phi-a}{\sigma}} + 1}, \qquad a \in \mathbb{R},\ b \in \mathbb{R},\ \sigma > 0, \quad \phi \in \mathbb{R}. \tag{3.1}$$

In Figure 3 the function f is plotted for $a = 0.9$, $b = 1.0$ and different values of σ. As the value of σ decreases, the function becomes more rectangular and in the limit, as $\sigma \to 0$, its support converges to $[a, b]$.
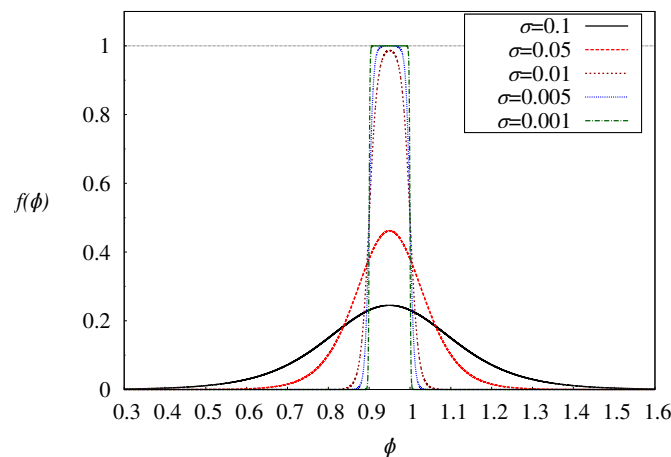

Figure 3. The impulse function, $f(\phi)$, plotted for $a = 0.9$ and $b = 1.0$ and different values of the parameter σ.

Including stochastic attachment in the model by adding the introduced impulse function to the biomass ODE (2.15) we obtain

$$\frac{dM_{i,j}}{dt} = \frac{1}{\Delta x}\left(J_{i+\frac{1}{2},j} + J_{i-\frac{1}{2},j} + J_{i,j+\frac{1}{2}} + J_{i,j-\frac{1}{2}}\right) + F_{i,j}M_{i,j} + \gamma_1 h(M_{i,j})C_{i,j}^{\gamma_2} f_{i,j}, \tag{3.2}$$

where $0 < \gamma_1 \ll 1, \gamma_2 \geq 1$ and

$$h(M) = \begin{cases} 1 - M - \epsilon, & M \leq 1 - \epsilon \\ 0, & M > 1 - \epsilon. \end{cases}$$

Except for the last one the terms were defined in (2.15) and (2.16), and

$$f_{i,j} := f(\phi_{i,j}) = (e^{\frac{\phi_{i,j}-b}{\sigma}} + 1)^{-1} - (e^{\frac{\phi_{i,j}-a}{\sigma}} + 1)^{-1}.$$

Here, we additionally introduced the factor $\gamma_1 h(M_{i,j})C_{i,j}^{\gamma_2}$ in the term for cell attachment. The factor $h(M_{i,j})$ ensures that bacterial cells attach preferably where no other cells are present, and the term $C_{i,j}^{\gamma_2}$ prevents attachment to carbon-free positions. The exponent γ_2 is included for a generic formulation, and γ_1 controls the intensity of bacterial attachment per unit time. In this parameter properties may be subsumed that are specific to the species, such as cell size, etc. We choose $\gamma_1 \ll 1$ such that the numerical realizations are in agreement with the experimental observations in [21] and [23]. The spatially discrete substrate ODE (2.17) remains unchanged.

In order to simulate the AF we consider in each grid cell (i, j) the Itô SDE

$$d\phi_{i,j} = \psi h(M_{i,j})C_{i,j}^{\nu}dW_t^{(i,j)}, \quad \nu \geq 1, \tag{3.3}$$

where $W^{(i,j)}$ denote independent standard scalar Wiener processes and $dW_t^{(i,j)}$ the corresponding Itô differentials. The exponent ν is included for the sake of generality. The constant coefficient ψ controls the magnitude of noise, which can be harnessed as a control parameter for the rate of attachment events. The stochastic fluctuations in the system are proportional to volume filling and available substrate. One could argue that the amplitude of the stochastic variation is linked to the available space and the amount of substrate in a grid cell. In fact, without mollification of the impulse function, i.e., as $\sigma \to 0$, attachment events occur when the value of $\phi_{i,j}$ in a grid cell falls in the support of the impulse $[a, b]$ (assuming that $\phi_{i,j}$ is initially outside of the support of $f_{i,j}$); the more frequently this happens, the higher is the rate of attachment. If a grid cell is free of biomass, then the higher the carbon concentration, the greater is the amplitude of the stochastic variation in (3.3), which increases the likelihood that $\phi_{i,j}$ attains a value within $[a, b]$. This, in turn, increases the chance of attachment of bacteria to that grid cell. As time progresses, either the biomass density reaches its saturation level or the carbon concentration tends to zero. In either case, the right-hand side of (3.3) becomes zero and $\phi_{i,j}$ reaches a steady state.

Recognizing the underlying mechanism of the SDE (3.3), we remark that even if the stochastic variation tends to zero as space is occupied and/or substrate is consumed, the asymptotic (steady state) value of $\phi_{i,j}$ can lie in the support of the impulse function, $[a, b]$. However, multiplying the impulse function by the factor $\gamma_1 h(M_{i,j})C_{i,j}^{\gamma_2}$ nullifies the lasting impulse that can be caused by these relatively rare events: it switches off the attachment of bacteria if either no carbon or no space is available in a grid cell.

A key question that we will answer in simulations below, is whether our mollification (3.1) approximates this behavior well enough.

The semi-discrete system for the model including random attachment is obtained via the lexicographical grid ordering (2.18) in the same way as (2.19). The difference is the addition of the impulse function to the biomass ODEs and the coupling to the system of Itô SDEs (3.3). The AF and the Itô differential are both evaluated at the centers of the grid cells, i.e.,

$$\phi_{i,j}(t) \quad := \quad \phi\left(t, x_i, y_j\right) \approx \phi\left(t, \left(i - \frac{1}{2}\right)\Delta x, \left(j - \frac{1}{2}\right)\Delta x\right), \tag{3.4}$$

$$dW_t^{(i,j)} \quad := \quad dW\left(t, x_i, y_j\right) \approx dW\left(t, \left(i - \frac{1}{2}\right)\Delta x, \left(j - \frac{1}{2}\right)\Delta x\right). \tag{3.5}$$

Using the mapping (2.18), we obtain two additional NM-component vectors for the AF and the Itô

differential : $\boldsymbol{\phi} = (\phi_1, ..., \phi_{NM})^T$ and $d\mathbf{W}_t = \left(dW_t^{(1)}, ..., dW_t^{(NM)}\right)^T$ with $\phi_p := \phi_{\pi(i,j)} = \phi_{i,j}$ and $dW_t^{(p)} := dW_t^{(\pi(i,j))} = dW_t^{(i,j)}$.

The whole system now comprises $3NM$ SDEs. In fact, the equations for the biomass and carbon concentration are regular ODEs that can be considered as special cases of Itô SDEs with zero stochastic terms. To simplify notations we formally write the entire model as system of SDEs,

$$d\mathbf{M} = \left[\mathfrak{D}\left(\mathbf{M}\right)\mathbf{M} + \mathfrak{F}\left(\mathbf{C}\right)\mathbf{M} + \mathfrak{I}\left(\mathbf{M}, \mathbf{C}\right)\mathbf{F}(\boldsymbol{\phi})\right] dt,$$

$$d\mathbf{C} = \left[-\mathfrak{G}\left(\mathbf{C}\right)\mathbf{M}\right] dt, \tag{3.6}$$

$$d\boldsymbol{\phi} = \mathfrak{T}\left(\mathbf{M}, \mathbf{C}\right) d\mathbf{W}_t,$$

where $\mathfrak{D}\left(\mathbf{M}\right)$, $\mathfrak{F}\left(\mathbf{C}\right)$ and $\mathfrak{G}\left(\mathbf{C}\right)$ are $NM \times NM$ matrices with the properties described in Remark 1. We introduced two new $NM \times NM$ matrices: $\mathfrak{I}\left(\mathbf{M}, \mathbf{C}\right)$ and $\mathfrak{T}\left(\mathbf{M}, \mathbf{C}\right)$. The vector $\mathbf{F}(\boldsymbol{\phi})$ has NM components and its entries are the impulse function evaluated in each grid cell.

Remark 2. *The matrices $\mathfrak{I}\left(M, C\right)$ and $\mathfrak{T}\left(M, C\right)$ are diagonal with the pth diagonal entries $i_{pp} = \gamma_1 h(M_p)C_p^{\gamma_2}$ and $t_{pp} = \psi h(M_p)C_p^\nu$, respectively. They vanish if $M_p = 1$ or $C_p = 0$ and attain their maximum values when $M_p = 0$ and $C_p = 1$. Therefore, we have $0 \leq i_{pp} \ll 1$ and $t_{pp} \geq 0$. The components of the vector $\mathbf{F}(\boldsymbol{\phi})$ are given by $f_p := f_{\pi(i,j)} = f_{i,j}$ such that $f_p = (e^{\frac{\phi_p - b}{\sigma}} + 1)^{-1} - (e^{\frac{\phi_p - a}{\sigma}} + 1)^{-1}$. By construction, we have $0 \leq f_p \leq 1$ and f_p vanishes when ϕ_p, the value of the pth entry of $\boldsymbol{\phi}$, lies outside the support of the impulse function.*

Due to the singularity in the diffusion coefficient (2.3), we cannot apply classical SDE theory to analyze the well-posedness and qualitative behavior of (3.6), e.g., the non-negativity and boundedness of solutions. Hence, as in the analysis of the underlying deterministic models [41, 42], we regularize the diffusion coefficient and consider

$$D_\epsilon(M) = \begin{cases} 0, & M < 0, \\ \delta\dfrac{M^\alpha}{(1-M)^\beta}, & 0 \leq M \leq 1 - \epsilon, \\ \delta\epsilon^{-\beta}(1-\epsilon)^\alpha, & M > 1 - \epsilon. \end{cases} \tag{3.7}$$

This yields the regularized SDE problem

$$d\mathbf{M}^\epsilon = \left[\mathfrak{D}_\epsilon\left(\mathbf{M}^\epsilon\right)\mathbf{M}^\epsilon + \mathfrak{F}\left(\mathbf{C}^\epsilon\right)\mathbf{M}^\epsilon + \mathfrak{I}\left(\mathbf{M}^\epsilon, \mathbf{C}^\epsilon\right)\mathbf{F}(\boldsymbol{\phi}^\epsilon)\right] dt,$$

$$d\mathbf{C}^\epsilon = \left[-\mathfrak{G}\left(\mathbf{C}^\epsilon\right)\mathbf{M}^\epsilon\right] dt, \tag{3.8}$$

$$d\boldsymbol{\phi}^\epsilon = \mathfrak{T}\left(\mathbf{M}^\epsilon, \mathbf{C}^\epsilon\right) d\mathbf{W}_t.$$

3.3. Non-negativity and boundedness of solutions

The drift and diffusion terms of the regularized SDE system (3.8) are locally Lipschitz continuous and hence, a unique local solution exists [30]. We now show that the biomass density and carbon concentration are non-negative and bounded. To this end we recall comparison principles and

invariance criteria for systems of SDEs in [35] and [43] that yield explicit necessary and sufficient conditions for the boundedness and non-negativity of solutions.

We consider a general system of SDEs

$$dX(t) = f(t, X(t))dt + g(t, X(t))dW_t, \quad t > 0,$$
$$X(0) = X_0,$$

(3.9)

where $f : [0, \infty) \times \mathbb{R}^n \to \mathbb{R}^n$ is a continuous, measurable function and $g : [0, \infty) \times \mathbb{R}^n \to \mathbb{R}^{n \times m}$ is a continuous, measurable matrix-valued mapping. Moreover, W is an m-dimensional adapted Wiener processes and dW_t denotes the corresponding Itô differential.

We assume that there exists a unique solution of (3.9).

Definition 3.1. *A subset $S \subset \mathbb{R}^n$ is called invariant for the stochastic system (3.9) if for every initial data $X_0 \in S$ the corresponding solution $X(t)$ almost surely remains in S for $t \geq 0$, i.e.,*

$$P(\{X(t) \in S, \ t \geq 0\}) = 1.$$

For the following theorem we refer to [35].

Theorem 3.2. *Let $a_p, b_p \in \mathbb{R} \cup \{\infty\}$, $b_p > a_p$ and $p \in I$ where $I \subseteq \{1, ..., n\}$ is non-empty. The set*

$$S := \left\{ X \in \mathbb{R}^n : a_p \leq X_p \leq b_p, \ p \in I \right\}$$

is invariant for the SDE system (3.9) if and only if, for all $t \geq 0$ and $p \in I$,

- $f_p(t, X) \geq 0$ *for $X \in S$ such that $X_p = a_p$,*
- $f_p(t, X) \leq 0$ *for $X \in S$ such that $X_p = b_p$,*
- $g_{p,r}(t, X) = 0$ *for $X \in S$ such that $X_p \in \left\{a_p, b_p\right\}$, $r = 1, ..., m$.*

This theorem extends the well-known tangential condition for systems of ODEs (see [48]) to SDE systems. In fact, if the stochastic term g is identically zero, the set S is invariant for the ODE system if and only if the conditions on f are satisfied.

For the following comparison principle for SDEs we refer to [43].

Theorem 3.3. *Let $I \subseteq \{1, ..., n\}$ be non-empty. We consider two SDE systems of the form (3.9) with drift terms $f, \tilde{f}$, diffusion terms $g, \tilde{g}$ and initial data $X_0, \widetilde{X}_0$. The following statements are equivalent:*

(i) If the initial data satisfy $(X_0)_p \geq (\widetilde{X}_0)_p$, $p \in I$, then

$$P\left(\left\{X_p(t) \geq \widetilde{X}_p(t), \ t \geq 0\right\}\right) = 1, \qquad p \in I.$$

(ii) The functions $f, \tilde{f}, g$ and $\tilde{g}$ satisfy

- $f_p(t, X) \geq \tilde{f}_p(t, \widetilde{X})$ *for $t \geq 0$,*
- $g_{p,r}(t, X) = \tilde{g}_{p,r}(t, \widetilde{X})$ *for $t \geq 0$, $r = 1, ..., m$,*

for all $X, \widetilde{X} \in \mathbb{R}^n$ such that $X_p = \widetilde{X}_p$ and $X_l \geq \widetilde{X}_l$ for $l \in I$.

This theorem extends classical comparison principles for ODE systems to SDEs. In particular, if the stochastic term is zero, the comparison principles hold if the ODE system is quasi-monotone (see [48]).

We now apply Theorem 3.2 and Theorem 3.3 to show that the biomass density and the carbon concentration of the regularized system (3.8) are non-negative and bounded. In the sequel, we denote by $\mathbf{1} \in \mathbb{R}^{NM}$ and $\mathbf{0} \in \mathbb{R}^{NM}$ the vectors $\mathbf{1} = (1, ..., 1)^T$ and $\mathbf{0} = (0, ..., 0)^T$. Furthermore, we write $\mathbf{U} \leq \mathbf{V}$ if the inequality holds componentwise, i.e., if $U_p \leq V_p$ for all $p \in \{1, ..., NM\}$.

Proposition 1. *Let $\mathbf{M}^\epsilon$ and $\mathbf{C}^\epsilon$ be the solutions of* (3.8) *and suppose that the initial data satisfy* $\mathbf{0} \leq \mathbf{M}^\epsilon(0) \leq \eta\mathbf{1}$ *with $\eta \in (0, 1)$ and $\mathbf{0} \leq \mathbf{C}^\epsilon(0) \leq \mathbf{1}$. Then, the corresponding solutions satisfy*

$$\mathbf{0} \leq \mathbf{C}^\epsilon \leq \mathbf{1} \quad \text{almost surely for all } t \geq 0,$$
$$\mathbf{0} \leq \mathbf{M}^\epsilon < \mu \quad \text{almost surely for } t \in [0, T],$$

where the constant μ depends on the initial values and T.

Proof. The stochastic terms in the equations for the biomass density and carbon concentration in (3.8) are zero, and hence, it suffices to verify the conditions on the drift term.

The initial data for the carbon concentration satisfies $\mathbf{0} \leq \mathbf{C}^\epsilon(0) \leq \mathbf{1}$. According to Remark 1, the matrix $\mathfrak{G}(\mathbf{C}^\epsilon)$ is diagonal and the *pth* component of the right hand side is

$$-(\mathfrak{G}(\mathbf{C}^\epsilon)\mathbf{M}^\epsilon)_p = -\Upsilon\frac{C_p^\epsilon M_p^\epsilon}{\kappa + C_p^\epsilon}.$$

For $C_p^\epsilon = 0$ the latter vanishes and for $C_p^\epsilon = 1$ it is non-positive if $M_p^\epsilon \geq 0$. Therefore, the equation for the carbon concentration satisfies the conditions in Theorem 3.2.

The equation for the biomass consists of three parts: the stochastic cell attachment, $\mathfrak{I}(\mathbf{M}^\epsilon, \mathbf{C}^\epsilon)\mathbf{F}(\phi^\epsilon)$, the biomass growth, $\mathfrak{F}(\mathbf{C}^\epsilon)\mathbf{M}^\epsilon$, and the diffusion, $\mathfrak{D}_\epsilon(\mathbf{M}^\epsilon)\mathbf{M}^\epsilon$. Recalling Eq (2.16), Remark 1 and Remark 2, we have

- *pth* component of $\mathfrak{I}(\mathbf{M}^\epsilon, \mathbf{C}^\epsilon)\mathbf{F}(\phi^\epsilon)$:

$$\mathfrak{if}_p = \gamma_1 h(M_p^\epsilon)C_p^{\epsilon\gamma_2}f_p^\epsilon,$$

- *pth* component of $\mathfrak{F}(\mathbf{C}^\epsilon)\mathbf{M}^\epsilon$:

$$\mathfrak{fm}_p = (\frac{C_p^\epsilon}{\kappa + C_p^\epsilon} - \lambda)M_p^\epsilon,$$

- *pth* component of $\mathfrak{D}_\epsilon(\mathbf{M}^\epsilon)\mathbf{M}^\epsilon$:

$$\begin{aligned}\mathfrak{dm}_p =\frac{1}{2\Delta x}&\Big[\Big(D_\epsilon(M_{p-M}^\epsilon) + D_\epsilon(M_p^\epsilon)\Big)M_{p-M}^\epsilon + \Big(D_\epsilon(M_{p-1}^\epsilon) + D_\epsilon(M_p^\epsilon)\Big)M_{p-1}^\epsilon \\ &- \Big(D_\epsilon(M_{p-M}^\epsilon) + D_\epsilon(M_{p-1}^\epsilon) + 4D_\epsilon(M_p^\epsilon) + D_\epsilon(M_{p+1}^\epsilon) + D_\epsilon(M_{p+M}^\epsilon)\Big)M_p^\epsilon \\ &+ \Big(D_\epsilon(M_{p+1}^\epsilon) + D_\epsilon(M_p^\epsilon)\Big)M_{p+1}^\epsilon + \Big(D_\epsilon(M_{p+M}^\epsilon) + D_\epsilon(M_p^\epsilon)\Big)M_{p+M}^\epsilon\Big].\end{aligned}$$

For $M_p^\epsilon = 0$, we have $0 \leq \mathfrak{if}_p$, $\mathfrak{fm}_p = 0$ and

$$\mathfrak{dm}_p =\frac{1}{2\Delta x}\Big[\Big(D_\epsilon(M_{p-M}^\epsilon)\Big)M_{p-M}^\epsilon + \Big(D_\epsilon(M_{p-1}^\epsilon)\Big)M_{p-1}^\epsilon$$

$$+ \left(D_\epsilon(M^\epsilon_{p+1}) \right) M^\epsilon_{p+1} + \left(D_\epsilon(M^\epsilon_{p+M}) \right) M^\epsilon_{p+M} \Big].$$

Since all terms are non-negative if M^ϵ_{p-M}, M^ϵ_{p-1}, M^ϵ_{p+1} and M^ϵ_{p+M} are non-negative, it follows that $\mathbf{M}^\epsilon \geq 0$.

Finally, to prove the upper bound for $\mathbf{M}^\epsilon$ we use Theorem 3.3 and construct a suitable barrier function. We consider the ODE system

$$\frac{d}{dt}\hat{\mathbf{M}}^\epsilon = \mathfrak{D}_\epsilon \left(\hat{\mathbf{M}}^\epsilon \right) \hat{\mathbf{M}}^\epsilon + \mathfrak{F}\left(\hat{\mathbf{C}}^\epsilon \right) \hat{\mathbf{M}}^\epsilon + \gamma_1 \mathbf{1},$$
$$\frac{d}{dt}\hat{\mathbf{C}}^\epsilon = -\mathfrak{G}\left(\hat{\mathbf{C}}^\epsilon \right) \hat{\mathbf{M}}^\epsilon. \tag{3.10}$$

Let $\mathbf{M}_{tot} = \sum_{p=1}^{NM} \hat{M}^\epsilon_p$ and $\mathbf{C}_{tot} = \sum_{p=1}^{NM} \hat{C}^\epsilon_p$. Then, adding up the equations we obtain

$$\frac{d}{dt}\left(\mathbf{M}_{tot} + \frac{1}{\Upsilon}\mathbf{C}_{tot} \right) \leq NM\gamma_1,$$

and consequently,

$$\left(\mathbf{M}_{tot}(t) + \frac{1}{\Upsilon}\mathbf{C}_{tot}(t) \right) \leq \left(\mathbf{M}_{tot}(0) + \frac{1}{\Upsilon}\mathbf{C}_{tot}(0) \right) + NM\gamma_1 t. \tag{3.11}$$

The non-negativity of the solutions $\hat{\mathbf{M}}^\epsilon$ and $\hat{\mathbf{C}}^\epsilon$ of the ODE system (3.10) follow by the same arguments as for the stochastic system. Hence, the inequality (3.11) implies that the biomass density is bounded. Finally, since

$$\mathfrak{f}_p = \gamma_1 h(M^\epsilon_p) C^{\epsilon\gamma_2}_p f^\epsilon_p \leq \gamma_1,$$

we observe that solutions of the ODE system (3.10) are upper solutions for the stochastic attachment model (3.8), which implies the boundedness of the biomass density.

□

Remark 3. *Since we impose homogeneous Neumann boundary conditions and allow for random cell attachment events, we cannot expect that the biomass density necessarily remains bounded by its maximum value for all times. While attachment ceases as $M = 1$ is reached, the deterministic growth part, depending on parameters and initial conditions might lead to M reaching unity somewhere in finite time. This is in agreement with the situation for the deterministic model with diffusive supply of substrates [41], in which case there is the possibility that one might find $M \to 1$ a.e. in finite time and a subsequent breakdown of the model. In our setting, contrary to the one studied in [41], substrate is not replenished but continuously depleted in locations where biomass is present, i.e., substrate degradation is faster. Where biomass is present, C is a strictly monotonically decreasing function and $C \to 0$ eventually. Biomass loss will start dominating locally when $C < \lambda\kappa/(1-\lambda) =: C_{crit}$ [N.B.: we tacitly assume here $\lambda < 1$, otherwise the cell loss rate would be stronger than the growth rate even under ideal conditions and the population is not able to establish itself], eventually leading to $M \to 0$. Strictly speaking, as in the case of the deterministic model, under homogeneous Neumann conditions the stochastic attachment model is valid as long as the biomass density remains below its physical upper bound. In all simulations the initial biomass density is low and did not attain values close to 1 anywhere in the domain, i.e., C dropped down to C_{crit} always before M reached unity, i.e., this restriction is not severe from a practical perspective.*

Table 1. Model parameters used for 2D simulations.

Parameter	Symbol	Value	Reference
Motility Coefficient	$\tilde{\delta}$	1.0×10^{-6}	[29]
Diffusion Coefficient Exponent 1	α	4.0	[29]
Diffusion Coefficient Exponent 2	β	4.0	[29]
Maximum Consumption Rate	$\tilde{\Upsilon}$	0.4	[29]
Half Saturation Concentration	$\tilde{\kappa}$	0.01	[29]
Cell Loss Rate	$\tilde{\lambda}$	0.42	[29]
Control Parameter of $f(\phi)$	σ	0.00001	assumed
Shift Parameters of $f(\phi)$	a & b	0.99 & 1.00	assumed
Density of an Attached Cell per Time Step	γ_1	0.01	assumed
Biomass Attachment Exponent	γ_2	1.0	assumed
Constant Coefficient in the AF SDE	ψ	0.04 & 0.08	assumed
Exponent in the AF SDE	ν	1.0	assumed
Initial Data for $\boldsymbol{\phi}$	$\boldsymbol{\phi_0}$	0.0	assumed
Starting Integer Seed for the PRNG	sd	10	assumed
Domain Length	L	1.0	assumed
Domain Width	W	1.0	assumed
Grid Size	$N \times M$	$2^9 \times 2^9$	assumed

4. Computational implementation

To numerically integrate the SDE (3.6) we use the Euler-Maruyama method, the probably simplest and most basic algorithm to solve Itô differential equations. It is known that the Euler-Maruyama method approximates the sample solutions of an SDE with the strong order of convergence equal to 0.5 [45]. To cope with the small time steps that are often required by the Euler-Maruyama method for spatial models, we use a restarted version, where the time-step length is newly determined in every integration step in an adaptive manner. The noise term in the Euler-Maruyama approximation is represented by increments of the Wiener process. These are simulated by known pseudo random number generator algorithms, which we have parallelized in OpenMP to reduce the overall simulation runtime. For more details on computational aspects we refer the reader to [47].

5. Simulation results and computational analysis

5.1. 2D simulations

We present some illustrative simulations of the model in the 2D setting. In a first scenario, the simulation is initiated with a clean domain, *i.e.*, the initial data for biomass density equals zero everywhere. In a second scenario, the simulation is initiated with a small established colony at the center of the domain. Our goal is to investigate qualitatively the effect of an established colony on the random attachment of free cells and the ensuing colony formations and developments. In both of these scenarios, we look into variations in the biofilm development that might arise from changing the

rate of bacterial cell attachment. In all these simulations the carbon concentration is set to its maximal value over the entire domain, i.e., we mimic a homogeneous substratum, e.g., paper chads that are often used in experimental studies, such as [40]. For each scenario, we consider two different rates of stochastic cell attachment. The simulation parameters for all these cases are given in Table 1.

5.1.1. Case 1: initially clean domain

We control the emerging characteristics of these simulations, namely the rate of stochastic cell attachment and the overall biofilm formation features, by adjusting ψ in (3.3); in simulation 2 this value is half of the one used in simulation 1 (see Table 1). The results of these simulations, for selected time points, are shown in Figures 4 and 5. Starting from the homogeneous initial conditions, as the simulations progress temporally, small colonies start to emerge gradually. Figure 4 for simulation 1 demonstrates that the seeds of the initial colonies begin to become visible around $t = 25$. These colonies grow radially and eventually manifest crater-like features, as a consequence of carbon depletion behind the moving front of the colony. These simulated communities resemble those of the cellulolytic biofilms observed experimentally, see Figure 1 for comparison. As the radial growth continues, some colonies coalesce and form communities. In turn, these communities grow and join other communities, which produce even larger communities.

A comparison of Figures 4 and 5, reveals the effect of the attachment rate ψ on colony formation. For simulation 2, colony formation becomes noticeable later than in simulation 1 (around $t \approx 45$ vs. $t \approx 25$) and more scarcely. As a consequence, the initial colony formations are patchier and it takes longer for these colonies to merge and form communities. The substrate supply lasts longer, to around $t \approx 80$, in simulation 2 to the point that some remaining spots of the substratum that still contain maximal carbon concentrations have become sites for the formation of new colonies, whereas in simulation 1 carbon is depleted at $t \approx 45$. This suggests that cell attachment, rather than growth can be the limiting factor in carbon utilization.

5.1.2. Case 2: initially colonized domain

As demonstrated in Figures 6 and 7, the central colony grows radially with a distinct circular front that sweeps through the domain. In the wake of this front, the substrate is consumed, which eventually renders the area within the central crater-like structure unavailable to random cell attachment.

As before, by changing the value of ψ the frequency of attachment events is controlled. For higher values, more random attachment events occur over the areas that are still replete with substrate and are not yet occupied by biomass. This is evidently the case in Figure 6 where by $t = 35$ numerous small colonies have formed and many newly emerged colonies are visible. The growth rate for these small peripheral colonies is such that by $t = 40$ many of them have merged and developed into communities. Subsequently, the biofilm communities surround the central colony from all directions and merge into it. This interaction distorts the sharp front of the central colony, which is characteristic for the underlying deterministic model [29], and suppresses its further growth toward the boundaries of the domain. The situation is very different when ψ is set to half the previous value. Figure 7 displays how the lower rate of random attachment can emphasize the usual characteristics that we expect from a deterministic model of a single-colony cellulolytic biofilm, i.e., attachment processes do not notably affect the biofilm dynamics.
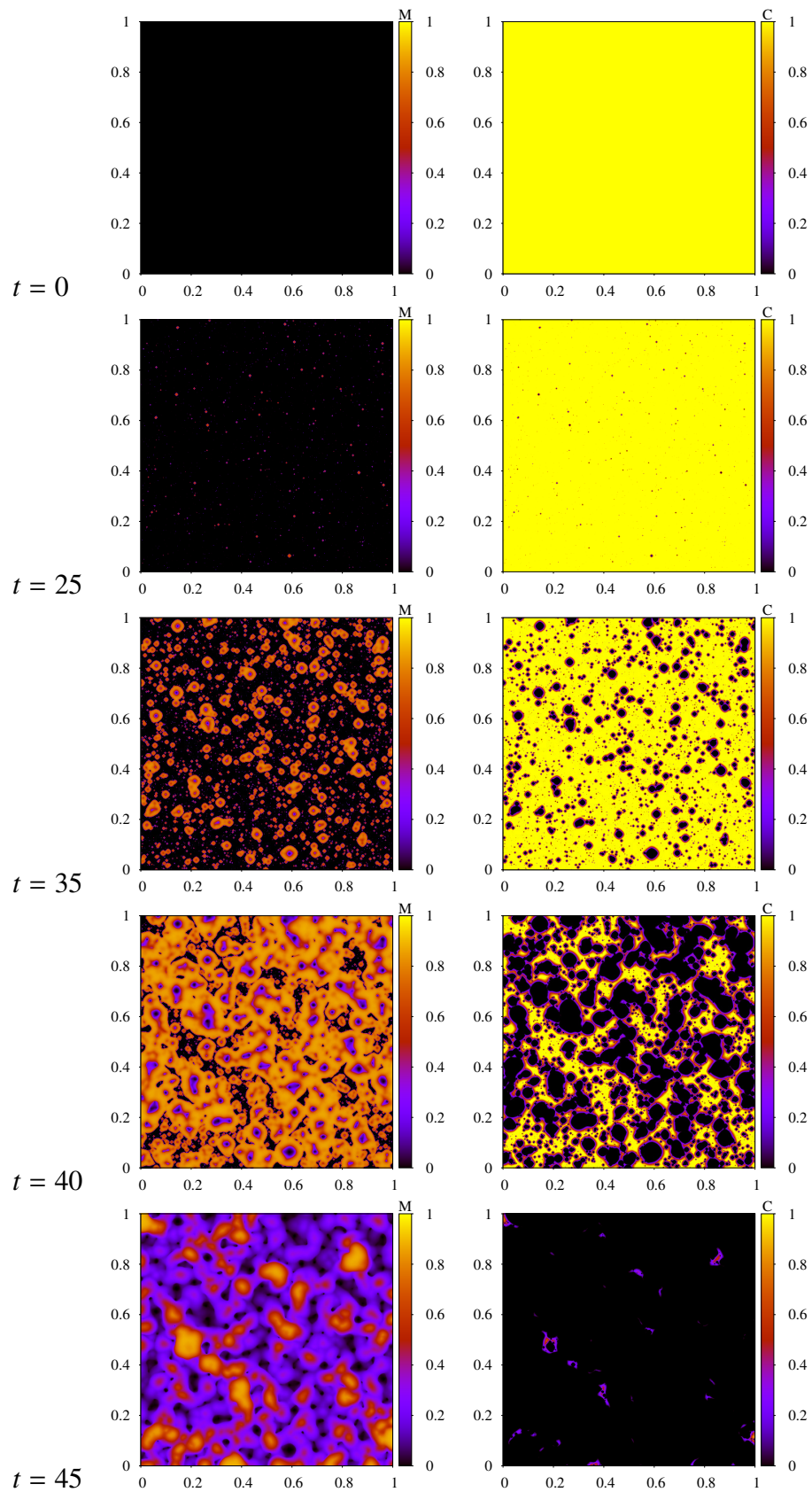
Figure 4. Simulation of (3.6) with $\psi = 0.08$ and initially $M \equiv 0$ everywhere. Shown are biomass density M (left) and carbon concentration C (right) at different times t.
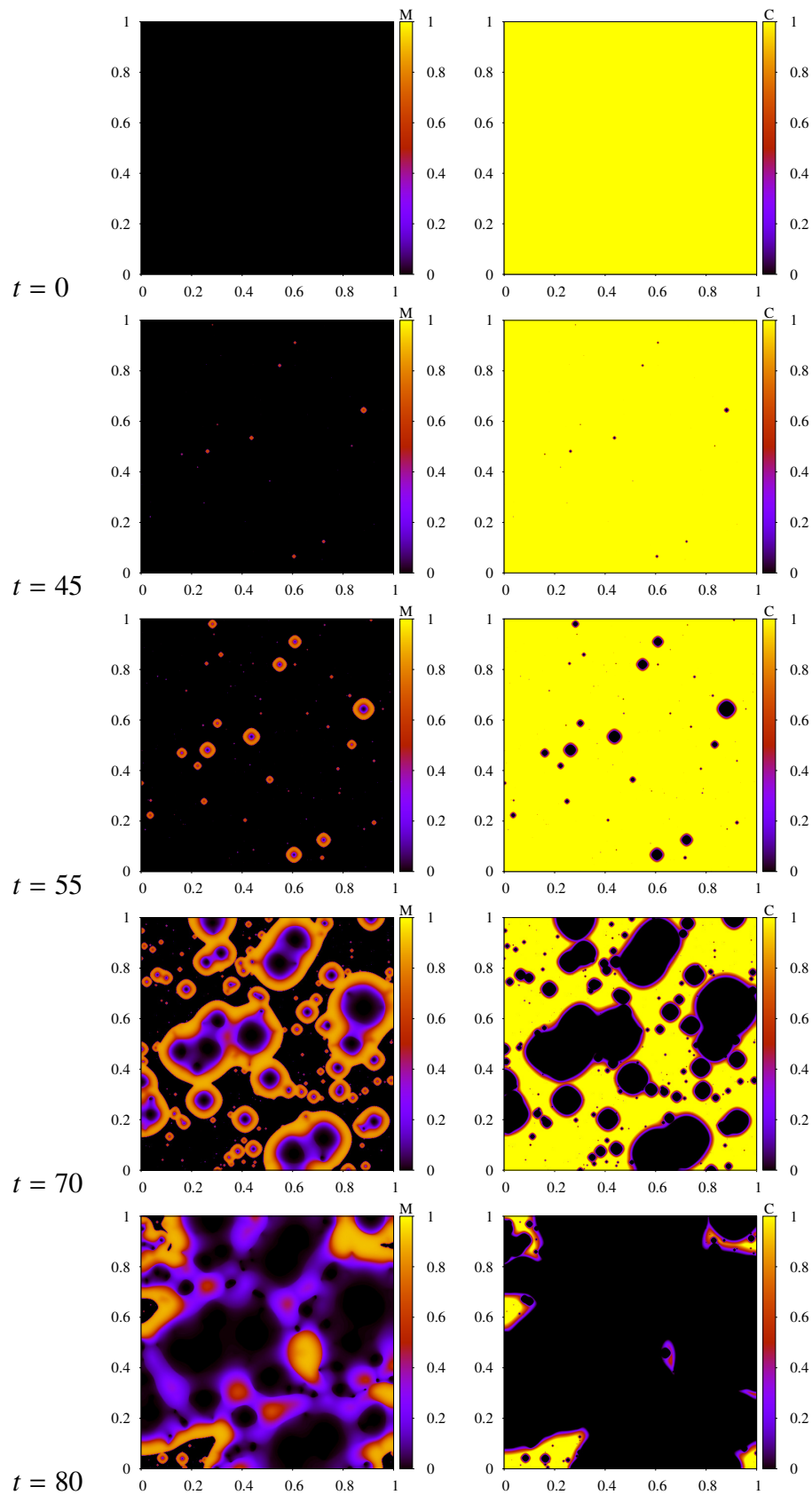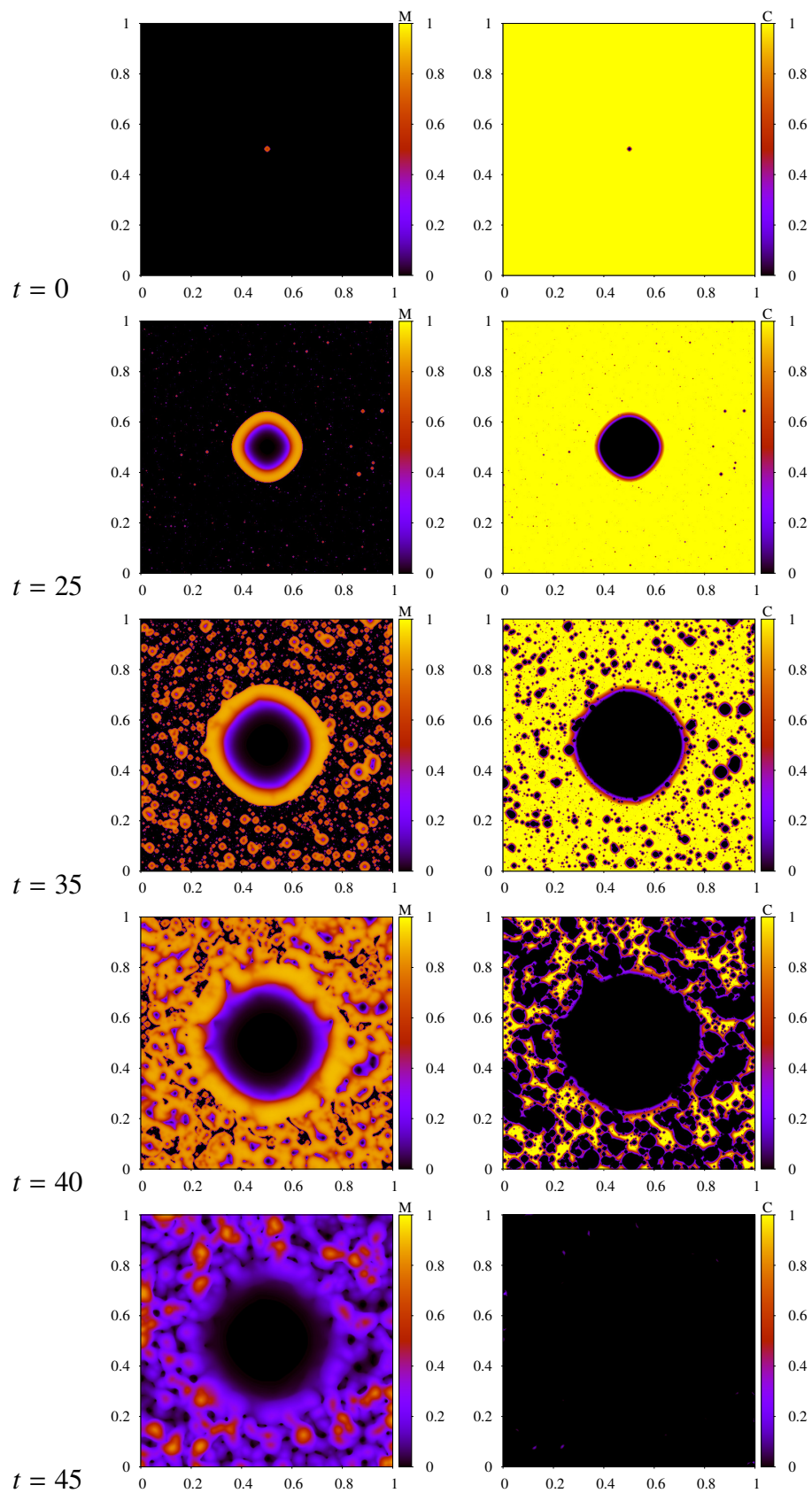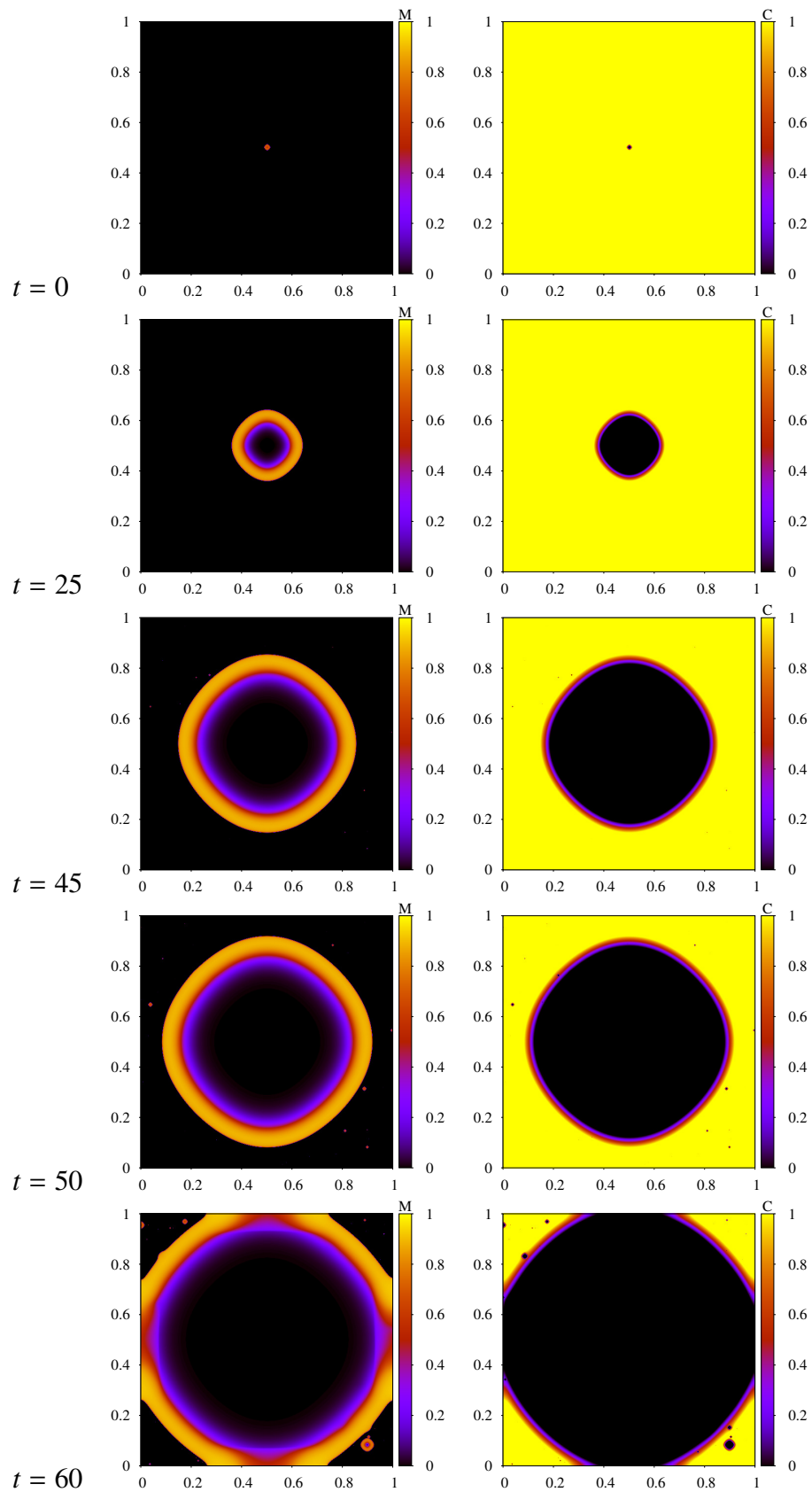
Figure 5. Simulation of (3.6) with $\psi = 0.04$ and initially $M \equiv 0$ everywhere. Shown are biomass density M (left) and carbon concentration C (right) at different times t.

Figure 6. Simulation of (3.6) with $\psi = 0.08$ and initial colony in the center. Shown are biomass density M (left) and carbon concentration C (right) at different times t.

Figure 7. Simulation of (3.6) with $\psi = 0.04$ and initial colony in the center. Shown are biomass density M (left) and carbon concentration C (right) at different times t.

Table 2. Model parameters used in 1D simulations in section 5.2.

Parameter	Symbol	Value	Reference
Constant Coefficient in the AF SDE	ψ	0.1 & 0.2	assumed
Starting integer seed for the PRNG	sd	1-100	assumed
Number of grid cells	N	2^4-2^{13}	assumed
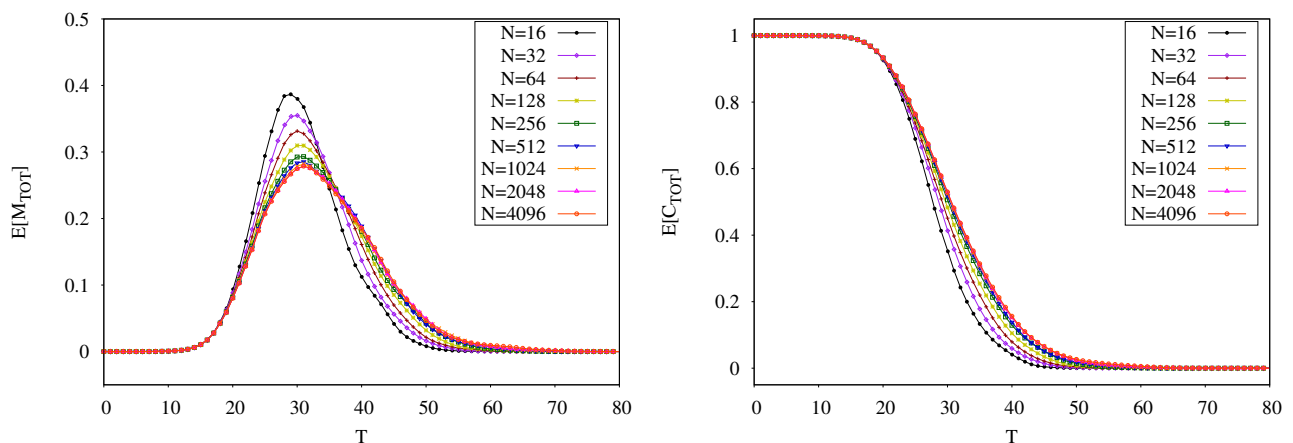
Figure 8. The time evolution of the total biomass M (left) and carbon C (right) for several grid sizes averaged over a hundred realizations. The spatial size of each random cell attachment is $\frac{1}{16}$ of the domain length. Here, $\psi = 0.2$.
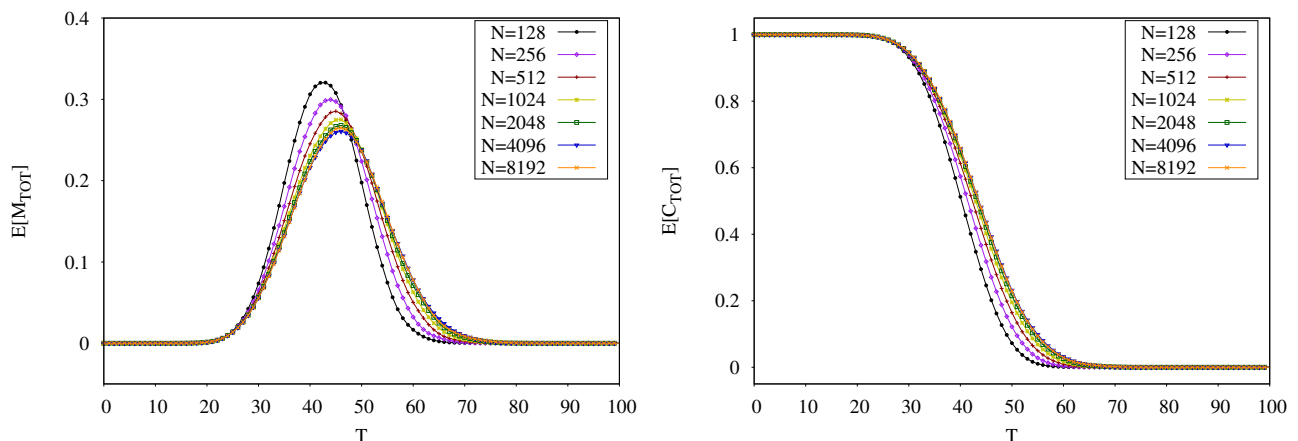
Figure 9. The time evolution of the total biomass M (left) and carbon C (right) for several grid sizes averaged over a hundred realizations. The spatial size of each random cell attachment is $\frac{1}{128}$ of the domain length. Here, $\psi = 0.1$.

5.2. Grid refinement studies

We have performed a spatial grid convergence study to investigate how the attachment model (3.6) behaves as the spatial discretization is refined. This has been carried out in 1D simulations because of the high computational costs incurred by 2D simulations. Each simulation in the study has been repeated several times with different initialization of the pseudo random number generator (PRNG) seed. The simulation parameters used for all 1D simulations are mostly the same as the ones provided in Table 1. The few changes are summarized in Table 2. For quantitative comparability of results between grid resolutions, we keep the size of attachment sites constant. We test two values, $\frac{1}{16}$ and $\frac{1}{128}$ of the domain length.

The primary quantities of interest that we track in our simulation experiment are the total biomass and the total carbon in the system,

$$M_{TOT}(t) := \int_{\Omega} M(t, x)dx = \Delta x \sum_{p=1}^{N} M_p, \tag{5.1}$$

and

$$C_{TOT}(t) := \int_{\Omega} C(t, x)dx = \Delta x \sum_{p=1}^{N} C_p, \tag{5.2}$$

where $\Omega = [0, 1] \subset \mathbb{R}$ is the spatial domain and N is the number of grid points. The sample mean and the sample standard deviation for these quantities of interest, as functions of t, taken over the repeats of the simulations are computed as

$$\text{sample mean} := E[X] = \frac{1}{n} \sum_{i=1}^{n} X_i, \tag{5.3}$$

$$\text{sample standard deviation} := SD[X] = \sqrt{\frac{1}{n-1} \sum_{i=1}^{n} (X_i - E[X])^2}, \tag{5.4}$$

$$\text{standard error of the sample mean} := SEM[X] = \frac{SD[X]}{\sqrt{n}}, \tag{5.5}$$

where X stands for M_{TOT} and C_{TOT} and n is the number of samples (realizations).

The sample means are plotted in Figures 8 and 9, taken over $n = 100$ simulations for each grid size in the convergence study. For both sizes of attachment sites the results are qualitatively different, and easily predictable: The amount of carbon decreases monotonically. During an initial phase, when little biomass is in the system the amount of carbon remains close to the initial amount. Decay becomes notable at the same time as the biomass in the system increases, and levels off eventually when all carbon is consumed. When the amount of carbon drops below half of the initial amount, the amount of biomass decreases again. This indicates that the loss terms in the biomass equation begin to dominate and that also attachment ceases, due to a lack of carbon in the system, i.e., a lack of favourable attachment sites. This qualitative behaviour is observed for all grid sizes and both sizes of attachment sites. Coarser grids predict a higher biomass peak, and both an earlier occurrence of the biomass peak and the leveling off of carbon. However, as the grid is refined, convergence of the sample mean is observed for both quantities of interest.
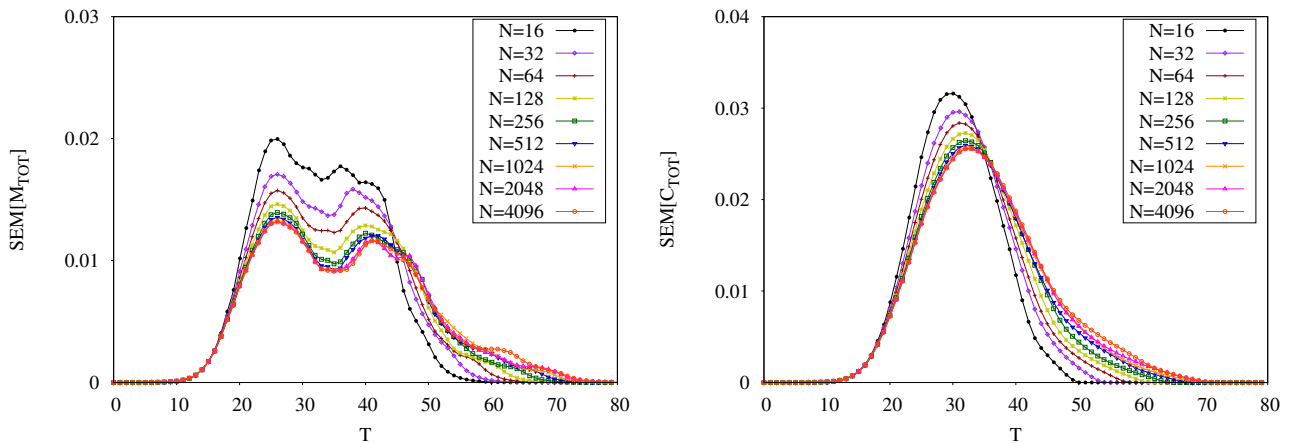
Figure 10. The biomass SEM (left) and carbon SEM (right) for several grid sizes obtained from a hundred realizations. The spatial size of each random cell attachment is $\frac{1}{16}$ of the domain length. Here, $\psi = 0.2$.
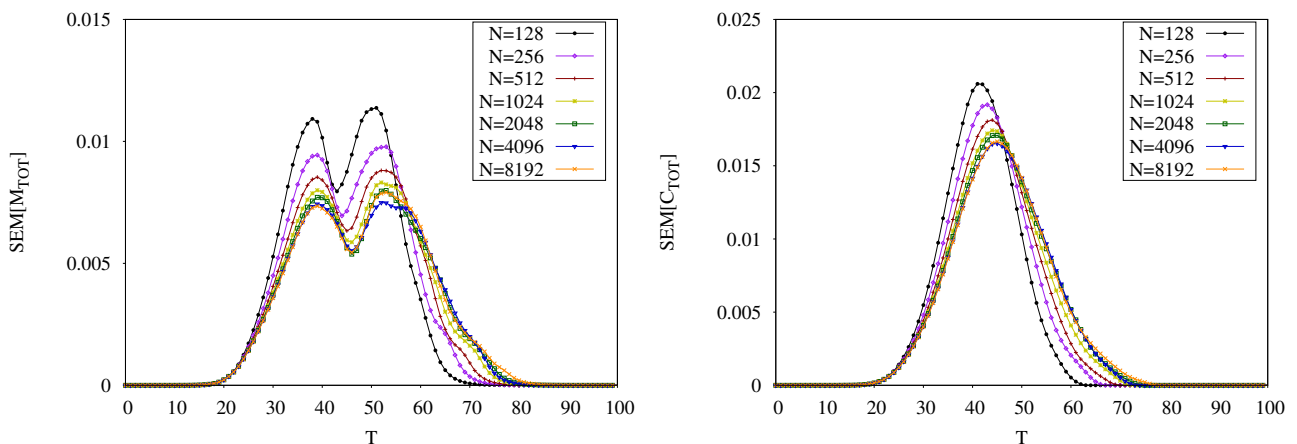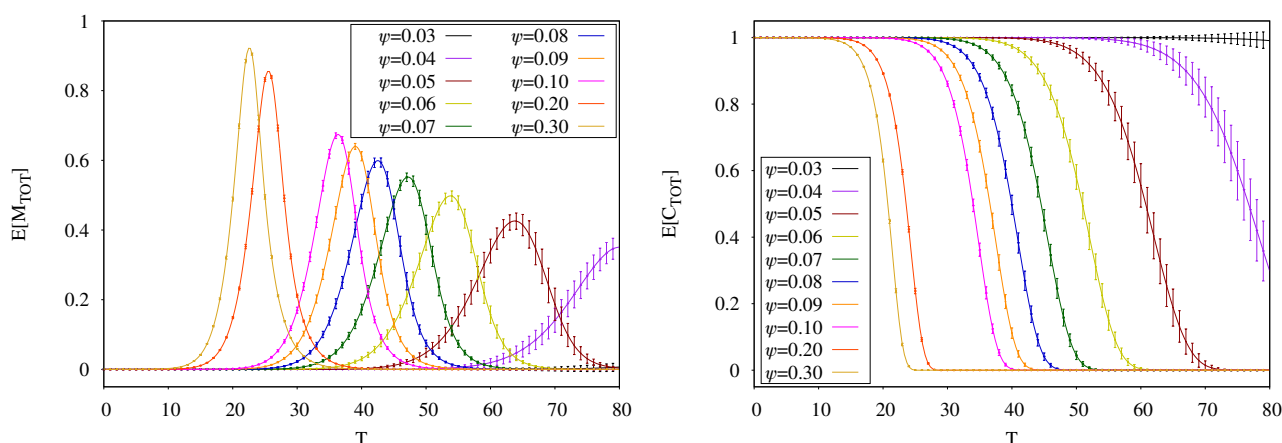

Figure 11. The biomass SEM (left) and carbon SEM (right) for several grid sizes obtained from a hundred realizations. The spatial size of each random cell attachment is $\frac{1}{128}$ of the domain length. Here, $\psi = 0.1$.

Figure 12. The mean values of the spatially integrated biomass density (left) and carbon concentration (right) plotted against time for several values of ψ. Here, $a = 0.99$, $b = 1.00$ and $\sigma = 0.00001$ for all the one hundred sampled simulations. The vertical bars represent the mean ± one standard deviation.

Also the standard error of the sample means converge as the grid is refined for both quantities of interest, cf. Figures 10, 11. The time course for these values reflects the time course of the sample means: For M_{TOT} we observe two maxima, one corresponding to the phase of increasing biomass, one to the decay phase, for C_{TOT} we have a single peak during the decay phase. For both quantities of interest the deviation from the mean is larger for the coarser grids.

5.3. Dependency on attachment parameters

In order to investigate the dependence of the model behaviour on the parameters that we introduced to describe cell attachment we carry out a series of simulations in which the values of a, b, σ, ψ are varied. The remaining parameters are given in Table 1. This simulation experiment is conducted on a square domain of 256×256 grid points, which represents a good compromise of grid resolution and computational cost. Each simulation is repeated several times and we again use the total biomass (5.1) and carbon (5.2) in the system as quantities of interest and compute the sample mean and sample standard deviation in accordance with (5.3), (5.4). For all sampled simulations, the substratum or spatial domain has been set to have zero biomass attached to it initially. As simulation time progresses, random cell attachment events occur, and subsequently establish small colonies, which then grow, consume the substrate and eventually merge to form larger communities (Figures 4 and 5 provide visual representations for this scenario).

The effect of varying ψ over one order of magnitude, from 0.03 to 0.3 are shown in Figure 12. As we elaborated in Section 5.1.1, the higher values of this parameter are associated with more instances of stochastic cell attachment that also begin to occur earlier in time. This model behavior is corroborated by the results shown here as well. Greater ψ values result in more random cell attachment and as more initial colonies form, the substratum is covered by more distributed amounts of biomass, which accelerates both, the growth phase and the consumption of the carbon substrate. As a result, the biomass curves become narrower and attain higher peaks, and the carbon curves become steeper during the consumption phase. As the value of ψ is increased, the standard deviation of both
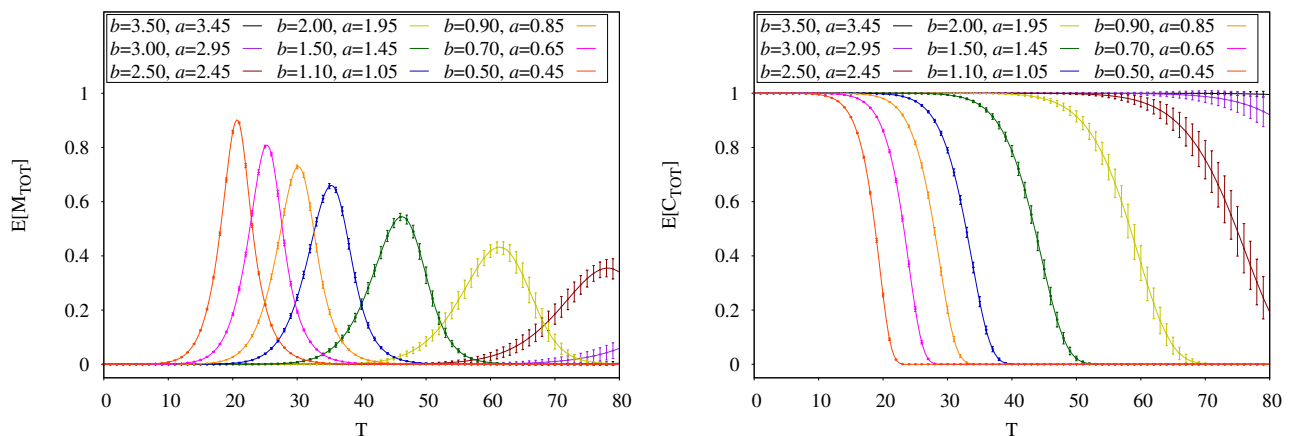
Figure 13. The mean values of the spatially integrated biomass density (left) and carbon concentration (right) plotted against time for several values of a and b. Here $\psi = 0.1$, $b - a = 0.05$ and $\sigma = 0.00001$ for all the one hundred sampled simulations. The vertical bars represent the mean $\pm$ one standard deviation.

quantities of interest decrease, i.e., as the number and size of colonies grows due to higher rates of random attachment the uncertainty in M_{TOT} and C_{TOT} curves is reduced. This is a consequence of the increased colonisation and thus the reduction of possible attachment sites.

Figure 13 shows the effect of varying a and b simultaneously. We kept the difference $b - a = 0.05$ constant and varied $0.45 \leq a \leq 3.45$. As the values of a and b are reduced, the maximum and minimum values of the biomass and carbon curves, respectively, are attained earlier in time. This behavior is similar to the one manifested by higher values of ψ, which induced more frequent random attachment. Since the value of ψ is the same for all the sampled simulations in this case, the reason that more attachment events occur for smaller values of a and b may be attributed to the higher likelihood of the event where the value of the AF falls within the impulse interval as the interval is shifted towards zero (the initial value of the AF over the whole domain). We also observe the same gradual decrease of the standard deviation in both quantities of interest. The reason is the same as the one discussed for the role of ψ, *i.e.*, more instances of random attachment reduce the subsequent possibility of attachment and slows down growth.

The situation is different for the dependency of σ, which we varied between 10^{-2} and 10^{-7}, see Figure 14. For the smaller values there is no substantial effect on the quantities of interest observed, only for the largest value small differences are noticed. The standard deviation in all cases remains small. This confirms that the mollification of the discrete impulse that was introduced via (3.1) indeed converges as $\sigma \to 0$, and that our approach is well suited to mimic discrete attachment events. The reason for the fairly low sensitivity observed here can be ascribed to the role of σ in the impulse function f, see Figure 3. For values of σ close to the ones used in our simulations, the impulse function maintains a rectangular shape between the values of a and b. Hence, the cut-offs for effective ϕ values are sharp and very close to the numerical values of a and b. For larger values of σ the rectangular shape gives way to a mound-like shape with variable width, which does not equal the length of $[a, b]$ everywhere. Therefore, as long as σ assumes values in the range that maintains somewhat sharp cut-offs for ϕ, the frequency of stochastic cell attachment is most sensitive to the values of a, b and ψ.
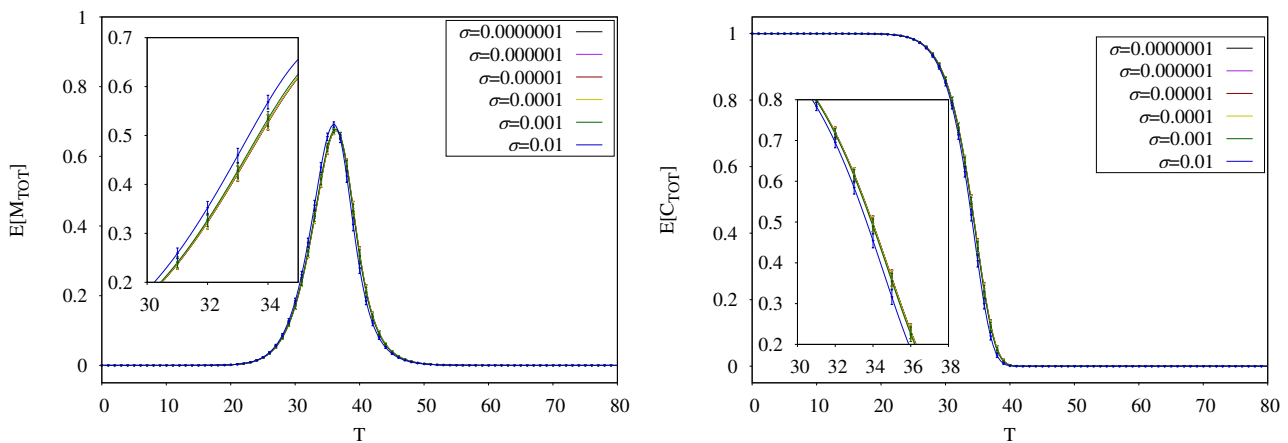
Figure 14. The mean values of the spatially integrated biomass density (left) and carbon concentration (right) plotted against time for several values of σ. Here $a = 0.99$, $b = 1.00$ and $\psi = 0.1$ for all the one hundred sampled simulations. The vertical bars represent the mean $\pm$ one standard deviation. The insets provide a zoomed-in portion of the graphs for clarification.

6. Discussion

Through performing many numerical simulations, we provided evidence that our model of stochastic attachment of free bacterial cells to the substratum is a viable mathematical framework that can be incorporated into deterministic continuum models. We especially focused on cellulolytic biofilms for which the underlying deterministic PDE-ODE system has been previously shown to model biomass growth and inverse colony formation in accordance with experimental findings [29].

The particular nature of random cell attachment and the fact that there are not sufficient experimental data in the literature to be used to parameterize possible underlying mechanisms, especially in the case of cellulolytic biofilms, make the deployment of a stochastic approach for modeling the attachment process adequate for now. However, future experiments and further analysis of their results may eventually provide a path towards a model that incorporates more physical details in a more deterministic manner.

The mere addition of noise to the underlying PDE model would not lend itself readily and directly to describe random cell attachment: while such additive noise would naturally lead to an Itô SDE for which established numerical methods exist, the inherent possibility that solutions attain negative values and the random magnitude of the stochastic part render this approach unsuitable for constructing random attachment with certain characteristics. Moreover, attachment described by such an approach will lose its discrete nature. These were the main reasons for us to decouple the stochasticity of attachment from the main model equations and to describe it via a separate variable.

To cast our model in the theoretically well-established formalism of stochastic differential equations is a step forward from the *ad-hoc* method used previously in [29]. Although this may be viewed as the most significant accomplishment of the present work, the underlying stochastic variable that we introduced, namely the attachment factor ϕ, which produces the required stochasticity in the system, is not physically on par with the other two state variables, the biomass density M and the carbon

concentration C. Since our main purpose has been utilizing an approach that is supported by an existing rigorous theoretical foundation, the introduced SDE formulation can be viewed as the first step towards modeling the stochastic attachment phenomenon. There is room for improving our approach, e.g., by attempting to reformulate a similar model as mathematically simpler Random Ordinary Differential Equation [44].

Our numerical simulations have shown that in a two-dimensional setting the proposed model produces results that exhibit similar characteristics as the cellulolytic biofilm formations documented in Figure 1. Furthermore, our 2D simulations may be viewed as computer experiments where different scenarios with various initial states can be replicated and consequently an improved understanding of the studied biological system could be acquired. Our computational analysis suggests the convergence of the spatially discrete model as the grid is refined.

The number and frequency of stochastic cell attachment is most sensitive to the constant coefficients in the equation for the attachment factor ψ, and the shift parameters of the impulse function, a and b. These three parameters summarize both the external physical conditions and reactor properties. Accordingly, and this is also supported by our simulations, the temporal and spatial component of attachment cannot be controlled separately in an easy manner by these parameters. Furthermore, our simulations have shown that the continuous mollification (3.1) of the discrete impulse function is able to mimick discrete attachment events, i.e., the mollified version behaves as one expects from the discrete version if $\sigma \to 0$.

As a direct consequence of the model being inherently stochastic, some of the simulation results and drawn conclusions have been treated statistically. The number of sampled simulations for these statistical treatments have been determined such that the level of stochasticity would not change dramatically as more samples were included in calculating the desired statistics. To determine this number, we have used the 1-norm of the standard deviation as a measure, which can be calculated from the data for both 1D and 2D simulations.

In the case of 1D simulations, we have both biomass data and carbon data for a number of grid sizes that have been used for the results in Section 5.2. The 1-norm is calculated easily by summing over standard deviations from several time points. Figure 15 shows the results of these calculations for one of the grid sizes.

As the sample size is increased, the 1-norm of the standard deviation for the biomass begins to show less variation above seventy samples. Therefore, for keeping the computational costs manageable, we have used a maximum of one hundred samples for calculating the desired statistics. The same type of behavior is somewhat evident from the carbon curves as well. This result holds true for other grid sizes, which we have not displayed here.

The 2D case, is handled similarly. The biomass and carbon data have been acquired from sensitivity analysis simulations as described in Section 5.3. Figure 16 shows the results for calculating the 1-norm of standard deviations for different sample sizes. Both, biomass and carbon curves, begin to show signs of levelling off after the number of samples exceeds eighty.

One of the most challenging aspects of our study has been the computational realization of the numerical simulations. This has manifested itself at different levels, *e.g.*, in the strict time step restriction of the Euler-Mayurama method, which however is also necessary for a sufficiently accurate resolution of the temporal stochasticity, and the process of random number generation. The algorithms we have utilized for random number generation are textbook methods with well-known
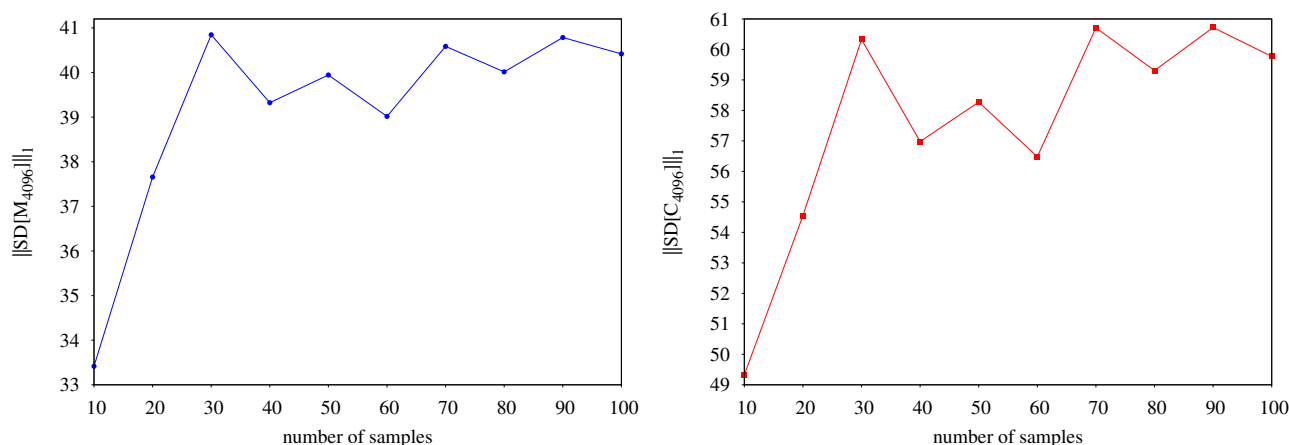
Figure 15. The 1-norm of standard deviation calculated using standard deviations at several time points for a 1D setting. The x-axis represents the number of samples. The grid size is $2^{14} = 4096$. The biomass curve is on the left and the carbon curve is on the right. Note that the numerical data from the spatially integrated state variables are used here. The simulation parameters are given in Table 1 and specifically here $\psi = 0.2$, $a = 0.99$, $b = 1.00$ and $\sigma = 0.00001$.
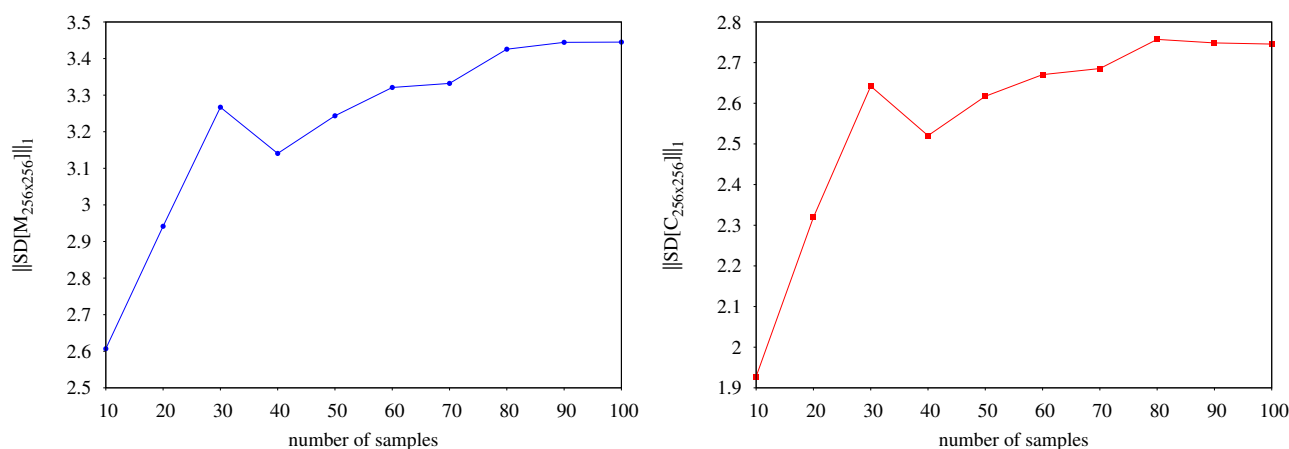

Figure 16. The 1-norm of standard deviation calculated using standard deviations at several time points for a 2D setting. The x-axis represents the number of samples. The grid size is 256×256. The biomass curve is on the left and the carbon curve is on the right. Note that the numerical data from the spatially integrated state variables are used here. The simulation parameters are given in Table 1 and specifically here $\psi = 0.08$, $a = 0.99$, $b = 1.00$ and $\sigma = 0.00001$.

properties. We used them because they are easy to implement and test, and well understood. Relying in a first proof of principle study on such proven and generally applicable methods, allows us to focus on the modeling question at hand and to separate it from questions of reliability of the numerical

methods used. While there are established and generally accepted methods for semilinear parabolic stochastic PDE extensions, e.g., [33], these often use the specific properties of the linear diffusion operator and cannot be easily extended and applied to quasilinear problems. In order to make the model that we presented suitable for large scale simulation experiments, however, this question of acceleration of the numerical simulations needs to be addressed.

Finally, our model development was presented for a two-dimensional setup that corresponds to the spread of cellulosic biofilms over very thin substrata. In this case, the entire domain is a potential attachment site. To apply the same ideas to a fully three-dimensional setup with a substratum of non-negligible depth, penetration of the colonies into the substratum needs to be accounted for. The computational domain in that case would be a section of the substratum, some part of the boundary of which interfaces with the (not explicitly considered) aqueous phase. Possible attachment sites would be initially restricted to that part of the boundary and, later on, as the wave fronts penetrate into the substratum, potentially also to the wake side of these fronts, i.e., the set of attachment sites would be a 2D surface in the 3D domain. Similarly, a two-dimensional setup describing vertical degradation of a thin slice of substratum could be handled, leading to a 1D curve in the 2D domain. Such a setup will bring with itself some additional technical challenges which to address we leave for a forthcoming study.

7. Conclusion

Discrete random cell attachment can be included in a deterministic continuous biofilm model by introducing an auxiliary stochastic process that generates attachment impulses. This formulation leads to a stochastic model that naturally preserves non-negativity, as opposed to the direct addition of white noise. Using a regularised version of the underlying partial differential equation, this model also preserves the boundedness of solutions. In contrast to an earlier more direct approach to include (nonlinear) stochastic impulses, this formulation leads to an Itô stochastic differential equation for which the established stochastic numerical methods are known to reproduce the stochasticity of the differential equation correctly. We used here the simplest numerical solver for a stochastic differential equation, namely the Euler-Maruyama method. This turned out to be highly expensive. More advanced stochastic time integration schemes should be explored in the future.

Whereas we developed our model here for the particular setting of cellulosic biofilms, we think that in principle this discrete stochastic attachment framework can also be incorporated into other two-dimensional or three-dimensional continuum models, including models of more traditional biofilms where nutrients are supplied by diffusion from the aqueous phase to the colony.

Conflict of interest

No conflicts are declared.

References

1. Z. Lewandowski, H. Beyenal, *Fundamentals of Biofilm research*, CRC Press, Boca Raton, 2007.

2. B. D'Acunto, L. Frunzo, V. Luongo, M. R. Mattei, Free boundary approach for the attachment in the initial phase of multispecies biofilm growth, *Zeitschrift für angewandte Mathematik und Physik*, **70** (2019), 91.

3. B. D'Acunto, L. Frunzo, M. R. Mattei, Continuum approach to mathematical modellingof multispecies biofilms, *Ric. di Mat.*, **66** (2017), 153–169.

4. A. Masic, H. J. Eberl, Persistence in a single species cstr model with suspended flocs and wall attached biofilms, *Bull. Math. Biol.*, **74** (2012), 1001–1026.

5. A. Masic, H. J. Eberl, A modeling and simulation study of the role of suspended microbial populations in nitrification in a biofilm reactor, *Bull. Math. Biol.*, **76** (2014), 27–58.

6. H. J. Gaebler, H. J. Eberl, A simple model of biofilm growth in a porous medium that accounts for detachment and attachment of suspended biomass and their contribution to substrate degradation, *Eur. J. Appl. Math.*, **29** (2018), 1110–1140.

7. E. Alkvist, I. Klapper, A multidimensional multispecies continuum modelfor heterogeneous biofilm development, *Bull. Math. Biol.*, **69** (2007), 765–789.

8. B. V. Merkey, B. E. Rittmann, D. L. Chopp, Modeling how soluble microbial products (smp) support heterotrophic bacteria in autotroph-based biofilms, *J. Theor. Biol.*, **259** (2009), 670–683.

9. N. G. Cogan, Two-fluid model of biofilm disinfection, *Bull. Math. Biol.*, **70** (2008), 800–819.

10. J. Dockery, I. Klapper, Finger formation in biofilms, *SIAM J. Appl. Math.*, **62** (2002), 853–869.

11. J. P. Ward, J. R. King, Thin-film modelling of biofilm growth and quorum sensing, *J. Eng. Math.*, **73** (2012), 71–92.

12. H. J. Eberl, D. F. Parker, M. Van Loosdrecht, A new deterministic spatio-temporal continuum model for biofilm development, *Comput. Math. Methods Med.*, **3** (2001), 161–175.

13. K. A. Rahman, R. Sudarsan, H. J. Eberl, A mixed-culture biofilm model with cross-diffusion, *Bull. Math. Biol.*, **77** (2015), 2086–2124.

14. M. Whidden, N. Cogan, M. Donahue, F. Navarrete, L. D. L. Fuente, A two-dimensional multiphase model of biofilmformation in microfluidic chambers, *Bull. Math. Biol.*, **77** (2015), 2161–2179.

15. F. El-Moustaid, A. Eladdadi, L. Uys, Modeling bacterial attachment to surfaces at and early stage of biofilm development, *Math. Biosci. Eng.*, **10** (2013), 821–841.

16. H. J. Eberl, H. Khassehkhan, L. Demaret, A mixed-culture model of a probiotic biofilm control system, *Comput. Math. Methods Med.*, **11** (2010), 99–118.

17. A. Carroll, C. Somerville, Cellulosic biofuels, *Annu. Rev. Plant Biol.*, **60** (2009), 165–182.

18. M. H. Langholtz, B. J. Stokes, L. M. Eaton, 2016 billion-ton report: Advancing domestic resources for a thriving bioeconomy, available from: `https://bit.ly/2LLjdBS`, 2016, URL `https://doi.org/10.2172/1271651`, (accessed 12-June-2018).

19. J. G. Linger, A. Darzins, *Consolidated Bioprocessing*, 267–280, Springer New York, New York, NY, 2013.

20. B. G. Schuster, M. S. Chinn, Consolidated bioprocessing of lignocellulosic feedstocks for ethanol fuel production, *BioEnergy Res.*, **6** (2013), 416–435.

21. A. Dumitrache, G. Wolfaardt, G. Allen, S. N. Liss, L. R. Lynd, Form and function of *Clostridium thermocellum* biofilms, *Appl. Environ. Microbiol.*, **79** (2013), 231–239.

22. V. Mbaneme-Smith, M. S. Chinn, Consolidated bioprocessing for biofuel production: recent advances, *Energy Emission Control Technol.*, **3** (2015), 23–44.

23. Z.-W. Wang, S.-H. Lee, J. G. Elkins, J. L. Morrell-Falvey, Spatial and temporal dynamics of cellulose degradation and biofilm formation by caldicellulosiruptor obsidiansis and clostridium thermocellum, *AMB Express*, **1** (2011), 30.

24. M. van Loosdrecht, J. Heijnen, H. Eberl, J. Kreft, C. Picioreanu, Mathematical modelling of biofilm structures, *Antonie van Leeuwenhoek*, **81** (2002), 245–256.

25. D. G. Davies, M. R. Parsek, J. P. Pearson, B. H. Iglewski, J. W. Costerton, E. P. Greenberg, The involvement of cell-to-cell signals in the development of a bacterial biofilm, *Science*, **280** (1998), 295–298.

26. S. M. Hunt, E. M. Werner, B. Huang, M. A. Hamilton, P. S. Stewart, Hypothesis for the role of nutrient starvation in biofilm detachment, *Appl. Environ. Microbiol.*, **70** (2004), 7418–7425.

27. J. Zhu, J. J. Mekalanos, Quorum sensing-dependent biofilms enhance colonization in vibrio cholerae, *Dev. Cell*, **5** (2003), 647–656.

28. Y. H. An, R. J. Friedman, Concise review of mechanisms of bacterial adhesion to biomaterial surfaces, *J. Biomed. Mater. Res.*, **43** (1998), 338–348.

29. H. Eberl, E. Jalbert, A. Dumitrache, G. Wolfaardt, A spatially explicit model of inverse colony formation of cellulolytic biofilms, *Biochem. Eng. J.*, **122** (2017), 141–151.

30. P. E. Kloeden, E. Platen, *Numerical Solution of Stochastic Differential Equations*, Springer-Verlag, Berlin Heidelberg, 1992.

31. B. Øksendal, *Stochastic differential equations: An introduction with applications*, Springer, Berlin Heidelberg, 2003.

32. A. Dumitrache, G. Wolfaardt, G. Allen, S. N. Liss, L. R. Lynd, Tracking the cellulolytic activity of *Clostridium thermocellum* biofilms, *Biotech. Biofuels*, **6** (2013), 175.

33. G. J. Lord, C. E. Powell, T. Shardlow, *An introduction to computational stochastic PDEs*, 50, Cambridge University Press, New York, NY, 2014.

34. J. Cresson, M. Efendiev, S. Sonner, On the positivity of solutions of systems of stochastic pdes, *ZAMM - J. Appl. Math. Mech. / Zeitschrift für Angewandte Mathematik und Mechanik*, **93** (2013), 414–422.

35. J. Cresson, B. Puig, S. Sonner, Stochastic models in biology and the invariance problem, *Discrete Cont. Dyn- B*, **21** (2016), 2145–2168.

36. R. Fox, Stochastic versions of the Hodgkin-Huxley equations, *Biophys. J.*, **72** (1997), 2068–2074.

37. A. Kamina, R. W. Makuch, H. Zhao, A stochastic modeling of early HIV-1 population dynamics, *Math. Biosci.*, **170** (2001), 187–198.

38. H. C. Tuckwell, E. L. Corfec, A stochastic model for early hiv-1 population dynamics, *J. Theor. Biol.*, **195** (1998), 451–463.

39. H. J. Eberl, L. Demaret, A finite difference scheme for a degenerated diffusion equation arising in microbial ecology, *Electron. J. Differ. Equ*, **15** (2007), 77–95.

40. A. Dumitrache, H. J. Eberl, D. G. Allen, G. M. Wolfaardt, Mathematical modeling to validate on-line co2 measurements as a metric for cellulolytic biofilm activity in continuous-flow bioreactors, *Biochem. Eng. J.*, **101** (2015), 55–67.

41. M. A. Efendiev, S. V. Zelik, H. J. Eberl, Existence and longtime behavior of a biofilm model, *Commun. Pure Appl. Anal.*, **8** (2009), 509–531.

42. M. Ghasemi, H. J. Eberl, Time adaptive numerical solution of a highly degenerate diffusion–reaction biofilm model based on regularisation, *J. Sci. Comput.*, **74** (2018), 1060–1090.

43. J. Cresson, B. Puig, S. Sonner, Validating stochastic models: invariance criteria for systems of stochastic differential equations and the selection of a stochastic Hodgkin-Huxley type model, *Internat. J. Biomath. Biostat.*, **2** (2013), 111–222.

44. X. Han, P. E. Kloeden, *Random Ordinary Differential Equations and Their Numerical Solution*, Springer, Singapore, 2017.

45. P. E. Kloeden, E. Platen, H. Schurz, *Numerical Solution of SDE Through Computer Experiments*, Springer-Verlag, Berlin Heidelberg, 1994.

46. J. Monod, The growth of bacterial cultures, *Ann. Rev. Microbiol.*, **3** (1949), 371–394.

47. Y. Rohanizadegan, *A Stochastic Formulation of Bacterial Attachment in a Spatially Explicit Model of Cellulolytic Biofilm Formation*, Master's thesis, University of Guelph, Canada, 2018, available from: URL `https://atrium.lib.uoguelph.ca/xmlui/handle/10214/14673`.

48. W. Walter, *Ordinary Differential Equations*, Springer New York, 1998.