Mathematical analysis, modelling and simulation of microbial population dynamics

Vincent Quedeville

THÈSE

En vue de l'obtention du

DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE

Délivré par :

Institut National Polytechnique de Toulouse (Toulouse INP)

Discipline ou spécialité :

Dynamique des fluides

Présentée et soutenue par :

M. VINCENT QUEDEVILLE

le jeudi 4 juin 2020

Titre :

Mathematical analysis, modelling and simulation of microbial population dynamics

---

Ecole doctorale :

Mécanique, Energétique, Génie civil, Procédés (MEGeP)

Unité de recherche :

Laboratoire d'Ingénierie des Systèmes Biologiques et des Procédés (LISBP)

Directeur(s) de Thèse :

M. RODNEY FOX

M. JEROME MORCHAIN

Rapporteurs :

M. LAURENT BOUDIN, UNIVERSITE SORBONNE

Mme MAGALI RIBOT, UNIVERSITE D'ORLEANS

Membre(s) du jury :

Mme BEATRICE LAROCHE, INRA JOUY EN JOSAS, Président

M. JEROME MORCHAIN, INSA TOULOUSE, Membre

M. PHILIPPE VILLEDIEU, ONERA TOULOUSE, Membre

M. RODNEY FOX, IOWA STATE UNIVERSITY, Membre

Table des matières

Glossary

A_P PTS activity. 107, 109, 115, 116, 121, 123–128, 131, 133, 160

A_p permease activity. 106, 107, 109, 115–117, 120, 121, 123–128, 131–133, 143–145, 147, 148, 154, 160, 161

K_P PTS affinity to glucose. 103, 105, 107, 109, 116–118, 153

K_{conv} cell length to cell mass ratio, supposed equal to $10^3\rho\pi d^2/4$. 103, 112, 113, 115, 116, 118, 131, 133, 157

K_p permease affinity to glucose. 103, 105, 107, 109, 116–118, 133, 148

$N_{\max}$ maximum permease number. 107, 115–118, 123, 127–131, 143

N_p cell-scale permease number. 106, 107, 109, 115–117, 120, 121, 123–128, 131–133, 137, 143–145, 147, 148, 154, 160, 161

R_{NX} biomass to cell number ratio, largely ignored by the experimentalists, making any headway towards microscopical scale-population dynamics modelling a fantasy. 116, 117

$\mathcal{M}_{i,j,o,r}$ mesh cell. 127, 128

$\psi_{\max}$ maximum uptake capability per permease. 103, 105, 107, 109, 115, 116, 118, 133

τ_P PTS uptake characteristic time for a 3D PBM. 103, 105, 107, 116, 118, 131

τ_p permease uptake characteristic time for a 3D PBM. 103, 105

$q_{P_{\max}}$ maximum PTS uptake capability. 103, 105, 107, 109, 115–118, 128, 131, 133, 139, 148, 149, 156, 157, 159

q_P PTS uptake capability. 100–103, 105, 107, 110, 115–117, 120, 128, 131–133, 143, 148, 150, 151, 154, 161

q_p permease uptake capability. 100, 101, 103, 105–107, 116, 133, 143, 148, 150, 151

$v_p^{\max}$ maximum growth potential, supposed equal to $\frac{q_{PTS_{\max}}}{Y_{SX}10^3\rho\pi d^2/4}$. 115, 123, 124, 156

Abstract

Microbial cells are used in many industrial applications and geared towards user–set products of interest. These biological processes are conducted in aerated and agitated reactors, the suspended solid biotic phase being exposed to both liquid turbulence and gas flow.

In many occasions, scaling up the process is a key issue because the performances of the biocatalyst cannot be maintained as the reactor size increases. The main reason for this loss of performance is the spatial heterogeneity in the sugar, oxygen, pH and temperature fields. In this context, microbial cells travelling in large scale industrial bioreactors are frequently exposed to concentration fluctuations. These time–varying signals trigger various biological responses that can be conflated into "adaptation" or "stress". Roughly speaking, many peripheral metabolic reactions are activated; the living cost increases and the targeted biological transformation's yield is impacted. The ability to adapt to a fluctuating environment singularises biological systems in comparison with chemical ones. Taking into account this feature requires some insight into the dynamics of the biological processes. Indeed the biological machinery's rendition is a function of both the microbial cells' environment and their "physiological state". The latter is the result of the cells' history, in other words their trajectory inside the heterogeneous reactor. Microbiologists would call on a memory effect, chemical engineers would refer to the multiphase nature of the system along with non–equilibrium concepts while mathematicians describe this aspect as the Markovian feature of the cells' journey inside the fermenter.

One intriguing aspect of a cell's functioning is the uncoupling between uptake and growth rate. One of the fundamental laws in biology is the so called Monod law which states that the growth rate is a function of the limiting nutrient concentration. In fact, in a steady–state continuous culture, it is observed that the uptake rate and the growth rate are proportional, the proportionality constant being christened yield coefficient. It grossly reflects the idea that some cells are produced out of some nutrients. It is also known that the glucose to biomass yield is actually not constant but slightly depends on the growth rate: the higher the growth rate, the poorer the conversion rate. This can be understood as the necessity to optimise the conversion processes when the nutrients are scarce, whereas nutrient abundance allows the microbial cells to set in motion their comfort (less constraining, so to say) metabolic regime. Overall, both the number of internal bioprocess and their conversion efficiency are impacted by the nutrient availability. All these considerations result from experimental observations of steady—state continuous cultures. When exposed to a sudden change in their environment, the cells' transient response deviates from the said steady–state laws. For example, the instantaneous uptake rate in the following of a nutrient limitation relief can be five to ten times the maximum uptake rate in a nutrient rich environment: one shall then consider a non–local equilibrium between the microorganisms and their environment over short time scales. Although such a behaviour is well–known in the microbial reign, it is hardly ever taken into account in the modelling of microbial cell population dynamics.

Another key feature in microbial dynamics is the cell division, through which the cell number exponentially increases until resource depletion. This mechanism introduces some heterogeneity among the population of cells: the time lapse between two consecutive division events is not constant, neither is the cell birth– or division–size. Laws of conservation of matter apply when extensive quantities like mass or enzymatic content are allocated to n–th generation cells from their (n-1)–th generation ancestors. On the other hand, the distribution in intensive quantities is more chaotic and yields consequential fluctuations from mother– to daughter–cells and between sister–cells. A population's heterogeneity can then manifest in properties such as lengthening rate, genetic expression, ...

Modelling population dynamics requires to pick a few insightful properties which distribution best fits the outputs from real–life experiments. To this end, not only the number of potential candidates is large but the dynamics for these internal properties is not a priori known. Thus, a precise definition of the model's objectives is of primary importance in order to focus on a limited but meaningful set of internal cell properties. Analysiswise, most of these quantities have to be compactly supported given that a material balance in a finite size fermenter must lead to finite individual cell mass, enzymatic content, ... The mathematical theory of microbial populations consecrates a $\mathbb{R}_+$ setting though, what is a hazardous framework and a quite inconvenient assumption if numerical methods like finite volume are to be implemented to solve a population balance equation. Regarding this latter remark, some tricks are traditionally enforced by applied mathematicians like setting an artificial upper bound to the integration domain, giving rise to conservative or non–conservative truncations depending on the simulation's aims. A non–conservative truncation has proven handy to address the physics of gelation in fluid mechanics for example. However, this does not seem sound in the context of biological populations for which rupture must happen with probability 1 at a (rather small) finite size, prompting the need to formulate the population balance, not its truncation, over a bounded subset of $\mathbb{R}^p$, if p variables of interest are considered. In this case, the analysis is significantly complicated by the preliminary assumption, similarly as solving a Partial Differential Equation over a bounded subset in comparison with solving it over $\mathbb{R}^p$.

The modelling does not only consist in picking the most insightful variables and their domain, but also in formulating their dynamics as functions of both the microorganisms' "physiological state" and their environment. Multiple time scales are involved depending on the function of the tracked compound in the organisms' organisational structure: an enzyme that testifies to the cell–liquid transfer has to adapt way faster to environmental fluctuations than a cofactor which solely dictates the synthesis of growth–inducing fibers. In the former case, the enzyme's activity level can be seen "almost" at equilibrium with the extracellular medium. In the latter case, it would be quite dubious to claim that the cofactor's rate of activity is at any time a mere function of the substrate offer. With these considerations in mind, a comprehensive biological model must somehow report the wide range of time scales some parts of a bacterium adapt to their environment at, in line with the experimentalists' observations.

Advanced modelling and simulation of bioreactors considering cell population dynamics is therefore a multidisciplinary subject blending mathematics, microbiology and chemical engineering aspects.

The following work will be split into five chapters. The first one will consist in a presentation of bibliographic references pertaining to population dynamics from mathematical, numerical or biological perspectives. Reviewing the analysis contributions gives an insight into the current advancement in the theory of structured models, whereas the microbiology production aims at inspiring the formulation of the most relevant while economical model that captures the bacteria's reported dynamical behaviour when exposed to an ever–fluctuating environment.

The second part will consecrate the mathematical analysis of age– and size–structured models. Both frameworks are significantly different in the sense that nothing a priori bounds the cells' age whereas their mass, amongst others, is finite by assumption. The division process being first and foremost dictated by the cells' size, the divergence of the rupture function at a finite length is a compulsory assessment in the statement of any existence/uniqueness result regarding size–structured models. What singles out the existence and uniqueness result for the eigenelements associated with the transport–fragmentation equation lies in the use of the generational redistribution operator's regularising property that allows to conclude on the Lipschitz–regularity of the Malthus eigenfunction which shapes the steady–state solution. The age distribution can be calculated analytically and informs on the interdivision time distribution too. Rigourous relationships between their respective moments provide an answer to a 60 year–old debate that was still an unsettled issue in the biology community.

The third segment focuses solely on the multiscale modelling of biological reactors. The work bears upon picking the most insightful variables to reproduce at the least cost the standard microbiology experiments, along with their domain and time evolution. Two major factors impact the individuals' dynamical behaviour: their metabolism and the substrate availability. Tackling the former aspect involves a minimal set of biochemical reactions which numerical solution can be implemented using two methods, the "algebraic" one being much faster than the "logical" one. Taking into account the hydrodynamics as the fermenter is treated as one homogeneous compartment leads to an original sub–mesh refinement that precludes any numerical instability when microbial cultures are simulated in an open reactor.

The penultimate chapter is dedicated to the description of the numerical methods to solve the previous chapter's model. Lagrangian, eulerian, and deterministic/stochastic algorithms are confronted for a litmus test consisting in setting a continuous reactor to steady state, with identical dilution rate and substrate feed. The quicker code exhibit significant differences in comparison with the data retrieved by its two counterparts, leaving the question of the best price–quality ratio wide open.

To round off this thesis, a few microbiology experiments which conclusions can be found in the pertaining literature are numerically simulated. Four litmus tests provide as many discussions of the transient states following each of these perturbations and evidence the limits of chapter 3's model. Only qualitative features are evaluated in this work, the quantitative adequacy between numerics and experimental measurements falls within the competence of biological engineers.

Résumé

Résumé Des bactéries sont utilisées dans de nombreux dispositifs industriels dans l'optique de synthétiser un certain nombre de produits d'intérêt. Ces procédés biologiques se déroulent traditionnellement en réacteur aéré et agité, la phase biologique en suspension étant exposée à la fois à la turbulence de la phase liquide et à la circulation de la phase gazeuse.

Dans de nombreuses occasions, la montée en échelle du procédé est une question cruciale car les rendements de conversion de la phase biologique sont amenés à être affectés lorsque la taille du réacteur augmente. La principale raison de cette baisse de performance est l'hétérogénéité spatiale des champs de sucre, d'oxygène, de pH ou de température. Dans ce contexte, des micro–organismes transportés dans un réacteur industriel de grande capacité sont fréquemment exposés à des concentrations fluctuantes. Ces signaux, variables dans le temps, déclenchent diverses réponses biologiques que l'on peut qualifier d'"adaptation" ou de "stress". Plus spécifiquement, il est question de réactions métaboliques périphériques ; le coût énergétique de base de la vie de l'organisme augmente et le rendement de la réaction biochimique recherchée s'en trouve affecté. La capacité d'adaptation à un environnement changeant distingue les systèmes biologiques des systèmes chimiques, mais nécessite de ce fait de se pencher sur la réponse propre du vivant. En effet, la performance de l'appareil biologique dépend à la fois de l'environnement des cellules et de leur état de santé. Ce dernier est une manifestation de l'histoire des cellules, c'est à dire leur trajectoire dans le réacteur hétérogène. Ce phénomène est qualifié d'effet de mémoire dans la communauté des biologistes, alors que le génie chimique a consacré les termes de caractère multiphasique d'un système et de déséquilibre et que la notion de processus markovien est invoquée par les mathématiciens pour décrire le séjour des cellules dans la cuve agitée.

Un aspect intéressant du fonctionnement des cellules est le découplage entre leur capacités d'assimilation et de croissance. L'une des lois fondamentales de la biologie est due à Monod et stipule que le taux de croissance d'une population est une fonction de la concentration du nutriment limitant. Dans les faits, il est observé que les taux d'assimilation et de croissance sont proportionnels, ce qui traduit l'idée que le substrat assimilé est source de biomasse. Il est également acquis que le rapport sucre sur biomasse est légèrement moins bon lorsque le taux de croissance augmente. La cause de ce moindre taux de conversion peut s'expliquer par le besoin d'optimiser le fonctionnement des micro–organismes en cas de défaut de nutriments, tandis qu'une abondance de substrat leur donne toute latitude pour fonctionner selon leur régime métabolique préférentiel (c'est à dire le moins contraignant de leur point de vue). La disponibilité des métabolites impacte à la fois le nombre de processus internes mis en jeu par les bactéries et leur rendement. Bien sûr, ces observations sont issues de conclusions expérimentales portant sur des cultures continues en régime permanent. Lorsqu'elles sont exposées à une modification soudaine de leur environnement, les cellules se comportent autrement qu'elles le feraient à l'équilibre. Ainsi, la vitesse instantanée d'assimilation dans la foulée de la limitation d'un nutriment peut atteindre cinq à dix fois la valeur maximale d'assimilation dudit métabolite présent dans un système en excès : il est donc indispensable de rendre compte d'un équilibre non–local entre les organismes et leur environnement sur des échelles de temps courtes. Bien que ce comportement soit intrinsèque au règne animal, il est trop souvent négligé dans la modélisation de dynamiques de populations biologiques.

Un autre élément incontournable de la dynamique de populations microbiennes est la division cellulaire, à travers laquelle le nombre de cellules augmente exponentiellement jusqu'à épuisement des ressources offertes. Ce mécanisme introduit de l'hétérogénéité au milieu d'une population cellulaire : l'intervalle de temps entre deux événements consécutifs de division cellulaire n'est pas constant d'un individu à l'autre, pas plus que la taille à la naissance ou à la division. Les lois de conservation peuvent porter sur des quantités extensives telles que la masse ou le contenu en enzymes entre générations n et $n+1$. D'un autre côté, la redistribution des grandeurs intensives est plus anarchique et peut amener à des comportements sensiblement différents d'une cellule–mère à ses cellules–filles ainsi qu'entre deux cellules–sœurs. De la sorte, une population peut être distribuée en vitesse d'allongement, taux d'expression des gènes, ...

Modéliser la dynamique d'une population nécessite de choisir les variables dont la distribution reflète au mieux les données extraites de la littérature. Dans ce contexte, non seulement le nombre de candidats potentiels est élevé mais la loi d'évolution de ces propriétés internes relève de l'inconnu. Par conséquent, il est d'une importance capitale de définir précisément les objectifs du modèle pour se concentrer sur un nombre restreint et significatif de dimensions. Du point de vue analytique, la plupart de ces quantités se doivent d'être à support compact, étant entendu qu'un bilan matière dans un réacteur de capacité finie ne peut conduire à une masse ou un contenu enzymatique par individu infinis. Cependant, en analyse de populations microbiennes, le cadre naturel est $\mathbb{R}_+$ tout entier et soulève deux questions pratiques : le poids accordé à des tailles de cellules qui ne peuvent être encontrées au cours d'une expérience et le traitement du bord supérieur du support d'une distribution si l'équation de transport–fragmentation est amenée à être résolue, par exemple, par méthode de volumes finis. Ce dernier point trouve traditionnellement des réponses dans la communauté des mathématiciens, l'une d'elles étant l'imposition d'un bord supérieur artificiel au domaine d'intégration, ce qui donne lieu à des troncatures conservatives ou non–conservatives selon

les attentes de la simulation. C'est ainsi qu'une troncature non–conservative s'avère très efficace pour modéliser un processus de gélation en mécanique des fluides par exemple. En revanche, cette approche n'est pas saine dans le cas de populations de bactéries pour lesquelles la rupture se produit avec une probabilité 1 à une taille finie (voire microscopique), ce qui amène à formuler l'équation de bilan de population sur un compact de $\mathbb{R}^p$ (p désignant le nombre de variables), et non sur l'espace tout entier. L'hypothèse de support borné complique sensiblement l'analyse cependant, de même que résoudre une Équation aux Dérivées Partielles sur un borné est traditionnellement plus problématique que la résoudre sur $\mathbb{R}^p$.

Le travail de modélisation ne porte pas uniquement sur le choix des variables les plus à propos et leur domaine de définition ; il s'agit aussi de mettre en équation leur dynamique, qui dépend non seulement de l'état physiologique des cellules mais aussi de l'offre du milieu de culture. De nombreuses échelles de temps sont impliquées selon la fonction du composé ciblé dans le fonctionnement de la bactérie : une enzyme qui participe du transfert liquide–cellule doit s'adapter bien plus rapidement à des oscillations de concentration qu'un cofacteur impliqué dans la synthèse de fibres protéiques responsables de la croissance de la cellule. Dans le premier cas, le niveau d'activation de l'enzyme peut être supposé "presque" à l'équilibre avec son environnement immédiat. Dans le deuxième cas, il semble plus incertain d'affirmer que le taux d'activité du cofacteur d'intérêt est à chaque instant une fonction algébrique de la quantité de substrat disponible. Compte tenu de ces considérations, un modèle biologique tout–terrain doit rendre compte de l'étendue des échelles de temps impliquées dans l'adaptation d'une cellule à l'offre de son voisinage, sur la base de considérations de microbiologie.

Ainsi, un travail de pointe de modélisation et simulation de bioréacteurs incluant la dynamique de la phase biologique nécessite des connaissances dans les domaines des mathématiques, de la microbiologie et du génie chimique.

Le travail ci–après sera divisé en cinq parties. La première d'entre elles constitue une présentation de références bibliographiques traitant de dynamique de population d'un point de vue mathématique, numérique ou biologique. Les articles d'analyse permettent de se rendre compte de la compréhension actuelle des équations de transport–fragmentation, la production des microbiologistes servant à mettre en place la modélisation la plus sensée et économique permettant de rendre compte de leurs conclusions quant à la réponse dynamique de bactéries exposées à un environnement fluctuant.

Le deuxième chapitre consacre l'analyse mathématique de modèles structurés en taille et en âge. Les deux modèles sont fondamentalement différents au sens où rien ne borne a priori l'âge d'une cellule alors que la masse d'un organisme, entre autres, est finie par hypothèse. Le processus de division étant d'abord lié à la taille des particules, la divergence de la fonction de rupture en une longueur finie est une condition préalable à tout énoncé de résultat d'existence d'un modèle structuré en taille. L'originalité du résultat proposé d'existence et unicité des éléments propres associés à l'équation de transport–fragmentation tient en l'exploitation du caractère contractant de l'opérateur de redistribution générationnelle qui autorise à conclure de la Lipschitz–continuité de la fonction propre de Malthus, celle–là même qui donne sa forme géométrique à la distribution d'équilibre de la population. La distribution en âge possède une expression analytique et renseigne

également sur la distribution en temps d'interdivision d'une population biologique. Un lien rigoureux entre leurs moments respectifs permet de répondre à un débat vieux de 60 ans que les biologistes n'avaient pas réussi à trancher.

Le troisième volet traite uniquement d'aspects de modélisation multi–échelle de réacteurs biologiques. Le travail porte sur le choix de variables permettant de reproduire à moindre coût les expériences standard exposées dans la littérature, ainsi que sur leur domaine de définition et leur évolution temporelle. Deux facteurs influencent le comportement dynamique des microorganismes : leur métabolisme et la disponibilité de leur substrat de prédilection. Concernant le premier aspect, un nombre minimal d'équations biochimiques est intégré à un code de calcul via deux routines, le traitement "algébrique" du modèle étant considérablement plus rapide que son pendant "logique". La prise en compte de l'hydrodynamique dans un réacteur assimilé à un seul compartiment donne lieu à un raffinement original de sous–maille qui évite toute instabilité numérique lors de la simulation de la dynamique de populations en réacteur ouvert.

Le quatrième segment est dédié à la description des outils numériques mis en oeuvre pour résoudre le modèle explicité au chapitre 3. Une méthode lagrangienne, une méthode eulérienne, ainsi qu'un algorithme hybride déterministe/statistique sont confrontés dans le cas test de la mise en régime d'un réacteur ouvert à mêmes taux de dilution et alimentation en substrat à l'entrée du réacteur. La méthode la plus rapide témoigne de différences assez notables avec les données produites par les deux autres codes, ce qui laisse la question du meilleur rapport qualité–prix à l'appréciation de l'utilisateur.

Enfin, cette thèse se conclut sur la simulation d'expériences–type de microbiologie dont les conclusions peuvent être trouvées dans la littérature. Quatre cas test donnent lieu à autant d'interprétations des états transients issus de chacune des perturbations et montrent les limites du modèle exposé au chapitre 3. Seul le comportement est ici évalué, l'adéquation quantitative entre données numériques et mesures de la littérature relevant de la communauté de génie biologique.

Chapter 1

State of the art

1.1 Biological considerations

Here and below, interest will be taken in typical rod–shape cells such as *Escherichia coli*, *Bacillus subtilis*, *Saccharomyces pombe* which physiology has been the topic of a multitude of biology articles. In some cases, experimental observations made on the baker's yeast, *Saccharomyces cervisiae*, are also presented for they reveal microbiological behaviours that seem to be shared by many microbial strains. One constant difficulty is related to the absence of fundamental laws governing biological processes. A second specificity is that individual observations (at the single cell level) are extremely demanding and must be repeated or parallelised in order to gain statistically converged information at the macroscopic (or population) scale. The multiplicity of individual parameters making any cell different from each other makes for another complicating factor.

1.1.1 The cell cycle

It is well–known now that the cell cycle breaks down into 3 to 4 phases: quiescence (traditionally denoted G0), growth (G1), DNA replication and division (M, often preceded by a DNA repair phase called G2). Depending on the organisms under consideration, the cell cycle counts into tens of minutes to hours. The end of the cycle is characterised by the division of the mother–cell into two daughter–cells. The cell cycle duration is often referred as the interdivision time or generation time, that is the time elapsed between two consecutive divisions. On the other hand, the cells' growth is continuous throughout their lifetime, hinting at a significantly shorter time scale in comparison with the interdivision time. The said growth in mass has to be preceded by an uptake of organic compounds from the culture medium and their transformation through different sets of independent biochemical processes, the purpose thereof ranging from building–blocks formation for the membrane to DNA replication to energy supply. It emerges from these general considerations that several characteristic time scales usher the cell cycle dynamics.

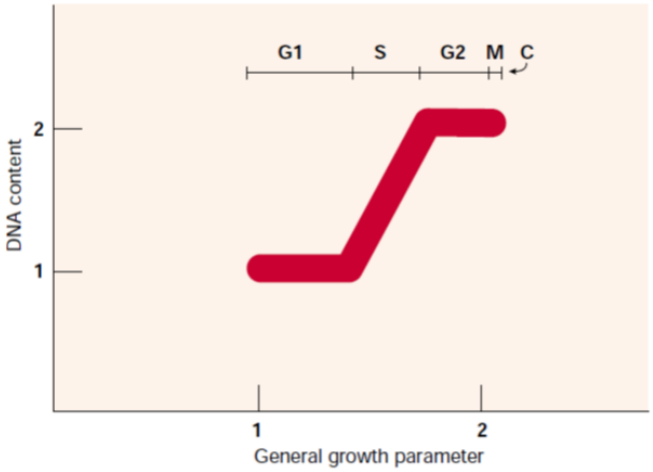

Fig. 3 | Two-parametric histogram of a cell culture growing under steady-state conditions. The cell cycle is composed of a G1 phase (G1), S phase (S), G2 phase (G2), mitosis (M) and cytokinesis (C).

Figure 1.1 – Evolution of a cell's DNA content through the cell cycle. By the time the DNA doubles any other extensive cell parameter will double but the time course of each process may be partly uncorrelated. Borrowed from [7].

According to Boye and Nordstrom, the vision of a cycle is not the most appropriate. It is more relevant to consider that several independent events have to take place before a cell divides and that checkpoints ensure that one particular event can not start before a given set of biochemical process is completed. In standard conditions, each step must be performed once before the cell divides but this does not mean that they are strictly consecutive as suggested from the cyclic representation. Moreover, some cell properties (such as mass) evolve continuously through the cell cycle whereas some others such as length and DNA content may be associated to one particular phase in the cycle. Several processes can take place simultaneously and might overlap. However the "cell cycle" will be maintained for the sake of convenience with reference to the repetitive nature of the entire process, keeping in mind that the duration of each step may vary from one cell to another. Indeed, some proteins and small signalling molecules have been identified as regulators of the cell cycle, however it is not clear whether they only perform as gatekeepers (allowing one particular step to begin) or contribute to the rate of the biochemical reactions. In the specific case of *Escherichia coli*, these authors explain that

> ...guanosine tetraphosphate (ppGpp) and cyclic AMP are involved in the regulation of the cell cycle which provides links between the cell cycle and the general nutritional status of the cell.

It will be shown in the following of this manuscript that the same molecules are also produced when *Escherichia coli* cells are exposed to changes in the nutrient availability. This indicates that when a cell perceives a modification in its environment it produces some compounds that will tend to alter the duration of the cell cycle.

1.1.2 The cells' morphology

It has been known for some time now that a distribution in cell size exists within a population of microbial cells, even under steady–state conditions. This distribution basically results from the fact that cells elongate and divide. A few insights into these two essential features of cell growth and multiplication will be provided hereafter. The cells are assumed cylindrical with constant diameter d and variable length l, so that both their surface πdl, their volume $\pi\frac{d^2}{4}l$ and mass [1] will be linear functions of the cell length.

Cell lengthening

The lengthening rate can take various shapes, but two options have been prominently considered thus far in the literature: linear (see Kubitschek & Friske's [60] for instance) and exponential (notably by Schaechter & al. [124]) growth throughout the cell cycle.

> The mean dry mass accumulation of *E. coli* has been reported as increasing linearly, and cell length growth has been described as linear, bilinear and trilinear, and exponential. The size of the budding yeast *S. cerevisiae* has been observed to increase exponentially by some approaches, but to have a nonexponential and cell cycle–dependent growth curve by others. [36]

According to Godin & al.'s [36] contribution featuring microfluidic experiments to track *E. Coli* cells with respect to time, both patterns are suitable and the behaviour is highly likely to be dictated by the culture conditions and the duration of the experiment (see Figure 1.2). Note that their conclusion arise from a direct measurement of cell buoyant mass $m_c = V_c(\rho_c - \rho_f)$, with V_c the cell volume and ρ_c, ρ_f the respective cell and fluid densities. Recent microscopic observations on the cell shape dynamics of *E. coli* cells indicate that elongation at the cell scale is most likely bilinear rather than exponential. This also points to the importance of the very definition of a cell's length, distance between its two caps or length of its cell membrane [118]. Nobs and Maerkl have reported a constant elongation rate in their study of *S. pombe* (a rod–shaped yeast) involving microfluidic devices and particle tracking techniques [102] (see figure 1.3). Robert & al. [119] have evidenced a sublinear growth pattern from crossing a certain critical length until rupture becomes inevitable. However, it seems that the gain in mass is continuous while the elongation may occur during a limited period of the total interdivision time only.

1. under the assumption of a constant density

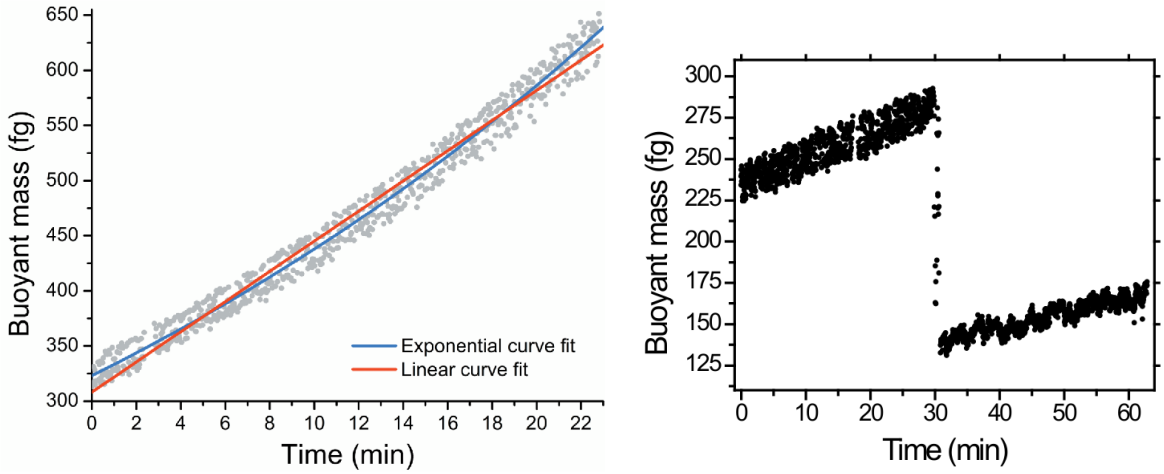

Figure 1.2 – Left: linear vs. exponential growth patterns for *E. Coli* throughout their cell cycle, compared with raw measurements. Over a doubling time, the difference between the two fits appears too tenuous to formulate an unassailable conclusion. Right: further supplementary data tend to point towards a straight line growth pattern. Borrowed from [36].

As a consequence, any modelling of the cell growth is suitable depending on the core experimental dataset and dwell upon these considerations from a quantitative point of view would in essence fuel a long lasting debate in the biophysical community. It is worth noticing that the exponential growth at the population scale does not depend on the growth features at the cell scale for it results from the division into two daughter cells in a finite time.

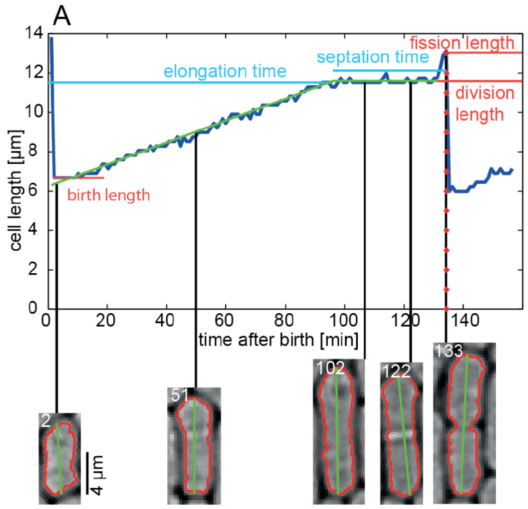

Figure 1.3 – Elongation is linear in time for *S. pombe* cells and makes for 77 % of the total cell–cycle duration. [102].

Cell division

When they divide, all mother–cells give birth to two daughter–cells with conservation of mass. All other quantities are assumed randomly distributed and some experimental data like Tanouchi & al.'s [135] tend to hint at the conclusion that the redistribution of a mother cell's volume has little to no influence on the lengthening rate of the daughter–cells. Several models have been proposed to relate the division frequency to the cell properties. One of them, named "timer model" relates the cell division to the cell age. Robert & al. have ruled out the "timer" initiator for the septation due to its poor matching to experimental data and concluded that the rupture process is in essence a size–related phenomenon. The takeaway from Campos & al.'s contribution [9], is that division is triggered by a length gain from a cell's size at inception, what the biologists have christened "adder mechanism". Taheri–Araghi & al.'s measurements for various strains of *E. Coli* [134] tend to confirm this claim. This conclusion comes out from the fact that a better correlation is found when considering the added length rather than the length itself. However, Taheri–Araghi & al.'s claim is highly unlikely to mesh with a compactly supported length distribution. Convergence towards a stable size distribution was actually obtained under the hypothesis that equal length partition takes place at division. However, should the redistribution of a cell's cytoplasm be unequal (what is an undeniable fact), lineages of ever–growing organisms for generations on end would inevitably manifest themselves. Along with the hypothesis of a linear relationship between individual growth rate and cell mass, extremely large cells would represent an ever growing fraction of the total cell population. To date, there is no real consensus regarding the very reason a cell divides.

Cells do not divide before some specific events have occurred, some of them being correlated with a gain in mass or length. Obviously the gain in mass is somehow related to their ability to uptake nutrients and proceed to their biochemical transformation. It comes from the considerations on cell lengthening and cell division that these two phenomena require an adequate and robust modelling in order to avoid ill–posed problems.

Individual and ensemble observations

Yesteryear out of reach, information at the cell's scale is now accessible through cytometry [15, 16, 43, 74, 128] and single cell culture systems [58, 72, 141, 148]. These new experimental facts confirm that it is unlikely that all individuals of the same species are in the same physiological state (see figures 1.4, 1.5 and 1.6). First of all, not all individuals have the same age, understood as age within the cell cycle, that is, the time elapsed since the last division. Then, two cells with the same age do not necessarily have the same mass or the same composition, and from there, different internal reaction rates may be observed even though they are in the same fluid medium. As a consequence, it is unlikely that individual organisms exhibit the same growth rate. However, a majority of experimental observations presented in the literature deal with averages over a large number of individuals with different ages and compositions.

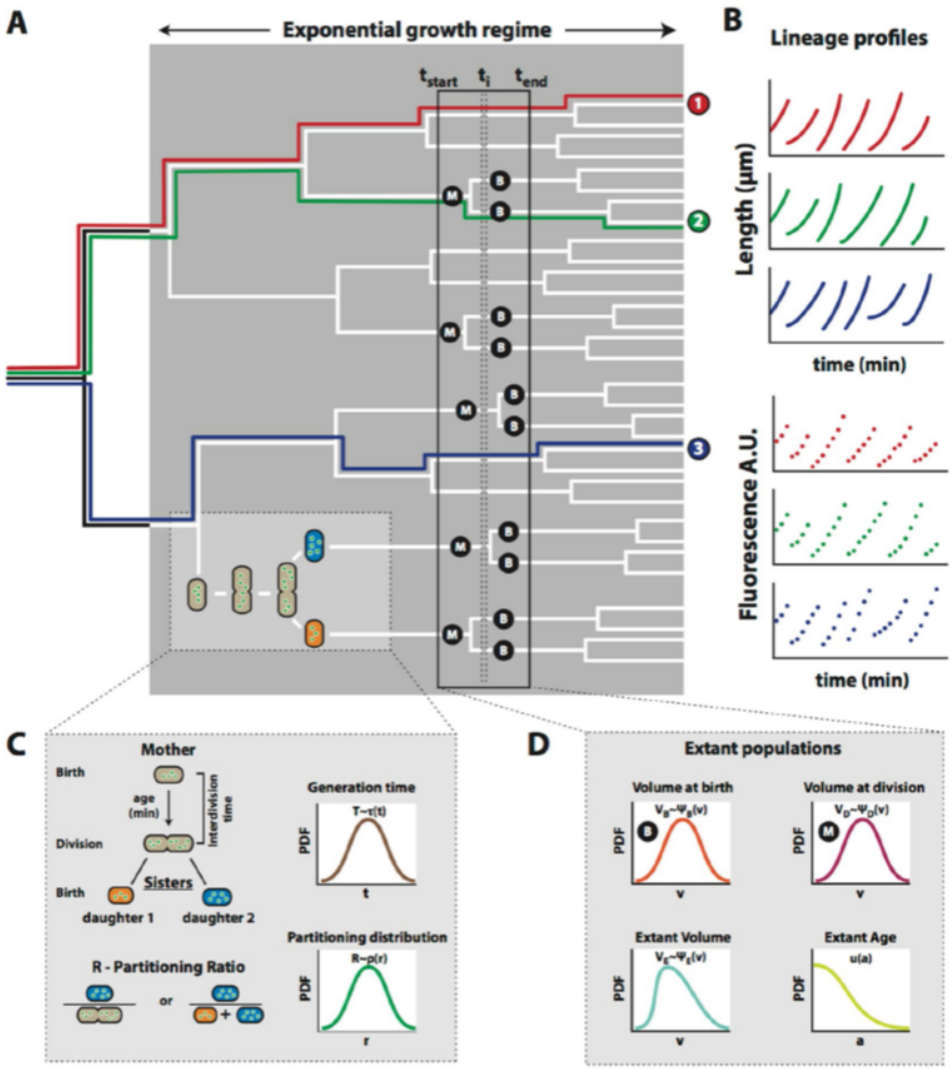

Figure 1.4 – The cell length at division varies from one cell to another amid the same lineage, along with the lengthening rate. The elongation rate is seemingly exponential in some cases or rather linear (B). When a cell divides, the internal content is not equally distributed among the two daughter cells (C). Observations performed over a collection of extant cells in the cultivation reveal a distribution in cell age, volume at birth, and volume at rupture (D) [141].

Cell cultures normally are heterogeneous due to factors such as the cell cycle, inhomogeneous cell microenvironments, and genetic differences. However, distributions of cell properties usually are not taken into account in the characterization of a culture when only population averaged values are measured. [47]

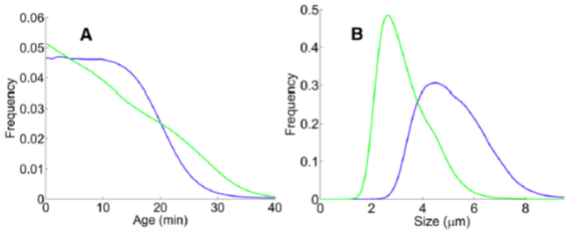

Figure 1.5 – Age (A) and size (B) distribution in two continuous culture of *E. Coli*. Borrowed from [119].

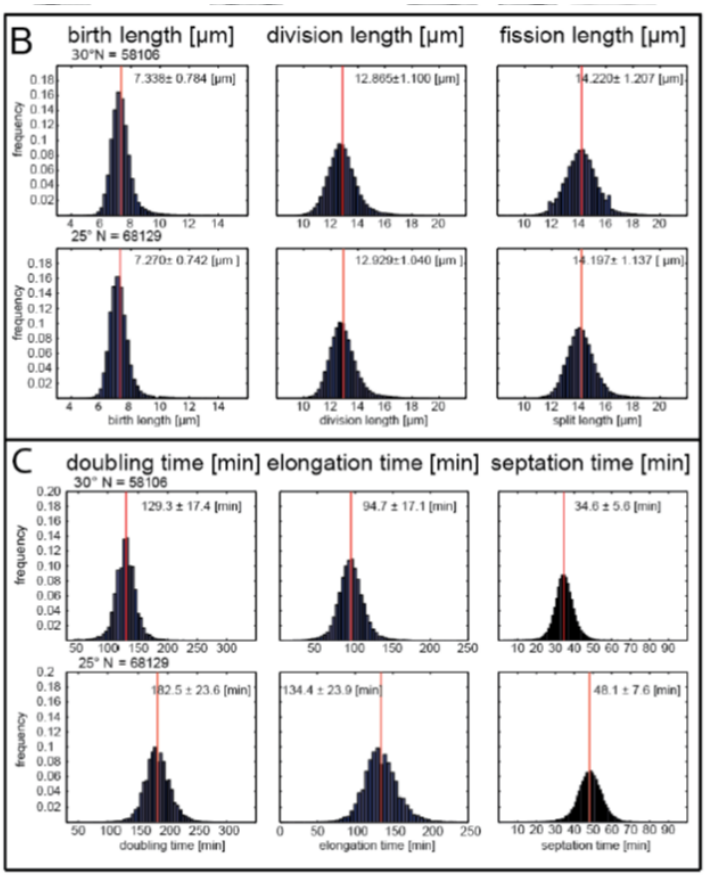

Figure 1.6 – Experimental evidence concerning the distribution in cell length (B) and cell cycle duration (C) in continuous cultures of *Saccharomyces pombe*. In each graph the upper and lower raws correspond to experimental results obtained at 30°C and 25°C respectively. Cell cycle durations (biochemical reaction rates) are impacted by a temperature change while physical properties and cell partitioning at division are not. Borrowed from [102].

From these experimental observations, a natural conclusion is that individuals are all different from

one another. Because of that, the use of the concept of population balance to describe the evolution of a population of living organisms might seem self-evident. Hatzis, Srienc and Fredrickson explain that

> "Although continuum models [i.e. models based on macroscopic properties] have been proven adequate for many practical situations, they do not constitute the natural framework for the description of microbial population phenomena. Fredrickson and co–workers [...] recognized that population models must acknowledge the segregated or corpuscular nature of microbial populations as well as the subcellular structure and composition of individual cells". [39]

In addition to its conceptual appeal to deal with the case of multiphase and heterogeneous systems, the population balance framework allows for a detailed examination of the rationale behind a population's evolution, the characteristic times of these physical and biological phenomena, with the aim of studying the dynamic interactions between a population and its environment. However, the extreme complexity of living systems, the multitude of variables required for describing them, and the difficulties in solving population balance equations will be met. This also raises the question of the trigger for the diversity observed within a population, leading to an investigation of the generational redistribution in intracellular content and the environment's fluctuations amongst others. The question of the mainstay of a cell's dynamic adaptation is still very open and encompasses a spectrum of intricated time scales, the founding principles thereof bearing upon the most constituent laws of fundamental biology. Evidently, a better description of the coupling between mass transfer, energy and biochemical transformation is required.

Definitions of growth:

The macroscopic definition:

Historically, the specific growth rate of a cell culture was defined with reference to its doubling time. This definition assumes that the culture has entered the so–called exponential growth phase, which in mathematical terms means that the distribution of any cell property is self–similar. [132]

$$\begin{aligned} X(t) &= X_0 e^{\mu t} \\ \mu &= \frac{\ln 2}{\tau_d} \end{aligned} \tag{1.1}$$

X being the cell concentration (number or mass per unit volume), X_0 its value at time $t = 0$, μ (time unit $^{-1}$) the population's specific growth rate, τ_d its doubling time. Provided that all necessary resources are available to the growing cells, one can identify a maximum specific growth rate (μ_{max}) at the population scale. Typical values are around $1h^{-1}$ for fast growing bacteria such as *Escherichia coli* and $0.5h^{-1}$ for yeast.

In practice, the population doubling time is estimated from the slope of a curve plotting the logarithm of the cell number (or mass) as a function of time, as shown in Figure 1.7. In the graph regarding the growth of *Bacilus subtilis* a lag–phase, corresponding to the equilibration of the population with its environment is clearly visible before $t = 250min$. Such a behaviour is very common, consitutive of microbial cells and reflects the time necessary for adapting the internal enzymatic reactions [91]. Microscopic observations confirm that the total length increases because of cell division.

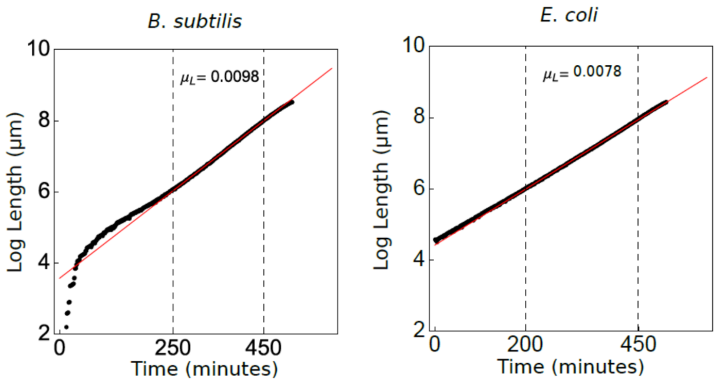

Figure 1.7 – Linear fit of the log transformed total cell length gives access to the population growth rate in the stationary/exponential growth regime [141]. *B. subtilis* and *E. coli* are rod–shaped cells. Under constant mean density hypothesis, the total elongation rate is proportional to the population growth rate in mass.

As illustrated in Figure 1.6, a distribution of interdivision time exists within a population and it is important to recognise that its doubling time is not the first moment of the normalised interdivision time distribution $p(\tau)$:

$$\tau_d \neq \int_0^\infty \tau p(\tau) d\tau \tag{1.2}$$

The reason for that was presented as early as 1956 by Powell [112], and will lead to further remarks later in this document. For the moment it may be sufficient to observe that the fraction of the population having a smaller interdivision time will produce more cells than the other ones in the time interval τ_d.

Cell scale and population scale definitions: It is also possible to propose a definition of a specific growth rate at the individual scale in terms of elongation or mass change, leading to:

$$\begin{aligned} v_l &= \frac{1}{l}\frac{dl}{dt} \\ v_m &= \frac{1}{m}\frac{dm}{dt} \end{aligned} \tag{1.3}$$

From there a population specific growth rate through a summation over all individuals can be defined, provided that the probability density function $p(l)$ is accessible.

$$\mu_l = \int v_l p(l) dl \tag{1.4}$$

1.1.3 The cells' biochemistry

All the cells' physiological processes (lengthening, rupture, ...) are highly energy–consuming, prompting a permanent need for available nutrients in their immediate environment. More than the organisms' length, their capability to uptake the available substrate and convert it into useful

growth–inducing proteins is the main cause for their lengthening, as shown for example by Lambert & Kussell [64]. The medium–cell transfer is the result of two mechanisms operating at drastically different scales: the user–set culture conditions, enforcing a reactor–scale availability of substrate, and the organisms' uptake strategy to extract the best of their vicinity.

Uptake of carbonaceous substrates

In the sequence of events leading to cell multiplication, the uptake of nutrient occurs ahead of cell growth. A specificity of living systems is the regulation of the liquid–cell mass transfer at the expense of energy consumption. This allows them to either select their nutrients or to upper–bound the rate of biochemical reactions, as a function of their needs, through the activation/deactivation of specific and non–specific transport systems. Ferenci's results [23, 25–27] on the assimilation mechanisms of carbonaceous substrates (mainly in *Escherichia coli* bacteria) are reproduced here. An illustration of these mechanisms is presented on figure 1.8.

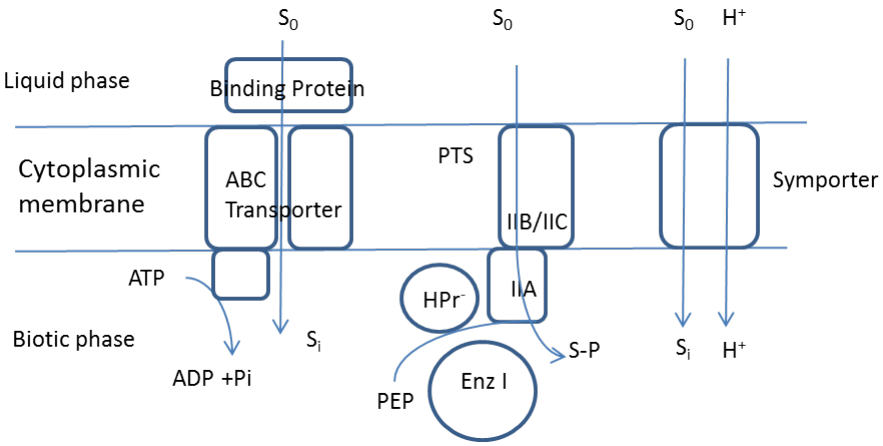

Figure 1.8 – The functioning of three uptake membrane transporters: permease (left), PTS (middle) and symporter (right) (reworked from Ferenci [24])

Three groups of mechanisms responsible for substrate transport have been identified in the case of Gram–negative bacteria:

1. A so–called high affinity system (efficient at low concentrations) drawing its energy from ATP hydrolysis and involving a binding protein located in the periplasm (the external part of the cell membrane), called permease. It is an active transport type system.
2. A PhosphoTransferase System (PTS) with intermediate affinity for the substrate. This system is controlled by the ratio of two compounds (phosphoenolpyruvate and pyruvate). Two types of modelling are proposed for this system:
   a. A Michaelis–Menten law

$$q_S = q_{S,\max}\frac{S}{K_S + S} \tag{1.5}$$

   q_S standing for the specific substrate uptake rate (in $g_S.g_X^{-1}.h^{-1}$), S the substrate concentration in the liquid phase, K_S an apparent constant, named affinity or half–saturation constant and $q_{S,\max}$ an experimental constant defining the maximum uptake rate.

b. A more precise formulation explicitly involves the concentration ratio of two intracellular compounds: (phosphoenolpyruvate / pyruvate), [10].

$$q_S = q_{S,\max} \frac{S \frac{C_{pep}}{C_{pyr}}}{\left(K_1 + (K_2 + S)\frac{C_{pep}}{C_{pyr}} + K_3 S\right)\left(1 + \frac{G_{G6P}}{K_4}\right)} \tag{1.6}$$

K_i being constants, $C_{pep}, C_{pyr}, G_{G6P}$ the concentration of three internal compounds: phosphoenolpyruvate, pyruvate and glucose–6–phosphate, respectively. Without going into the details, a low ratio *pep/pyr* impacts negatively the cell uptake capacity. On the other hand, a ratio of 1 makes this model rather similar to the previous one. The accumulation of glucose–6–phosphate reduces the cell uptake capacity.

3. A low affinity system (effective at high concentrations), called symporter, in which the transport of the substrate is coupled with an ion exchange through the membrane.

It should be noted that not all sugar transport systems are specific; a cell thus acquires the means to assimilate a multitude of carbon compounds [70]. This is especially true when the nutrient is scarce and the cell actively regulates its different uptake systems in order to scavenge nutrient renmants. In case of nutrient excess, the cell actually limits its uptake capacity, most probably because processing too large amounts of nutrients at the same time is detrimental to its fitness. It also appears that these systems work in parallel, with overlapping working ranges, in order to provide the most efficient system depending on the nutrient availability. A closer look into the main pathways which carbon sources are processed through will help understanding the main features of the regulation mechanisms.

A few words on central carbon metabolism

Central carbon metabolism uses a complex series of enzymatic steps to convert sugars into metabolic precursors. These precursors are then used to generate the entire mass of the cell [103]. Two illustrations of the central metabolism in *E. coli* and *S. cervisiae* are provided in figure 1.9 and 1.10. These chemical routes, or pathways, constitute the backbone of the cell metabolism and a minimal description is required to better understand the various modelling approaches used to describe them.

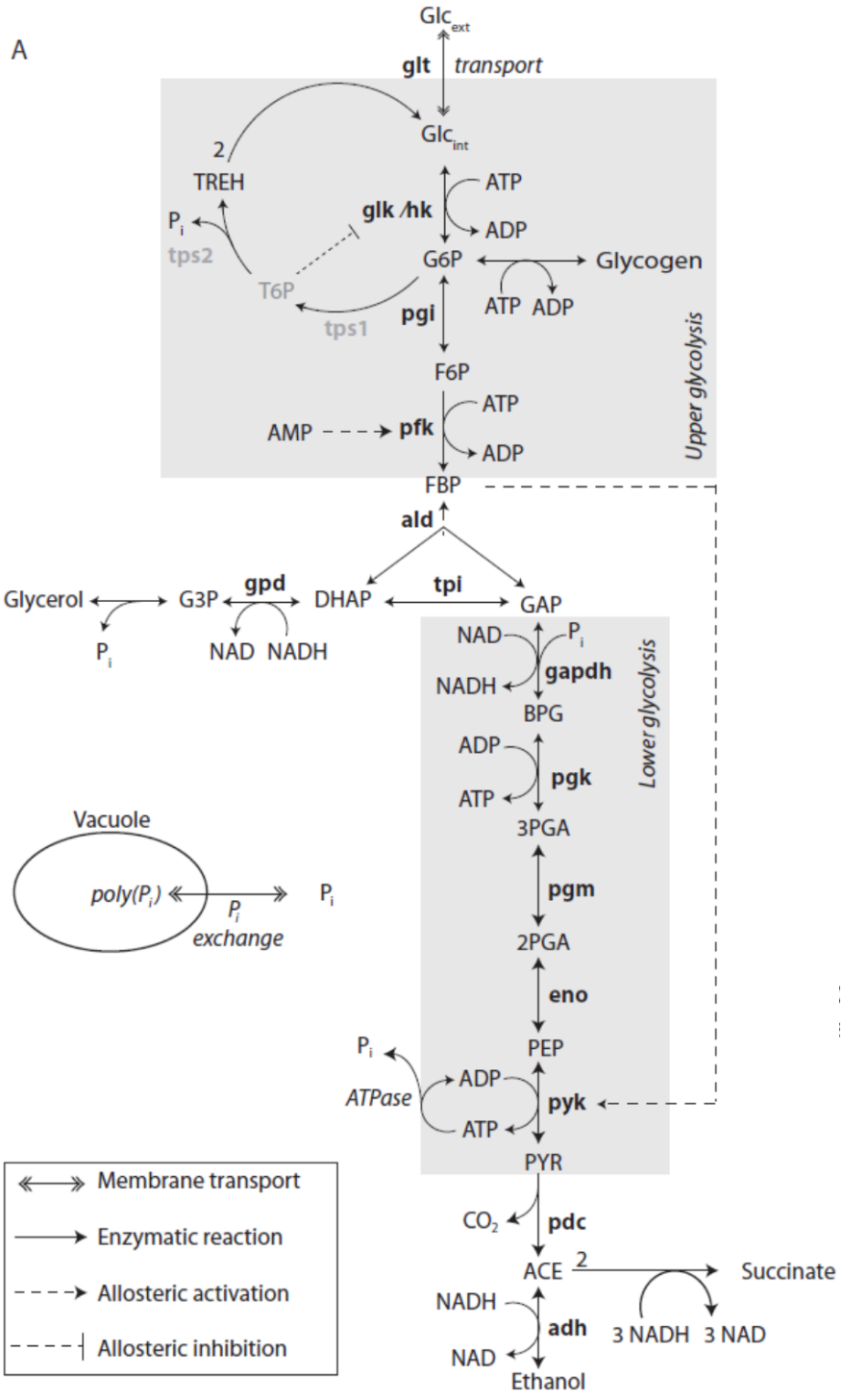

Figure 1.9 – Glycolysis pathway of *S. cerevisiae*. Metabolites feature in capital letters, enzymes in bold font. [142]

In *S. cerevisiae*, the upper and lower glycolysis pathways are connected at the FBP (Fructose–1,6 biphosphate)–GAP (Glyceraldehyde3–Phosphate) node. At equilibrium, the fluxes in the upper and lower parts are equal. A too massive flux in the upper part results in carbon excess to be processed towards glycerol production. This will also create an accumulation of all compounds in the upper glycolysis. Interestingly, the accumulation of FBP promotes an increase in the rate of the reaction converting phosphoenolpyruvate into pyruvate within the lower glycolysis (dashed arrow in Figure 1.9). This can be regarded as a way to improve the carbon processing capacity of the lower glycolysis pathway. In the meantime, the accumulation of G6P (another compound of the upper glycolysis) amplifies the accumulation of intracellular glucose (Glc_{int}) through the inhibition of the enzymes, glk/hk, which in the end negatively impacts the glucose uptake rate. This constitues a negative feed back loop aiming a reducing the flux entering the upper glycolysis if the processing capacity of the lower part is not sufficient. A noticeable point is that the three processes involved in the response to an excessive carbon flux in the upper glycolysis leads to the production of a common compound P_i that is polymerised and stored within the cell.

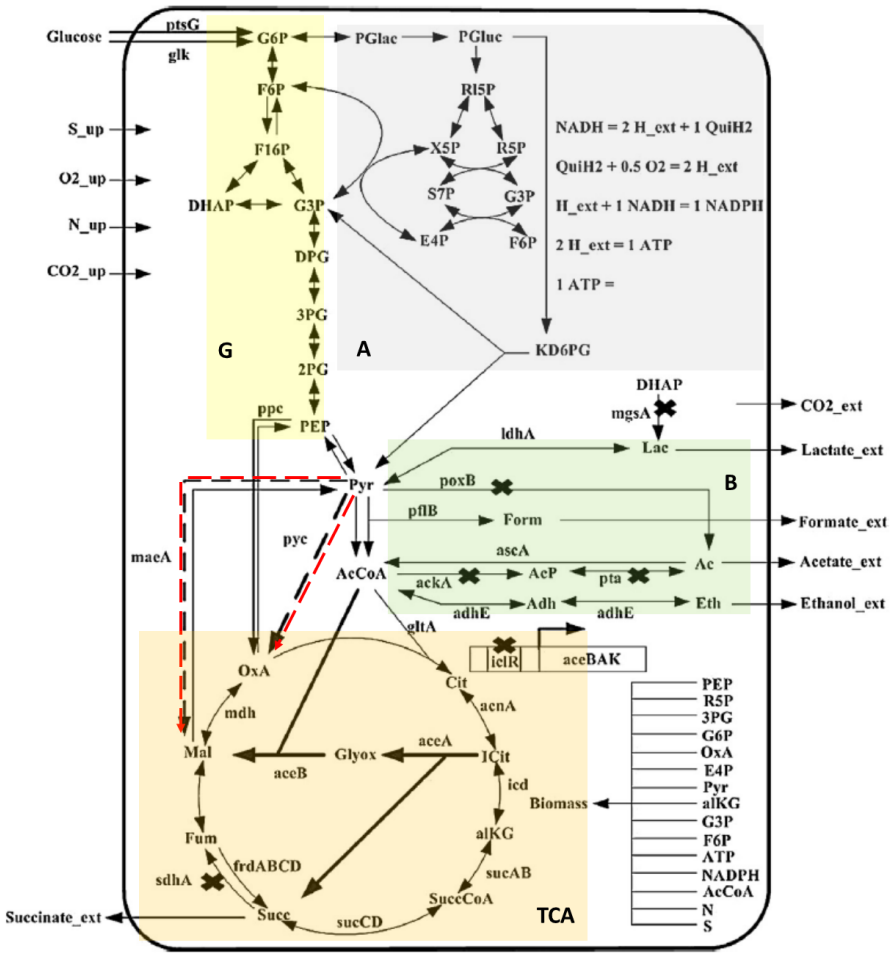

Figure 1.10 – Central carbon metabolism of *E. coli*. The upper and lower glycolysis (G) appear in yellow, the Tricarboxylic Acid Cycle (TCA) in orange, the pentose phosphate pathway (A) in grey and the mixed acid pathway in green (B). Red dashed line arrows between the lower glycolysis and the TCA cycle indicate a regulation mechanism that controls the orientation of carbon fluxes at the pyruvate note. Adapted from [147]

A broader view of the central metabolism of *E. coli* is presented in figure 1.10. The glycolysis looks similar to what has just been described for a yeast, these pathways being assumed universal. As a result, a pyruvate accumulation is similarly responsible for a decrease in the uptake capacity as reported for the yeast, what justifies the model for the glucose uptake rate presented earlier. Now, a focus is made on the connection with two other groups of reactions: the TCA cycle which partakes in the production of many building blocks essential for the cell growth and the mixed acid pathway (B) that manages the consequences of a disequilibrium between carbon fluxes in the glycolysis and TCA cycle. Here it is visible that pyruvate occupies a key position as it connects the three pathways. In this figure, red dashed arrows reveal a positive forward loop: higher amounts of pyruvate activate the TCA cycle. On the other hand, the excess of pyruvate resulting from a disequilibrium between the glycolytic flux and the TCA cycle flux can be diverted in the mixed

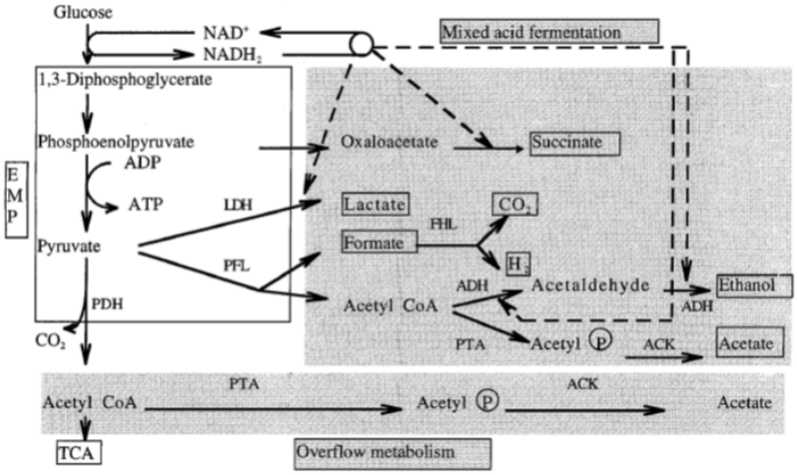

Figure 1.11 – Details of mixed–acid fermentation and overflow metabolic pathways in *E. coli* [146]

acid pathway leading to the production of various acids (lactate, formate and mainly acetate). In the general case of a non modified strain (wild type), acetate re–consumption might take place if the carbon flux from the glycolysis is not sufficient to satisfy the requirements for the TCA cycle ([142]).

Further details on acetate production in *E. coli* are presented in Figure 1.11. This bacterium is able to produce its energy using oxygen (aerobic metabolism). Under aerobic conditions, some acetate may be produced through overflow when acetyl–CoA accumulates as a result of an imbalance between glycolytic and TCA fluxes. *E. coli* is also able to produce its energy through fermentation (without oxygen), a process which final products include acetate, ethanol, succinate, lactate and carbon dioxide. Overall, experimental data hint at the conclusion that, at equilibrium, a fraction of the assimilated sugar turns into biomass at a roughly constant yield. When suddenly unhinged though, the organisms lose in efficiency in comparison and manifest their discomfort by excreting by–products (mainly acetate) in the medium. These less energetic sources of carbon can be reconsumed if S proves in default. All in all, a flurry of biochemical pathways continuously contribute to an apparent growth rate, only a few of them being actually accessible in real time to the experimentalists.

To sum up, the cell uptake can be regarded as the input of the upper glycolysis while the cell growth is the output of the TCA cycle. Some signalling molecules are involved in regulation loops either directly or through a modification of the gene activity which produces enzymes controlling the reaction rates (see figure 1.12). A negative feed–back loop is present in the upper glycolysis and results in a decrease in the uptake capacity if the assimilated carbon flux exceeds the cell needs. In *E. coli*, this relative excess is detected through the accumulation of pyruvate. Within the glycolysis pathway a positive forward loop increases the reaction rate of puyruvate synthesis in the lower part when the upper part faces an excessive carbon flux. Another positive forward loop, starting from pyruvate, controls the reaction rates in the TCA cycle. Owing to these three loops, the bacterium is able to regulate its uptake as a function of its needs:

— Insufficient uptake rate causes a pyruvate depletion which, in turn, reduces the growth capacity. In the meantime, this low level of pyruvate acts as a signal to try and improve the

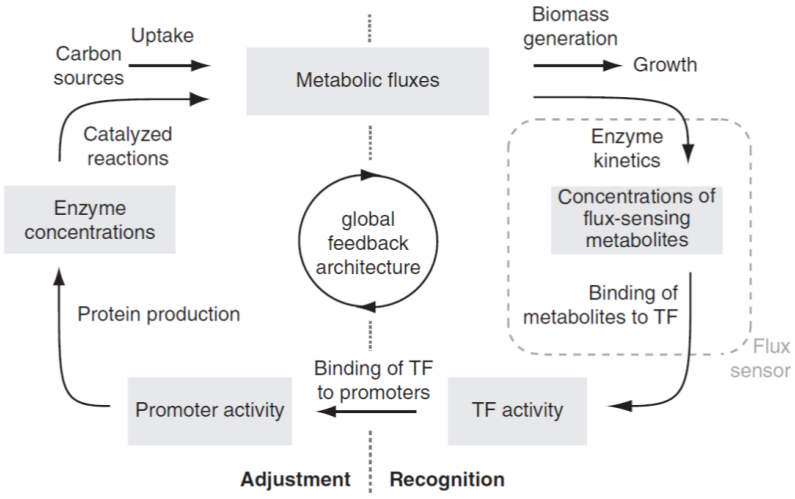

Figure 1.12 – Global feed–back loop architecture connecting growth to uptake through metabolism. [56]

uptake capacity.

— Excessive uptake rate causes a pyruvate accumulation which acts as a signal for increasing the activity in the TCA cycle (higher growth capacity). In the meantime, a too high level of pyruvate will inhibit the uptake capacity.

— Depending on the availability of oxygen, the difference between the mass fluxes processed the glycolysis and the TCA cycle, is diverted into by products through mixed acid pathways or overflow metabolism.

Many examples of regulation loops can be found in the microbiological literature. In general, multiple loops act in a coordinated way to ensure some stability at the cell level (homeostasis) and therefore imply a limited number of key compounds (such as pyruvate). However, interconnected loops can also play a filtering or switching role in the cell's machinery. Brandman explains that linking fast and slow positive feedback loops creates a "dual-time" switch that is both rapidly inducible and resistant to noise in the upstream signaling system [8]. In the end, one may keep in mind the illustration presented in figure 1.12 which shows that at the global level, the metabolic activity is regulated in order to adapt the biochemical reaction rates to the incoming nutrient fluxes through the modulation of enzymes concentrations.

1.2 Modelling Bioreactions

The fact that some reactions might be turned on and off whilst their rate constant is depending on the concentration of an enzyme makes the description of biological reactions extremely complex. The sophistication of bioreaction models has evolved in parallel with the experimental techniques, from black box model based on macroscopic conversion of nutrients into new cells up to genome–scale models encompassing all levels of internal biochemistry (genes, transcription factors, proteins, metabolites).

1.2.1 Modelling a continuous bioreactor using an unstructured kinetic model

Biological populations dynamics in a fermenter are traditionally considered in a close (batch) or open (chemostat) system. In a continuous culture, some substrate and cells are drained out of the tank at the same flow rate as the inlet feed to keep the liquid volume constant. At the inlet, some substrate is poured into the reactor and mixed in order that nutrients are available to the cells.

On the face of it, the simplest model consists in writing mass balances for both biomass and substrate leading to a system of ordinary differential equations:

$$\begin{aligned} \frac{dS}{dt} &= D(S_f - S) - q_S X \\ \frac{dX}{dt} &= (\mu - D)X \end{aligned} \tag{1.7}$$

S ($g_S \cdot L^{-1}$) standing for the substrate (usually glucose) concentration and X ($g_X \cdot L^{-1}$) for the biological phase concentration. S_f ($g_S \cdot L^{-1} \cdot h^{-1}$) stands for the substrate concentration in the feed and D (h^{-1}) for the flow–to–volume or dilution rate that ushers the renewal of both the biotic and the abiotic phase. μ ($g_X \cdot g_X^{-1} \cdot h^{-1}$) is the population's "growth rate" in mass calculated at the reactor scale, that filters all the underlying mesoscopic/microscopic phenomena. At the scale this modelling is aimed at, it must be assumed that growth in mass is tantamount to growth in cell number. q_S ($g_S \cdot g_X^{-1} \cdot h^{-1}$) is the global uptake rate once again at the macroscopic scale. The standard modelling thereof is engineered with the idea of fitting steady–state experimental data.

The analysis of a collection of steady–state experimental observations leads to the following conclusions:

i. A different couple, $\{X, S\}_D$, is found depending on the dilution rate
ii. The growth rate equals the dilution rate
iii. The uptake rate is proportional to the growth rate
iv. A simple relationship can be established between the growth rate and the residual sugar concentration, a customary fit consisting in a Monod law featuring strain–dependent saturation μ_{max} ($g_S \cdot g_X^{-1} \cdot h^{-1}$) and affinity K_S constants.

These observations have been translated into mathematical expressions such as :

$$\begin{aligned} \mu &= D \\ \mu &= \frac{q_S}{Y_{SX}} \\ \mu &= \mu_{max} \frac{S}{K_S + S} \end{aligned} \tag{1.8}$$

where Y_{SX} ($g_S \cdot g_X^{-1}$) is a reactor–scale substrate–to–cell ratio, once again dreamed up from steady–state mass balances.

$$Y_{SX} = \frac{S_f - S}{X} \tag{1.9}$$

However, in equation (1.8), writing $\overline{\langle \mu \rangle}_D$ rather than μ would be more appropriate since the identified growth rate is actually an average growth rate at the population scale (hence the $\langle\rangle$ notation), observed at steady state (hence the bar over the ensemble average), for one particular value of the dilution rate (hence the subscript D). Similarly, q_S and Y_{SX} in equation (1.8) and (1.9) are also experimental, population averaged, steady–state quantities.

Unfortunately, these subtleties have not been considered with enough care, or misunderstood, and the set of relationships (1.8) and (1.9) has been routinely used as if they were fundamental laws describing the macroscopic dynamics of a cell population. The system's equilibrium depicts a manifold indexed by the dilution rate D and the following relationships strictly hold only if $S = S_D$, $X = X_D$ the respective substrate and biomass concentrations at steady state:

$$\begin{aligned} \mu_{max}\frac{S_D}{K_S + S_D} = D \Leftrightarrow S_D &= \frac{DK_S}{\mu_{max} - D} \\ D(S_f - S_D) = Y_{SX}DX_D \Leftrightarrow X_D &= \frac{1}{Y_{SX}}\left(S_f - \frac{DK_S}{\mu_{max} - D}\right) \end{aligned} \tag{1.10}$$

A rigorous examination of the facts indicates that these are only correlations between experimental data identified at steady state. Whether the steady–state (1.8) framework holds at transient state is debatable, the Monod equation implying that the biological response to an ever–fluctuating medium is instantaneous and solely dictated by the said medium.

Similarly, it also emerges from that approximation that the cell number is completely determined by the total biomass in the reactor, neglecting that a cell's mass or volume is significantly impacted by its history or its age in the cell cycle. Even if a constant cell number per gram seems, on average, a reasonable assumption at steady state, any perturbation like a substrate injection in the fermenter shall have progressive, noticeable repercussions on the population's behaviour, due to the nature of the cell cycle. Indeed, the cell cycle operates at a time scale that cannot compare, for example, with the substrate uptake or the lengthening throughout the cell's lifetime. Hence, the system 1.7 and its closure laws, Y_{SX} and $\mu = f(S)$, that filters too much information below the macroscopic threshold is not suited to address the intrication of processes which manifest when experiences are performed.

1.2.2 Unstructured kinetic models

Unstructured models describe biological transformations using constant conversion yields and a description of reaction rates based on the concentration of nutrients in the liquid phase.

$$\begin{aligned} r_X &= \mu(S)X \\ r_S &= -\frac{1}{Y_{SX}}r_X \\ r_P &= \frac{1}{Y_{SP}}r_S \end{aligned} \tag{1.11}$$

X is the cell concentration (in general mass per unit volume), S the concentration of the carbon source and P a product of the cell metabolism. Y_{ij} is a constant conversion yield in grams of j produced par gram of i consumed (<1) and μ is the population specific growth rate (as defined in equation 1.1, page 8). No information regarding the intracellular bioreactions is present in this type of model and the biological phase is thus treated as a black box which converts nutrients into new cells and products. The relationship between the growth rate and the nutrient concentration, as well as the conversion yields, must be regarded as correlations issuing from a collection of experimental data collected in batch or chemostat culture at steady state. *Caution must be paid to the fact that the conversion yields, Y_{ij}, do not correspond to real stoichiometric coefficients.* One of the most

famous example is the so–called Monod law :

$$\mu = \mu_{max} \frac{S}{K_S + S} \frac{O_2}{K_O + O_2} \tag{1.12}$$

where μ is the growth rate and has a dimension of a time inverse, K_S is named half–saturation or affinity constant and μ_{max} is the maximum growth rate. The product of two hyperbolic functions expresses the fact that both sugar S and oxygen O_2 are required for growth. In practice however none of the parameters are actually constant and their value vary from one study to another because they are essentially fitting parameters and also because they are impacted by any non–modelled aspect of the intracellular functioning.

This crude description is not applicable to the study of detailed or transient microbial response, it is limited to steady–state or pseudo steady–state studies. However this type of model, because of its simplicity and similarity with classical description of chemical reactions, is widely used.

1.2.3 Structured kinetic models

Structured models incorporate some information related to the biological phase. This information may involve physical properties, composition, rate of intracellular processes and physiology. Thus, it becomes possible to distinguish between individuals with respect to their internal properties. We will first think in terms of non–segregated structured kinetic models: same set of dynamic internal properties for all individuals, also named *average cell approach*.

From a fundamental point of view, these models consist in writing mass balances over the biotic phase incorporating intracellular reactions for most metabolites and additional uptake and excretion fluxes for a few of them actually involved in the mass transfer between the cell and its medium. As one might expect, the intracellular concentrations are coupled through non–linear equations because of the numerous consecutive, competitive or cyclic reaction schemes. Such models are formulated in terms of a set of ordinary differential equations describing the evolution of the mass of internal metabolites. It has been known for sometimes now that non–linear dynamic systems can exhibit complex dynamic behaviours such as steady–state multiplicity, bifurcations or sustained oscillations.

Owing to this approach, the rate of the biochemical transformations now depends on both the external conditions and the cell's physiological state. Indeed, the definition of inter–facial mass fluxes typically involves both extracellular and intra cellular concentrations [2, 42]. Also, the list of all intracellular concentrations can be considered as a vector of internal state variables which define the *physiological state* of the cell. This type of model can be used to describe the average concentration representing the behaviour of an *average cell*. It can also be used as a single cell model along with cell ensemble approach to get a population–scale description. However, Henson showed that the long term asymptotic behavior leads to a identical state vector for all cells. The convergence in time is related to the initial distribution and the time constants of the kinetic model. In order to produce a real population model, it was proposed to sample the maximum reaction rate (k_j in figure 1.13 from a gaussian distribution. Thus, all cells share the same set of reactions but with slightly different rate constants.

The accuracy in the cell's dynamics description increases with the number of internal reactions but so does also the dimension of the problem and the number of kinetic laws and parameters to set or identify. Indeed, in practice it is impossible to describe comprehensively the internal biochemistry through a set of elementary reactions with known stoichiometry. Alternatively, some equivalent reactions are build up considering an ensemble of elementary processes: energy production, building

block synthesis, etc. A simple example is illustrated in figure 1.13: uptake, energy production though oxidative or fermentation pathways (catabolism), ethanol production (P_2) and new cell synthesis (anabolism) are described in terms of five chemical transformations. Note that equation is not chemically balanced and the α, β, γ coefficient incorporate some experimental knowledge like the ATP production for the glucose oxidation.

$s \rightarrow S$	$r_s = k_s \dfrac{s}{s + K_s}$
$S \rightarrow \beta_{11} P_1 + \gamma_{11} X_{ATP}$	$r_1 = r_s - \alpha_{41} r_4$
$P_1 + \alpha_{22} O_2 \rightarrow \beta_{23} CO_2 + \gamma_{21} X_{ATP}$	$r_2 = k_2 \dfrac{P_1}{P_1 + K_2}$
$P_1 \rightarrow \beta_{32} P_2 + \beta_{33} CO_2$	$r_3 = k_3 \dfrac{P_1}{P_1 + K_3}$
$\alpha_{41} S + \beta_{41} P_1 + \gamma_{41} X_{ATP} \rightarrow X$	$r_4 = k_4 X_{ATP}$

Figure 1.13 – The kinetic structured model for aerobic growth of *Saccharomyces cerevisiae*. The first equation describes sugar uptake, the second and third are energy producing reactions, the fourth one yields ethanol (the product P_2) and the last one defines the biomass growth rate [2].

In comparison to the unstructured approach, structured kinetic models are superior insofar as they will produce different apparent conversion yields $Y_{i,j}$ depending on the actual fate of the substrate within the cell various pathways. However, the usually large number of internal variables makes it difficult to use them in conjunction with transport equations for the fluid phase. Also, the algebraic relationship between external and internal sugar concentration which is often used in these models is questionable.

1.2.4 Metabolic models

The general idea of metabolic models is to avoid the calculation of metabolite concentrations and estimate directly the mass fluxes in the different pathways of a metabolic network. The unknowns of the problem are the reaction rates which obey a linear system of equations issuing from mass balances. This system results from the underlying hypothesis that no accumulation takes place for most metabolites which are only intermediates. The dimension of this system is generally high because a large number of reactions are considered and the system itself is under–determined. Its resolution involves an optimisation step along with the definition of objectives (generally the maximization of growth rates) and constraints (upper and lower bounds for the mass fluxes).

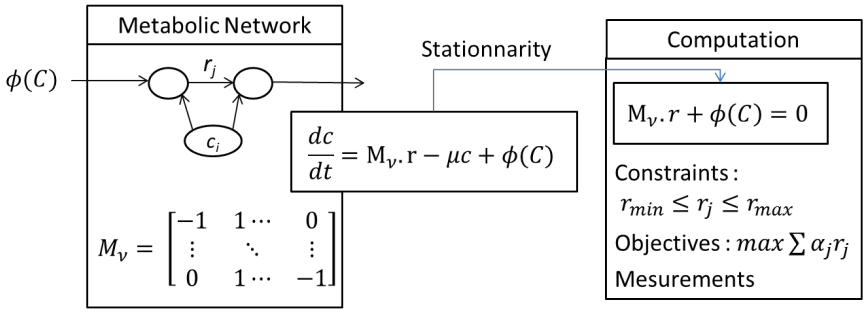

Figure 1.14 – Principle of a metabolic model based on stoichiometry. At the growth scale (over a long period of time), the hypothesis of non–accumulation (or stationarity of the composition) leads to a linear system of equations dealing with rates. Generally under–determined, this system is solved via an under–constrained optimisation procedure. Reproduced from Llaneras and Pico [73].

The level of details is very high compared with lumped kinetic models but this non–accumulation hypothesis seriously limits the ability of such model to describe the cell dynamics:

> The key distinction between conventional chemical reaction systems and metabolic networks, which is often missing in kinetic metabolic network models, is the influence of regulation and control. In conventional chemical reaction systems, knowledge of the kinetics completes the treatment of the system. In biological systems, however, all levels of metabolic function (i.e. transcription, translation, and catalytic activity) are tightly integrated and coordinated with the global environment of the organism, hence yielding adaptability in the face of changing conditions. Thus, the 'conventional' mathematical treatment of a metabolic network, encompassing only the kinetics and stoichiometry, is often hard pressed to correctly predict system adaptation because it lacks a description of the forces driving the adaptation. This capability, however, is exactly what is required for the reengineering of physiology [143]

In order to tackle these issues (under–determination and adaptability), Pigou and Morchain [111] proposed to a metabolic mode–structured model. The cell metabolism is described in terms of modes, each one implying several chemical reactions and being balanced mass– and energy–wise. Beside the definition of the metabolic modes, the cell state is defined by its maximum potential growth rate μ^p. The knowledge of that rate lowers the level of indetermination of the metabolic model that can now be solved using an algebraic procedure. The dynamic is obtained through an equation describing the adaptation of the cell growth capacity:

$$\frac{d\mu^p}{dt} = \frac{1}{T_{\mu_c}} \left(\mu^\star(\mathbf{C_L}) - \mu^p \right) \tag{1.13}$$

T_{μ^p} being a time constant for growth adaptation and $\mu^\star(\mathbf{C_L})$ the mean growth rate of a population being at equilibrium with its environment defined by $\mathbf{C_L}$, the vector of liquid phase–concentrations (Equation 1.12 is an example).

This approach was used in combination with a hydrodynamic model to study the effects of large scale heterogeneities on the metabolic response of a cell population [110]. The whole metabolic rate calculation depends on the calculation of the uptake rates $\phi(C)$ in figure 1.14. In most published

works, it is assumed that the sugar uptake rate of each cell is equal to the population–averaged uptake at equilibrium

$$q_S = q_{S,\max} \frac{S}{K_S + S} \tag{1.14}$$

However, the fate of that sugar is different because each cell requirements for growth is different. In a population, depending on their potential growth rate, some cells may *percieve* the environment as limiting while some other will *percieve* that nutrients are present in excess (relative to their needs).

The advantages of this approach are:

— Cell dynamics is considered through the adaptation of each cell maximum potential growth rate.

— The effective growth rate is either limited by the cell potential or by the external mass transfer of nutrients [95].

— The difference between the amount of uptaken sugar and the maximum sugar consumption for growth provides a quantification of the disequilibrium between a cell and its environment. It allows the calculation of the rate of side–reactions and serves as the driving force for the growth rate's adaption.

The drawbacks of this approach are:

— It is difficult to justify how each cell can sense the average growth rate of the entire population.

— The uptake rate is identical for all cells and calculated from the concentrations in the environment from population–scale information.

1.3 Dynamic responses of a cell population

In this section are presented some experimental results illustrating the essential features regarding the response of a microbial cell population to some environmental changes. The intrinsic complexity of the biological system is such that the responses may conflate a multitude of various forms. A comprehensive review of these multiple responses is virtually out of reach and the talk will therefore focus on the time characteristics of a few phenomena that require a particular interest in the context of bioreactor modelling.

1.3.1 Growth rate

Before getting into the details, it is worth recalling the words of Jacques Monod inspired by Hinshelwood, regarding the question of bacterial adaptation. Early works on bacterial growth focused on the so–called lag phase, a period of time during which the population growth rate progressively increases before it reaches a stable value revealing the stationary growth phase.

> A broader approach to the problem of relations between lag and enzymatic adaptation should also be considered. As emphasized by Hinshelwood, the lag and acceleration phases represent essentially a process of equilibration, the functioning of a regulatory mechanism, by virtue of which certain enzyme balance inside the cell is attained. That such a mechanism must exist is obvious, since in its absence, the cell would not survive even slight variations of the external environment. [91]

Therefore, the stationary growth phase should rather be regarded as the state of equilibrium of a dynamic system that manifests at the cell population scale. Even though an algebaric relationship is found between the population growth rate and the residual substrate concentration in a chemostat

culture at steady state, there is no reason why this equilibrium law should apply at the cell scale or under transient conditions. This logical reasoning is confirmed by the theoretical work of Perret [107] and evidenced through many experimental observations [38, 48].

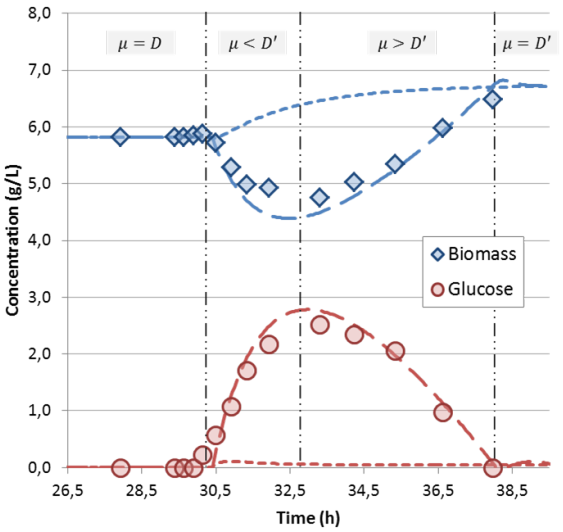

Figure 1.15 – Response to a sudden increase of the dilution rate in a chemostat culture, 0.1 to 0.42 h^{-1}. Comparison of experimental data along with unstructured (dotted lines) and structured kinetic models (dashed lines). Experimental data from [48], simulation data from [110].

A dilution rate step–up leads to a sudden increase in the nutrient availability and a reduction of the average residence time in the reactor. On the one hand more resources are available (and the sugar concentration generally increases from a few mg/L to several g/L) but on the other hand the slower growing cells are washed out and the overall cell mass in the reactor decreases for some time before the cells adapt their growth rate and recover. It is clear that in the transient phase, the population's growth rate is first lower then higher than the (new) dilution rate. The usual unstructured approach that relates algebraically the growth rate to the sugar concentration predicts a rapid gain in biomass concentration as well as a practically constant sugar concentration. These predictions do not match with experimental observations. In contrast, assuming some inertia in the growth rate adaptation (as proposed in equation 1.13) allows a much better description of the transient response.

Another interesting example is provided for the yeast *Saccharomyces cerevisiae* by the experimental data from Guillou & al. regarding a step–up in dilution rate [38]. The step increase of the dilution rate is rather moderate entailing a sugar concentration increase from 10 to 60 mg/L) and in front of this limited perturbation, cells mobilise their internal storage. The authors explain that mobilisation of storage carbohydrates participates in the progression of the cell cycle by providing a surplus of ATP (energy) required at the bud emergence. Indeed, the fraction of budding cells increases significantly after the dilution shift. Observing a higher oxygen consumption and carbon dioxide production, they concluded that the consumption of internal storage provides an apparent increased capacity to convert the glucose surplus into biomass by increasing the carbon flow into respiration. It is important to note that the dilution shift is moderate, leading to a three fold

increase in the glucose influx which does not result in a saturation for the glycolytic pathway.

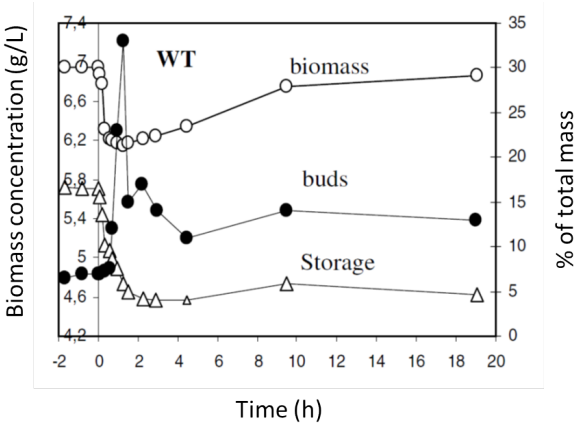

Figure 1.16 – Response to a sudden increase of the dilution rate, 0.05 to 0.15 h^{-1}, in a chemostat culture. The transient decrease in biomass concentration coincides with the consumption of internal carbohydrate storage accompanied by an increase in the number of new cells as reflected by the budding index. Experimental data from [38]

To sum up, the magnitude of the dilution upshift is critical in the analysis of the cell response. A moderate increase of the glucose influx can be damped through a fast mobilisation of internal storage allowing rapid energy production whereas a larger step–up leads to the saturation of the cell capacity to process glucose that accumulates significantly in the liquid phase.

1.3.2 Uptake rate

Basically, the uptake rate should be regarded as the mass transfer rate between the liquid phase and the biotic phase. As already stated, this mass transfer does not depend on the thermodynamic properties and concentration solely. In microbial systems, the mass transfer is regulated by the cell itself at the expense of energy consumption. However, whatever the cell uptake capacity, the actual uptake rate can be limited by the rate of external transport towards the cell membrane.

Uptake dynamics with finite resources

We consider here the experiments facts reported by Notley–McRobb & al. [104] regarding the adaptation of sugar transport systems in batch culture. In this system, cells are cultivated in a closed reactor and consume an initial amount of sugar until its depletion. The rate of change in the sugar concentration results from the uptake by the microbial cells only. In the experiments, these authors tracked the induction of genes, measured the uptake rates and the production of some internal metabolites such a cAMP inside and outside the cell. Their results are presented in figure 1.17 for two different *E. coli* strains. Sugar and biomass concentrations are visible on the upper graphs, the intracellular and extracellular level of cAMP on the lower graphs. It is remarkable that cAMP production takes place well before the glucose exhaustion ($0.3mMor56mg/L$) and is accompanied by a massive increase in the cell potential uptake rate (middle graph, left column). The measure of the potential uptake rate was performed independently, via cell sampling and exposure to

a glucose rich medium. A very surprising and questioning point is that the threshold concentration that triggers the activation of the alternate transport system, about $50mg.L^1$, is far above the usual value of the K_S constant of the PTS system. In other words, the PTS system is still fully active, capable of uptaking sugar at its maximum rate, and yet the cell perceives that it is time to activate scavenging uptake capacities. Note also that these extra capacities are not exploited right away since no change in the slope of the glucose concentration curve is visible. It can be remembered that cAMP was identified as a regulator of the cell cycle.

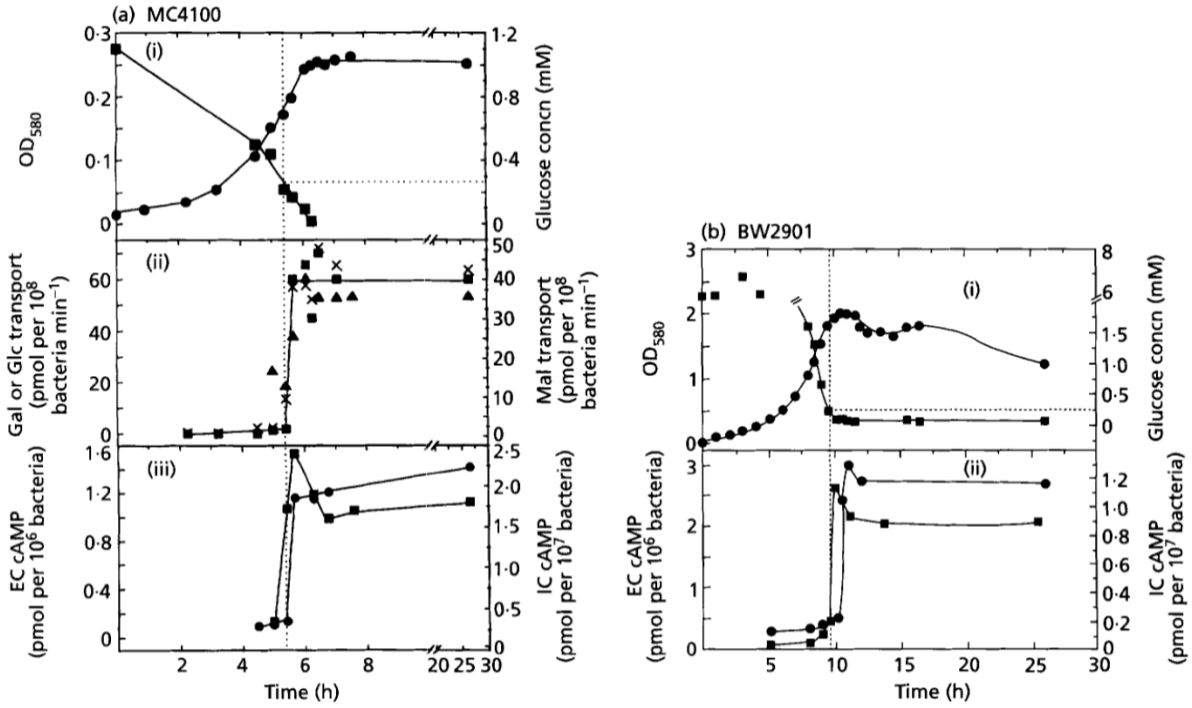

Figure 1.17 – A massive spike in cAMP goes with a slump in extracellular substrate, hinting at a change in the cells' uptake strategy (borrowed from Notley–McRobb & al. [104]). Indeed, as mentioned in [104], "cAMP contributes to the improved scavenging ability of bacteria growing on micromolar concentrations of glucose".

Scale-down experiments

In the field of biochemical engineering, one of the most popular experiment for the study of the cell response to spatio–temporal fluctuations is the so–called scale–down system consisting in a two–stage bioreactor. In order to mimic the concentration changes as experienced by microbial cells in large scale bioreactors, George and his co–workers imagined a loop made of two bioreactors [35].

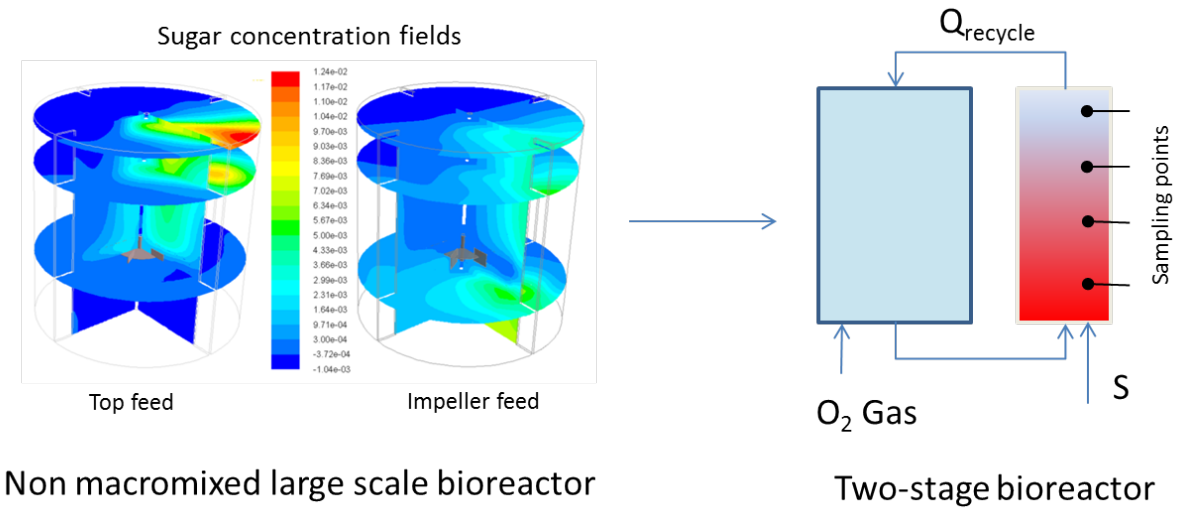

Figure 1.18 – Two–stage bioreactor: the scale down experiment to simulate large scale gradients in heterogeneous bioreactors. Highly concentrated zones appear in red, depletion zones in dark blue.

In the first one, the volume thereof benig very large compared to the second, the sugar concentration is low (see figure 1.18 right). The carbonaceous substrate is fed at the inlet of the second reactor, generally a tubular reactor equipped with static mixers for intense mixing between the cell suspension and the concentrated sugar feed. Hence a sudden increase in the sugar concentration is experienced by a cell which enters the tubular reactor. The residence time in this reactor is relatively small (around 2 minutes). As cells are sent back to the first reactor, the sugar concentration suddenly turns limiting–low due to dilution. As a result, this type of reactor produces a bimodal concentration distribution that resembles that observed in large scale bioreactors. As they circulate in the loop, cells experience high concentration events of two minutes each interrupted by starvation periods which duration goes with the exponential residence time distribution in the large bioreactor. Although this system was originally designed to study the consequences of imperfect macromixing, it was further improved to investigate the glucose uptake rate dynamics through the addition of sampling ports in the tubular reactor. Indeed, Neubauer, Häggström and Enfors used this reactor to quantify the glucose consumption of *E. coli* cells in the wake of a sudden increase in its concentration [99]. The main experimental results are visible in Figure 1.19 where the concentration profiles for glucose, acetate and formate along the tubular reactor at different instants of the fed–batch culture are presented. The cultivation started in batch mode (only the first reactor is used) and switched to fed–batch mode through opening the loop and feeding in the tubular reactor. In the present case, a constant sugar feeding rate was used. Each column of the figure corresponds to successive observations performed shortly and then every two hours after the switch. On the top right corner of each graph, is reported the growth rate and the cell concentration (in $g.L^{-1}$). The growth rate decreases with respect to time because the feed rate is constant while the number of cell keeps increasing. To end up with the description of these graphs, the right column corresponds to experiments where oxygen enriched air was injected in the tubular reactor in order to prevent oxygen exhaustion. The red dashed lines indicate the expected concentration profile assuming the

steady–state uptake law apply.

$$q_S = q_{S,\max}\frac{S}{K_S + S} \tag{1.15}$$

It is remarkable that the instantaneous uptake rate in the 25 seconds (0.4 minute) following the glucose shoot is in fact much larger than the observed maximum value when cells are continuously exposed to high concentration (in a batch culture, for example). This initial overshoot characterised by $q_S(t) > q_{S,\max}$ is followed by a relapse period where $q_S(t) \approx 0 < q_{S,\max}$. Analysing the existing literature at that time, the authors postulated that such a behavior could be explained by the existence of multiple transport systems. The repeated exposure to starvation periods would activate additional transporters responsible for the over–assimilation, above the observed value when only one type of transporter is active.

The temporal resolution of the experimental sampling system has to be sufficiently high in order to actually detect the two-step uptake process: the final glucose concentration at the outlet of the tubular reactor (after 2 minutes) is quite similar to the prediction that could be obtained using the steady–state uptake law (red dashed line).

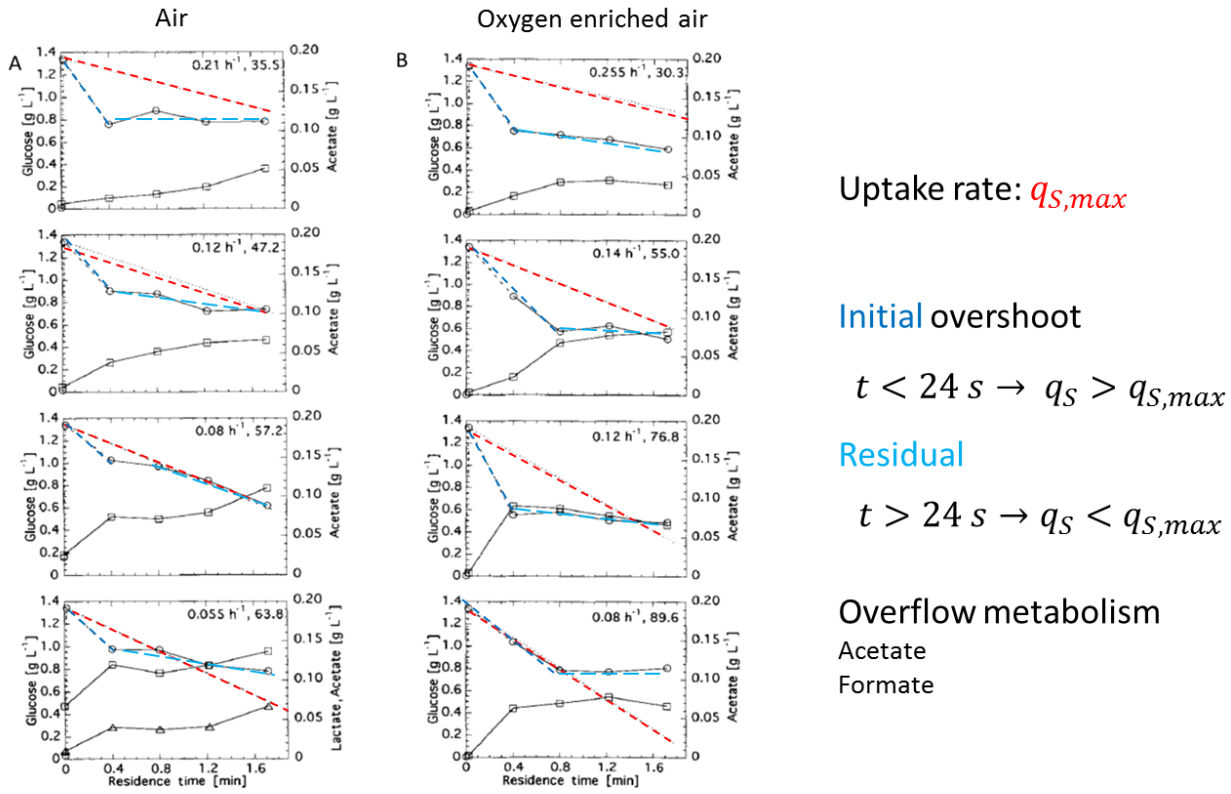

Figure 1.19 – Dynamic response of *E. Coli* cells to a sudden exposure to high glucose concentrations measured in a two–stage bioreactor operated in fed–batch mode

Several other studies [70,96] showed that the sugar uptake rate of various cells (not only *E. Coli*) suddenly exposed to high concentrations does not depend on the growth rate. In these studies,

cells cultivated in a chemostat (at steady state) were exposed to high glucose levels. Uptake and growth during the first minute following the change were measured and plotted against the growth rate prior to the perturbation (see figure 1.20 for an illustration). In Neubauer's experiment, the microbial cells were forced to lower their growth rate because of ever increasing limitation in terms of glucose availability (because of the constant feed). So, in Neubauer's experiment, the microbial population is not at steady state when it faces the glucose step–up. The bacteria's growth rate is only their current growth rate and is not stationary. Thus the apparent contradiction in the reported observations is indeed related to the history of the culture. It reveals that multiple dynamics are at work: in a transient regime both the growth rate and the uptake rate are responding to the external concentration changes, but they do so on distinct time scales.

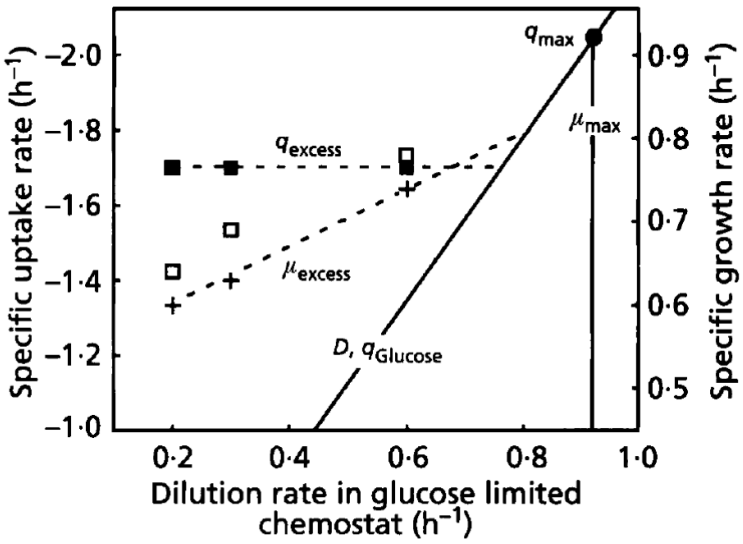

Figure 1.20 – Instantaneous uptake rate of chemostat–cultivated *E. coli* cells in the wake of a sudden exposure to a $5mg.L^{-1}$ glucose concentration measured in a batch system.

It is of paramount importance to observe that the uptake rate is proportional the growth rate in a chemostat culture at steady state whereas this is no longer the case when the equilibrium is disrupted. The excess uptake rate is independent on the sugar concentration, the excess growth rate is not proportional to the growth rate prior to the perturbation neither is the uptake rate proportional to the growth rate. In other words, the steady–state uptake law (1.15) is not applicable to the transient response of a microbial population to glucose concentration changes.

The duration of the high concentration events is constant whereas the time lapse in between is not. Thinking in terms of dynamic systems, it is likely that population heterogeneity results from the convolution of the residence time distribution in the large reactor and the dynamic response of the biological system. Note also that the flow rate between the two reactors impacts the frequency of high concentration events, the duration of the said events but also the homogeneity of the entire system. Indeed, the higher the flow rate in the loop, the more spatially homogeneous the multistage reactor. The complex interplay between the spatial and temporal dimensions makes the experiments very difficult to design and their modelling a very challenging task.

1.3.3 Metabolic response

In their 1993 article on scale–down reactor, George & al. concluded that *...repeated (short) residence times in zones of high sugar concentration have an influence on the microbial metabolism and thereby on the bioreactor performance.* From the previously mentioned results, it is clear that the uptake rate in highly concentrated zone increases far above the cell's requirements for growth. Depending on the magnitude of this difference the perturbation is damped or not. In this section, several experimental results related to the metabolic response following a strong perturbation will be examined.

The BioScope experiment

BioScope is the name of an experimental device allowing a higher temporal resolution of the cell response to a sudden change in the external glucose concentration [66]. Cells are sampled from a continuous culture at steady state (D = 0.1 h^{-1}), quickly mixed with some sugar and transported at a known flow rate in a silicon tube, permeable to oxygen and carbon dioxide. Multiple sampling points are distributed along the tube in order to measure the concentration of external metabolites, produced by the cells and excreted in the liquid phase (see figure 1.21). Typical results are presented

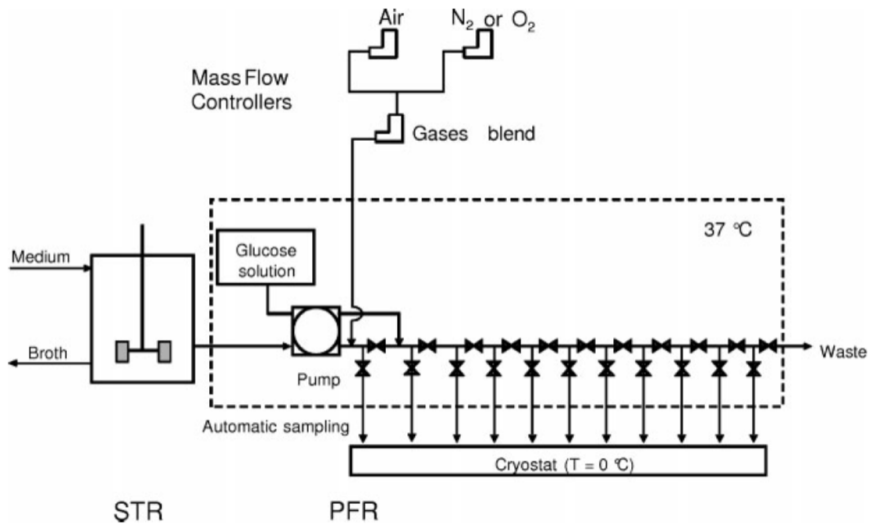

Figure 1.21 – Principle of a bioscope: cells cultivated in a continous stirred reactor are sampled, mixed with glucose and flushed into a tube. Sampling is performed along the tube to collect and measure the cell response at fixed times after the perturbation [66]

in figure 1.22. The authors identified two different phases in the response, the first one corresponding to an increase in the acetate and formate production rates and the second one being characterised by a much lower production rate. In the aerobic case, the oxygen concentration was maintained at a high level during the experiment. Minor amounts of acetate and formate were produced. In the anaerobic case, the oxygen concentration was zero in the tube and the results presented in Figure 1.23 show that the production of mixed acids was in that case continuous and much more significant. This metabolic response is typical of mixed acid fermentation which aim is to provide the organism

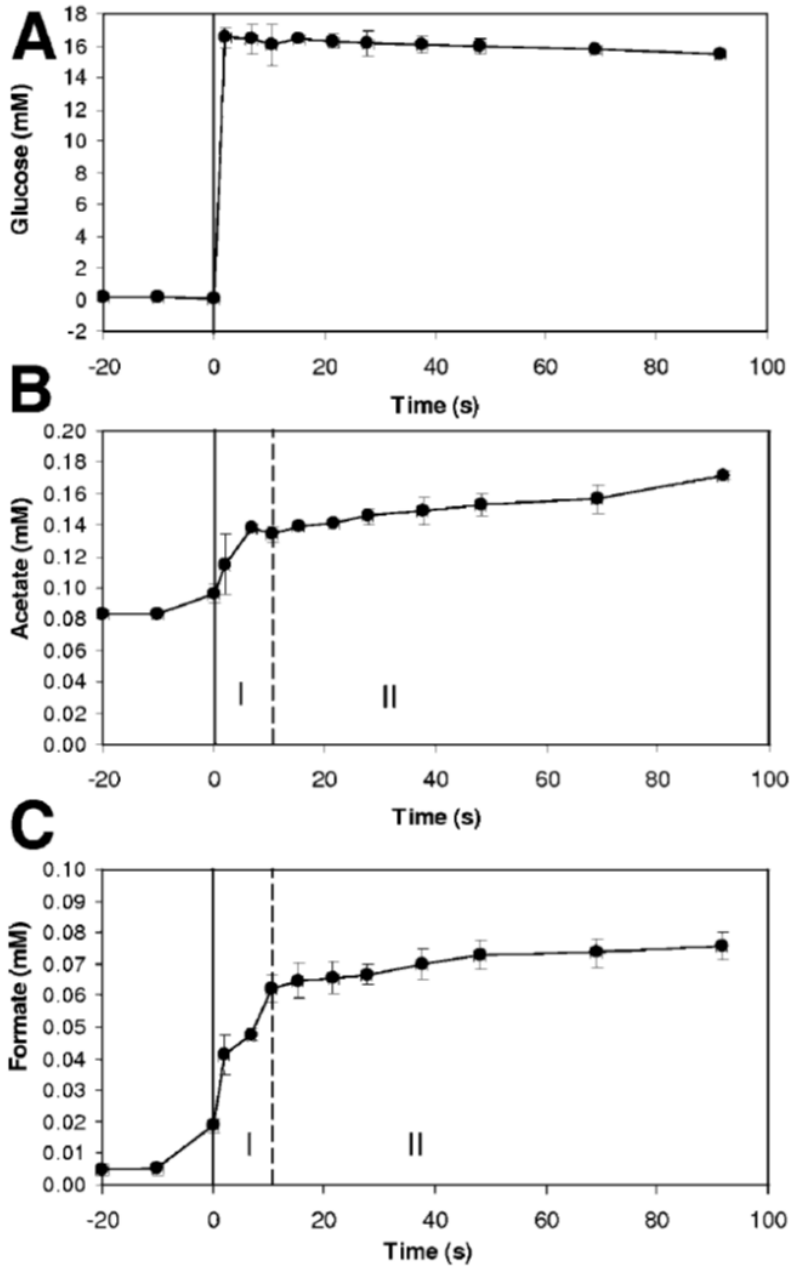

Figure 1.22 – Instantaneous response of *E. Coli* cells to a sudden exposure to a $16mM$ glucose concentration ($2.88\ g.L^{-1}$) under aerobic conditions. The instantaneous glucose uptake rate is given by the slope of the glucose profile. Two phases are visible: the first one is characterised by a significant production of acetate and formate, the second one is marked by a much lower production of these compounds.

with energy. The authors could measure the instantaneous glucose and oxygen uptake rate right after the glucose pulse, the results featuring on Table 1.1. The steady–state values are in line with

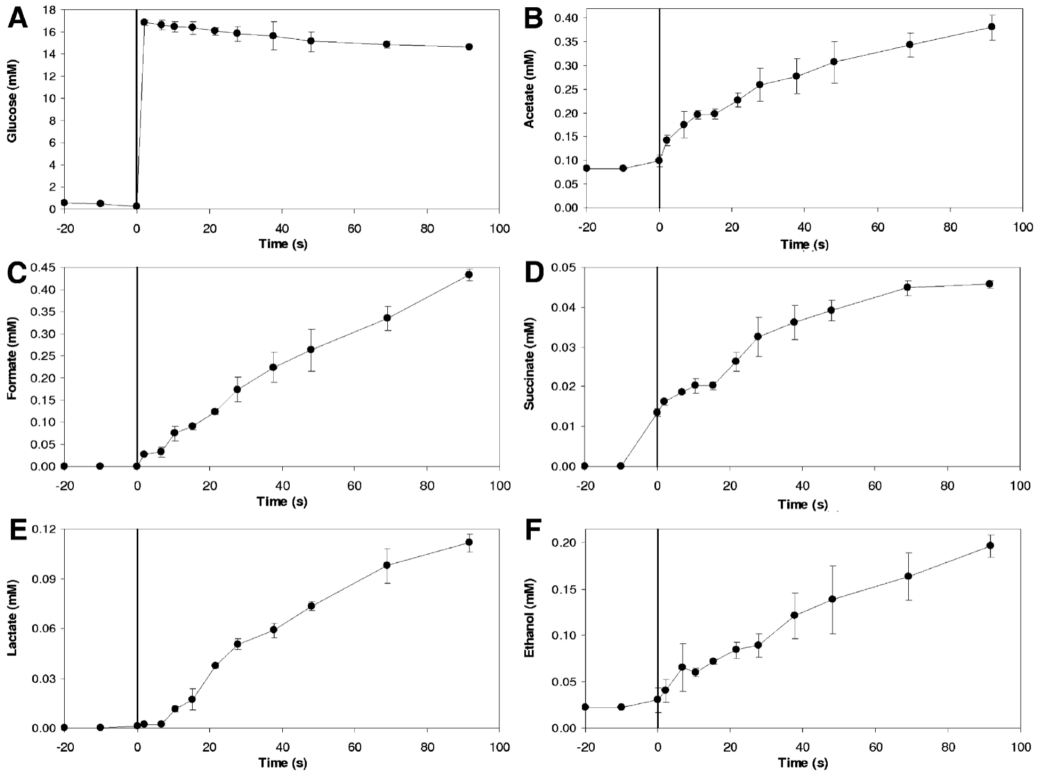

Figure 1.23 – Instantaneous response of *E. Coli* cells to a sudden exposure to a $16mM$ glucose concentration ($2.88\ g.l^{-1}$) under anaerobic conditions. The instantaneous glucose uptake rate is given by the slope of the glucose profile. The production of mixed acids is in that case continuous [66]

previously published data [49]. The glucose uptake rate after the aerobic pulse is significantly higher (50 %) than the maximum value that can be measured in a batch culture ($q_{S,batch} \approx 2g_S.g_X^{-1}.h^{-1}$ assuming a maximum growth rate of $1h^{-1}$ and $Y_{SX} \approx 0.5g_X.g_S^{-1}$ [49,99]). In the anaerobic pulse the instantaneous uptake rate is even twice larger than in the aerobic experiment indicating that the cells' actual uptake capacity can be as high as several grams of sugar per gram of biomass and per hour as also reported in the Neubauer's experiment. Clearly such high values can not be explained if a single transport system is considered. It is very likely that multiple transporters having different maximum uptake capacities are active at the same time. Depending on the culture conditions, they contribute to various fractions of the total uptake.

Table 1.1 – Specific glucose and oxygen uptake rates before and after the addition of sugar [66]

	$q_S(g_S.g_X^{-1}.h^{-1})$	$q_{O_2}(g_{O_2}.g_X^{-1}.h^{-1})$
Chemostat [1]	0.24	0.12
Aerobic pulse [1]	2.7	0.51
Anaerobic pulse [1]	5.58	-
Two–stage reactor [2]	10	-
Batch [2]	2	-

The very significant difference between the BioScope results and Neubauer's is that in the latter case cells were repeatedly exposed to long lasting starving condition ($27min$ in the CSTR) between two substrate pulses. Thus Neubauer writes:

> This raises the question of whether the oscillating changes from low to high glucose concentration in the two–compartment reactor support the enrichment of intracellular components which are responsible for the uptake of glucose [99].

Glucose pulse in a chemostat

At the Toulouse Biotechnology Institute, Sunya, Delvigne (Louvain, Belgium), Uribelarrea, Molina–Jouve and Gorret studied the transient responses of *Escherichia coli* to a glucose pulse of various intensities in an aerated chemostat culture [133]. Dissolved oxygen, pH and metabolites concentrations in the liquid phase as well as offgas composition were measured. The intensity of the glucose pulses was $0.08g.L^{-1}$ (green curves), $0.4g.L^{-1}$ (blue curves) and $1g.L^{-1}$ (red curves).

Figure 1.24 depicts the evolution of glucose, acetate and formate in the wake of the pulse. Considering the short duration of the transient period (maximum 25 min) compared to the residence time ($>$ 6h), these data can be regarded as the response of the metabolism only, disregarding the wash–out due to dilution. Thus, one observes that the uptake and production rates are constant and similar in all experiments. Secondly, acetate re–consumption takes place once the surplus of glucose was assimilated. This indicates that the cells complement their need for carbon through acetate uptake. The produced formate is not re–consumed.

The examination of the dissolved oxygen profile in Figure 1.25 reveals that the addition of sugar entails a sudden drop in the liquid phase oxygen concentration, which remains almost zero until glucose is depleted. Thus, the whole response of the culture is under influence of a limiting oxygen flux. Indeed contrary to what a concentration–based approach would suggest, an oxygen concentration close to zero does not mean that the oxygen uptake rate is null. Reasoning in terms

1. Reference [65]
2. Reference [99]

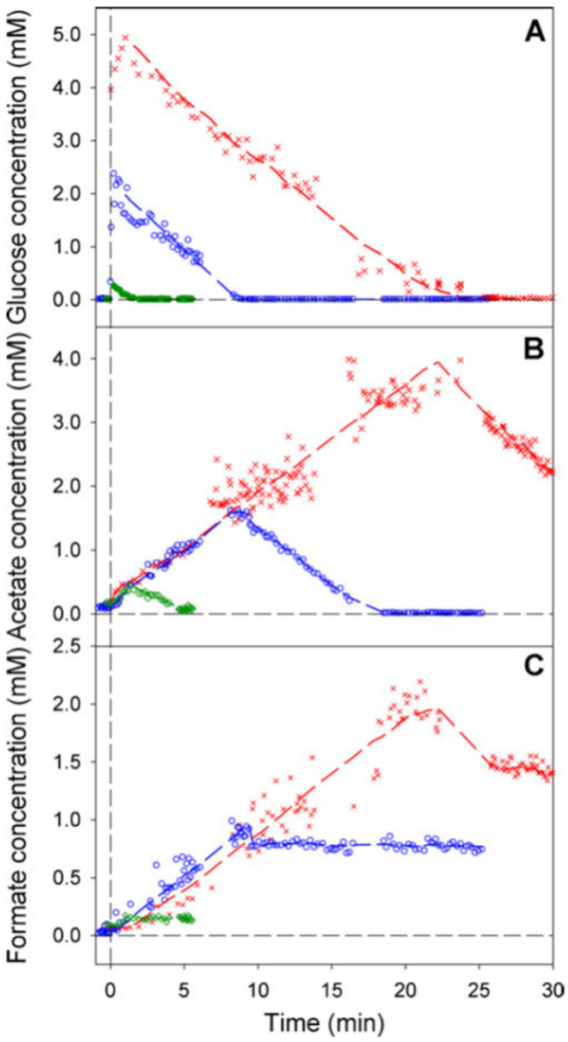

Figure 1.24 – Time evolution of the glucose, acetate, and formate in the wake of a pulse of various intensities, $0.08g.L^{-1}$ (green curves), $0.4g.L^{-1}$ (blue curves) and $1g.L^{-1}$ (red curves), imposed in a steady–state chemostat ($D = 0.15h^{-1}$) (borrowed from [133]).

of fluxes is more appropriate: in the wake of the glucose pulse the oxygen demand increases so much that the oxygen uptake rate becomes limited by the gas–liquid transfer rate. After glucose exhaustion, the oxygen concentration rises up to the steady state value at a different rate depending on the amount of acetate present in the liquid phase. This indicates that acetate is probably engaged in an oxidative pathway. The slump in oxygen is accompanied by a decrease in the oxygen concentration and an increase in the carbon dioxide concentration in the gas phase.

In order to identify which metabolic pathways are activated, the respiratory ratio, defined as the rate of carbon dioxide production to the rate of oxygen consumption is often used. A value of one is indicative of oxidative metabolism, whereas a value higher than one indicates that fermentation takes place, glucose being used in the mixed–acid pathway to produce energy. The results presented on figure 1.26 indicates that the metabolism is purely oxidative prior to the pulse. Right after the

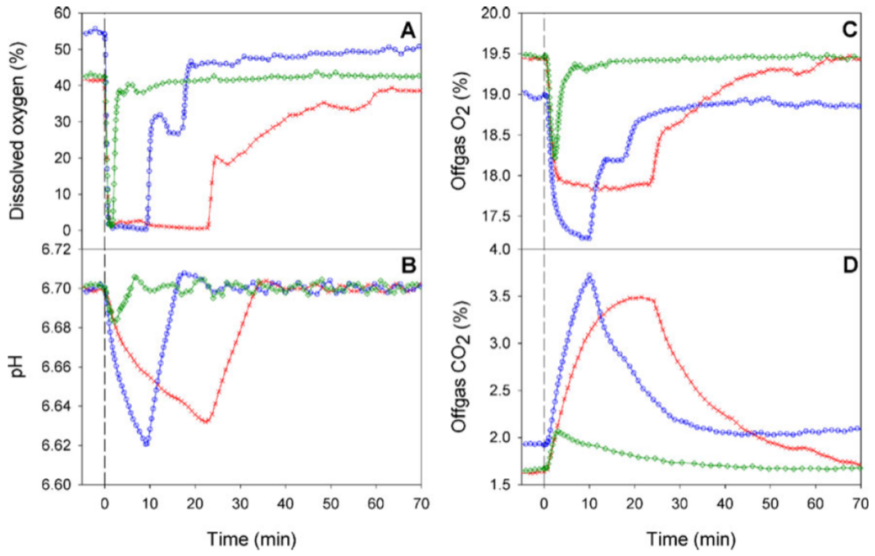

Figure 1.25 – Time evolution of dissolved oxygen, pH, O_2 and CO_2 in the exhaust gas (borrowed from [133]).

pulse, the oxygen uptake rate correlates with the carbon dioxide production rate, the latter being more pronounced. This reveals as the authors state that a combination of overflow and fermentation is at work. From these data, the authors could estimate the instantaneous glucose uptake rate. In that case the glucose uptake rate was only 71% of the maximum uptake rate in unlimited culture. However, it seems that this value is an average over the entire pulse rather that the instantaneous uptake rate in the few second following the glucose addition as in Lara's and Neubauer's works. Nevetheless, this result indicates that the availability of oxygen plays an important role in the regulation of the glucose uptake.

1.3.4 Conclusion on biological considerations

The individuals' adaption to an ever–fluctuating environment is in essence a membrane feature. From experimental grounds like Neubauer's [99], different types of transmembrane enzymes are thought to contribute to the cells' substrate uptake capability. A handful of such transporters were evidenced in the 1990s, which are classified into two main categories: molecule–specific, high–affinity and non–specific, low–affinity systems. To relieve the reader of some modelling burden, only one specific (called PTS throughout this work) and one non–specific (christened permease from now on) transporters will be addressed, without loss of generality considering the observed half–saturation constant for all these enzymatic compounds. PTS are considered the routine uptake system. They allow the well–nourished cells to cherry–pick their favorite source of carbon when the individuals are free to put all other substrates aside. Permeases are considered the scavenger uptake system. They operate like porins and are built to complement the total uptake rate should the PTS fall short of the backbone machinery's needs, what happens when the substrate is scarce at the cell's neighbourhood. It is understood that the PTS activation exerts a negative feedback on the permeases' functioning at high S, this knock–on effect waning as S is decreasing. The non–specific system is deemed less desirable to the organisms than its prime counterpart, as it lets other compounds entering the

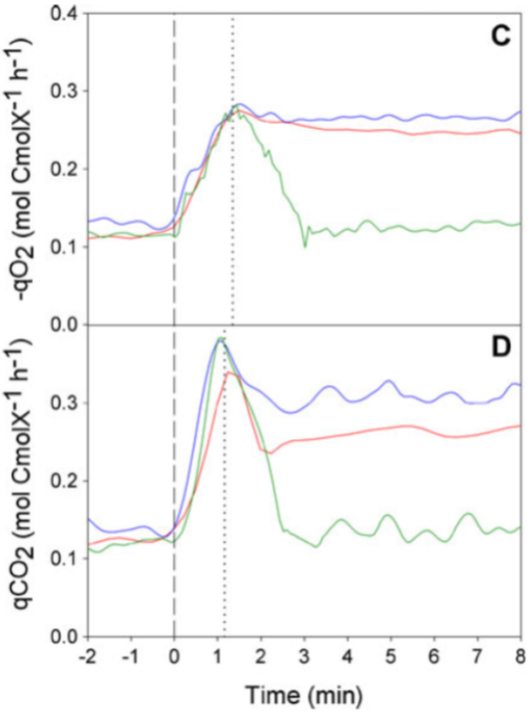

Figure 1.26 – Time evolution of the specific oxygen and carbon dioxide rate during the first 8 minutes following the pulse ($D = 0.15h^{-1}$) (borrowed from [133]).

cytoplasm. Amplitudewise though, the permease capacity is thought to overreach (by a factor of 5 according to Neubauer [99]) the largest glucose–unlimited (that is batch) PTS uptake rate. This behavior is probably induced by repeated exposure to starvation conditions. Following the sudden relief of starvation conditions, the glucose uptake rate becomes much larger than the cell needs and the whole metabolism can be affected over several generation times. The magnitude of the metabolic disorder is highly dependent on the actual physiological state of the cell, the magnitude of the perturbation and the availability of oxygen.

1.4 Population Balance Equations

1.4.1 Overview

The concept of Population Balance Equations (PBEs) originates from the field of fluid mechanics, Smoluchowski [127] establishing the coagulation equation in his seminal 1916 publication. It is not only the master equation when it comes to rupture/coalescence of liquid droplets and bubbles but also a standard framework in social sciences (cf. Kuhn & al. [61]), chemistry (polymerisation for instance, see Lebaz & al. [68]), physics (see Hulburt & Katz [45] for an application in crystallisation), meteorology (for example Scott's [125]) and biology, with special reference to cell cycle physiology (the latter will be further detailed in an appropriate section).

PBEs aim at discriminating a population with respect to inner coordinates which are assumed a representative singularisation of each and every individual. Concretely, to address a variety of bubbles (further pertaining information features, for example, in Laupsien's [67] work) or flocks (such an approach is provided, amongst others, by Guérin & al. [37]) that emerges from an experiment, each element of the set is supposed comprehensively determined by their characteristic size and morphology. When living organisms are considered instead, the modelling requires a shot at the interaction between the biotic and abiotic phases and no unicist formulation has been set in stone thus far. On the contrary, these questions are still under construction, as recent literature is rife with model developments which aim at shedding light on experimental observations (cf. Stamatakis & Zygourakis [130], Fadda & al. [22], Morchain & al. [95] for a non-exhaustive overview). From a historical point of view, the quantities of interest have been age and size which knowledge was the only information at the reach of the experimentalists. It was well understood in the microbiology community (notably by Powell [112] or Koch [51]) that the redistribution of a dividing cell's content affects the maturity of the subsequent daughter–cells, potentially distributing the organisms' growth rate as a result. Such a cell–scale feature could not be captured with macroscopic sets of ODEs like (1.7), prompting the need for more powerful mathematical tools aiming at blending growth and redistribution.

1.4.2 General formulation of PBE

The standard PBE is a transport equation that can capture, depending on the case study, non–linear (like coagulation) or linear (like fragmentation) processes which affect the description of a population. In all generality, such an equation reads:

$$\frac{\partial}{\partial t}n(t,x,\xi) + \nabla_x(\dot{x}n(t,x,\xi)) + \nabla_\xi(\dot{\xi}n(t,x,\xi)) = \int_{\Omega_\xi} G(x,\xi',n(t,x,\xi'))n(t,x,\xi')d\xi' \tag{1.16}$$

Hereinafter, n is called number density function and reports how a population is distributed into a certain number of inner variables which comprise the vector ξ. x stands for the vector of spatial coordinates. By definition, $n(t,x,\xi)d\xi$ stands for the number of *particles* per unit volume (of the physical space), having their properties in an infinitesimal domain, $d\xi$, of the state space Ω_ξ. The entire space domain consists in a N_d–dimensional rectangular domain ($\Omega_\xi = \bigotimes_{i=1}^{N_d}\Omega_{\xi_i}$).

Thus, the second term on the left–hand side of equation (1.16) represents the transport in the physical space, the third one stands for the transport in the space of inner variables (particle properties). The right–hand side term depends on the specific processes whereby particles appear and disappear from the system. G the gathering of all other (linear or non–linear) physical phenomena which can modify the fate of the particles under consideration. Further appropriate boundary conditions are usually inferred from physical considerations; in crystallisation for example, negative– or null–volume individuals are supposed non–existent.

In the so–called homogeneous case, the number density function does not depend on the space coordinates burt only on the organisms' inner properties. In this context, an integration of (1.16) over the whole reactor leads to:

$$\frac{\partial}{\partial t}n(t,\xi) + \nabla_\xi(\dot{\xi}n(t,\xi)) + \int_{\Sigma_x}\dot{x}n(t,x,\xi)d\Sigma_x = \int_{\Omega_\xi} G(\xi',n(t,\xi'))n(t,\xi')d\xi' \tag{1.17}$$

The third term on the left hand side of equation (1.17) stands for the fluxes through the boundaries, Σ_x, of the physical volume Ω_x; it is null if a closed system is considered, but not if the culture operates in a continuous fermenter for which the draining plays a role in the population balance.

For a closed cultivation system the equation reads:

$$\frac{\partial}{\partial t}n(t,\xi) + \nabla_\xi(\dot{\xi}n(t,\xi)) = \int_{\Omega_\xi} G(\xi', n(t,\xi'))n(t,\xi')d\xi' \tag{1.18}$$

and for an open cultivation system with a dilution rate D (in h^{-1}) it turns into:

$$\frac{\partial}{\partial t}n(t,\xi) + \nabla_\xi(\dot{\xi}n(t,\xi)) + Dn(t,\xi) = \int_{\Omega_\xi} G(\xi', n(t,\xi'))n(t,\xi')d\xi' \tag{1.19}$$

1.4.3 Analytical results for PBEs

Many pathways towards the solution of such PDEs have been explored over the last 50 years, albeit analytical solutions come out of highly restrictive assumptions only, the latter having little physical meaning. Most of the work has born upon linear biological PBEs, no solution to (1.16) being known in the general framework. Thus, the mathematical work for these equations has focused on the properties of the transient and steady–state solutions such as the functional space they belong to. Analysiswise, the first existence and uniqueness result for Smoluchowski's equation (in $\mathcal{C}^0(\mathbb{R}_+, L^1(\mathbb{R}_+))$, provided that the rupture and agregation functions are continuous and bounded) was proven in 1956 by Melzak [86]. A similar result to Melzak's for the system (1.22) was stated in the mid-1980s by Dawidowicz and Loskot [14].

Similarly, the mathematics literature is fraught with technical results for the eigenproblem which purpose is a better understanding of the asymptotic solution. In this context, a primal equation is satisfied by the eigenfunction (with no loss of generality, this function will be christened N throughout this work) that carves the geometrical shape of the steady–state solution, and the corresponding eigenvalue λ (the so–called Malthus parameter) dictates the rate of convergence towards the aforementioned eigenvector. A dual equation is contrived in the distribution sense, the solution thereof involving the same eigenvalue as the primal equation and an eigenfunction ϕ which definition is loosened by a mere integrability condition. The standard procedure consists in applying Krein–Rutman's theorem to a regularised problem which solution breaks down into a manifold of eigenelements and ensuring that limits can be taken with the help of a priori estimates. Such results have been published in the context of equal redistribution (see Perthame & Ryzhik's 2004 paper [109]) and constant ($N \in \mathcal{S}(\mathbb{R}_+)$, $\phi \equiv 1$) or bounded ($N \in W^{1,\infty}(\mathbb{R}_+)$, $\phi \in \mathcal{C}^1(\mathbb{R}_+)$) rupture function, unequal redistribution (cf. Michel's 2005 work [87]) and locally bounded, locally integrable rupture function (where $N \in L^1(\mathbb{R}_+) \cap L^\infty_{loc}(\mathbb{R}_+)$), with constant time evolution of the inner variables. In Doumic & Gabriel's 2010 contribution [20], the latter assumption is loosened without a hitch, and $\dot{\xi}N \in W^{1,1}(\mathbb{R}_+)$, $\phi \in L^\infty(\mathbb{R}_+, (1+\xi^k))$ for a certain $k > 0$ can be proven.

1.4.4 Numerical resolution of the PBEs

Beyond the cases in point which serve as insights more than descriptions of physical phenomena, no analytical solutions to PBEs are available for the time being, making the development of light and accurate numerical algorithms a mandatory step to investigate applied problems.

From numerical perspectives, deterministic and statistical tools have been implemented to solve (1.22). Two approaches (lagrangian and eulerian) complement each other to the extent that neither can capture the population–scale information in its entirety. Thus far, the most popular methods to solve (1.22) have been the Finite Volume, Finite Element and Monte–Carlo procedures. A brief excerpt is presented for the sake of clarity.

The Finite Volume method

This deterministic eulerian method relies on the so–called conservative form of (1.22):

$$\frac{\partial}{\partial t} n(t,\xi) + \nabla_\xi \cdot \left[\dot{\xi} n(t,\xi)\right] + \gamma(\xi) n(t,\xi) = (S_n)$$

Over a volume O of the phase space Ω_{xi}, an integration of (1.22) from time t to $t+dt$ reads:

$$\begin{aligned}\int_O n(t+\delta t,\xi)d\xi = \int_O n(t,\xi)d\xi - \int_t^{t+\delta t}\int_O \nabla_\xi \cdot \left[\dot{\xi} n(t,\xi)\right] d\xi ds \\ - \int_t^{t+\delta t}\int_O \gamma(\xi) n(t,\xi) d\xi ds + \int_t^{t+\delta t}\int_O (S_n) d\xi ds \end{aligned} \tag{1.20}$$

An application of Green's theorem turns the integral of the transport term into a difference of fluxes at ∂O. The evaluation of this very term is the determinant of the Finite Volume method theory, the literature overflowing with schemes of miscellanious orders to solve any type of PDE.

The method's stability is guaranteed should the CFL condition:

$$\delta t \sum_{i=1}^{N_d} \frac{\dot{\xi}_i}{\Delta \xi_i} \leq 1$$

be met, with $\Delta\xi_i = |OO'|$ the distance between the center of volume O and its counterpart belonging to a neighbouring O' volume.

It can happen that a reasonable accuracy for those algorithms is obtained with a relaxation of the condition touching upon the conservation of the moments of the solution. For instance, if M_0 stands for the integral $\int_{\Omega_\xi} n(\xi)d\xi$ and M_1 for $\int_{\Omega_\xi} \xi n(\xi)d\xi$, M_0 only or M_1 only conserving schemes can be implemented with satisfying precision to solve various physical processes.

The subtleties of all these algorithms have been explained in all generality by LeVeque [71] for hyperbolic PDEs. In the context of PBEs, Finite Volume methods for Smoluchowski's equation have been introduced by Filbet & Laurençot [28] in one dimension, their algorithm consisting in a second–order resolution in both t and ξ to simulate the coagulation of particles of finite size into a particle of "infinite" size. The said algorithm was extended to a two–dimensional PBE by Qamar & Warnecke [114]. A very simple M_0 and M_1 conserving Finite Volume scheme for a pure breakage PBE was thereafter published by Saha & al. [123], using an appropriate weighting for both the redistribution and the breakage terms. It was shown to be second–order accurate in ξ regardless the mesh. A meshing along the characteristic curves associated to (1.22) was also engineered by Abia & al. [1] with proven second–order accuracy should the lengthening rate be 3 times differentiable and the rupture and redistribution functions 2 times differentiable.

The Finite Element method

This deterministic eulerian method relies on the so–called weak form of (1.22): for any $\varphi \in \mathcal{D}(\Omega_\xi)$, an integration over the phase space reads:

$$\int_{\Omega_\xi} \frac{\partial}{\partial t} n(t,\xi)\varphi(\xi)d\xi - \int_{\Omega_\xi} \dot{\xi} n(t,\xi)\varphi'(\xi)d\xi + \int_{\Omega_\xi} \gamma(\xi) n(t,\xi)\varphi(\xi)d\xi = \int_{\Omega_\xi} \varphi(\xi) \text{ Source } d\xi \quad (1.21)$$

using the fact that φ has compact support to ditch the boundary term emerging from the use of Green's theorem in the transport term.

Once a suitable shape–regular mesh Ω_h has been built and the right Sobolev space to solve the weak form of the PDE has been identified (say $W^{k,p}$, $p \geq 1$), a finite element space is defined as the finite–dimentional subspace of $W^{k,p}$ that is spanned by a user–set number of L^2–orthonormal basis functions (usually polynomials) on Ω_h. The order of the method is given by the cardinality of the polynomial basis.

This method has been extensively implemented to solve PBEs. A Finite Elements algorithm using (1.21) along with a RK4 time integration was proposed by Mantzaris & al. [79]. In Ganesan's contribution [34], a coercive form of the weak formulation of a transport–diffusion equation for crystallisation was engineered with the help of local stabilisation parameters and solved with an operator–splitting Galerkin/SUPG method.

The Monte–Carlo method

This statistical lagrangian method consists in tracking fictitious particles which inner coordinates are updated using the dynamical laws for $\dot{\xi}$. It allows a circumvention of the (1.22) PDE at each time step and is the best–suited method for high–dimensional PBEs. In detail, the PDF is approached by the sum:

$$n(t,\xi) \approx \sum_{i=1}^{I} \frac{1}{I} \delta_{\xi - \xi_i(t)}$$

where $\xi_i(t)$ is the markovian consequence of the i-th particule's history in the system from its inception at time t_{i0}. I can be set constant, meaning that no death process is implemented and each birth event in the system goes with the removal of one of the extant particles. In the case of birth–and–death processes though, I is another markovian marker of the state of the population at a certain time.

Applied to PBEs, this method would break down into two main processes. The transport term is deterministically computed as it comes down to an integration along the characteristic curves passing through the state at birth. The time evolution of the particle number is the result of a stochastic treatment of the physical phenomena the PBE aims at reporting (rupture, coalescence, dilution, ...). More precisely, the drawing of appropriate random numbers for each individual can determine whether or not the particle will be affected by the said phenomena.

This method was engineered as early as 1949 by Kendall [50] and its first reference in the context of biological population dynamics features in Shah & al.'s 1976 article [126]. From the central limit theorem, the method's accuracy is well–known to scale as $I^{-1/2}$ only, requiring a substantial sample to ensure reliable numerics. However, the avoidance of an eulerian integration of (1.22) provides numerical diffusion–free outputs, what is quite commendable since the method's uncertainty is only determined by the user–controlled population number I, enabling a consequential gain in accuracy at the reasonable expense of a rise in I. Furthermore, its numerical cost scales as IN_d when N_d variables are tracked among the population, what makes it quite affordable as a large number of variables have to be dealt with, in comparison with a Finite Volume algorithm which meshing can be viewed as a tensor product of N_d contributions for instance.

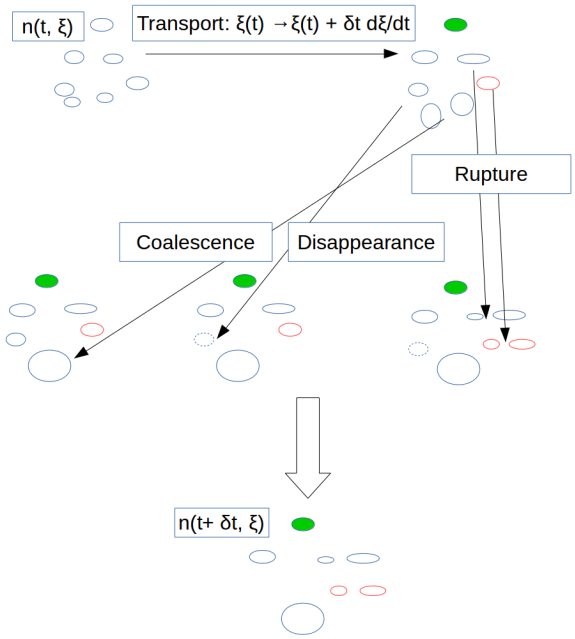

Figure 1.27 – Illustration of a general Monte–Carlo procedure to treat Smoluchowski's equation

Other methods

The Finite Difference method is a deterministic, eulerian method based on the strong form of the (1.22) PDE. The solution is updated from the knowledge of n at the nodes of a N_d–dimensional

mesh and Taylor's formula to calculate, for any ξ_i:

$$n(t+\delta t,\xi_i) = n(t,\xi_i) + \delta t\frac{\partial}{\partial t}n(t,\xi_i) + \frac{1}{2}\delta t^2\frac{\partial^2}{\partial t^2}n(t,\xi_i) + ...$$

with the definition of $\partial^k n/\partial t^k$ emerging from Leibniz's rule. Just like the Finite Volume case, the method's order is dictated by the truncation and a CFL condition must be satisfied to guarantee the algorithm's stability. Such a numerical solution to a mass–structured model was presented by Mantzaris [80] and invokes half–time step intermediate solutions between the grid points, the computation thereof resulting from the Lax–Friedrichs scheme. More of the same was laid out by Mantzaris & al. [77], where the aforementioned hybrid scheme is compared with higher–order explicit and implicit Finite Difference algorithms, computational time– and stability–wise. From theoretical perspectives, convergence and error estimates for this method have been thoroughly investigated by Kostova [55], the case study being an age–structured model.

A similar approach, albeit less time–consuming, is the Lattice Boltzmann method. It comes down to tracking a particle density which, at time t and position x, is geared towards a predefined number of neighbours x_q, $q \in \{1,...,Q\}$, ecah jump occuring at velocity $v_q = (x_q - x)/\delta t$. It has been proven (see Dubois's [21] analysis and references therein) that this process is tantamount to an upwind differencing scheme for a standard transport equation. Numerical simulations making use of this technique were performed by Majumder & al. [76] in the context of size–structured PBEs for crystallisation.

Spectral methods were implemented by Mantzaris & al. [78] to solve PBEs. In this case, n is expanded onto a L^2–orthogonal (typically Legendre or Tchebychev) polynomial basis:

$$n(t,\xi) = \sum_{i\geq 1}\omega_i(t)p_i(\xi)$$

An I–th order truncation thereof turns (1.22) into a set of I differential equations involving the ω_i at each time step. To this end, each differential equation is multiplied by $p_j/\int p_i p_j$ and integrated with respect to ξ to isolate the time evolution of the ω_i s by virtue of the orthogonality relationship between the p_i s. In [78], spectral Galerkin and similar tau methods are compared with pseudo–spectral collocation methods which require that the I–th order approximation of the solution to (1.22) satisfies the PBE at I so–called collocation points.

The method of classes (cf. Kumar & Ramkrishna's 1995 articles [62, 63] or Morchain & al.'s work [95]) involves once again a N_d–dimensional meshing of the phase space, over which (1.22) can be approached by a weighted Dirac function centered on a (fixed or moving) pivot. In order to compute both the weights and the pivots at each time step, discrete equations for two integral quantities (number and mass, typically) have to be derived from the PBE. It has been proven a reliable asset inasmuch as both the particle number and the total mass are conserved, with the advantage of flexibility regarding the meshing (cartesian, geometric, nonregular) of the phase space.

The method of moments consists in a priori assuming the shape of n (beta, Weibull, exponential, ...) and tracking as many moments of the distribution as necessary to update all the shape/scale parameters that characterise the presumed distribution. This quite efficient method poses numerical

problems when it comes to the computation of the normalising factor, for instance in the beta case where a $\beta(75,75)$ distribution cannot be restituded by the current state–of–the–art processers. A way around this untoward hindrance could consist in a Quadrature Method of Moments, that is an approximation of the solution to (1.22) by a sum of Dirac functions:

$$n(t,\xi) = \sum_{j=1}^{J} \omega_j \delta_{\xi-\xi_j}$$

with exact calculation of the ω_j, ξ_j from the knowledge of the $2J$ first moments of n, whatever its shape. Numerical tools have been developed to this end, notably Wheeler's [144] algorithm, and an exhaustive review of this technique is provided by Marchisio & Fox [84]. However, one distribution is defined by no less than an infinite sequence of moments, prompting a cautious need to ensure the realisability of the outputs of the numerics (see Nguyen & al. for such an example [100]).

1.4.5 Takeaway

One dimensional structured models have been thoroughly investigated over the last decades, resulting in a luxuriant literature pertaining to the PBE analysis. With no analytical solution to (1.22) at the researchers' reach, the implementation of numerical methods has made up the bulk of the scientific production since the 1970s. These algorithms are engineered to simulate ever more complex (non–linear, multivariate, ...) problems emerging from the field of physics, chemistry, or biology. Before focusing on the use of population balance equations to describe microbial populations, some experimental considerations have to be explained beforehand.

1.5 Population Balances in biology

Two different formulations are found for biological population dynamics depending on whether the cell age is considered or not. A standard PBE in biology would take the following shape:

$$\begin{aligned} &\frac{\partial}{\partial t} n(t,\xi) + \nabla_\xi \cdot \left(\dot{\xi} n(t,\xi)\right) + \gamma(\xi) n(t,\xi) + D(\xi) n(t,\xi) = S_n, \quad \xi \in \Omega_\xi \\ &n(t,\xi)_{\xi \in \Sigma_\xi} = (\mathrm{BC}) \\ &n(0,\xi) = n_0(\xi) \ , \ \xi \in \Omega_\xi \end{aligned} \tag{1.22}$$

$D(\xi)$ (h^{-1}) is an overall disappearance rate combining cellular (death) and environmental (hydraulic dilution) factors. Here, Σ_ξ is a part of Ω_ξ's boundary and γ (time unit $^{-1}$) is the rupture function or cell division frequency.

S_n and (BC) take different shape depending on the chosen inner coordinates. If one element of ξ is the cell age, then $S_n = 0$ and the boundary condition reads

$$(\mathrm{BC}) = 2 \int_{\Omega_\xi} \gamma(\xi') P(\xi,\xi') n(t,\xi') d\xi'$$

with P the redistribution kernel that models the probability that mother–cells of state ξ' give birth to daughter–cells of state ξ. The factor 2 signals that one dividing cell gives birth to two daughter–cells.

On the other hand, if ξ does not take the age into account, S_n will be a redistribution integral:

$$S_n = 2\int_{\Omega_\xi} \gamma(\xi')P(\xi,\xi')n(t,\xi')d\xi'$$

and (BC) a Neumann or Dirichlet condition.

1.5.1 Analytical results

Age–structured models

Biological age–structured models are shaped from the McKendrick–von Foerster equation:

$$\begin{aligned} \frac{\partial}{\partial t}n(t,a) + \frac{\partial}{\partial a}n(t,a) &= -d(a)n(t,a)\ ,\ f:\mathbb{R}_+ \to \mathbb{R}_+ \\ n(t,0) &= \int_0^\infty \gamma(a)n(t,a)da \end{aligned} \tag{1.23}$$

for which analytical solutions exist using the method of characteristics (cf. Trucco's seminal [137] article). d (time unit $^{-1}$) is the death function.

As early as 1983, an existence and uniqueness result in $L^\infty(\mathbb{R}_+, L^1(\mathbb{R}_+))$ for age-dependent population dynamics was derived by Chipot [11] when the rupture function is measurable. Existence and uniqueness of a solution (in $\mathcal{C}^0(\mathbb{R}_+, L^1(\mathbb{R}_+))$) and eigenelements to the age PBE have been proposed by Clairambault & al. [13]. It has also been demonstrated by Bartlomiejczyk & al. [3] that the redistribution operator is a contraction with respect to a Bielecki norm, provided the rupture function is bounded in the space of the inner coordinates. Overall, the mathematical theory for age–structured models is now wrapped up, taking into account that the organisms' age belongs to $\mathbb{R}_+$ as a whole and has no reason to be restricted to a bounded interval. Indeed, as explained by Hjortso & Nielsen [44], a cell will not divide if substrate is not abundant enough to allow its preliminary growth, protracting its division age with no limit.

Modellingwise, age–PBEs represent a massive step forward in describing the behaviour of biological populations in reactors to the extent that bacteria age profiles can be retrieved from (1.23). Indeed, the only reported time scale at the reach of the 1.7 set of ODEs is given by the dilution rate D, prompting the need to formulate any cell–scale characteristic time as a function of D.

A quantity which analysis has proven fruitful for is the interdivision time (or age at division), even though this variable does not explicitly feature in the equations. The definition of growth introduced at the beginning of the manuscript along with the unstructured kinetic model for a continuous bioreactor would lead to the following calculation for a population's doubling time

$$2X = X\exp(D\tau_d) \Leftrightarrow \tau_d = \frac{\ln(2)}{D} \tag{1.24}$$

However, from physical grounds, it seems evident that as soon as interdivision times are distributed in a population, some organisms will divide more than once in an interval of Lebesgue measure τ_d whereas some will not generate any descendent in the meantime. In other words, the healthier cells are to contribute more to a population's doubling, in such a way that a recorded mean interdivision time must be less than or equal to τ_d. The equality would be tantamount to a non-distribution of the interdivision times in the population that does not seem plausible. The population's doubling time must not be conflated with the mean interdivision time.

Before the formal statement of PBEs, Powell conjectured from his own measurements that a population's mean interdivision time has to be less than τ_d [112,113]. From physical grounds, when interdivision times are measured in a culture, the healthier cells are granted the larger weight and their consecutive division events must sway the interdivision time distribution to the left. However, these connections between mean interdivision time and τ_d have stirred a protracted debate since Powell provided his experimental conclusions in 1956. Painter & Marr [105,106], emulating Powell's reasoning, concluded in an attempt to link age– and interdivision time–profiles that the first moment of the interdivision time distribution must be greater than τ_d. Few eyebrows have been raised since, apart Tyson & Hannsgen's [139], but their reasoning over an isolated generation of organisms holds true in a closed system only and is far from universal.

Agewise, it was proven as early as 1963 by Fredrickson & Tsuchiya that the age distribution is decreasing on its support $[0, \infty[$ [32]. Further developments have been derived by Lebowitz & Rubinow [69] from a generation expansion and Ramkrishna [117] using Laplace transform to solve the Volterra integral equation satisfied by the newborn cell density function. Comparisons with experimental measurements for *C. črescentus* (which division process amount to swarming) feature in Jafarpour's 2018 work [46].

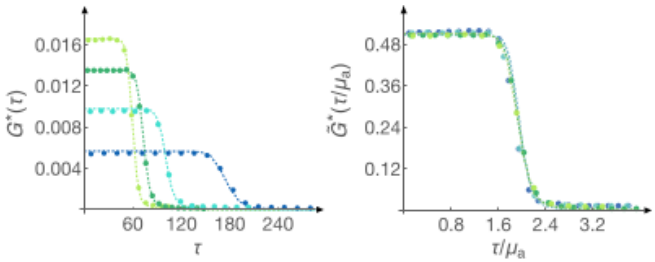

Figure 1.28 – Left: experimental (dotted lines) and numerical (full lines) age PDFs for *C. črescentus*. Right: Age over mean age PDFs for the same bacteria. Borrowed from [46].

Size–structured models

In his book on PBEs for biology, Perthame claims :

> For unicellular organisms the renewal equation does not apply, mainly because age is not the most relevant parameter that determines mitosis (the reproduction stage). The mass of the cell, its length, its DNA content, the level of certain proteins as cyclins or some other biological parameters are often more relevant. [90]

Studying size–structured models for biological populations was topical in the biophysics community in the 1960s, analytical solutions emerging from heavily simplifying hypotheses only. Typically, formulae for equal redistribution and linear or exponential growth have been derived by Kubitschek [59] from probabilistic arguments and by Beyer [5] from a generation expansion.

When the growth function is not prescribed, Diekmann & al. [18] have derived an existence and uniqueness result in $\mathcal{C}^0([0,T]\times[l_{\min}, l_{\max}])$ $\forall T > 0$ in the context of equal redistribution, where the support $[l_{\min}, l_{\max}]$ of the size distribution is shaped by the assumption that the minimal length

at rupture is $2l_{\min}$ and an arbitrary definition of $l_{\max}$ that forces an exponential shape of the rupture function.

It must be put to the reader's attention that to close a size–structured PBE:

$$\frac{\partial}{\partial t}n(t,\xi)+\frac{\partial}{\partial \xi}\left(\dot{\xi}n(t,\xi)\right)+\gamma(\xi)n(t,\xi)=2\int_{\xi}^{\infty}\gamma(\xi')P(\xi,\xi')n(t,\xi')d\xi'$$
$$n(t,0)=0=\lim_{\xi\to\infty}n(t,\xi)$$

an experimental information, $P(\xi,\xi')$, is introduced in the modelling, what comes down to a somehow ill–formulated problem when it comes to calculating analytical solutions. The only exception is the equal redistribution case that turns the right–hand side of the PBE into $4\gamma(2\xi)n(t,2\xi)$, the riddance of P allowing Kubitschek's developments.

It results from this impediment that asymptotic size profiles have been another research topic in the biophysics community. Bell & Anderson's seminal 1967 article [4] features an eigenproblem which solution is made up of analytical formulae over an interval of interest and extrapolations from physical assumptions.

Age in the cycle and maturity rate 2-D structured model

Age is sometimes ditched from PBEs in favour of the bounded degree of maturity (see Rubinow's [122]) that is understood by Rotenberg [120] as the normalised age in the cell cycle. For this very internal variable, slight differences from the classic McKendrick–Von Foerster equation are put to the reader's attention: the cell division translates into a source from cells of maximal age (which, without loss of generality, can be set to 1) to cells of age 0 and the redistribution integral concerns the change in the rate of maturity throughout the cell cycle. Rotenberg's equation explicitly reads:

$$\frac{\partial}{\partial t}n(t,\mu,v)+v\frac{\partial}{\partial \mu}n(t,\mu,v)=\int_{0}^{\infty}r(\mu,v,v')n(t,\mu,v')-r(\mu,v',v)n(t,\mu,v)dv'+\delta_{\mu}S_n(t,v)$$
$$S_n(t,v)=vn(t,\mu,v)|_{\mu=0}$$

with μ the age in the cell cycle, v its time derivative, and r the transition rate in v.

Modellingwise, this framework is particularly well suited to the synchronisation of tumorous cells as part of chemotherapy treatments (see [120] and references therein), and, among other results, a Green function is retrieved by Rotenberg [121] from the assumption that the time evolution of the cells' maturity is bounded and the transition rate of one rate of maturity to any other is constant. This formulation has even drawn the attention of mathematicians: a characterisation of the Malthus eigenelements was published by Mischler & al. in 2004 [90] under the assumption that the rupture function is uniformly bounded. Rotenberg's formalism sheds light on the different perspectives of modelling the cell division and incorporating it in a PBE. When the age is normalised to the extent that rupture always occurs at age = 1, the time derivative of this so–called degree of maturity is obviously not equal to 1, meaning that cell division occurs at any maturity rate and maturity rate transitions occur at any time. In other words, when the cell cycle duration is set, the maturity velocity is not anymore: both variables can be seen as "dual".

1.5.2 A review of popular structured PBE

In this section will be provided a brief description of increasingly complex formulations of a structured PBE. In particular, special emphasis will be placed on the coupling between the PBE for the biological population and the mass conservation of nutrients. These classes of models are referred to as structured models including extra–cellular environment. Clearly, the introduction of a coupling with the liquid phase considerably increases the complexity of the whole model since it involves a two–way coupling. Each cell responds with its own dynamics to the changes in the external concentration whereas the nutrients concentration in the liquid phase results from the contribution of all cells.

Size–structured models

$$\begin{aligned}
&\frac{\partial}{\partial t}n(t,\xi) + \frac{\partial}{\partial \xi}\left(r(\xi,S)n(t,\xi)\right) + \left(\gamma(\xi,S) + D\right)n(t,\xi) = 2\int_{\xi}^{\infty}\gamma(\xi',S)P(\xi,\xi')n(t,\xi')d\xi' \\
&\frac{\partial S(t)}{\partial t} = D(S_f - S(t)) - \int \nu(\xi,S)r(\xi,S)n(t,\xi)d\xi \\
&n(t,\xi=0) = 0 \\
&S(t=0) = S_0
\end{aligned} \tag{1.25}$$

ξ standing for the cell size (length, volume or mass), D for the dilution rate (h^{-1}), and S_f for the sugar concentration in the feeding current. $r(\xi,S)$ is the individual cell growth rate (in the unit of $\dot{\xi}$) and $\nu(\xi,S)$ plays a similar role as Y_{SX} in unstructured models. D is equal to Q/V for a chemostat, that is a reactor of volume V flushed by the flow rate Q. Setting D to zero and S_0 to a positive value leads to a set of equations for a batch culture.

An illustration of such a model is proposed by Subramanian, Ramkrishna, Fredrickson and Tsuchiya in their 1970's paper, where ξ corresponds to the cell mass [132]. Using the notation of the present manuscript, their model turns into (1.25) with

$$r(\xi,S) = \phi_{\xi}\frac{S}{K_S + S}\xi \tag{1.26}$$

$$\nu(\xi,S) = 0.75 \tag{1.27}$$

At the cell level exponential growth is considered and the substrate–to–biomass yield is constant. In Perthame's presentation of this type of model $\nu(\xi,S)$ is implicitly set to one meaning that a gain in mass of the suspended biological phase results from a corresponding mass loss in the liquid phase. The real situation is more complex and this will actually constitute one of this work's key issues. In general, $\nu(\xi,S)$ is not equal to one because cells consume and excrete some products (for example microbial cells consume sugar and oxygen to produce new cells and carbon dioxide). Subramanian and co–workers considered this fact in the simplest possible manner using a constant value for $\nu(\xi,S)$ but the crucial point is that the ratio between the substrate uptake rate and the growth rate (rate of mass change) is not constant in the transient regime as shown in the first part of this chapter (Section 1.3.2, page 27).

The rupture and mass distribution functions were expressed as:

$$\gamma(\xi, S) = f(\xi_c, S).r(\xi, S) \tag{1.28}$$

$$P(\xi, \xi') = 30\frac{\xi^2 (\xi' - \xi)^2}{\xi'^5} \tag{1.29}$$

where ξ_c is a data based constant which indicates the mean cell mass at division. In the same vein, the exponents in the mass distribution law are extracted from experimental data or simply hypothesised. Size–structured PBE remain the most popular framework in the biochemical engineering community in particular because the cell size is experimentally accessible. The main difficulty lies in relating the cells' metabolic behaviour to their size which is obviously not the most relevant quantity to do so [40].

Age in the cycle (and size) –structured models

A few papers were dedicated to the description of the cell cycle with $\xi = a$ the age in the cycle [13] or $\xi = \{a, m\}$ age and cell mass [39]. In that case, I number density functions are described $i \in \{1, I\}$ and an equivalent number of transition functions from one phase of the cycle to the next one have to be defined. Note that the transition function from phase I to phase 1 can be regarded as a boundary condition.

$$\begin{aligned} &\frac{\partial}{\partial t} n_i(t,\xi) + \frac{\partial}{\partial \xi}\left(r_i(\xi) n_i(t,\xi)\right) + \left(\gamma_i(\xi) + D_i\right) n_i(t,\xi) = 0 \\ &n_i(t,\xi)|_{\xi=0} = \int_0^\infty \gamma_{i-1}(\xi) n_{i-1}(t,\xi) d\xi \quad 2 \le \xi \le I-1 \\ &n_1(t,\xi)|_{\xi=0} = 2\int_0^\infty \gamma_I(\xi) n_I(t,\xi) d\xi \end{aligned} \tag{1.30}$$

In each phase of the cycle, cells start at age 0 and are aging with speed r_i. More details on the identification of γ_i from experimental data and analytical solution can be found in [6].

The particular case treated by Hatzis consists in taking the age into account in the first and last phases of the cycle only and the sole mass in the intermediate phase. Accordingly the transition from phase 1 to 2 is age–dependent, the second being a function of the organisms' mass only and the third one an age–dependent function. Although these transition functions can be identified from experimental data, Hatzis pointed out that these measurements are made under steady–state conditions and more than likely will not reflect the transient population features. Further investigations aiming at giving a general form to the model are provided in [29, 30]

Cell composition–structured models

As previously stated, the cell growth rate depends on the availability of some nutrients in the extra–cellular environment. A biology–inspired refinement is to consider that the growth rate is the consequence of many biochemical reactions involving intra–cellular compounds (namely metabolites and enzymes). Therefore a more realistic model would also consider intracellular composition to formulate an expression for the growth and uptake rates, leading to $\xi = \{c, e\}$ where c stands for a vector of metabolites concentration and e a vector of enzymes that control the rate of intracellular reaction. The dimension of this very problem increases with the number of internal metabolites to the point where in practice, the continuous approach is no longer feasible. Ordinary differential

equations are written for the real or cybernetic variables and population effects are accounted for through the addition of randomness in the model parameters (maximum reaction rates and affinity constants) [41,54,131,149].

Growth rate–structured models

In face of the difficulty to establish the constitutive laws for the dynamics of all intracellular variables, and considering the extremely large number of internal compounds that should be taken into account, Morchain proposed to use the growth rate as an internal variable of the PBE [93,95]. Note that, here again, the cell size is no longer part of the inner variables:

$$\begin{aligned}
&\frac{\partial}{\partial t}n(t,\mu) + \frac{\partial}{\partial \mu}\left(\dot{\mu}n(t,\mu)\right) + (\gamma(\mu,S)+D)n(t,\mu) = \int_{\mu}^{\infty}\gamma(\mu',S)P(\mu,\mu')n(t,\mu')d\mu' \\
&\frac{\partial S(t)}{\partial t} = D(S_f - S(t)) - \int_0^{\infty} q_S(\mu,S)n(t,\mu)d\mu + S_0 \\
&n(t,\mu=0) = 0 \\
&S(t=0) = S_0
\end{aligned} \tag{1.31}$$

A constant value was used for $\nu(\mu,S)$ and the partition function $P(\mu,\mu')$ was picked among log–normal distributions (see [95]) in line with Yasuda's experimental data [148] regarding steady–state interdivision time distributions in a perfused reactor. With this formulation, the difficulty establishing a sounded relationship between cell mass and metabolic reaction rates is circumvented since the known specific growth rate serves as an input for the metabolic model. By the way, the degree of freedom is lowered by one which allows for a direct calculation of the metabolic rates. Hence, one avoids a time–consuming optimisation step along with the definition of objective functions that often invoke the fact that cell would tend to optimise their growth rate.

As already mentioned when presenting metabolic models, the previously proposed formulation for $\dot{\mu}$ incorporates some information regarding the average population growth rate at equilibrium.

$$\dot{\mu} = \frac{1}{T_\mu}\left(\mu^\star(S) - \mu\right) \tag{1.32}$$

T_μ being a time constant for growth adaptation and $\mu^\star(S)$ the mean growth rate of a population being at equilibrium with its environment.

Another limitation of this formulation is that the cell–scale uptake rate is algebraically retrieved from the growth rate whereas uptake rate and growth rates are not coupled in the transient regime.

1.6 Synthesis and Objectives

1.6.1 Synthesis

Biological aspects

It is quite obvious to anyone observing a population that individuals are all different. Their interactions with their environment determine the macroscopic behaviour of the population. For years this has been the only information accessible to the microbiologists. In the late 1990s headway in experimental techniques (microscopy and flux cytometry) allowed to evidence the distribution of

any biological trait within a microbial cell population. From an experimental point of view, this opened a new field of investigation and it was shown that age, size, biochemical content, elongation rate, reaction rates are dynamic properties of any cell. Three main sources of *evolution* can be identified:

1. The partitioning of the cell content at division.
2. The interactions with the environment (exchanges or signal)
3. The genetic mutations [2]

As a consequence, cell properties may vary from one cell cycle to the other and they may also vary within the duration of a cell cycle. The macroscopic behaviour is the result of a multitude of coupled biochemical reactions and it is observed that the response to external perturbations can not always be predicted using population–averaged properties. The response is highly dependent on the magnitude of the perturbation, the initial conditions and the culture's history. In some cases, spontaneous oscillations of macroscopic properties are observed. Clearly, non–linear dynamics are at work in biological systems and several time scales have to be considered. Because of that, many general principles in the standard biology have to be reexamined:

— The definition of growth rate.
— The stationary feature of a microbial population.
— The rate of exchange between a cell and its environment.
— The scope of application of ensemble– and time–averaged general *laws of Biology*.

Compiling the advances in various fields it seems necessary to take into account the existence of multiple transport systems working in parallel. These transporters determine the uptake capacity of the cell and thus the mass fluxes fueling the cell metabolism. The uptake rate is tightly connected to the rate of some metabolic reactions and the whole machinery is responding in a synchronised way to external fluctuations and internal requirements.

In multiphase systems, thermodynamic laws of equilibrium can be used to connect the concentration on both sides of the interface; no such constitutive laws have been formulated for biological systems and a mixed approach combining mass and energy conservation principles, biophysical, physiological considerations along with experimental observations is the only way to go ahead.

Because the number of internal properties is considerable and the interactions with each other incomplete, it is not possible to proceed to a direct simulation of the individual cell behaviour.

PBE modelling

From their inception in the first half of the 20th century, PBEs were addressed to treat measurable features of a population as continua. These variables of interest had to be traceable (size or age) in order to compare the numerics with the experiments. As the transport–fragmentation framework was getting quite common in the biophysics community in the 1960s, the very first age–size–structured model was developed by Bell & Anderson [4] and steady–state cell volume profiles were deemed satisfying in view of the accuracy of the then–available experimental measurements. This paved the way for similar applications to age–size PBEs in closely related research fields (see Shah and Ramkrishna's [126] article for an application to droplet morphological profiles). How to capture the aforementioned biological complexity in a PBE is still an open problem though, mostly because the bacteria's adaption to their medium is still experimentally intangible. Predictions of

2. These will be put aside in this work

the composition of a medium go beyond an accurate description of the cells' mechanics, prompting the need to somehow introduce some biochemistry in a population balance model. To this end, different strategies were carried out to make headway towards the modelling of industrial bioreactors, the accomplishments thereof making up the next section. Size–structured models have been a substantial source of inspiration in the bioreactor modelling community since the 1960s Tsuchiya & al. [138] and Fredrickson & al.'s [31] seminal articles. When *E. Coli* cells are addressed, size can be refered to as mass (cf. Subramanian & al.'s [132]), volume (as in Bell & Anderson's 1968 article [4] or Stamatakis's 2009 work [129]), or even cytoplasmic content (most notably by Mantzaris, whom proposes in [83] a differentiation with respect to one generic variable that can define RNA or protein content). In this context, both growth and rupture functions have been picked among polynomials [82], enforcing the assumption that the size distribution is supported over the whole $\mathbb{R}_+$.

In (1.22), the right–hand side's integral involves a redistribution kernel which formulation is an attempt to fit experimental measurements. Inductive reasonings in addition to ever more sensitive data have fueled the chemical engineering literature dealing with population balances. Different redistribution kernels have been tested: dirac (in [4] for example), beta (see Hatzis & al.'s [39] article or Fadda & al.'s [22] contribution) and even normal (notably by Henson [41] and Zhang & al. [151], although a gaussian law is well–known to be supported over $\mathbb{R}$ as a whole, giving some weight to negative values of a positive quantity (mass in Henson's case).

When age is considered, a standard refinement of the classical (1.22) equation is a multistaged approach involving a set of coupled PBEs pertaining to the different phases of the cell cycle (cf. Hatzis & al.'s [39] approach). In Billy & al.'s [6], experimental and numerical data have led to the conclusion that the duration of the consecutive phases is well modelled using a gamma law. On the other hand, formulating an age–dependent rupture function has been deemed an inaccuracy by Robert & al. [119] when compared with a size–dependent counterpart, the current trend being a turning away from age only–PBE.

As mentioned by Subramanian & al. [132], "a single parameter such as cell mass or cell age cannot adequately represent the physiological state of the organism if it is desired to take into account the fact that the response of a cell to the external environment depends on the cell's origin". Researchers have been well aware that more biology needs to be introduced in the modelling to allow a multiphase simulation of industrial fermenters, different pathways having been explored over the last decades.

The coupling with the culture medium

In the perspective of coupling the biotic phase with the abiotic phase, one idea would lie in so–called cybernetic modelling (see Kompala & al.'s [53, 54] and Young & al.'s [149]) and consist in predicting the dominance of complementary/competing metabolic pathways in a certain culture condition among a user–set amount of flux modes with the aim of optimising an objective function. More precisely, cybernetic variables (traditionally the fraction of resources allocation and the activation of the said resources, as mentioned by Straight & Ramkrishna [131]) aim at favouring the fluxes with the highest return on investment in terms of growth rate. This strategy is connected to

advances in genome–scale metabolic models and has to be devoted a substantial amount of computational time in any numerical implementation of PBEs in biology.
Another option would be a stochastic treatment of this biochemical burden that provides degrees of freedom on the modelling of unmeasurable quantities like gene expression levels (cf. Marguet & al.'s work in progress [85]). In this case, engineering evolution equations for each of these unmeasurable coordinates is still in the to–do list, most of the variability in these quantities bearing upon the sole generational redistribution.

Finally, the derivation of data–fitting single-cell models has been so far the most popular attempt at reducing this biochemical complexity. In this case, the structuring of the population involves a user–set number of intracellular metabolites which mass balance is thought to massively impact the cell cycle or provide an insight into experimental observations. Mantzaris's [81] for example, consecrates a lac–permease expression level–structured model to study the dynamics of the distribution of lac operon activity. Henson's [42] application to yeasts respiration consists in tracking intracellular chemical compounds which dynamics are dictated by interphase fluxes, the fitting parameters emerging from experimental measurements. This viewpoint has been further deepened by Stamatakis [129] with the aim of reducing the computational time of a N_d–PBE featuring size and $N_d - 1$ chemical compounds through a transformation into a size–PBE and a transport equation satisfied by the intracellular concentrations which solution can be calculated with the method of characteristics. Traditionally, a metabolic model must be formulated in order to close the PBE, the stoichiometry thereof coming from macroscopic mass balances with little consideration to the chemical feature of the reactions. A historical review of this strategy is presented by Nielsen & Villadsen [101].

Overall, such approaches would be excessively time–consuming if the vector of inner coordinates was to make room for all the growth–inducing chemical compounds; indeed, ξ would then have to include the sources of carbon (glucose, lactose, acetate), of phosphore (ATP/ADP), the co–enzymes, ... Here, the coupling with transport equations describing the motion of biological particles in the physical space is out of reach beacuse the dimension of the vector of inner properties is too large. In this context, further model reductions aiming at reporting reactor–scale outcomes of these metabolic maps through ad hoc quantities have been of seminal importance lately. One of these quantities (see Morchain & Fonade's [93], Morchain & al.'s [94] and references therein) is the commonly–named specific growth rate and is defined as the population–averaged gain in mass, what can be seen as an attempt at intertwining the cell and population scales. In [94], the circumvention of the microscopic complexity allows an integration of a one-dimensional PBE with a minimum number of classes and a large enough time step to implement (1.22) in a CFD code.

Another source of interest in the chemical engineering community pertains to the influence of the level of mixing on the substrate availability to the biotic phase. A continuous reactor is traditionally fed at one point by a highly concentrated solution which monomers diffuse throughout the system in a second phase. In this context, the macromixing consists in an evaluation of the abiotic phase's scattering in the fermenter and the micromixing in its homogeneity in an infinitesimal domain of the system. Cellwise, the available substrate at an individual's neighbourhood can be enough to satisfy its needs, what will be christened "biological regime" or a bad level of macro/micromixing can cause local scarcities and hamper the cells' cruise speed: this is labelled "physical regime". A thorough topical discussion can be found in [92]. Even though mixing issues are unlikely to cause

problems in a 1L bioreactor, their massive influence on 10^5L industrial–scale cultures are significant enough to require a particular attention in any modelling of biological reactors.

1.6.2 Objectives

The general objective in this work is to propose a population balance equation based modeling framework which combines a description of the nutrient uptake dynamics at the cell scale, the prediction of growth and products rates based on a metabolic model and the adaptation of internal cell properties in response to major changes in the cell environment.

Attention will be paid to the predictive capacity of the whole model. In other words, one must avoid or eliminate any population–based information when defining the cells' dynamics. In this context, a fundamental distinction between the cell potential and effective rates will be introduced. Rather than the usual formulation, based on concentrations in each phases, previous works showed than reasoning in terms of fluxes is much more sounded. Heterogeneous catalysis concepts such as external (physical) vs. internal (biological) limitation, introduced in previous works of the group, [95, 111] allow a general definition of limitation without referring to arbitrary constants.

The second chapter is based on an inductive mathematical approach to the topic. From general modelling assumptions the literature has consecrated (compactly supported distribution in size, singular rupture function), the well–posedness of the problem is first addressed. The order relation between measured mean doubling time and τ_d is also proven, along with exact relationships linking the moments of the distributions in age and interdivision time in batch and continuous culture.
The third chapter features an attempt to capture this biological burden into an all–terrain model. The bacteria's potential uptake rate and potential growth rate are no longer directly coupled to the environment. A sub–grid mixing model is used to formulate a closure law at the cell–liquid interface. Once the effective uptake rate is calculated, the fate of the uptaken carbon is obtained via a rudimentary metabolic model. The resulting cell–scale model is submitted to litmus tests inspired by the literature.
The fourth chapter is devoted to the numerical tools that are implemented from the perspective of simulating the abovementioned tests. Lagrangian and Eulerian methods are both required to cover a variety of experiments aiming at putting the model to the test.
The fifth chapter contains the simulation of various types of classical microbiology experiments which serve as a validation of the proposed approach. When possible a quantitative comparison is made. In other cases, simulations reveal some trends which require a dedicated experiment to assess their validity.

Chapter 2

Mathematical analysis of different structured models

2.1 Age–structured models (reference [115])

The main analytical results will be provided with the joint article. In order to introduce it, ξ will be a two–dimensional vector $\xi = (l, a)$, with $a \in \mathbb{R}_+$ and l upper–bounded to report the absence of macroscopic size bacteria. Imposing such a physical limit does not impact the following results in any way, the subtleties behind this modelling assumption being the topic of the next section. The (1.22) system reads in this context:

$$\begin{aligned} &\frac{\partial}{\partial t}n(t,l,a) + \frac{\partial}{\partial l}\left[\dot{l}n(t,l,a)\right] + \frac{\partial}{\partial a}n(t,l,a) + \gamma(l,a)n(t,l,a) + Dn(t,l,a) = 0 \\ &n(t,l,0) = 2\int_0^\infty \int_l^{\bar{l}} P(l,l',a')\gamma(l',a')n(t,l',a)dl'da' \ , \ n(t,l,a)|_{a=\infty} = 0 \text{ and } \dot{l}n(t,l,a)_{l\in\{0,\bar{l}\}} = 0 \\ &n(0,l,a) = n_0(l,a) \end{aligned} \tag{2.1}$$

From the PDE, one can define two PDFs. Firstly, the age distribution f corresponds to the frequency of cells of a certain age in a population:

$$f(t,a) = \frac{\int n(t,l,a)dl}{\iint n(t,l,a)dlda} = \frac{N(t,a)}{N(t)} \tag{2.2}$$

where $N(t,a) = \int n(t,l,a)dl$ and $N(t) = \iint n(t,l,a)dlda$. Its moments will be christened $\langle a^k \rangle$, $k \geq 1$.

The interdivision time distribution time g is also understood as a frequency, but the latter is taken as the fraction of cells which divide at a certain age among the cells which divide at any time in the population. γ being the cell-division frequency, the definition of g becomes obvious:

$$\begin{aligned} g : \mathbb{R}_+ &\to \mathbb{R}_+ \\ a \mapsto g(a) &= \frac{\int \gamma(l)n(t,l,a)dl}{\iint \gamma(l)n(t,l,a)dlda} \end{aligned} \tag{2.3}$$

Its moments will be denoted by $\langle \tau^k \rangle = \int_{\mathbb{R}_+} a^k g(a), da$, $k \geq 1$. The interdivision time is defined as the cell age at rupture, which corresponds to the measured cell-cycle duration in experiments. In other

words, $g(a)da$ is the probability that a cell which divides in the reactor does so between age a and $a+da$, not the probability that a cell divides between age a and $a+da$, the latter being tantamount to a life expectation distribution.

Setting h the life expectancy distribution, the differences between g and h will manifest themselves in the article, and a protracted debate that regards the treatment of experimental measurements will be put to an end. Analytical formulae derived from the relationship between f and g will further enhance the theory of age–structured models.

A takeaway from this publication is the crucial impact of a reactor's dilution rate on a population's observed biological outputs. The higher the dilution rate, the lower the mean interdivision time, what triggers a massive selection of the extant cells according to the health criterion. In particular, two continuous cultures of the same strain at a different dilution rate filter different fractions of the population and make any comparison quite dubious.

A critical analysis of Powell's results on the interdivision time distribution

Vincent Quedeville[1,2], Jérôme Morchain[2,*], Philippe Villedieu[3,4], and Rodney O. Fox[5]

[1]FERMAT, Université de Toulouse, CNRS, INPT, INSA, UPS, Toulouse, France
[2]Laboratoire d'Ingénierie des Systèmes biologiques et des Procédés, INSA de Toulouse, 135 Avenue de Rangueil, 31400 Toulouse, France. * Correspondence should be adressed to J.M. (email: jerome.morchain@insa-toulouse.fr)
[3]Institut de Mathématiques de Toulouse, Université de Toulouse, F-31062 Toulouse, France
[4]ONERA/DMPE, Université de Toulouse, F-31055 Toulouse, France
[5]Department of Chemical and Biological Engineering, Iowa State University, 618 Bissell Road, Ames, Iowa 50011-1098, USA

October 11, 2019

Abstract

The cell-age and interdivision-time probability density functions (PDFs) have been extensively investigated since the 1940s due to their fundamental role in cell growth. The pioneering work of Powell established the first relationship between the interdivision-time and cell-age PDFs. In the literature, two definitions for the interdivision-time PDF have been proposed. One stands for the age-at-rupture PDF and is experimentally observable, whereas the other is the probability density that a cell divides at a certain age and is unobservable. From Powell's results pertaining to the unobservable interdivision-time PDF, Painter and Marr derived an inequality that is true but is incorrectly used by experimentalists to analyse single-cell data. Unfortunately, the confusion between these two PDFs persists. To dissipate this confusion, exact relationships between the cell-age and the interdivision-time PDFs are derived in this work from an age-structured model, which can be used by experimentalists to analyse cell growth in batch and continuous culture modes.

1 Background

Understanding biological population dynamics in a fermenter has been of crucial importance in bioprocess engineering and many other fields, such as pharmacology, that require a mass production of metabolic by-products. A strain will grow differently in a batch or continuous fermenter, to the extent that one population will exhibit different characteristics depending on the culture conditions and the observed features cannot be compared. In particular, in an open system, the fermenter dilution rate, D, will determine the ensemble-averaged behaviour, such as the mean age or mean interdivision time, in other words the cell-cycle duration. As early as 1956 Powell [14] hinted at the seminal conclusion that a continuous culture's observed mean interdivision time, $\langle\tau\rangle$, must be less that the so-called population doubling time, meaning that as soon as the interdivision-time

distribution is asymmetric, the healthier cells will contribute more to maintaining a steady-state cell number than their less active counterparts. To date, that article has been cited in 372 research works, with significant interest from mathematicians [24], physicists [9], chemists [10] and biologists [1] on a variety of perspectives pertaining to the cell-cycle dynamics and the marginal distributions in different observable properties such as age, size or cell content. In the last decade, the development of microfluidic devices has broadened the biologists' horizons and given more accurate statistical information regarding the cell-cycle processes [11,22,25,26], allowing modelling assumptions to be tested against experimental results. However, Powell's logical reasoning leading to $\langle\tau\rangle \leq \ln 2/D$ is not a consensus view in the mathematical modelling community; indeed, in 1967 Painter & Marr [12] demonstrated the exact opposite inequality starting from Powell's work and no one has, to the authors' knowledge, questioned this assessment to date. If anything, Painter & Marr's demonstration has paved the way for experimental and analytical work, i.e., [3,19,25], attempting to consolidate Painter & Marr's viewpoint.

The very notion of interdivision-time distribution can embrace different quantities in spite of a common definition of the concept (i.e. the time elapsed between two consecutive division events of an observed organism), and no consensus has been reached to date on the relationship between these quantities. Consequently, some semantics are required to provide a framework for the analytical results presented in this work and for their comparison with experimental data.

- A cell's age is defined as the time elapsed since the division event that produced it (that is the age in the cycle), entailing that the quantity is reset to zero after each recorded rupture. It does not encompass the lineage's longevity that will be called the "age in the system". The latter is tantamount to the abiotic phase's lifespan because this time interval is just the residence time in the fermenter.

- An observable interdivision-time distribution, g, stands for a collection of recorded cell-cycle durations for which labels such as "generation time" [14] or "doubling time" [11] exist in the literature. The vocabulary is here borrowed from [26] so that interdivision time will be synonymous with the cell's age at rupture. As a consequence, g refers to Powell's so-called "carrier distribution" and $g(a)da$ defines the conditional probability that a cell that has divided has done so between age a and $a+da$. It is essential to notice that g captures the Markovian nature of the cell cycle process and reports the fact that a cell has reached age a in the system. Throughout the article, $\int ag(a)da$ will be called τ_{obs}.

- The unobservable interdivision-time distribution h, used by Powell [14], is such that $h(a)da$ is the probability that a newly formed organism will have a generation time in the range $[a, a+da]$. Its first moment $\int ah(a)da$ will be called τ_{uno}.

Powell sheds light on the dissimilitudes between these two interdivision-time distributions. This distinction is relevant for both batch and continuous conditions for different reasons. In an open fermenter, a cell's biological development must be considered along with its residence time so an organism's interdivision time could refer to an unobservable and, hence, unmeasurable event (for instance from statistical considerations, from any steady-state group of tracked particles, half of them will be washed out before dividing). In batch culture, the younger elements outnumber their ancestors due to the biotic phase's exponential growth and the statistical extra weight conferred on the less probable quicker interdivision times over a much larger share of the population pushs the age-at-rupture distribution to the left. Experimentally, the available data regarding interdivision

time pertain to the age at rupture, and the PDF's first moment is well approximated by the data set's arithmetic mean, provided the collected data set is large enough. In 1956, Powell [14] claimed he did produce an experimental equivalent for h and fitted the histogram with a Pearson type-III distribution. However, he remarked in 1964 [15] that "the generation times of the organisms which have, at a given time, completed their life span during the previous history of the culture do not compose" h; "they compose the carrier distribution" g. In fact, only information regarding the cell's age at rupture is available to experimentalists and, hence, it cannot be interpreted with analytical results intended for the unobservable interdivision-time PDF h. In 1967, Painter & Marr [12] extracted a lower bound on the first moment of Powell's interdivision-time distribution τ_{uno} and confused it with the observable mean interdivision time τ_{obs}, prompting some equivocal assertions (a very good recent example being [25]) by lack of consensus. Keeping these considerations in mind, this work aims to reconcile the persistent misunderstanding about these distinct paradigms, and on presenting exact analytical results regarding the observable interdivision-time distribution that are accessible to experimentalists. To illustrate the most important points, numerical examples are provided from Monte–Carlo simulations of a population balance model. Furthermore, unlike [13] or [22] where the so-called "timer" or "adder" models are given prominence, by virtue of [18] the rupture process will be assumed to be a function of the cell length.

The first part of the paper presents the general framework of population balance modelling in the context of microbial populations, the particular population balance equation (PBE) chosen for the present study and the mathematical definition of the age and observable interdivision-time distributions. The second part is devoted to analytical results leading to relationships valid for the age and interdivision-time PDFs that are observable from batch and chemostat experimental measurements. These analytical results are further underpinned by numerical simulations using a Monte–Carlo algorithm. In the discussion, the results from the previous section are compared to experimental data from the literature. A resolution of the seemingly contradictory conclusions in Powell's and Painter and Marr's works is provided.

2 Mathematical background and definitions

2.1 General formulation of a PBE for biological populations

Beginning with work in the 1960s [5, 20], PBEs have provided a general framework to describe the biological response to a user-defined experimental set-up. In this context, an inner coordinate is understood as a marginal variable and its law is retrieved through integration with respect to all other dimensions. When biological modelling of a continuous fermenter is addressed, the age-structured PBE takes the form

$$\begin{gathered}\frac{\partial}{\partial t}n(t,a,\boldsymbol{\xi})+\frac{\partial}{\partial a}n(t,a,\boldsymbol{\xi})+\boldsymbol{\nabla}\cdot\left[\dot{\boldsymbol{\xi}}n(t,a,\boldsymbol{\xi})\right]+\gamma(a,\boldsymbol{\xi})n(t,a,\boldsymbol{\xi})+D\,n(t,a,\boldsymbol{\xi})=0\\ n(t,0,\boldsymbol{\xi})=2\int_{\Omega_\xi}\int_0^\infty\gamma(a',\boldsymbol{\xi'})K(\boldsymbol{\xi},\boldsymbol{\xi'},a')n(t,a',\boldsymbol{\xi'})da'd\boldsymbol{\xi'}\\ \dot{\boldsymbol{\xi}}n(t,a,\boldsymbol{\xi})|_{\boldsymbol{\xi}\in\partial\Omega_\xi}=0,a\in[0,+\infty[\end{gathered}\tag{1}$$

with $\boldsymbol{\xi}\in\Omega_\xi\subset\mathbb{R}^n$, $n\geq 1$, the vector of inner coordinates, $\dot{\boldsymbol{\xi}}$ their rate of change, K the redistribution kernel, and $n(t,\boldsymbol{\xi})d\boldsymbol{\xi}$ the cell number in an infinitesimal domain of Lebesgue measure $d\boldsymbol{\xi}$. In (1) γ (time unit^{-1}) is the rupture function, or simply the cell-division frequency, and D (time unit^{-1})

stands for the so-called dilution rate that drives both the input feed and cell washout to maintain the medium volume. In general, $\dot{\boldsymbol{\xi}}$ is a function of both $\boldsymbol{\xi}$ and the organisms' environment, but its formulation has no impact on the section's results. Hereinafter, D will be assumed constant and washout is assumed uniform with respect to any inner coordinate (i.e., the fermenter is perfectly mixed).

In a continuous fermenter, an equilibrium will be reached when the time derivative in (1) vanishes and this condition is referred to as steady state. In a batch fermenter, the absence of a washout term will allow the cell number to grow at will as soon as the initial conditions have faded away. This equilibrium is thoroughly discussed in [20] and will be referred to as self-similar exponential growth. The time derivative in (1) does not vanish in this case, and the stability property will relate to the marginal distribution's geometrical shape. In other words, the scaled quantity

$$\frac{n(t,a,\boldsymbol{\xi})}{\int_{\Omega_\xi} n(t,a,\boldsymbol{\xi})d\boldsymbol{\xi}}$$

will be constant (or self similar) for any $\boldsymbol{\xi} \in \Omega_\xi$.

2.2 Application to *E. coli* population dynamics

Without loss of generality, this section will consecrate a two-dimensional PBE, $\boldsymbol{\xi}$ standing for the cells' length $l \in [0, \bar{l}[$ (m), where $\bar{l}$ is the maximum possible cell length before division. In this section, no laws for $\dot{l}$ are yet required. Nonetheless, it is understood that such a process must be a decreasing function of l since it involves the internal transport of membrane proteins from the cytoplasm, which takes longer as the cell grows larger. Indeed, as noted by Nobs & Maerkl [11], synthesis of cell-membrane components could be one such factor setting limits on the cell-doubling time, what seems universal enough to feature in any biological population modelling. Other patterns are conceivable though (as mentioned in [25]) but the lack of experimental data makes any consensus unattainable.

Also, the redistribution kernel obeys $\int_0^{l'} P(l,l',a')dl = 1$, which is the mathematical counterpart of the biological hypothesis that a given cell give birth to only two daughter cells during the division process.

Hence, (1) reads in this case:

$$\begin{gathered}\frac{\partial}{\partial t}n(t,l,a) + \frac{\partial}{\partial l}\left[\dot{l}n(t,l,a)\right] + \frac{\partial}{\partial a}n(t,l,a) + \gamma(l,a)n(t,l,a) + D\,n(t,l,a) = 0 \\ n(t,l,0) = 2\int\int_{l'>l} \gamma(l',a')P(l,l',a')n(t,l',a')dl'da' \\ \dot{l}n(t,l,a)|_{l=0} = 0 = \dot{l}n(t,l,a)|_{l=\bar{l}}\end{gathered} \tag{2}$$

In the system (2), a null-flux condition is assumed at the domain boundary in length, what is tantamount to the claim that no cell can grow beyond a certain length that challenges its biomechanical structure. From physical grounds, it will similarly be assumed that no cells will reach infinite age.

2.3 Definition of PDFs

Considering that $n(t,l,a)$ refers to the number density of cells with length l, age a at time t in a continuous reactor and γ is the cell-division frequency, the function

$$\begin{aligned} g : \mathbb{R}_+ &\to \mathbb{R}_+ \\ a &\mapsto g(a) = \frac{\int \gamma(l,a) n(t,l,a) dl}{\iint \gamma(l,a) n(t,l,a) dl da} \end{aligned} \tag{3}$$

will designate the interdivision-time PDF as it is observed in experimental measurements. Its moments are denoted by $\langle \tau^k \rangle = \int a^k g \, da$, $k \geq 1$, $\langle \tau^1 \rangle$ coinciding with τ_{obs}. It is brought to the reader's attention that $g(a)da$ is not the probability that a cell divides between age a and $a + da$. Instead, g denotes what Powell called the carrier distribution $\mathcal{C}_D$ in his 1956 article [14] and corresponds to the observed cell-cycle duration.

Furthermore, the cell-age PDF f can be retrieved by integrating (2) with respect to l:

$$f(t,a) = \frac{\int n(t,l,a) dl}{\iint n(t,l,a) dl da} = \frac{N(t,a)}{N(t)} \tag{4}$$

where $N(t,a) = \int n(t,l,a) dl$ and $N(t) = \iint n(t,l,a) dl da$. Thus $f(a)da$ is the probability that a cell in the reactor has an age between a and $a + da$, and f is therefore tantamount to Powell's ϕ in his 1956 article [14].

Comparing (3) to (4), we observe that the interdivision-time PDF is weighted by the cell-division frequency, while the cell-age PDF is not. At steady state, or under self-similar conditions, both PDFs will be independent of t.

3 Analytical and numerical results

In this section, we establish exact results concerning the interdivision-time and cell-age PDFs arising from the solution to the PBE introduced above.

3.1 Steady-state relation between f and g in a continuous fermenter

From the definition of f provided in (4), one gets:

$$\frac{\partial}{\partial t} f(t,a) = \frac{1}{N(t)} \frac{\partial}{\partial t} N(t,a) - \frac{f(t,a)}{N(t)} \frac{\partial}{\partial t} N(t) \tag{5}$$

The first term on the right-hand side of (5) is obtained through an integration of (2) with respect to l, i.e.

$$\int \frac{\partial}{\partial t} n(t,l,a) dl + \int \frac{\partial}{\partial l} \left[\dot{l} n(t,l,a) \right] dl + \int \frac{\partial}{\partial a} n(t,l,a) dl + \int \gamma(l,a) n(t,l,a) dl + D \int n(t,l,a) dl = 0 \tag{6}$$

In (6), the first term on the left-hand side designates $N(t,a)$'s time derivative and the null-flux boundary condition forces the second term to vanish. Furthermore,

$$\int \frac{\partial}{\partial a} n(t,l,a) dl = \frac{\partial}{\partial a} N(t,a) \quad \text{and} \quad D \int n(t,l,a) dl = D \, N(t,a)$$

Hence,

$$\frac{\partial}{\partial t}N(t,a) + \frac{\partial}{\partial a}N(t,a) + \int \gamma(l,a)n(t,l,a)dl + D\,N(t,a) = 0 \tag{7}$$

The second term on the right-hand side of (5) is retrieved from the double integral of (2):

$$\iint \frac{\partial}{\partial t}n(t,l,a)dlda + \iint \frac{\partial}{\partial l}\left[\dot{l}n(t,l,a)\right]dlda + \iint \frac{\partial}{\partial a}n(t,l,a)dlda \\ + \iint \gamma(l,a)n(t,l,a)dlda + D\iint n(t,l,a)dlda = 0$$

The same reasoning as before entails the conclusion that the first term on the left-hand side is in fact $N(t)$'s time derivative and the second term is null. Use of Fubini's theorem and the fact that there is no cell with an infinite age in the system turns the third term into

$$\int\int \frac{\partial}{\partial a}n(t,l,a)dadl = \int [n(t,l,a)]_{a=0}^{\infty}\,dl = -2\int\iint \gamma(l',a')P(l,l',a')n(t,l',a')dl'da'dl \\ = -2\iint \gamma(l,a)n(t,l,a)dlda \tag{8}$$

Consequently,

$$\frac{\partial}{\partial t}N(t) - \iint \gamma(l,a)n(t,l,a)dlda + D\,N(t) = 0 \tag{9}$$

Combining (9) and (7) in (5) and referring to the definition of f in (4) yields an equation for the time evolution of f :

$$\frac{\partial}{\partial t}f(t,a) = -\frac{\partial}{\partial a}f(t,a) - \frac{\int \gamma(l,a)n(t,l,a)dl}{N(t)} - \frac{f(t,a)}{N(t)}\iint \gamma(l,a)n(t,l,a)dlda \tag{10}$$

A steady-state relationship between f and g can be derived from this equation. Indeed, at steady state, f's derivative with respect to time vanishes (removing the time dependence) and it also follows from (9) that

$$\iint \gamma(l,a)n(l,a)dlda = DN \tag{11}$$

Therefore, from (10)

$$\frac{df}{da}(a) = -\frac{\int \gamma(l,a)n(l,a)dl}{N} - Df(a) = -\frac{\int \gamma(l,a)n(l,a)dl}{\iint \gamma(l,a)}n(l,a)dlda\frac{\iint \gamma(l,a)n(l,a)dlda}{N} - Df(a) \tag{12}$$

The last step consists in using (3), the definition of g, and (11) in (12) to get the desired relationship:

$$f'(a) = -Dg(a) - Df(a) \tag{13}$$

3.2 An analytical solution for the cell-age distribution at steady state in a continuous fermenter

The differential equation (13) can be solved using Duhamel's formula and yields

$$f(a) = f(0)e^{-Da} - De^{-Da}\int_0^a e^{Da'}g(a')da' \tag{14}$$

One only needs to provide $f(0)$ to completely define f. From the definition of f, given in (4), the null-age relation provided in (2) and the steady-state relation (11), the boundary condition reads:

$$f(0) = \frac{\int n(l,0)dl}{\iint n(l,a)dlda} = \frac{2\int\iint \gamma(l',a')P(l,l',a')n(l',a')dl'da'dl}{N} = \frac{2DN}{N} = 2D \tag{15}$$

Consequently the cell-age distribution at steady state reads:

$$f(a) = 2De^{-Da} - De^{-Da}\int_0^a e^{Da'}g(a')da' \tag{16}$$

This result extends Ramkrishna's [17] work dealing with analytical and numerical solutions of age and size PBMs in a closed bioreactor. It is worth mentioning that (15) and (16) echo Powell's equation (9) [14]. Both derivations complement each other since the cell-age PDF definitions are in fact identical (Powell's ϕ is equivalent to our f). However, Powell's results involve an interdivision-time distribution that is unobservable from experiments contrary to our g.

3.3 For a continuous fermenter at steady state $\tau_{obs} \leq \ln(2)/D$.

This result is obtained from rearranging (16) and taking the limit $a \to \infty$. One can first check that the application $a :\mapsto \exp(Da)f(a)$ is strictly decreasing on $\mathbb{R}_+$. Indeed:

$$\frac{d}{da}(\exp(Da)f(a)) = D\exp(Da)f(a) + \exp(Da)(-Df(a) - Dg(a)) = -D\exp(Da)g(a) < 0$$

since g is strictly positive on $\mathbb{R}_+$. As 0 is an obvious lower bound to $a \mapsto \exp(Da)f(a)$, the latter converges to a finite limit $\lambda \geq 0$. As a consequence:

$$\int_0^a e^{Da'}g(a')da' = 2 - \frac{1}{D}e^{Da}f(a) \Rightarrow \lim_{a\to\infty}\int_0^a e^{Da'}g(a')da' = 2 - \frac{\lambda}{D}$$

Then, developing the exponential into a power series and making use of Jensen's inequality leads to:

$$\sum_{k\geq 0}\frac{D^k}{k!}\int_0^\infty a'^k g(a')da' = \sum_{k\geq 0}\frac{D^k}{k!}\langle\tau^k\rangle = 2 - \frac{\lambda}{D} \geq \sum_{k\geq 0}\frac{D^k}{k!}\tau_{obs}^k$$

and thus to

$$2 - \frac{\lambda}{D} \geq e^{D\tau_{obs}} \Leftrightarrow \tau_{obs} \leq \frac{\ln(2-\frac{\lambda}{D})}{D} \leq \frac{\ln(2)}{D} \tag{17}$$

The last inequality does not prevent τ_{obs} from being equal to $\ln(2)/D$, which would happen if all moments $\langle\tau^k\rangle$ were equal to τ_{obs}^k. This would basically force g to be a Dirac delta function: $\delta_{a-\ln(2)/D}$. In this case though, the observable and unobservable distributions are identical and mirror the behaviour of an unstructured model. A preliminary conclusion was first formulated by Tyson & Hannsgen [24], but the authors missed Powell's [15] remark pertaining to the difference between the two interdivision-time distributions, preventing their result from being applicable to actual experimental data.

An additional conclusion that stems from (14) is that the outlet-age profile (that must be tantamount to the fermenter's because of the uniform washout assumption) differs significantly from the liquid phase's (i.e., De^{-Da}), because the biological phase renewal is a consequence of two competing phenomena: dilution and cell division. A graphic comparison between the two residence time distributions is shown in figure 1.

3.4 For a continuous fermenter at steady state $\langle a\rangle + \tau_{obs} = 1/D$.

Taking the first moment of (16) yields

$$\begin{aligned}
\langle a\rangle = \int_0^\infty a f(a) da &= 2D\int_0^\infty a e^{-Da} da - D\int_0^\infty a e^{-Da}\int_0^a e^{Da'} g(a') da' da \\
&= 2\int_0^\infty e^{-Da} da - D\int_0^\infty e^{Da'} g(a')\int_{a'}^\infty a e^{-Da} da da' \\
&= \frac{2}{D} - \int_0^\infty e^{Da'} g(a')\left(-a e^{-Da}\big|_{a'}^\infty + \int_{a'}^\infty e^{-Da} da\right) da' \\
&= \frac{2}{D} - \int_0^\infty a g(a) da - \int_0^\infty \frac{1}{D} e^{-Da} e^{Da} g(a) da \\
&= \frac{2}{D} - \tau_{obs} - \frac{1}{D}
\end{aligned}$$

and thus,

$$\langle a\rangle + \tau_{obs} = \frac{1}{D} \tag{18}$$

Using (17), an upper bound to τ_{obs} is obtained:

$$\tau_{obs} \leq \frac{\langle a\rangle \ln(2)}{1-\ln(2)} \approx 2.259\langle a\rangle \tag{19}$$

The same reasoning yields a relation between the second-order moments of f and g:

$$\begin{aligned}
\langle a^2\rangle &= \int_0^\infty a^2 f(a) da \\
&= 2D\int_0^\infty a^2 e^{-Da} da - D\int_0^\infty a^2 e^{-Da}\int_0^a e^{Da'} g(a') da' da \\
&= -2a^2 e^{-Da}\big|_0^\infty + 2\int_0^\infty 2a e^{-Da} da - \int_0^\infty e^{Da'} g(a')\int_{a'}^\infty D a^2 e^{-Da} da da' \\
&= \frac{4}{D^2} - \int_0^\infty e^{Da'} g(a')\, -a^2 e^{-Da}\big|_{a'}^\infty da' - \int_0^\infty e^{Da'} g(a')\int_{a'}^\infty 2a e^{-Da} da da' \\
&= \frac{4}{D^2} - \langle\tau^2\rangle - \int_0^\infty e^{Da'} g(a')\, -\frac{2}{D} a e^{-Da}\Big|_{a'}^\infty da' - \int_0^\infty e^{Da'} g(a')\int_{a'}^\infty \frac{2}{D} e^{-Da} da da' \\
&= \frac{4}{D^2} - \langle\tau^2\rangle - \frac{2}{D}\tau_{obs} - \frac{2}{D^2} \\
&= \frac{2}{D^2} - \langle\tau^2\rangle - \frac{2}{D}\tau_{obs}
\end{aligned}$$

yielding:

$$\langle a^2\rangle = \frac{2}{D}\langle a\rangle - \langle\tau_{obs}^2\rangle \tag{20}$$

Consider now an age-synchronised population, i.e., no variance in age is observed ($\langle a^2\rangle = \langle a\rangle^2$). Then using (18) and (20) one can determine whether a non-zero variance can exist in the interdivision-time distribution.

$$\langle\tau^2\rangle - \tau_{obs}^2 = \frac{2}{D}\langle a\rangle - \langle a^2\rangle - \frac{1}{D^2} - \langle a\rangle^2 + \frac{2}{D}\langle a\rangle = -2\langle a\rangle^2 + \frac{4}{D}\langle a\rangle - \frac{1}{D^2}$$

The second-order polynomial would vanish for $D\langle a\rangle \in \left\{1-\sqrt{2}/2, 1+\sqrt{2}/2\right\}$, the latter value being impossible given that $1+\sqrt{2}/2 > 1$. However, if $\langle a\rangle$ were equal to $(1-\sqrt{2}/2)D$, the mean interdivision time τ_{obs} would be $\frac{\sqrt{2}}{2D} > \frac{\ln(2)}{D}$, which is not possible according to (17). In other words, an age-synchronised steady-state population has to exhibit some variance in its interdivision time. As a consequence it can not remain age synchronised in a continuous fermenter, a result that was already conjectured by Yasuda [26]. This well-known result was also thoroughly discussed in [2,8].

3.5 For a self-similar batch fermenter $\tau_{obs} \leq \langle a\rangle \ln(2)/(1-\ln 2)$.

In a closed fermenter, (2) does not have a washout term and, as a consequence, reads

$$\frac{\partial}{\partial t}n(t,l,a) + \frac{\partial}{\partial l}\left[\dot{l}n(t,l,a)\right] + \frac{\partial}{\partial a}n(t,l,a) + \gamma(l,a)n(t,l,a) = 0$$

with the same boundary condition. Hence, $N(t)$'s dynamics take the form

$$\frac{dN(t)}{dt} = \iint \gamma(l,a)n(t,l,a)dlda$$

and $N(t,a)$ follows from the same reasoning as in the previous section:

$$\frac{\partial}{\partial t}N(t,a) + \frac{\partial}{\partial a}N(t,a) + \int \gamma(l,a)n(t,l,a)dl = 0$$

entailing f's dynamics:

$$\frac{\partial}{\partial t}f(t,a) = -\frac{\partial}{\partial a}f(t,a) - \frac{\int \gamma(l,a)n(t,l,a)dl}{N(t)} - \frac{f(a)}{N(t)}\iint \gamma(l,a)n(t,l,a)dlda \tag{21}$$

For self-similar growth,

- f must be independent of t, which forces (21)'s left-hand side to vanish.
- $\iint \gamma(l,a)n(t,l,a)dlda/N(t)$ reaches a constant value that was called ν_m by Powell.

Thus, for self-similar growth, (21) reads

$$\frac{df}{da}(a) = -\nu_m f(a) - \frac{\int \gamma(l,a)n(t,l,a)dl}{N(t)}$$

and the initial condition takes the form

$$f(0) = \frac{1}{N(t)}\int 2\iint \gamma(l',a')P(l,l',a')n(t,l',a)dl'da = 2\nu_m$$

Once again, by virtue of Duhamel's theorem,

$$f(a) = 2\nu_m e^{-\nu_m a} - \frac{1}{N(t)}e^{-\nu_m a}\int_0^a e^{\nu_m a'}\int \gamma(l,a')n(t,l,a')dlda' \tag{22}$$

The similarity between (22) and (16), with ν_m playing the same role as D, allows the immediate conclusion

$$\langle a\rangle + \tau_{obs} = \frac{1}{\nu_m} \tag{23}$$

and is accessible as soon as the cell-age and interdivision-time PDFs are measured. Furthermore, the same reasoning as in the previous paragraph yields the conclusion

$$\tau_{obs} \leq \frac{\ln(2)}{\nu_m} \Leftrightarrow \tau_{obs} \leq \frac{\langle a \rangle \ln(2)}{1 - \ln(2)} \tag{24}$$

which is the same relation between τ_{obs} and $\langle a \rangle$ as (19). The equality would hold if all cells were equally "healthy". If this situation cannot be strictly ruled out, it was not observed experimentally by Powell and is highly unlikely to occur.

3.6 Numerical examples.

In this part, which deals with *E. coli*, all cells will be assumed cylindrical with constant diameter d (m) (in accordance with [26]), so that both a cell's surface and volume are functions of l only. The same assumptions regarding the cell geometrical feature can also be made for *Bacillus subtilis* (as discussed in [25]). In order to put our results to the test, a comprehensive model must be formulated and simulated using either Eulerian or Lagrangian methods. We draw the reader's attention to the fact that Lagrangian methods allow the removal of the cell age from the PBE (1) because this very feature is accessible as soon as a cell is tracked in time. In fact, Monte–Carlo methods make the model one dimension smaller and, as a result, are preferable from a computational perspective. Our Monte–Carlo simulation aims at illustrating more complex metabolic features and involves more than two variables. Notwithstanding, this has no influence on the section's results dealing with age-related PDFs, because these extra variables can always be taken out through partial integrations. In the model used in our Monte–Carlo simulations, $\dot{\boldsymbol{\xi}}$ reads

$$\begin{aligned} \dot{l} &= \frac{q_S}{\rho V Y_{SX}} \left(1 - \frac{l}{\bar{l}}\right)^{\eta} \\ q_S &= q_{S_1} + q_{S_2} \quad \text{with dynamics:} \\ \dot{q_{S_1}} &= \frac{1}{\tau_1} \left[f_1(S) - q_{S_1}\right] \\ \dot{q_{S_2}} &= \frac{1}{\tau_2} \left[f_2(S) f_3(q_{S_1}) - q_{S_2}\right] \end{aligned}$$

where S is the substrate concentration in the fermenter, ρ a cell's mass density ($\sim 10^3$ kg/m^3), V its volume (a linear function of l), Y_{SX} (g/g) a (constant) substrate-to-mass ratio, and τ_1, τ_2 (hr) characteristic times of the respective mechanism's adaptation. The functions f_1 and f_2 are of Monod shape and associate S to respective q_{S_1} and q_{S_2}. The function f_3 serves at a restricting factor that aims at accounting for q_{S_1}'s inhibiting influence over q_{S_2} in accordance with [4]. This refinement aims at uncoupling the substrate uptake and lengthening at the cell scale, but is not needed for steady-state conditions. The model for $\dot{l}$ ensures that a cell divides before crossing the $l = \bar{l}$ border and the close-to-zero exponent guarantees that the lengthening phenomenon is almost linear with respect to l for most of the cell cycle.

Furthermore, the division frequency model is

$$\gamma(l) = \begin{cases} \frac{1}{T} \frac{(\bar{l}-l)^{\kappa} - (\bar{l}-l_{\text{inf}})^{\kappa}}{(\bar{l}-l_c)^{\kappa} - (\bar{l}-l_{\text{inf}})^{\kappa}} & \text{if } \; l_{\text{inf}} \leq l < \bar{l} \\ 0 & \text{if } \; l \notin [l_{\text{inf}}, \bar{l}[\end{cases}$$

Table 1: Parameter used in the simulations.

Parameter	Value	Description	Reference
D	$0.15\ \mathrm{hr}^{-1}$	Dilution rate	From experiment
l_{inf}	7×10^{-6} m	Minimal length at rupture	[11]
l_c	11×10^{-6} m	Standard length at rupture	[11]
$\bar{l}$	18×10^{-6} m	Maximal length at rupture	[11]
T	2 hr	Time scale in the cell division rate	Assumed
Y_{SX}	$1/0.42 \approx 2.38$ g/g	Substrate-to-mass ratio	[21]
τ_1	25 s	q_{S_1} characteristic time	Assumed
τ_2	5 s	q_{S_2} characteristic time	Assumed
d	10^{-6} m	Cell diameter	Assumed
η	0.05	Shape parameter	Assumed
κ	-0.96	Parameter	Assumed

with T (hr) a time constant, l_{inf} (m) the minimal length at rupture, and l_c (m) a characteristic division length.

The idea that γ depends only on l is borrowed from Robert & Al. [18]. Other assumptions have been investigated recently in the literature, such as an "adder" model [22], which seems less convenient from a numerical simulation perspective. Indeed, due to the non-equivalent redistribution in length at rupture, such a mechanism could allow fractions of the population to grow more and more for generations on end until non-physical cell lengths are encountered.

For completeness, the redistribution kernels in l and q_S are assumed independent, beta and symmetric. To explain the first hypothesis, it is inferred from raw experimental data for two different *E. coli* strains [23] that the growth rate and the length at birth are relatively independent . With little appropriated cell-scale information to the authors' knowledge, full uncorrelation was considered, easing the analysis of the model's sensibility to this factor. The parameters employed are given in table 1. It can be demonstrated that the inequality $-\kappa + \eta > 1$ entails the mathematical well-posedness of the problem. From physical grounds, this condition ensures that the rupture process overtakes lengthening as the cell length approaches the upper bound $\bar{l}$.

Other elongation rate formulations, including linear or exponential laws can be found in the literature [7, 18, 22, 25]. These laws are generally based on fitting single-cell measurements. In general none of these formulations suits the data better than the others [7]. Furthermore, Robert et al. evidence a sublinear elongation as the cell length approaches a critical value, what seems reasonable considering that it turns increasingly difficult for any organism to maintain their growth rate as feeding an ever-growing cell membrane at a constant rate would likely end up mustering more resources than is available to them. Also, from a practical point of view, it is worth noticing that any experimental device introduces a bias against the older cells that are also most probably the longest.

Modelling-wise, the linear and exponential formulations imply that nothing restrains the cell elongation. In any case, the choice of the lengthening rate model must be consistent with the division frequency in order to prevent the production of cells with an infinite length. Our $\dot{l}$ and γ respect this constraint, even though other combinations are valid as long as the above restriction is met. In the end, however, the analytical results derived in this work do not depend on any particular choice.

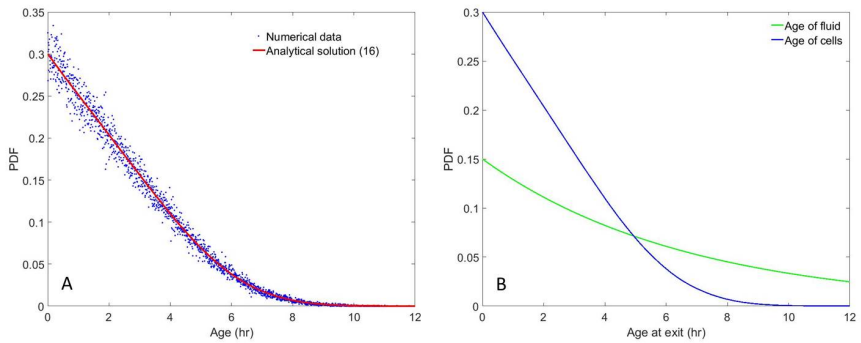

Figure 1: Steady-state results for continuous culture with $D = 0.15$ hr^{-1}. Left: Cell-age PDF f from the Monte–Carlo simulation (blue points), compared with the analytical solution (black line) (16) where $\langle a \rangle \approx 2.360$ hr. Right: Cell-age PDF (black line) compared with the fluid residence-time PDF (green line).

With these considerations in mind, the algorithm consists in tracking the cell's inner coordinates with respect to time, from random clipped-Gaussian initial samples. Then the division and washout events are determined by sampling two random numbers u, x:

- Let $u \sim \mathcal{U}_{[0,1]}$ sampled for each cell at each time step: mitosis occurs if $1 - e^{-\gamma(l)\delta t} < u$
- Let $x \sim \mathcal{E}(\frac{1}{D})$ sampled for each cell at birth: washout occurs should the cell's age be greater than x.

When a cell divides, its inner properties are redistributed according to the kernel K, and each new cell is given a residence time drawn from $\mathcal{E}(\frac{1}{D})$. The cell age is reset to zero for one of the daughter cells, making room for a new lineage in the fermenter, whereas the other daughter keeps the record of the mother-cell's lineage. All algorithms are coded in C++11 and the data are processed with Matlab R2016a.

3.7 Comparison between analytical and Monte-Carlo simulations results

The Monte–Carlo simulation reaches a steady state after 4 to 5 times the slowest characteristic time $D^{-1} \approx 6.667$ hr. From this point onwards, consecutive division events are recorded for 1,003,306 cells over the course of 37.5 hr. Around 50% (501,322) divide at least twice and 25% (250,402) divide three times or more. This substantial database yields a numerical accuracy of approximately 10^{-3} for estimating averages. As can be seen from figure 1, the steady-state cell-age PDF matches many well-known results (see [6,19] for instance), and its first moment is $\langle a \rangle \approx 2.360$ hr. In figure 2, the corresponding interdivision-time and length-at-division PDFs are provided, and it can be seen that both PDFs exhibit a right-skewed shape. Furthermore, the mean interdivision time can be retrieved and is approximately $\tau_{obs} \approx 4.314$ hr. It is worth noting that

- $\langle a \rangle + \tau_{obs} \approx 6.673$ hr. This value differs from $1/D$ by less than 0.1%.
- $\ln(2)/D \approx 4.621$ hr $> \tau_{obs}$ and $\langle a \rangle \ln(2)/(1 - \ln(2)) \approx 5.331$ hr $> \tau_{obs}$.

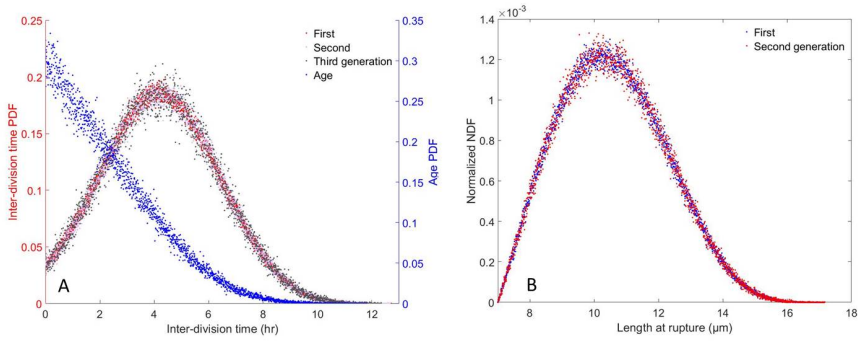

Figure 2: Monte–Carlo simulation results for continuous culture with $D = 0.15$ hr^{-1}. Left: Steady-state cell-age (red points) and interdivision-time g (cyan points) distributions where $\langle\tau\rangle \approx 4.314$ hr. Right: Length distribution at division.

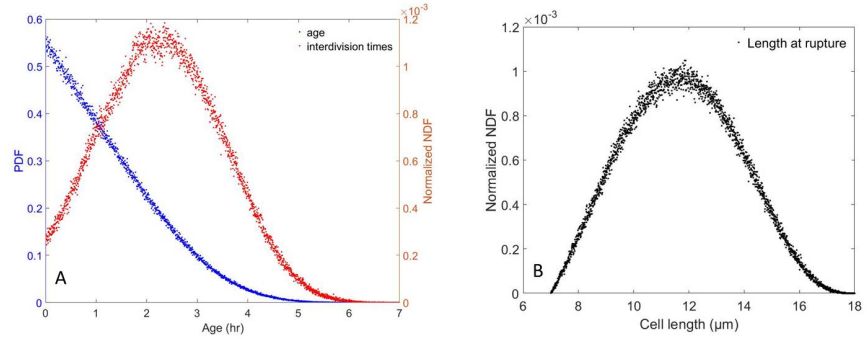

Figure 3: Batch culture results in exponential-growth regime from Monte–Carlo simulation. Left: PDFs for cell age (blue points) and interdivision time, i.e. g (red points). Right: PDF for length at rupture.

Batch-culture simulations (cf. figure 3) exhibit a fairly similar pattern once exponential growth is reached. In this context, ν_m is retrieved from the population's growth in mass (that is tantamount to its growth in number as mentioned in [15,16]) over a certain time interval:

$$\nu_m = \frac{\ln\left(\frac{m(t+\Delta t)}{m(t)}\right)}{\Delta t} = \frac{\ln\left(\frac{N(t+\Delta t)}{N(t)}\right)}{\Delta t}$$

with t, $t+\Delta t$ belonging to the so-called "log phase". In the Monte–Carlo simulation, $\Delta t = 3.28925$ hr, cell mass was multiplied by 2.471 to three decimal places, and $\nu_m \approx 0.276$ hr^{-1}. The mean cell age and interdivision time satisfy the properties:

- $\tau_{obs} \approx 2.327$ hr $< \ln(2)/\nu_m$
- $\langle a\rangle \approx 1.302$ hr $> (1 - \ln(2))/\nu_m$
- $\langle a\rangle + \tau_{obs} \approx 3.629$ hr.

In comparison, $1/\nu_m \approx 3.636$ hr, which differs from $\langle a\rangle + \tau_{obs}$ by less than 0.2%.

4 Discussion

4.1 Powell's analytical results in a continuous, well-mixed fermenter.

Before PBE tools were developed to address a population's variability in different inner properties, the marginal distributions were retrieved from infinitesimal computations and Cauchy problems were extracted to be solved analytically and confronted with experimental data. Powell's seminal article [14] is no exception, and a relation coupling the cell-age and interdivision-time PDFs is discussed for both batch and continuous fermenters. However, (16) is not exactly the formula Powell retrieved from his own infinitesimal calculus, because the two interdivision-time PDFs do not share the same definition. Indeed, g is the conditional probability that, given a cell divides, it does so at age a, whereas Powell's interdivision-time PDF relates to the probability that a cell divides at age a, the latter being less convenient in practice because it disregards the available memory from cells that reach age a and do not fully embrace the Markovian nature of the cell-cycle process. In the following, Powell's f distribution will be labelled h for the sake of clarity.

A Powell-like differential equation can be devised with our definition of g, by starting from a set of N cells of which $Nf(a)da$ belong to the age interval $[a, a+da]$ at time t. Then, during an interval of Lebesgue measure δt, $Nf(a)(1-e^{-D\delta t})$ cells are washed out and

$$\int_a^{a+\delta t}\int \gamma(l)n(t,l,a')dlda' = \int_a^{a+\delta t} g(a')da' \iint \gamma(l)n(t,l,a)dlda \tag{25}$$

produce daughter cells of age zero (from the (3) definition of g). Consequently, after division by N on both sides,

$$\begin{aligned} f(a+\delta t)-f(a) &= -f(a)(1-e^{-D\delta t}) - D\int_a^{a+\delta t} g(a')da' \\ &= -\delta t D f(a) - D\left(\int_a^{\infty} g(a')da' - \int_{a+\delta t}^{\infty} g(a')da'\right) + o(\delta t) \\ &= -\delta t D f(a) + \delta t D\frac{d}{da}\int_a^{\infty} g(a')da' + o(\delta t) \\ &= -\delta t D f(a) - \delta t D g(a) + o(\delta t) \end{aligned}$$

One immediately obtains $f'(a) = -Df(a) - Dg(a)$, which mirrors (13), showing that this result is independent from the calculation methodology. It is remarkable that the reference to g eliminates the need to compute any conditional probability.

However, if Powell's definition of h, which contains all interdivision times, is used to establish a conservation equation for the number of cells with age a in a reactor then one has to check that a cell has actually reached that age a in the system. This leads to a conditional probability and Bayes' theorem leads to

$$\begin{aligned} P(\text{interdivision time} \le a+\delta t | \text{age} \ge a) &= \frac{P(\text{interdivision time} \le a+\delta t \cap \text{interdivision time} \ge a)}{P(\text{interdivision time} \ge a)} \\ &= \frac{\int_a^{\infty} h(a')da' - \int_{a+\delta t}^{\infty} h(a')da'}{\int_a^{\infty} h(a')da'} = 1 - \frac{\int_{a+\delta t}^{\infty} h(a')da'}{\int_a^{\infty} h(a')da'} \end{aligned}$$

In order to reach age $a+\delta t$, any cell has to reach age a, remain in the system for at least δt and not divide between a and $a+\delta t$. An infinitesimal calculation using Taylor's formula entails:

$$\begin{aligned} f(a+\delta t) &= f(a)\exp(-D\delta t)\frac{\int_{a+\delta t}^{\infty} h(a')da'}{\int_a^{\infty} h(a')da'} \\ &= f(a)\left[(1-\delta tD+o(\delta t))(\frac{\int_a^{\infty} h(a')da' + \delta t\frac{\partial}{\partial a}\int_a^{\infty} h(a')da' + o(\delta t)}{\int_a^{\infty} h(a')da'})\right] \end{aligned}$$

Developing and simplifying the second-order terms leads to:

$$\begin{aligned} f(a+\delta t) &= f(a)\left[1-\delta tD+\delta t\frac{-h(a)}{\int_a^{\infty} h(a')da'}+o(\delta t)\right] \\ \Leftrightarrow \frac{f(a+\delta t)-f(a)}{\delta t} &= f(a)\left[-D-\frac{h(a)}{\int_a^{\infty} h(a')da'}+o(1)\right] = f(a)\left[-D+\frac{\partial}{\partial a}\ln(\int_a^{\infty} h(a')da')+o(1)\right] \\ \underset{\delta t\to 0}{\Leftrightarrow} f'(a) &= f(a)\left[-D+\frac{\partial}{\partial a}\ln(\int_a^{\infty} h(a')da')\right] \end{aligned} \tag{26}$$

Straightforward computations then result in Powell's proposed law for the relationship between the cell-age PDF and h, which is indeed consistent given his memoryless function h.

$$\begin{aligned} f(a) &= 2D\exp(-Da)\exp(\ln(\int_a^{\infty} h(a')da') \\ &= 2De^{-Da}\left[1-\int_0^a h(a')da'\right] \end{aligned} \tag{27}$$

While all relations described in [14] are true in both batch and continuous culture, it is of crucial importance to draw the reader's attention to a fallacious reasoning involving Powell's definition of h. The latter aims at evaluating the probability that a cell's interdivision time is more or less than its residence time, which is determined by the relation coupling f and h in (27):

$$\begin{aligned} \langle a\rangle &= \frac{2}{D} - 2D\int_0^{\infty} h(a')\int_{a'}^{\infty} ae^{-Da}dada' \\ &= \frac{2}{D} - 2D\int_0^{\infty} h(a)\left(\frac{1}{D}ae^{-Da}+\frac{1}{D^2}e^{-Da}\right)da \\ &= \frac{2}{D} - 2\int_0^{\infty} ae^{-Da}h(a)da - \frac{2}{D}\int_0^{\infty} e^{-Da}h(a)da \end{aligned}$$

305 The latter basically results in the equalities

$$\int_0^{\infty} e^{-Da}h(a)da = \frac{1}{2} \quad \text{and} \quad \int_0^{\infty} ae^{-Da}h(a)da = \frac{\tau_{obs}}{2}$$

This must be compared to Powell's assumption that a cell has a probability of 1/2 of yielding two daughter cells before washout occurs, and the same probability that a cell is washed out before it begins a division event. Indeed, given that the residence time t_{res} in a well-mixed fermenter obeys an exponential law:

$$t_{res}(t) = De^{-Dt}$$

it follows that

$$P(\text{interdivision time} < \text{residence time}) = \int_0^\infty De^{-Dt} \int_0^t h(a)dadt$$

Then, using once again Fubini's theorem,

$$P(\text{interdivision time} < \text{residence time}) = \int_0^\infty h(a) \int_a^\infty De^{-Dt}dtda = \int_0^\infty e^{-Da}h(a)da = \frac{1}{2}$$

306 which is consistent with Powell's result based on physical grounds.

Furthermore, the mean interdivision time is obviously not equal to $\int_0^\infty ah(a)da$ because Powell's definition of h does not match the observable interdivision-time PDF. To convince oneself, the relations coupling the cell-age PDF with g (16) or h (27) entail the conclusion immediately:

$$\left.\begin{aligned} f(a) &= 2De^{-Da} - De^{-Da}\int_0^a e^{Da'}g(a')da' \\ f(a) &= 2De^{-Da} - 2De^{-Da}\int_0^a h(a')da' \end{aligned}\right\} \Leftrightarrow g(a) = 2e^{-Da}h(a)$$

307 which provides the relation between g and h in a well-mixed fermenter. $g(a)$ is conspicuously greater
308 than $h(a)$ if $e^{-Da} > 1/2 \Leftrightarrow a < \ln(2)/D$, the reverse inequality holding if $a > \ln(2)/D$.

309 To conclude the discussion of continuous cultures, the differences between g and h are shown in
310 figure 4. Because h records all interdivision times, it lends weight to cells that are highly unlikely
311 to divide in a fermenter. The observable interdivision-time PDF g references actual rupture events,
312 these divisions being less and less likely as a approaches $\ln(2)/D$. This physical reasoning testifies
313 to the inequalities $\tau_{obs} < \ln(2)/D < \int_0^\infty ah(a)da$, and the relation $g(a) = 2e^{-Da}h(a)$ allows the
314 conclusion $g(a) > h(a)$ for $a \in [0, \ln(2)/D[$, the inverse relation being satisfied for $a > \ln(2)/D$.

315 4.2 Painter & Marr's inequality for the unobservable PDF.

In their 1967 article [12] addressing the interdivision-time PDF in a continuous, well-mixed fermenter, Painter & Marr incorrectly extracted the inequality $\tau_{obs} \geq \ln(2)/D$ from Powell's relation $1 = 2\int_0^\infty e^{-Da}h(a)da$. From their point of view, developing the exponential as a power series after factoring $2e^{-D\tau_{obs}}$ would reduce to

$$2e^{-D\tau_{obs}} \int_0^\infty e^{-D(a-\tau_{obs})}h(a)da = 1$$

and, using the fact that the exponential function is convex,

$$2e^{-D\tau_{obs}} \int_0^\infty [1 - D(a - \tau_{obs})]h(a)da \geq 1$$

316 Then, Painter & Marr erroneously stated that $\int_0^\infty ah(a)da = \tau_{obs}$ to conclude. However, $\int_a^\infty ah(a)da$
317 is not τ_{obs} but refers to $\tau_{uno} \geq \tau_{obs}$ instead. Once again, the confusing definition of the interdivision-
318 time distribution lends artificial weight to zero-measure fractions of a population.

319 Referring to Painter and Marr's work, van Heerden and co-workers produced a slightly biased
320 fit of their experimental interdivision-time PDF. Hence, their data analysis procedure involving h
321 instead of g, lead $\langle a \rangle + \tau_{fit}$ to be greater than D^{-1} by a significant 7% margin and τ_{fit} to be

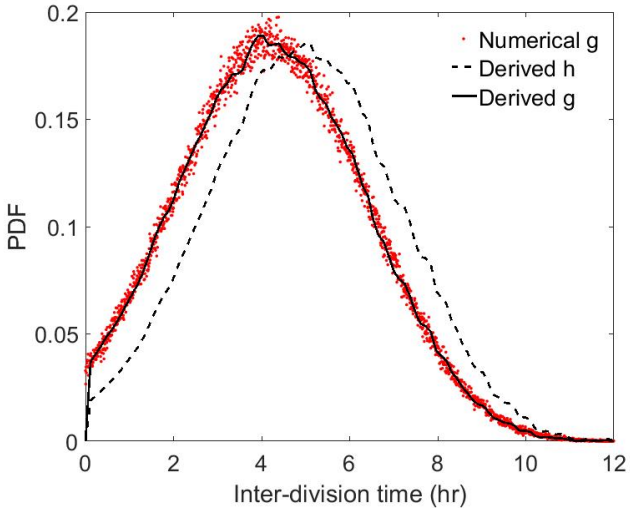

Figure 4: Distributions of inter-division time in Powell's formalism: h (black dashed line) and its measurable counterpart g (black line). The numerical data retrieved from the Monte-Carlo code (red points) are shown for comparison. In general, h lends more weight to the older cells than g, so that $\int_0^\infty ag(a)da < \ln(2)/D < \int_0^\infty ah(a)da$.

greater than $2.259\langle a\rangle$. However, using their raw data for *B. subtilis*, we find that $\tau_{obs}, \langle a\rangle$ and D agree with both (18) and (19). Moreover, their Supplementary Data regarding *E. coli* are in complete agreement with $\langle a\rangle + \tau_{obs} = D^{-1}$. This analysis confirms that analytical, numerical and experimental results are in perfect agreement provided that equation (16) is used instead of (27) when dealing with a set of measured interdivision times. To conclude this discussion, it is pointed out that the experimental procedure itself affects the observed interdivision-time distribution. In Yasuda's experiments using an optical tweezer to remove cells from the growth chamber *following their division*, no cell is washed out before dividing. Therefore, an interdivision-time distribution from such measurements resembles h more than the one stemming from a continuous system.

5 Concluding remarks

The exact results developed in this work throw light on the equivocal interpretations of the notion of interdivision time appearing in the literature where two different PDFs were considered from the analytical and experimental perspective. Starting from a PBE, rigorous mathematical results for the observable interdivision-time distribution have been established (complementing recent work by Jafarpour et al. [9] for instance), and numerical examples are provided to supplement the theoretical results. As expected, the steady-state PDFs from the Monte–Carlo simulations proved to be in accordance with the analytical expressions. This paradigm is more suitable than Painter and Marr's when it comes to experimental data treatment. Indeed their conclusions were based on the first moment of the unobservable cell interdivision-time distribution. The relationships provided in this work match the experimental data by van Heerden and co-workers regarding *E. coli* and *B. subtilis*.

Analysis-wise, no expression for the PDF of the cell length is accessible because the integral $\int \gamma(l')P(l,l')n(t,l')dl'$ has no specific shape. Furthermore, with two relations pertaining to P:

1. $P(l, l') = P(l' - l, l')$

2. $\int_0^{l'} P(l, l')dl = 1$

one can extract the dynamics of the length distribution's zeroth and first-order moments only, with the help of integrations by parts and Fubini's theorem. However, no additional formulae are available if no other relations constrain P.

Acknowledgements

The authors gratefully acknowledge the financial support from the French National Research Agency through the 2015 "Attractivity Chair" program, Idex UNITI-"Biological, Reacting, Multiphase Flows (BIREM)" conv-ANR-11-Idex-0002-02

References

[1] David Bates, Jessica Epstein, Erik Boye, Karen Fahrner, Howard Berg, and Nancy Kleckner. The Escherichia coli baby cell column: a novel cell synchronization method provides new insight into the bacterial cell cycle. *Molecular Microbiology*, 57(2):380–391, July 2005.

[2] Frédérique Billy, Jean Clairambault, Franck Delaunay, Céline Feillet, and Natalia Robert. Age-structured cell population model to study the influence of growth factors on cell cycle dynamics. *Mathematical Biosciences and Engineering*, 10(1):1–17, February 2013.

[3] Hans Bremer and Gordon Chuchward. Age Fractionation in Bacteria by Membrane Elution: Relation Between Age Distribution and Elution Profile. *Journal of Theoretical Biology*, 74:69–81, 1978.

[4] Thomas Ferenci. Adaptation to life at micromolar nutrient levels: the regulation of Escherichia coli glucose transport by endoinduction and cAMP, FEMS. *Microbiol. Rev.*, 18:301–317, July 1996.

[5] A. G. Fredrickson, D. Ramkrishna, and H. M. Tsuchiya. Statistics and Dynamics of Procaryotic Cell Populations. *Mathematical Biosciences*, 103:327–374, 1967.

[6] A. G. Fredrickson and H. M. Tsuchiya. Continuous Propagation of Microorganisms. *AIChE Journal*, 9(4):459–468, July 1963.

[7] Michel Godin, Francisco Feijó Delgado, Sungmin Son, William H Grover, Andrea K Bryan, Amit Tzur, Paul Jorgensen, Kris Payer, Alan D Grossman, Marc W Kirschner, and Scott R Manalis. Using buoyant mass to measure the growth of single cells. *Nature Methods*, 7(5):387–390, May 2010.

[8] Christos Hatzis, Friedrich Srienc, and A. G. Fredrickson. Multistaged corpuscular models of microbial growth: Monte Carlo simulations. *BioSystems*, 36:19–35, 1995.

[9] Farshid Jafarpour, Charles S. Wright, Herman Gudjonson, Jedidiah Riebling, Emma Dawson, Klevin Lo, Aretha Fiebig, Sean Crosson, Aaron R. Dinner, and Srividya Iyer-Biswas. Bridging the Timescales of Single-Cell and Population Dynamics. *Physical Review X*, 8, 2018.

[10] Nikos V. Mantzatis. Stochastic and deterministic simulation of heterogeneous cell population dynamics. *Journal of Theoretical Biology*, 241:690–706, 2006.

[11] Jean-Bernard Nobs and Sebastian J. Maerkl. Long-Term Single Cell Analysis of S. pombe on a Microfluidic Microchemostat Array. *PLOS ONE*, 9(4):1–11, April 2014.

[12] Page R. Painter and Allen G. Marr. Inequality of Mean Interdivision Time and Doubling Time. *Journal of General Microbiology*, 48:155–159, 1967.

[13] Benoît Perthame. *Transport Equations in Biology*. Springer Verlag, Birkhäuser, 2007.

[14] E. O. Powell. Growth Rate and Generation Time of Bacteria, with Special Reference to Continuous Culture. *Journal of General Microbiology*, 15:492–511, 1956.

[15] E. O. Powell. A Note on Koch & Schaechter's Hypothesis about Growth and Fission of Bacteria. *Journal of General Microbiology*, 37:231–249, 1964.

[16] Vincent Quedeville, Hicham Ouazaite, Bastien Polizzi, Rodney O. Fox, Philippe Villedieu, Pascal Fede, Fabien Létisse, and Jérôme Morchain. A two-dimensional population balance model for cell growth including multiple uptake systems. *Chemical Engineering Research and Design*, 132:966–981, 2018.

[17] Doraiswami Ramkrishna. Statistical Models of Cell Populations. *Adv. Biochem. Eng.*, 11:1–47, 1979.

[18] Lydia Robert, Marc Hoffmann, Nathalie Krell, Stéphane Aymerich, Jérôme Robert, and Marie Doumic. Division in Escherichia Coli is triggered by a size-sensing rather than a timing mechanism. *BMC Biology*, 12(17), 2014.

[19] Evgeny B. Stukalin, Ivie Aifuwa, Jin Seob Kim, Denis Wirtz, and Sean X. Sun. Age-dependent stochastic models for understanding population fluctuations in continuously cultured cells. *J. R. Soc. Interface*, 10, 2013.

[20] G. Subramanian, D. Ramkrishna, A. G. Fredrickson, and H. M. Tsuchiya. On the mass distribution model for microbial cell populations. *Bull. Math. Biophys.*, 32:521–537, 1970.

[21] Sirichai Sunya, Frank Delvigne, Jean-Louis Uribelarrea, Carole Molina-Jouve, and Nathalie Gorret. Comparison of the transient responses of Escherichia coli to a glucose pulse of various intensities. *Applied Microbiol Biotechnol*, 95:1021–1034, 2012.

[22] Sattar Taheri-Araghi, Serena Bradde, John T. Sauls, Norbert S. Hill, Petra Anne Levin, Johan Paulsson, Massimo Vergassola, and Suckjoon Jun. Cell-Size Control and Homeostasis in Bacteria. *Current Biology*, 25:385–391, February 2015.

[23] Yu Tanouchi, Anand Pai, Heungwon Park, Shuqiang Huang, Rumen Stamatov, Nicolas E. Buchler, and Lingchong You. A noisy linear map underlies oscillations in cell size and gene expression in bacteria. *Nature*, 523(7560):357–360, July 2015.

[24] John J. Tyson and Kenneth B. Hannsgen. Cell growth and division: a deterministic/probabilistic model of the cell cycle. *Journal of Mathematical Biology*, 23:231–246, 1986.

[25] Johan van Heerden, Mannus Kempe, Anne Doerr, Timo Maarleveld, Niclas Nordholt, and Frank Johannes Bruggeman. Statistics and simulation of growth of single bacterial cells: Illustrations with B. subtilis and E. coli. *Scientific Reports*, 7(1):1–11, December 2017.

[26] Kenji Yasuda. Algebraic and Geometric Understanding of Cells: Epigenetic Inheritance of Phenotypes Between Generations. In Susann Müller and Thomas Bley, editors, *High Resolution Microbial Single Cell Analytics*, volume 124 of *Advances in Biochemical Engineering / Biotechnology*, pages 55–81. Springer Berlin Heidelberg, January 2011.

Author contributions statement

V.Q. produced the mathematical results, wrote the manuscript, the Monte–Carlo code and performed numerical simulations.
J.M. co-developed the Monte–Carlo code, analysed the numerical experiments, wrote and reviewed the manuscript
P.V. produced the mathematical results, analysed the results, wrote and reviewed the manuscript
R.O.F. produced the mathematical results, analysed the results, wrote and reviewed the manuscript

Additional information

The authors declare no competing interests.

2.2 Size–structured models

Throughout this section, $\xi = l \in]0, \bar{l}[$ will denote the cells' inner coordinate and (1.22) will read:

$$\frac{\partial}{\partial t} n(t,l) + \frac{\partial}{\partial l}\left[\dot{l} n(t,l)\right] + \gamma(l) n(t,l) + D n(t,l) = 2 \int_l^{\bar{l}} \gamma(l') P(l,l') n(t,l') dl'$$
$$n(t,l)|_{l \in \{0,\bar{l}\}} = 0$$

2.2.1 Growth in mass or growth in number?

The very notion of growth rate comes as a spurious spin to the (1.7) system that cannot discriminate between cell number and cell mass, both quantities being supposed completely proportional at such a macroscopical scale the set of ODEs is designed to. However, an integration of the 0-th order moment of (1.22) boils down to:

$$\int_0^{\bar{l}} \frac{\partial}{\partial t} n(t,l) dl + \int_0^{\bar{l}} \gamma(l) n(t,l) dl + D \int_0^{\bar{l}} n(t,l) dl = 2 \int_0^{\bar{l}} \int_l^{\bar{l}} P(l,l') \gamma(l') n(t,l') dl' dl$$

A use of Fubini's theorem allows to switch integrals in the right–hand side term to yield:

$$\int_0^{\bar{l}} \int_l^{\bar{l}} P(l,l') \gamma(l') n(t,l') dl' dl = \int_0^{\bar{l}} \int_0^{l'} P(l,l') dl \gamma(l') n(t,l') dl' = \int_0^{\bar{l}} \gamma(l') n(t,l') dl'$$

Then, if $N(t)$ stands for the total cell number at time t, its dynamics is determined by:

$$\frac{dN(t)}{dt} = -DN(t) + \int_0^{\bar{l}} \gamma(l) n(t,l) dl \tag{2.4}$$

That is, if

$$\mu_N := \frac{\int_0^{\bar{l}} \gamma(l) n(t,l) dl}{\int_0^{\bar{l}} n(t,l) dl}$$

stands for the population's growth rate in number:

$$\frac{dN(t)}{dt} = (\mu_N - D) N(t) \tag{2.5}$$

On the other hand, since it was assumed that a cell's mass is linearly correlated to its length, integrating the first–order moment of (1.22) will yield the dynamics of the reactor's biomass:

$$\int_0^{\bar{l}} l \frac{\partial}{\partial t} n(t,l) dl - \int_0^{\bar{l}} \dot{l} n(t,l) dl + \int_0^{\bar{l}} l \gamma(l) n(t,l) dl + D \int_0^{\bar{l}} l n(t,l) dl$$
$$= 2 \int_0^{\bar{l}} l \int_l^{\bar{l}} P(l,l') \gamma(l') n(t,l') dl' dl$$

given that the Dirichlet condition at $l \in \{0, \bar{l}\}$ forces the boundary term to vanish. Another use of Fubini's theorem will transform the right–hand side into:

$$\int_0^{\bar{l}} \int_l^{\bar{l}} l P(l,l') \gamma(l') n(t,l') dl' dl = \int_0^{\bar{l}} \int_0^{l'} l P(l,l') dl \gamma(l') n(t,l') dl'$$
$$= \int_0^{\bar{l}} \int_0^{l'} (l' - l) P(l,l') dl \gamma(l') n(t,l') dl' = \int_0^{\bar{l}} l' \gamma(l') n(t,l') dl' - \int_0^{\bar{l}} \int_0^{l'} l P(l,l') dl \gamma(l') n(t,l') dl'$$
$$\int_0^{\bar{l}} \int_0^{l'} l P(l,l') dl \gamma(l') n(t,l') dl' = \frac{1}{2} \int_0^{\bar{l}} l' \gamma(l') n(t,l') dl'$$

Therefore, the first moment of the length distribution obeys the equation:

$$\frac{d}{dt}\int_0^{\bar{l}} l n(t,l)dl = -D\int_0^{\bar{l}} l n(t,l)dl + \int_0^{\bar{l}} \dot{l} n(t,l)dl \tag{2.6}$$

and, as previously, a population–scale growth rate in mass μ_m can be defined to set the equivalence with the population's growth in number:

$$\mu_m = \frac{\int_0^{\bar{l}} \dot{l} n(t,l)dl}{\int_0^{\bar{l}} l n(t,l)dl}$$

$$\text{Hence: } \frac{d}{dt}\int_0^{\bar{l}} l n(t,l)dl = (\mu_m - D)\int_0^{\bar{l}} l n(t,l)dl \tag{2.7}$$

These results hint at previous conclusions from Doumic's analysis regarding mass–structured models [19], that is the cell number can only grow by division, whatever the elongation function, whereas the cell mass can only grow by lengthening. Of course, the conservation of mass at division is a key assumption leading to the (2.6) equation, allowing the simplification of the rupture–related terms and the (2.7) definition of μ_m.

Equations (2.4) and (2.6) are extremely similar, so much so growth in mass and growth in number will prove equivalent when μ_N and μ_m are actual doppelgangers. This happens when steady–state is reached in a chemostat, respectively turning (2.4) and (2.6) into $\mu_N = D$ and $\mu_m = D = \mu_N$. Also, exponential growth in batch culture is notably characterised by the invariance of the population's mean length with respect to time. In other words, $\mu_N = \mu_m = \mu$, which definition comes this time mainly from physical grounds. A comprehensive division pertaining to the approximation $\int_0^{\bar{l}} \dot{l} n(t,l)dl \approx \mu_m \int_0^{\bar{l}} l n(t,l)dl$ is provided in the appendix, section 4.

As informative as (2.4) and (2.6) are, the equations are geared towards macroscopic features only. A size profile is not available from (1.22) unless the integrals $\int_0^{l'} l^k P(l,l')dl$ can be calculated for $k \geq 2$, what is not an option if P is not given. On the other hand, even if P was known beforehand, say $P \sim \beta(p,p)$, $p \geq 1$, the following recursive equation for the k–th moment of the size distribution would read:

$$\frac{\partial}{\partial t}\int_0^{\bar{l}} l^k n(t,l)dl - k\int_0^{\bar{l}} l^{k-1}\dot{l} n(t,l)dl + D\int_0^{\bar{l}} l^k n(t,l)dl = \left[2\prod_{j=0}^{k-1}\frac{p+j}{2p+j} - 1\right]\int_0^{\bar{l}} l^k \gamma(l) n(t,l)dl$$

could be analytically calculated for certain $\dot{l}$ and γ functions only. For instance, the different terms are moments of the length NDF if and only if $\dot{l}$ and γ are polynomials in l, what has no reason to be true in all generality.

These hypotheses being too restrictive to make for a convenient modelling, the following section will concentrate on more general analytical results that do not require any information on the kernel P.

2.2.2 Existence and uniqueness result

As was mentioned in the first section, the size–PBE (1.22) is traditionally (cf. Doumic's [19]) solved in the $\mathcal{C}^0(\mathbb{R}_+, L^1(\mathbb{R}_+))$ functional space provided that the rupture function belongs to $L^\infty(\mathbb{R}_+)$

and the growth function is $\mathcal{C}_b^1$ over its domain. Modellingwise, this case study is not a satisfying working assumption, being agreed that no cell of macroscopic size is to be observed by microbiologists. A remedy could consist in the definition of an artificial $\bar{l}$ upper bound to the cell length and the formulation of appropriate γ and $\dot{l}$ functions, like:

$$\gamma(l) = C\left(1 - \frac{l}{\bar{l}}\right)^{\alpha-1} \text{ and } \dot{l} = K\left(1 - \frac{l}{\bar{l}}\right)^{\beta}$$

Obviously, γ has to be singular to express the absence of ever–growing cells, enforcing the $\alpha < 1$ assumption. Similarly, $\dot{l}$ is forced to converge to 0 as $l \to \bar{l}$ to annihilate any possibility of growth beyond $\bar{l}$, in other words one expects $\beta > 0$. Hereinafter, for the sake of convenience, the rest of the talk will involve a normalised variable $x = l/\bar{l} \in [0,1[$, turning the definition of $\dot{l}$ into $\dot{x} = K'(1-x)^{\beta}$, $K' = K/\bar{l}$.

As counter–intuitive as it seems, not every ($\alpha < 1$, $\beta > 0$) combination is suitable within this framework because a cell cluster could form at the impassible $\bar{l}$ boundary should the γ function be not steep enough. The following theorem gives meaning to a solution of (1.22) when the α and β coefficients are adequately chosen.

Theorem 2.2.1. *Consider the PBE*

$$\begin{aligned}
&\frac{\partial}{\partial t}n(t,x) + \frac{\partial}{\partial x}[\dot{x}n(t,x)] + \gamma(x)n(t,x) + Dn(t,x) = 2\int_x^1 \gamma(x')P(x,x')n(t,x')dx',\ (t,x) \in \mathbb{R}_+ \times [0,1] \\
&n(0,x) = n_0(x),\ x \in [0,1] \\
&n(t,0) = 0,\ t \geq 0
\end{aligned} \tag{2.8}$$

Suppose that P is a continuous probability density kernel such that $P(0,x) = 0\ \forall x \geq 0$ and $n_0 \in \mathcal{C}^0([0,1[)$. If γ and $\dot{x}$ are such that:

$$\gamma(x) = C(1-x)^{\alpha-1} \text{ and } \dot{x} = K'(1-x)^{\beta}$$

with $0 < \alpha < 1$, $0 < \beta$ and $\beta > \alpha$, then a unique solution to (1.22) exists in $\mathcal{C}^0(\mathbb{R}_+ \times ([0,1]))$. Furthermore, for any $t \geq 0$, $n(t,x) \underset{x\to 1^-}{\to} 0$.

Proof.

Sketch of the proof:

Without loss of generality, t_0 can be set equal to 0 (at the expense of a change $\tau = t - t_0$ of variables).
The result will rely on a splitting operator scheme, consisting in injecting a solution of the linear equation:

$$\frac{\partial}{\partial t}n(t,x) + \frac{\partial}{\partial x}[\dot{x}n(t,x)] + \gamma(x)n(t,x) + Dn(t,x) = 0 \tag{2.9}$$

with initial condition $n(t_0,\cdot) = n_0(\cdot) \in L^{\infty}([0,1])$
and boundary condition $n(\cdot,0) = 0$

in the equation:

$$\frac{\partial}{\partial t}n(t,x) = 2\int \gamma(x')P(x,x')n(t,x')dx'. \tag{2.10}$$

(2.9) can be solved using the method of characteristics, yielding a semi–group S_t that satisfies $n(t,\cdot) = S_t(n_0(\cdot))$. In other words, S_t is the operator that transports a solution along the characteristic curve passing through the associated initial condition. This part of the proof follows closely Doumic's strategy (cf. [19]) to demonstrate the existence of a Malthus eigenfunction in $\mathcal{C}_b^1$ as soon as $\dot{x}$ is a $\mathcal{C}^2$ function and $\gamma \in \mathcal{C}_b^1$.

Concluding on the existence/uniqueness of the solution to (2.8) comes from an application of Picard's fixed–point theorem. First, the solution to (2.10) is forthcoming through a use of Duhamel's rule:

$$n(t,\cdot) = S_t(n_0)(\cdot) + \int_0^t S_{t-s}\left(I(n)\right)(s,\cdot)ds$$

if I stands for the integral operator $I : n(t,x) \mapsto 2\int_x^1 \gamma(x')P(x,x')n(t,x')dx'$.

Then, the crucial step to ensure the problem's wellposedness lies in I's continuity (equivalently, its boundedness) property as a $\mathcal{C}^0 \to \mathcal{C}^0$ operator. More precisely, from Hölder's inequality:

$$I(n)(t,x) \leq \|\gamma\|_{L^1} \max_{0\leq x\leq x'<1} P(x,x')\|n(t,\cdot)\|_{L^\infty([0,1])}$$

and the application of Picard's fixed–point theorem follows as soon as it is proven that $S_t : \mathcal{C}^0 \to \mathcal{C}^0$ is continuous. Indeed, it will entail the existence of a real $T > 0$ such that the operator $A : n(t,\cdot) \mapsto S_t(n_0(\cdot)) + \int_0^t S_{t-s}(I(n)(s,\cdot))ds$ is a contraction on $[0,T]$. Applying the same reasoning to the $[nT,(n+1)T]$ intervals with $n \geq 1$ will be straightforward.

Solving (2.9)

For $\dot{x} = K'(1-x)^\beta$, $\gamma(x) = C(1-x)^{\alpha-1}$, (2.9) comes down to :

$$\frac{\partial}{\partial t}n(t,x) - \beta K'(1-x)^{\beta-1}n(t,x) + K'(1-x)^\beta\frac{\partial}{\partial x}n(t,x) + C(1-x)^{\alpha-1}n(t,x) + Dn(t,x) = 0$$

Using the method of characteristics, the relation between x and t is given by the Cauchy system:

$$\frac{dX}{dt} = K'(1-X)^\beta$$
$$X(0) = x_0$$

$$\Rightarrow \left[-\frac{(1-X)^{1-\beta}}{1-\beta}\right]_{x_0}^{X(t,x_0)} = K't \Leftrightarrow (1-\beta)K't = (1-x_0)^{1-\beta} - (1-X(t,x_0))^{1-\beta}$$

$$\text{Finally, } X(t,x_0) = 1 - \left[(1-x_0)^{1-\beta} - (1-\beta)K't\right]^{\frac{1}{1-\beta}}, \text{ for } t \leq \frac{(1-x_0)^{1-\beta}}{K'(1-\beta)} \tag{2.11}$$

The (2.2.2) equation cuts the $\mathbb{R}_+ \times [0,1]$ plan into two parts:

- If (t,x) is such that $x < 1 - [1-(1-\beta)K't]^{\frac{1}{1-\beta}}$: then no characteristic curve passes through the (t,x) point since it would involve a negative x_0. Therefore n can only be identically zero.

- If (t,x) is such that $x \geq 1-\left[1-(1-\beta)K't\right]^{\frac{1}{1-\beta}}$: then one characteristic curve passes through (t,x) and its origin $x_0 = X^{-1}(t,x)$ is unique.

If $\tilde{n}$ stands for the solution to (2.8), that is $\tilde{n}(t,x_0) = n(t,X(t,x_0))$, tracking it along the characteristics passing through $(0,x_0)$ leads to:

$$\frac{d\tilde{n}}{dt} = \frac{\partial}{\partial t}n(t,X(t,x_0)) + K'\left[(1-x_0)^{1-\beta} - (1-\beta)K't\right]^{\frac{\beta}{1-\beta}}\frac{\partial}{\partial x}n(t,X(t,x_0))$$

that is, given the (2.2.2) definition of $X(t,x_0)$:

$$\frac{d\tilde{n}}{dt}(t,x_0) = \frac{\partial}{\partial t}n(t,X(t,x_0)) + K'(1-X(t,x_0))^{\beta}\frac{\partial}{\partial x}n(t,X(t,x_0))$$

A substitution into equation (2.9) comes down to:

$$\frac{d\tilde{n}}{dt}(t,x_0) = \left(-D + \beta K'(1-X(t,x_0))^{\beta-1} - C(1-X(t,x_0))^{\alpha-1}\right)\tilde{n}(t,x_0)$$

Expressing the $X(t,x_0)$ flow from (2.2.2) turns the latter differential equation into:

$$\frac{d\tilde{n}}{dt}(t,x_0) = \left(-D + \frac{\beta K'}{\left[(1-x_0)^{1-\beta} - (1-\beta)K't\right]} - C\left[(1-x_0)^{1-\beta} - (1-\beta)K't\right]^{\frac{\alpha-1}{1-\beta}}\right)\tilde{n}(t,x_0)$$

Since this ODE is linear, an integration yields immediately:

$$\tilde{n}(t,x_0) = \tilde{n}(0,x_0)e^{-Dt}\exp\left(\int_0^t \frac{\beta K'}{\left[(1-x_0)^{1-\beta} - (1-\beta)K's\right]} - C\left[(1-x_0)^{1-\beta} - (1-\beta)K's\right]^{\frac{\alpha-1}{1-\beta}}ds\right)$$

This formula leads to the analytical solution for $n(t,x)$ once the integrals have been calculated:

$$\begin{aligned}\int_0^t \frac{\beta K'}{\left[(1-x_0)^{1-\beta} - (1-\beta)K's\right]}ds &= -\frac{\beta}{1-\beta}\left[\ln((1-x_0)^{1-\beta} - (1-\beta)K's)\right]_0^t \\ &= \frac{\beta}{1-\beta}\ln\left(\frac{(1-x_0)^{1-\beta}}{(1-x_0)^{1-\beta} - (1-\beta)K't)}\right)\end{aligned}$$

$$\text{and } \int_0^t -C\left[(1-x_0)^{1-\beta} - (1-\beta)K's\right]^{\frac{\alpha-1}{1-\beta}}ds = \frac{C}{(\alpha-\beta)K'}\left[\left((1-x_0)^{1-\beta} - (1-\beta)K's\right)^{\frac{\alpha-\beta}{1-\beta}}\right]_0^t$$
$$= \frac{C}{(\beta-\alpha)K'}\left[(1-x_0)^{\alpha-\beta} - \left((1-x_0)^{1-\beta} - (1-\beta)K't\right)^{\frac{\alpha-\beta}{1-\beta}}\right]$$

It remains to make use of the definition of $x(t)$ along the characteristic curve to remove x_0 from the solution. Hence:

$$n(t,x) = n_0(x_0)\exp(-Dt)\frac{\left((1-x(t))^{1-\beta} + (1+\beta)K't\right)^{\frac{\beta}{1-\beta}}}{(1-x(t))^{\beta}}\frac{\exp\left(\frac{C}{(\beta-\alpha)K'}\left((1-x(t))^{1-\beta} + (1+\beta)K't\right)^{\frac{\alpha-\beta}{1-\beta}}\right)}{\exp\left(\frac{C}{(\beta-\alpha)K'}(1-x(t))^{\alpha-\beta}\right)} \tag{2.12}$$

The only remaining part of the proof consists in ensuring the continuity (that is, the boundedness) of the S_t operator over a certain $[0,T]$ time interval. In other words, an evaluation of $n(t,x)$ for

$x \in [0,1[$ and its $x \to 1$ limit have to be retrieved from (2.12). To this end, the variations of the $x \to (1-x)^{-\beta}\exp(-C/(K'(\beta-\alpha))(1-x)^{\alpha-\beta})$ function are studied:

$$\frac{d}{dx}\left[(1-x)^{-\beta}\exp\left(-\frac{C}{K'(\beta-\alpha)}(1-x)^{\alpha-\beta}\right)\right] = \left[\beta(1-x)^{-1-\beta} - \frac{C}{K'}(1-x)^{\alpha-2\beta-1}\right]\exp\left(-\frac{C(1-x)^{\alpha-\beta}}{K'(\beta-\alpha)}\right)$$

$$= 0 \text{ iff } x = x^\star = 1 - \left(\frac{\beta K'}{C}\right)^{\frac{1}{\alpha-\beta}} < 1$$

$$< 0 \text{ iff } x > x^\star$$

As a consequence, the $x \to (1-x)^{-\beta}\exp(-C/(K'(\beta-\alpha))(1-x)^{\alpha-\beta})$ function is bounded, ensuring the boundedness of (2.12)'s right–hand side. Furthermore, given that $(1-x)^{-\beta}\exp(-C/(K'(\beta-\alpha))(1-x)^{\alpha-\beta})$ is of the form $1/X\exp(-1/X)$, its $X \to 0$ limit is well–known, meaning that the convergence to 0 of the $(1-x)^{-\beta}\exp(-C/(K'(\beta-\alpha))(1-x)^{\alpha-\beta})$ quantity is guaranteed for $x \to 1$. The only uncertainty pertains to the case x_0 $(= (1-x(t))^{1-\beta} + (1+\beta)K't) = 1$ for which the characteristic curve is restricted to the $(0, x_0)$ singleton. In the latter configuration, the final three factors on (2.12)'s right–hand side are equal to 1 though, as the left–hand side is obviously equal to $n_0(x_0)$.

At the end of the day, a positive C_T constant such that $\|S_t(n_0)\|_{L^\infty([0,1])} \leq C_T\|n_0\|_{L^\infty([0,1])}$ for any $t \in [0,T]$ is accessible, wrapping the proof of n's boundedness up.

Solving (2.10)

This section is a mere development of the steps mentioned in the sketch of the proof. It is mentioned as a reminder that, following Duhamel's principle, the operator–splitting induced solution of (1.22) reads :

$$n(t,x) = S_t(n_0) + \int_0^t S_{t-s}(I(n)(s,\cdot))$$

with S_t the semi–group associated to the solution to (2.9) and I the integral operator defined as $I : n(t,x) \mapsto 2\int_x^1 \gamma(x')P(x,x')n(t,x')dx'$. From the previous paragraph's results, S_t continuously maps $\mathcal{C}^0([0,1])$ into $\mathcal{C}^0([0,1])$: there exists a certain $C_T > 0$ such that $\|S_t(n_0)\|_{L^\infty} \leq C_T\|n_0\|_{L^\infty}$. Therefore, a solution to (1.22) exists in $\mathcal{C}^0(\mathbb{R}_+ \times [0,1[)$ if the redistribution integral is a contraction in $L^\infty([0,1])$.
Indeed, using the linearity of operator A defined as:

$$A(n)(t,x) = S_t(n_0) + \int_0^t S_{t-s}(I(n))(s,\cdot)ds$$

for two solutions n_1, n_2 to the size–PBE, the difference $n = n_1 - n_2$ satisfies the equation:

$$A(n)(t,x) = \int_0^t S_{t-s}(I(n))(s,\cdot)ds$$

The theorem is proven if, for a certain $T > 0$ there exists $\eta < 1$ such that

$$\left|\int_0^t S_{t-s}(I(n))(s,\cdot)ds\right| \leq \eta \max_{0\leq t\leq T}\|n(t)\|_{L^\infty([0,1])})$$

Applying Hölder's inequality will entail:

$$\begin{aligned}\left|\int_0^t S_{t-s}(I(n))(s,\cdot)ds\right| \leq & C_T\int_0^t 2\int_x^1 \gamma(x')P(x,x')n(s,x')dx'ds\\ &\leq 2TC_T\|\gamma\|_{L^1}\max_{0\leq x\leq x'<1}P(x,x')\max_{0\leq t\leq T}\|n(t)\|_{L^\infty}\\ &\leq \eta\max_{0\leq t\leq T}\|n(t)\|_{L^\infty} \text{ iff } T\leq \frac{\eta}{2C_S\|\gamma\|_{L^1}\max\limits_{0\leq x\leq x'<1}P(x,x')}\end{aligned}$$

what proves that I is a continuous operator and the integral defining I is convergent for any $x\in[0,1]$. Also, from the $P(0,x)=0\ \forall x\in[0,1]$ assumption, one retrieves $I(n)(t,0)=0$. P being uniformly continuous probability density kernel, $I(n)(t)$ is continuous over $[0,1]$.

Hence, by virtue of Picard's fixed point theorem, a unique solution to (2.8) can be infered on $[0,T]$. Repeating this iterative reasoning over the interval $[nT,(n+1)T]$, $n\in\mathbb{N}$, wraps up the proof of the existence and uniqueness result in $\mathcal{C}^0(\mathbb{R}_+\times[0,1])$. □

Discussion

- The restricting condition $\beta>0$ has major implications when it comes to modelling an organism's growth rate. If the lengthening is assumed linear with respect to time (that is $\dot{x}=K'$), defining γ using the equation:

$$\gamma(x)=\frac{C}{(1-x)^{\frac{1}{2}}}$$

entails the obvious characteristics:

$$X(t,x_0)=x_0+K't$$

and the exact same calculations as above would yield:

$$\begin{aligned}\tilde{n}(t,x_0)&=n_0(x_0)\exp(-Dt)\exp(-C\int_{t_0}^t\left(1-(x_0+K'(s-t_0))\right)^{-\frac{1}{2}}ds)\\ n(t,x)&=n_0(x_0)\exp(-Dt)\exp\left(\frac{2C}{K'}\left(\sqrt{1-x}-\sqrt{1-x+K't}\right)\right)\end{aligned}$$

the latter converging to the positive quantity $n_0(x_0)\exp(-Dt)\exp\left(-\frac{2C}{K'}\sqrt{K't}\right)>0$ as x tends to 1. This nonsensical condition is tantamount to an accumulation of the population at the $x=1$ value. In other words, such a modelling assumption will inevitably turn n to a δ_1 NDF.
More generally, unbounded rupture functions as proposed by Mantzaris [83] or Chrysinas [12] are forbidden to ensure that the support of the solution to (1.22) is relatively compact. The assumption of sublinear lengthening comes from Robert & al.'s raw experimental data and seems the most relevant when population balance equations are considered to describe the cell cycle. Indeed, it is thought increasingly difficult for any organism to maintain their growth rate constant as feeding an ever–growing cell membrane at a constant rate would likely end up mustering more resources than is available to them..

• If n is defined on a compact subset of $\mathbb{R}_+$, the result $n(t,\cdot) \in L^\infty([0,1[)\ \forall t$ is stronger than $n \in L^1([0,1[)$, because obviously $\|n(t,\cdot)\|_{L^1_{[0,1[}} \leq \bar{l}\|n(t,\cdot)\|_{L^\infty_{[0,1}}$. However, it does not contradict Perthame's [108] or Gabriel's [33] in any way. Furthermore, this stronger result does not require some of the assumptions featuring in Doumic & Gabriel's [20]. Firstly, $\dot{x}$ vanishes at $x = 1$. Also, there exists no real k_γ such that $\gamma(x)(1-x)^{-k_\gamma}$ is bounded on $[0,1[$. It remains to understand how the framework of the theorem affects the solution to the size–structured eigenproblem.

• Consider now an exponential growth framework, that is $\dot{x}$ reads:

$$\dot{x} = K'(1-x)^{-\beta}$$

with $\beta > 0$. The exact same calculations as enforced above yield:

$$K't = \frac{1}{\beta+1}\left((1-x_0)^{\beta+1} - (1-x(t))^{\beta+1}\right) \Leftrightarrow X(t,x_0) = 1 - \left((1-x_0)^{\beta+1} - (\beta+1)K't\right)^{\frac{1}{\beta+1}}$$

Then, with the same notations:

$$\begin{aligned}\frac{d\tilde{n}}{dt}(t,x_0) &= \frac{\partial}{\partial t}n(t,X(t,x_0)) + K'\left[(1-x_0)^{\beta+1} - (\beta+1)K't\right]^{-\frac{\beta}{\beta+1}}\frac{\partial}{\partial x}n(t,X(t,x_0)) \\ &= \frac{\partial}{\partial t}n(t,X(t,x_0)) + K'(1-x(t))^{-\beta}\frac{\partial}{\partial x}n(t,X(t,x_0))\end{aligned}$$

And the following linear differential equation will be solved:

$$\frac{d}{dt}\tilde{n}(t,x_0) = -D\tilde{n}(t,x_0) - \frac{\beta K'}{(1-x_0)^{\beta+1} - (\beta+1)K't}\tilde{n}(t,x_0) - \gamma(X(t,x_0))\tilde{n}(t,x_0)$$

It remains to evaluate the conditions under which the rupture process trumps the cell growth. If γ simply reads $\gamma(x) = C(1-x)^\alpha$, then:

$$n(t,x) = n_0(x_0)\exp(-Dt)\exp\left(-\int_0^t \frac{\beta K'}{(1-x_0)^{\beta+1} - (\beta+1)K's} + C\left[(1-x_0)^{\beta+1} - K's(\beta+1)\right]^{\frac{\alpha}{\beta+1}} ds\right)$$

In the right–hand side's integral, the first term will once again evidence a logarithm term:

$$\exp\left(-\int_0^t \frac{\beta K'}{(1-x_0)^{\beta+1} - (\beta+1)K's}ds\right) = \left(\frac{1-x(t)}{1-x_0}\right)^\beta$$

and the second term yields:

$$\exp\left(-C\left[(1-x_0)^{\beta+1} - K's(\beta+1)\right]^{\frac{\alpha}{\beta+1}} ds\right) = \exp\left(\frac{C}{K'(\alpha+\beta+1)}\left[(1-x(t))^{\alpha+\beta+1} - (1-x_0)^{\alpha+\beta+1}\right]\right)$$

At the end of the day, the right–hand side will be of form $X\exp(-X)$ (consequently tamed over the $[0,1]$ interval) if and only if $\alpha < -1-\beta$. In this case, the rupture function does not belong to L^1 though, making it impossible to apply Hölder's inequality to conclude on I's contracting property using Picard's fixed point theorem. If $\alpha \geq -1-\beta$ on the other hand, an inevitable accumulation of cells at $x = 1$ will lead to an infinite cell number at this particular value.

This remark echoes a previous brief aside mentioned in Diekmann & al.'s [17], but it is shown here that the $\alpha < 0$ condition is not a necessary assumption for the existence of a solution to (1.22).

2.2.3 Existence and uniqueness of Malthus eigenelements

Consider again the transport equation:

$$\begin{aligned} &\frac{\partial}{\partial t}n(t,x) + \frac{\partial}{\partial x}\left[\dot{x}n(t,x)\right] + \gamma(x)n(t,x) + Dn(t,x) = 2\int_x^1 \gamma(x')P(x,x')n(t,x')dx' \\ &n(t,0) = 0 = n(t,1) \end{aligned} \tag{2.13}$$

For $t \to \infty$, the solution to (2.13) will be aligned with the eigenvector corresponding to the transport–fragmentation operator's largest eigenvalue (as known as Malthus parameter). An existence and uniqueness result can be proven relying on Krein–Rutman's theorem for a regularised problem (allowing the strict positivity of the transport–fragmentation operator) and passing to the limit once convenient a priori bounds are extracted to this end. Two features of the rupture function must be gingerly addressed. Firstly, it diverges for $x \to 1$, requiring a truncation to ensure that Krein–Rutman's theorem holds. Secondly, its support will not be $[0,1[$ in practice, rather it will be assumed that no division is possible below a certain strictly positive threshold, preventing some of these events from being recorded at microscopic lengths. γ would read, for instance:
$\begin{cases} 0 & \text{if } x \leq x_\varepsilon \\ C\left[(1-x)^{\alpha-1} - (1-x_\varepsilon)^{\alpha-1}\right] & \text{if } x \geq x_\varepsilon \end{cases}$, $x_\varepsilon > 0$.

In this case, it must be proven that the distribution does not concentrate in the outside of the support of γ to such an extent that the NDF turns to a Dirac delta $n(t,l) = \delta_l$.
In a general way, the eigenelement problem breaks down into a primal and a dual equation:

$$\begin{aligned} &\frac{\partial}{\partial x}\left[\dot{x}N(x)\right] + \gamma(x)N(x) + (D+\lambda)N(x) = 2\int_x^{1-\zeta} \gamma(x')P(x,x')N(x')dx', 0 \leq x \leq x' \leq 1-\zeta \\ &N(0) = 0, \int_0^{1-\zeta} N(x)dx = 1,\ N \geq 0 \end{aligned} \tag{2.14}$$

$$\begin{aligned} &-\dot{x}\frac{\partial}{\partial x}\phi(x) + \gamma(x)\phi(x) + (D+\lambda)\phi(x) = 2\gamma(x)\int_0^x P(x',x)\phi(x')dx', 0 \leq x' \leq x \leq 1-\zeta \\ &\int_0^{1-\zeta} N(x)\phi(x)dx = 1,\ \phi \geq 0 \end{aligned} \tag{2.15}$$

Without loss of generality, one can set $\tilde{\lambda} = \lambda + D$. The following theorem will be proven in the rest of the section:

Theorem 2.2.2. *Consider the system* (2.14) - (2.15). *Suppose that P is a probability density kernel. If γ reads:*

$$\gamma(x) = C(1-x)^{\alpha-1}$$

and $\dot{x}$ reads:

$$\dot{x} = K(1-x)^{\beta}$$

with $0 < \alpha < \beta < \min(2\alpha, 1)$*, then the system admits a unique triplet* (λ, N, ϕ) *solution, and we have:*

$$N \in L^1 \cap \mathcal{C}^1([0,1[)$$
$$\phi \in L^1 \cap L^\infty([0,1[)$$
$$\tilde{\lambda} \leq \dot{x}(0)$$

Proof.

Sketch of the proof

The proof will involve the regularised problem:

$$\frac{\partial}{\partial x}\left[\dot{x}N_\varepsilon^\zeta(x)\right] + \gamma_\varepsilon^\zeta(x)N_\varepsilon^\zeta(x) + (D + \lambda_\varepsilon^\zeta)N_\varepsilon^\zeta(x) = 2\int_x^{1-\zeta} \gamma_\varepsilon^\zeta(x')P(x,x')N_\varepsilon^\zeta(x')dx', 0 \leq x \leq x' \leq 1-\zeta \tag{2.16}$$

$$N_\varepsilon^\zeta(0) = 0, \int_0^{1-\zeta} N_\varepsilon^\zeta(x)dx = 1,\ N_\varepsilon^\zeta \geq 0$$

$$-\dot{x}\frac{\partial}{\partial x}\phi_\varepsilon^\zeta(x) + \gamma_\varepsilon^\zeta(x)\phi_\varepsilon^\zeta(x) + (D + \lambda_\varepsilon^\zeta)\phi_\varepsilon^\zeta(x) = 2\gamma_\varepsilon^\zeta(x)\int_0^x P(x',x)\phi_\varepsilon^\zeta(x')dx', 0 \leq x' \leq x \leq 1-\zeta \tag{2.17}$$

$$\phi_\varepsilon^\zeta(1-\zeta) = 0,\ \int_0^{1-\zeta} N_\varepsilon^\zeta\phi_\varepsilon^\zeta(x)dx = 1,\ \phi_\varepsilon^\zeta \geq 0$$

with γ_ε^ζ a strictly positive function on the interval $[0, 1-\zeta]$. A rupture function with support on $[0, 1-\zeta]$ could consist, for instance, in $\gamma_\varepsilon^\zeta(x) = \gamma(x) - \gamma(x_\varepsilon)$, $0 \leq x \leq 1-\zeta$, γ being defined as in the previous section, $x_\varepsilon < 0$. It is worth mentioning that this refinement has absolutely no impact on its integrability: $\|\gamma_\varepsilon^\zeta\|_{L^1} \leq \|\gamma\|_{L^1}$. By the way, such a strategy has already been employed by Doumic in [19]. Finally, $D + \lambda_\varepsilon^\zeta$ can be renamed $\tilde{\lambda_\varepsilon^\zeta}$.

Applying Krein–Rutman's theorem in the L^∞ space has already been achieved by Doumic & Gabriel (cf. [20]) for a very similar eigenproblem as (2.16)-(2.17), meaning that despite the differences in the respective frameworks, it should be done with no difficulties in the beginning of the proof. In order to estimate the $W^{1,\infty}$ norm of the difference between two solutions of (2.16), more regularity has to be somehow retrieved from the said difference, the idea consisting in exploiting the regularising feature of the redistribution integral to upper–bound it with more precision than Hölder's inequality allows.

Then, to take the limits in ε and ζ in both (2.16) and (2.17), very similar estimates are retrieved from the equations and Ascoli's theorem will be used in both cases to conclude. The $\beta < 2\alpha$ inequality will manifest from the necessity to bound $(1-x)^{2\alpha-\beta}$ on the $[0,1[$ interval that appears as a mandatory step to estimate the solution to (2.16) in $W^{1,\infty}$. Such an estimate is out of reach of the solution to (2.17) that is only equicontinuous, what does not prevent from using Ascoli's theorem anyway. Finally, given that the operator

$$R : N \to \frac{\partial}{\partial x}N$$

is linear, $R(N_\varepsilon^\zeta)$ is continuous if and only if it is bounded, meaning that proving N_ε^ζ's $\mathcal{C}^1$ regularity comes down to the proof of $N_\varepsilon^\zeta \in W^{1,\infty}$.

Demonstrating the uniqueness of the solutions exploits the Generalised Relative Entropy (GRE) principle and the proof is in fact identical to Michel & al.'s contribution in the context of closely related biological systems (cf. [87]).

Application of Krein–Rutman's theorem

The first step of the proof consists in ensuring that Krein–Rutman's theorem applies to the regularised problem (2.16). The framework is here the functional space $W^{1,\infty}[0,1-\zeta]$, that is obviously equipped with the norm: $\|\cdot\|_{W^{1,\infty}[0,1-\zeta]} = \max_{0\le x\le 1-\zeta}(\cdot(x)) + \max_{0\le x\le 1-\zeta}(\dot{\cdot}(x))$. Consider to this end an operator $O : M \mapsto N_\varepsilon^\zeta$ such that:

$$\frac{\partial}{\partial x}\left[\dot{x}N_\varepsilon^\zeta(x)\right] + \gamma_\varepsilon^\zeta(x)N_\varepsilon^\zeta(x) + (\mu+D)N_\varepsilon^\zeta(x) = 2\int_x^{1-\zeta}\gamma_\varepsilon^\zeta(x')P(x,x')M(x')dx',\ 0\le x\le x'\le 1-\zeta \tag{2.18}$$

$$N_\varepsilon^\zeta(0) = 0,\ N_\varepsilon^\zeta(x)\ge 0\ \forall x\in[0,1-\zeta],\ \int_0^{1-\zeta}N_\varepsilon^\zeta(x)dx = 1$$

A unique solution to (2.18) exists if and only if O is a contraction, by virtue of Picard's fixed point theorem. O being linear, if $N_{\varepsilon,1}^\zeta = O(M_1)$, $N_{\varepsilon,2}^\zeta = O(M_2)$ are two distinct solutions to (2.18), the difference $N_{\varepsilon,1}^\zeta - N_{\varepsilon,2}^\zeta$ satisfies the equation:

$$\begin{aligned}\frac{\partial}{\partial x}\left[\dot{x}(N_{\varepsilon,1}^\zeta - N_{\varepsilon,2}^\zeta)(x)\right] + \gamma_\varepsilon^\zeta(x)(N_{\varepsilon,1}^\zeta - N_{\varepsilon,2}^\zeta)(x) + (\mu+D)(N_{\varepsilon,1}^\zeta - N_{\varepsilon,2}^\zeta)(x)\\ = 2\int_x^{1-\zeta}\gamma_\varepsilon^\zeta(x')P(x,x')(M_1-M_2)(x')dx'\end{aligned}$$

with $(N_{\varepsilon,1}^\zeta - N_{\varepsilon,2}^\zeta)(0) = 0$, and the absolute value $|N_{\varepsilon,1}^\zeta - N_{\varepsilon,2}^\zeta|$ will obey the inequation:

$$\begin{aligned}\frac{\partial}{\partial x}\left[\dot{x}|N_{\varepsilon,1}^\zeta - N_{\varepsilon,2}^\zeta|(x)\right] + \gamma_\varepsilon^\zeta(x)|N_{\varepsilon,1}^\zeta - N_{\varepsilon,2}^\zeta(x)| + (\mu+D)|N_{\varepsilon,1}^\zeta - N_{\varepsilon,2}^\zeta|(x)\\ \le 2\int_x^{1-\zeta}\gamma_\varepsilon^\zeta(x')P(x,x')|M_1-M_2|(x')dx'\end{aligned}$$

with obviously $|N_{\varepsilon,1}^\zeta - N_{\varepsilon,2}^\zeta|(0) = 0$. One can integrate this inequality making use of Grönwall's lemma and Duhamel's formula:

$$\begin{aligned}\frac{\partial}{\partial x}|N_{\varepsilon,1}^\zeta - N_{\varepsilon,2}^\zeta|(x) \le -\frac{1}{\dot{x}}\left(\frac{\partial\dot{x}}{\partial x} + \gamma_\varepsilon^\zeta(x) + \mu + D\right)|N_{\varepsilon,1}^\zeta - N_{\varepsilon,2}^\zeta|(x)\\ +\frac{2}{\dot{x}}\int_x^{1-\zeta}\gamma_\varepsilon^\zeta(x')P(x,x')|M_1-M_2|(x')dx'\\ \Rightarrow |N_{\varepsilon,1}^\zeta - N_{\varepsilon,2}^\zeta|(x) \le |N_{\varepsilon,1}^\zeta - N_{\varepsilon,2}^\zeta|(0)\exp\left(\int_0^x -\frac{1}{\dot{x}}\left(\frac{\partial\dot{x}}{\partial x} + \gamma_\varepsilon^\zeta(x') + \mu + D\right)dx'\right) \qquad (2.19)\\ + \int_0^x \exp\left(\int_{x'}^x -\frac{1}{\dot{x}}\left(\frac{\partial\dot{x}}{\partial x} + \gamma_\varepsilon^\zeta(x'') + \mu + D\right)dx''\right)\frac{2}{\dot{x}}\int_{x'}^{1-\zeta}\gamma_\varepsilon^\zeta(x'')P(x',x'')|M_1-M_2|(x'')dx''dx'\end{aligned}$$

On (2.2.3), the right–hand side's first term vanishes due to the boundary condition. Given our hypotheses pertaining to γ and $\dot{x}$, the second term can be bounded from the following steps:

- $\int_{x'}^{x} -\frac{1}{\dot{x}(x'')}\frac{\partial \dot{x}(x'')}{\partial x}dx'' = \ln\left(\frac{\dot{x}(x')}{\dot{x}(x)}\right)$
- $\int_{x'}^{x} -\frac{\gamma_\varepsilon^\zeta(x'')}{\dot{x}(x'')}dx'' = \frac{C}{K'(\beta-\alpha)}\left[(1-x')^{\alpha-\beta} - (1-x)^{\alpha-\beta}\right] + \int_{x'}^{x}\frac{\gamma(x_\varepsilon)}{\dot{x}(x'')}dx''$

  $\int_{x'}^{x} -\frac{\gamma_\varepsilon^\zeta(x'')}{\dot{x}(x'')}dx'' \leq \frac{C}{K'(\beta-\alpha)}\left[(1-x')^{\alpha-\beta} - (1-x)^{\alpha-\beta}\right]$
- $\int_{x'}^{x} -\frac{\mu+D}{\dot{x}(x'')}dl'' = \frac{(\mu+D)}{K'(1-\beta)}\left[(1-x)^{1-\beta} - (1-x')^{1-\beta}\right]$
- $\int_{x'}^{1-\zeta}\gamma_\varepsilon^\zeta(x'')P(x',x'')|M_1-M_2|(x'')dx'' \leq \int_0^{1-\zeta}\gamma_\varepsilon^\zeta(x')P(x,x')|M_1-M_2|(x')dx'$

  $\leq A||M_1-M_2||_{L^\infty}||\gamma_\varepsilon^\zeta||_{L^1}$

with $A = \sup_{x'\leq x'', 0\leq x''\leq 1-\zeta} P(x',x'')$, provided that P is a probability density kernel. Therefore:

$$|N_{\varepsilon,1}^\zeta - N_{\varepsilon,2}^\zeta|(x) \leq \frac{2A||M_1-M_2||_{L^\infty}||\gamma_\varepsilon^\zeta||_{L^1}}{K'}(1-x)^{-\beta}\int_0^x \exp\left(\frac{C}{K'(\beta-\alpha)}\left[(1-x')^{\alpha-\beta}-(1-x)^{\alpha-\beta}\right]\right)$$
$$\exp\left(\frac{(\mu+D)}{K'(1-\beta)}\left[(1-x)^{1-\beta}-(1-x')^{1-\beta}\right]\right)dx'$$

Hereinafter, the following notations will be used:

$$E(x) = \exp\left(\frac{C}{K'(\beta-\alpha)}(1-x)^{\alpha-\beta}\right)$$
$$F(x) = \exp\left(\frac{(\mu+D)}{K'(1-\beta)}(1-x)^{1-\beta}\right)$$

meaning that one can rewrite:

$$|N_{\varepsilon,1}^\zeta - N_{\varepsilon,2}^\zeta|(x) \leq \frac{2A||M_1-M_2||_{L^\infty}||\gamma_\varepsilon^\zeta||_{L^1}}{K'}(1-x)^{-\beta}\frac{F(x)}{E(x)}\int_0^x \frac{E(x')}{F(x')}dx'$$

Then, multiplying and dividing by $(1-x')^\beta$ yields:

$$|N_{\varepsilon,1}^\zeta - N_{\varepsilon,2}^\zeta|(x) \leq \frac{2A||M_1-M_2||_{L^\infty}||\gamma_\varepsilon^\zeta||_{L^1}}{K'}(1-x)^{-\beta}\frac{F(x)}{E(x)}\int_0^x (1-x')^\beta E(x')(1-x')^{-\beta}\frac{1}{F(x')}dx'$$
$$\leq 2A||M_1-M_2||_{L^\infty}||\gamma_\varepsilon^\zeta||_{L^1}\max_{0\leq x'\leq x}\frac{(1-x')^\beta E(x')}{(1-x)^\beta E(x)}F(x)\int_0^x \frac{1}{K'F(x')}(1-x')^{-\beta}dx'$$

once Hölder's inequality has been used in the right–hand side's integral. Firstly, the resulting integral can be analytically calculated:

$$\int_0^x \frac{1}{K'F(x')}(1-x')^{-\beta}dx' = \frac{1}{\mu+D}\left[\frac{1}{F(x)} - \frac{1}{F(0)}\right]$$

Then, one has to evaluate the critical points of the function:

$$f : x \mapsto (1-x)^{\beta} E(x)$$

Indeed, if f admits a minimum on the interior of the $[0,1[$ interval, the upper bound to $\max_{0 \leq x' \leq x} \frac{(1-x')^{\beta} E(x')}{(1-x)^{\beta} E(x)}$ will be immediately established. For instance, if f is strictly increasing on $[0,1[$, for any $x' \leq x$, then $f(x')/f(x)$ will be less than 1. In this context, the calculation of f's derivative follows:

$$f'(x) = E(x)\left[\frac{C}{K'}(1-x)^{\alpha-1} - \beta(1-x)^{\beta-1}\right]$$

and vanishes at $x = x^{\star} = 1 - \left(\frac{C}{\beta K'}\right)^{\frac{1}{\beta-\alpha}}$. It is worth mentioning that:

- $f(0) = \exp\left(\frac{C}{K'(\beta-\alpha)}\right) > 0$
- $f'(0) = \exp\left(\frac{C}{K(\beta-\alpha)}\right)\frac{C - K'\beta}{K'}$

which sign depends on the very quantity $(C - K'\beta)$ that determines the sign of $x^{\star}$. In this context, two options have to be considered:

- $C \geq K'\beta$: in this case, $x^{\star} \leq 0$ and f is increasing on $[0,1[$ in all generality. Therefore:

$$x' \leq x \Leftrightarrow f(x') \leq f(x) \text{ and: } \max_{0 \leq x' \leq x} \frac{f(x')}{f(x)} = 1$$

- $C < K'\beta$: in this case, $0 < x^{\star} < 1$ and the function is decreasing on $[0, x^{\star}]$ then increasing on $[x^{\star}, 1[$. Therefore:

$$\max_{0 \leq x' \leq x} \frac{f(x')}{f(x)} = \frac{f(0)}{f(x^{\star})}$$

and one can conclude that:

$$|N^{\zeta}_{\varepsilon,1} - N^{\zeta}_{\varepsilon,2}|(x) \leq \frac{2A\|M_1 - M_2\|_{L^{\infty}} \Xi \|\gamma^{\zeta}_{\varepsilon}\|_{L^1}}{\mu + D}\left[1 - \frac{F(x)}{F(0)}\right] \leq \frac{2A\|M_1 - M_2\|_{L^{\infty}} \Xi \|\gamma^{\zeta}_{\varepsilon}\|_{L^1}}{\mu + D}$$

because $0 \leq F(x)/F(0) \leq 1$, where Ξ is defined as:

$$\Xi = \max\left(1, \frac{f(0)}{f(x^{\star})}\right)$$

Hence, there exists a large enough μ to retrieve the inequality:

$$\|N^{\zeta}_{\varepsilon,1} - N^{\zeta}_{\varepsilon,2}\|_{L^{\infty}} \leq \frac{1}{2}\|M_1 - M_2\|_{L^{\infty}} \tag{2.20}$$

As a consequence, only the L^{∞} estimate on the derivative $\|\nabla_x (N^{\zeta}_{\varepsilon,1} - N^{\zeta}_{\varepsilon,2})\|$ is missing to conclude on the application of Krein–Rutman's theorem. In this case, a tighter upper bound will be required beforehand:

$$|N^{\zeta}_{\varepsilon,1} - N^{\zeta}_{\varepsilon,2}|(x) \leq \frac{2}{K'}(1-x)^{-\beta}\frac{F(x)}{E(x)}\int_0^x \frac{E(x')}{F(x')}\int_{x'}^1 C(1-x'')^{\alpha-1} P(x', x'')|M_1 - M_2|(x'')dx''dx'$$

because the redistribution integral will have to be estimated with more finesse than Hölder's inequality can afford. Without loss of generality, one can consider that $M_1 - M_2$ vanishes at $1-\zeta$: indeed, O being a linear operator, replacing say, M_2, with $M_2 \times M_1(1-\zeta)/M_2(1-\zeta)$ would have absolutely no consequence on the estimation of $\|N^{\zeta}_{\varepsilon,1} - N^{\zeta}_{\varepsilon,2}\|_{W^{1,\infty}}$. This assessment being agreed, keeping in mind that $M_1 - M_2$ is Lipschitz, one can infer:

$$(M_1 - M_2)(x) + \int_x^{1-\zeta} \nabla_x (M_1 - M_2)(x')dx' = 0 \Leftrightarrow |(M_1 - M_2)(x)| \leq L(1-x) \forall \zeta > 0,\ x \in [0,1[$$

with L the Lipschitz constant for $M_1 - M_2$. Hence:

$$\int_{x'}^{1} C(1-x'')^{\alpha-1} P(x', x'') |(M_1 - M_2)(x'')| dx'' \leq \frac{CLA}{\alpha+1}(1-x')^{\alpha+1} \tag{2.21}$$

By the way, this calculation entails the estimate on the last term of:

$$\nabla_x |N^{\zeta}_{\varepsilon,1} - N^{\zeta}_{\varepsilon,2}|(x) \leq -\left[\ln(\dot{x}(x)) + \frac{\gamma^{\zeta}_{\varepsilon}(x)}{\dot{x}(x)} + \frac{\mu + D}{\dot{x}(x)}\right] |N^{\zeta}_{\varepsilon,1} - N^{\zeta}_{\varepsilon,2}|(x) + \frac{2}{\dot{x}} \int_x^1 \gamma(x') P(x, x') |M_1 - M_2|(x') dx'$$

To treat the first term of the said estimate, it can be seen as a sum of powers of $(1-x)$ multiplied by n, so much so taming its most constraining contribution is enough to conclude. To this end, see that:

$$(1-x)^{\alpha-1-\beta} |N^{\zeta}_{\varepsilon,1} - N^{\zeta}_{\varepsilon,2}|(x) \leq \frac{2}{K'}(1-x)^{\alpha-1-2\beta} \frac{CLA}{\alpha+1} \frac{F(x)}{E(x)} \int_0^x (1-x')^{\alpha+1} \frac{E(x')}{F(x')} dx'$$

Then, multiplying and dividing by $(1-x')^{2\beta-2\alpha}$ and using once again that $x' \leq x$ yield:

$$(1-x)^{\alpha-1-\beta} |N^{\zeta}_{\varepsilon,1} - N^{\zeta}_{\varepsilon,2}|(x) \leq \frac{2}{K'}(1-x)^{\alpha-2\beta-1} \frac{CLA}{\alpha+1} \frac{F(x)}{E(x)}$$
$$\int_0^x (1-x')^{1+2\beta-\alpha} (1-x')^{2\alpha-\beta} (1-x')^{-\beta} \frac{E(x')}{F(x')} dx'$$

Once again, applying Hölder's inequality would boil down to:

$$(1-x)^{\alpha-1-\beta} |N^{\zeta}_{\varepsilon,1} - N^{\zeta}_{\varepsilon,2}|(x) \leq \frac{2CLA}{K'(\alpha+1)} F(x) \max_{0 \leq x' \leq x} \frac{(1-x')^{1+2\beta-\alpha} E(x')}{(1-x)^{1+2\beta-\alpha} E(x)} \tag{2.22}$$
$$\int_0^x (1-x')^{2\alpha-\beta} (1-x')^{-\beta} \frac{1}{F(x')} dx'$$

The crucial point of the proof consists in remarking that the quantity $(1-x')^{2\alpha-\beta}$ is upper bounded by 1 on the $[0,x]$ interval if and only if $2\alpha - \beta > 0$. Then, proceeding as above to calculate the integral:

$$\int_0^x (1-x')^{-\beta} \frac{1}{F(x')} dx' = \frac{K'}{\mu + D}\left[\frac{1}{F(x)} - \frac{1}{F(0)}\right]$$

the bulk of the calculation will bear upon the maximum of the remaining factor of (2.22)'s right–hand side. A similar calculation as above would yield:

$$\frac{d}{dx}\left[(1-x)^{1+2\beta-\alpha} E(x)\right] = (1-x)^{1+2\beta-\alpha} E(x)\left[-(1+2\beta-\alpha)(1-x)^{-1} + \frac{C}{K'}(1-x)^{\alpha-\beta-1}\right]$$

Obviously, the inflection point is this time $x^\star = 1 - \left(\frac{C}{K'(1+2\beta-\alpha)}\right)^{\frac{1}{\beta-\alpha}}$, which sign will define the behaviour of the quantity of interest. A similar maximum Ξ' can be found, turning the L^∞ estimate into:

$$(1-x)^{\alpha-1-\beta}|N^\zeta_{\varepsilon,1} - N^\zeta_{\varepsilon,2}|(x) \leq \frac{2CLA\Xi'}{(\mu+D)(\alpha+1)}$$

Hence, there exists a large enough μ to retrieve the inequality:

$$\|N^\zeta_{\varepsilon,1} - N^\zeta_{\varepsilon,2}\|_{W^{1,\infty}} \leq \eta\|M_1 - M_2\|_{W^{1,\infty}},\ \eta < 1$$

meaning that Krein–Rutman's theorem guarantees the existence of a solution to the regularised eigenproblem.

Passing to the limit in the primal equation

In order to take the limits $x_\varepsilon \to 0$ and $\zeta \to 0$ to conclude on the existence of solutions to (2.16) and (2.17), one has to extract appropriate a priori bounds on the solution. In this case, for $x_\varepsilon \to 0$, $\zeta \to 0$, N^ζ_ε has to be uniformly upper–bounded in ε and ζ in the L^∞ space along with $\nabla_x N^\zeta_\varepsilon$.

First, an upper bound to λ comes from an integration of (2.16)'s first–order moment:

$$(1-\zeta)\dot{x}(1-\zeta)N^\zeta_\varepsilon(1-\zeta) - \int_0^{1-\zeta}\dot{x}(x)xN^\zeta_\varepsilon(x)dx + \tilde{\lambda}^\zeta_\varepsilon\int_0^{1-\zeta}xN^\zeta_\varepsilon(x)dx = 0$$
$$\tilde{\lambda}^\zeta_\varepsilon\int_0^{1-\zeta}xN^\zeta_\varepsilon(x)dx \leq \int_0^{1-\zeta}x\dot{x}N^\zeta_\varepsilon(x)dx \leq \dot{x}(0)\int_0^{1-\zeta}xN^\zeta_\varepsilon(x)dx \Rightarrow \tilde{\lambda}^\zeta_\varepsilon \leq \dot{x}(0) = K$$

Secondly, an estimate for the integral $\int_0^{1-\zeta}\gamma^\zeta_\varepsilon(x)N^\alpha_\varepsilon(x)dx$ comes from the use of the zeroth–order moment of (2.16):

$$\begin{aligned}\dot{x}(1-\zeta)N^\zeta_\varepsilon(1-\zeta) + \int_0^{1-\zeta}\gamma^\zeta_\varepsilon(x)N^\zeta_\varepsilon(x)dx + \tilde{\lambda}^\zeta_\varepsilon\int_0^{1-\zeta}N^\zeta_\varepsilon(x)dx \\ = 2\int_0^{1-\zeta}\int_{x'}^{1-\zeta}\gamma^\zeta_\varepsilon(x'')P(x',x'')N^\zeta_\varepsilon(x'')dx''dx' \\ \Leftrightarrow \int_0^{1-\zeta}\gamma^\zeta_\varepsilon(x)N^\zeta_\varepsilon(x)dx = \tilde{\lambda}^\zeta_\varepsilon + \dot{x}(1-\zeta)N^\zeta_\varepsilon(1-\zeta) \leq \tilde{\lambda}^\zeta_\varepsilon + \zeta^\beta N^\zeta_\varepsilon(1-\zeta)\end{aligned}$$

N^ζ_ε being positive and integrable by assumption on the compact $[0, 1-\zeta]$, $N^\zeta_\varepsilon(1-\zeta)$ is necessarily a finite quantity to the point where one can set a real $\Lambda > 0$ such that:

$$\int_0^{1-\zeta}\gamma^\zeta_\varepsilon(x)N^\zeta_\varepsilon(x)dx \leq \Lambda$$

Then, the uniform L^∞ bound on N^ζ_ε is an immediate consequence of the properties satisfied by the rupture and lengthening functions. Indeed, proceeding as above:

$$\frac{\partial\dot{x}}{\partial x}N^\zeta_\varepsilon(x) + \dot{x}\frac{\partial N^\zeta_\varepsilon(x)}{\partial x} + \gamma^\zeta_\varepsilon(x)N^\zeta_\varepsilon(x) + \tilde{\lambda}^\zeta_\varepsilon N^\zeta_\varepsilon(x) = 2\int_x^{1-\zeta}\gamma^\zeta_\varepsilon(x')P(x,x')N^\zeta_\varepsilon(x')dx'$$
$$\frac{\partial N^\zeta_\varepsilon(x)}{\partial x} = -\frac{1}{\dot{x}(x)}\left(\frac{\partial\dot{x}}{\partial x} + \gamma^\zeta_\varepsilon(x) + \tilde{\lambda}^\zeta_\varepsilon\right)N^\zeta_\varepsilon(x) + \frac{2}{\dot{x}}\int_x^{1-\zeta}\gamma^\zeta_\varepsilon(x')P(x,x')N^\zeta_\varepsilon(x')dx'$$

If P is a probability density kernel (once again, set $A = \max_{0\leq x\leq x'\leq 1-\zeta} P(x,x')$), the upper bound to the right–hand-side's second term immediately comes from Hölder's inequality. Then, another use of Grönwall's lemma along with Duhamel's formula yields:

$$\frac{\partial N_\varepsilon^\zeta(x)}{\partial x} \leq -\frac{1}{\dot{x}(x)}\left(\frac{\partial \dot{x}}{\partial x} + \gamma_\varepsilon^\zeta(x) + \tilde{\lambda_\varepsilon^\zeta}\right) N_\varepsilon^\alpha(x) + \frac{2AC}{\dot{x}(x)}\Lambda. \text{ Hence, after integration}$$

$$N_\varepsilon^\zeta(x) \leq N_\varepsilon^\zeta(0)\exp\left[-\int_0^x \frac{1}{\dot{x}(x')}\left(\frac{\partial \dot{x}}{\partial x} + \gamma_\varepsilon^\zeta(x') + \tilde{\lambda_\varepsilon^\zeta}\right)dx'\right]$$

$$+\int_0^x \frac{2AC}{\dot{x}(x')}\Lambda\exp\left[-\int_{x'}^x \frac{1}{\dot{x}(x'')}\left(\frac{\partial \dot{x}}{\partial x} + \gamma_\varepsilon^\zeta(x'') + \tilde{\lambda_\varepsilon^\zeta}\right)dx''\right]dx'$$

$$\Rightarrow N_\varepsilon^\zeta(x) \leq \frac{2AC\Lambda}{\dot{x}(x)}\frac{G(x)}{E(x)}\int_0^x \frac{E(x')}{G(x')}dx' \tag{2.23}$$

using the definition of G:

$$G : x \mapsto \exp\left(\frac{\tilde{\lambda_\varepsilon^\zeta}}{(1-\beta)K'}(1-x)^{1-\beta}\right)$$

On the model of the previous proof, the L^∞ estimate to N_ε^ζ is immediate once (2.2.3)'s right–hand side is multiplied and divided by $(1-x')^\beta$. The same approach amounts to integrate $(1-x)^{-\beta}/G(x)$ and study the quantity $(1-x)^\beta E(x)$, which leads to the definition of a real Ξ such that:

$$N_\varepsilon^\zeta(x) \leq 2AC\Lambda\Xi \tag{2.24}$$

The next step consists in bounding the quantities $\frac{1}{\dot{x}}N_\varepsilon^\zeta$ and $\frac{\gamma_\varepsilon^\zeta}{\dot{x}}N_\varepsilon^\zeta$ in $L^\infty([0,1-\zeta])$. As previously remarked, these quantities amount to mere $(1-x)^{\cdot}N_\varepsilon^\zeta$ expressions. Hence, only the estimate for $\gamma_\varepsilon^\zeta/\dot{x}N_\varepsilon^\zeta$ will be retrieved. To this end, from the same starting point:

$$\frac{\gamma_\varepsilon^\zeta}{\dot{x}}N_\varepsilon^\zeta(x) \leq \frac{2C}{K'}(1-x)^{\alpha-2\beta-1}\frac{G(x)}{E(x)}\int_0^x \frac{E(x')}{G(x')}\int_{x'}^{1-\zeta}(1-x'')^{\alpha-1}P(x',x'')N_\varepsilon^\zeta(x'')dx''dx'$$

One first treats the last integral with Hölder's inequality:

$$\int_{x'}^{1-\zeta}(1-x'')^{\alpha-1}P(x',x'')N_\varepsilon^\zeta(x'')dx'' \leq 2A^2C\Lambda\Xi\frac{1}{\alpha+1}(1-x')^{\alpha+1}$$

and injects this upper bound into the previous inequality constraining $\gamma_\varepsilon^\zeta/\dot{x}N_\varepsilon^\zeta$:

$$\frac{\gamma_\varepsilon^\zeta}{\dot{x}}N_\varepsilon^\zeta(x) \leq \frac{4C^2A^2\Lambda\Xi}{K'(\alpha+1)}(1-x)^{\alpha-2\beta-1}\frac{G(x)}{E(x)}\int_0^x \frac{E(x')}{G(x')}dx'$$

As was previously done, a further multiplication/division by $(1-x')^{2\alpha-2\beta}$ will lead to the definition of a constant Ξ' such that:

$$\frac{\gamma_\varepsilon^\zeta}{\dot{x}}N_\varepsilon^\zeta(x) \leq \frac{4C^2A^2\Lambda\Xi\Xi'}{K'(\alpha+1)} \tag{2.25}$$

Using the same reasoning to bound $\frac{1}{\dot{l}}\frac{\partial \dot{l}}{\partial l}N_\varepsilon^\zeta$ would come down to

$$\frac{1}{\dot{l}}\frac{\partial \dot{l}}{\partial l}N_\varepsilon^\zeta(l) \leq \frac{4C^2A^2\Lambda\Xi\Xi''}{K'(\alpha+1)}$$

with Ξ'' defined as:

$$\Xi'' = \max_{0\leq x'\leq x} \frac{(1-x')^{1+\beta}E(x')}{(1-x)^{1+\beta}E(x)}$$

Therefore, the $W^{1,\infty}$ bound to N_ε^ζ is guaranteed given that this quantity is the sum of bounded quantities. Furthermore, none of these depends on x_ε, to the point where taking the limit $x_\varepsilon \to 0$ will be straightforward.

On the other hand, the $\zeta \to 0$ limit will be taken without a hunch either, because in this case, the estimate for $\int_0^1 \gamma^\zeta(x)N^\zeta(x)dx$ will be more amenable once the primal equation is integrated:

$$0 - 0 + \int_0^1 \gamma^\zeta(x)N^\zeta(x)dx + \tilde{\lambda^\zeta} = 2\int_0^1 \gamma^\zeta(x)N^\zeta(x)dx \Leftrightarrow \int_0^1 \gamma^\zeta(x)N^\zeta(x)dx = \tilde{\lambda^\zeta} \tag{2.26}$$

and the same calculations as above can be carried out with no difficulty.

The family $(\tilde{\lambda_\varepsilon^\zeta})_{\varepsilon\geq 0,\zeta\geq 0}$ is relatively compact in $\mathbb{R}_+$, therefore Bolzano–Weierstrass's theorem guarantees that it admits a subsequence that converges to the transport–fragmentation operator's (2.14) largest eigenvalue $\tilde{\lambda}$. The same reasoning applies to the $(N_\varepsilon^\zeta)_{\varepsilon\geq 0,\zeta\geq 0}$ family, using its uniform boundedness in $W^{1,\infty}[0,1-\zeta]$: by virtue of Ascoli's theorem, there exists a subsequence that converges to a limit N that is also Lipschitz continuous. N is by construction a solution to (2.14) on $[0,1[$.

Passing to the limit in the dual equation

Regarding the dual (2.15) equation, the same type of estimate can be retrieved from the dual equation (2.17), and 4 steps are required to conclude:

- $\phi_\varepsilon^\zeta \in L^\infty([0,1[)$: from the definition of (2.17), one gets:

$$\frac{\partial}{\partial x}\phi_\varepsilon^\zeta = \frac{\gamma_\varepsilon^\zeta}{\dot{x}}\phi_\varepsilon^\zeta + \frac{\tilde{\lambda}}{\dot{x}}\phi_\varepsilon^\zeta - 2\frac{\gamma_\varepsilon^\zeta}{\dot{x}}\int_0^x \phi_\varepsilon^\zeta(x')P(x',x)dx'$$

meaning that $\frac{\partial}{\partial x}\phi_\varepsilon^\zeta|_{x=0} > 0$. Therefore, if ϕ_ε^ζ had a singularity at $x=0$, the condition $\int_0^1 N_\varepsilon^\zeta\phi_\varepsilon^\zeta dx = 1$ could not be met: the integral would be divergent. This forces ϕ_ε^ζ to be L^∞ (and consequently L^1).

- Using once again Duhamel's rule to integrate (2.17), an analytical expression for ϕ_ε^ζ is forthcoming:

$$\phi_\varepsilon^\zeta(x) = \chi\exp\left(\int_0^x \frac{\tilde{\lambda}}{\dot{x}(x')} + \frac{\gamma_\varepsilon^\zeta(x')}{\dot{x}(x')}dx'\right)$$
$$-\int_0^x \exp\left(\int_{x'}^x \frac{\tilde{\lambda}}{\dot{x}(x'')} + \frac{\gamma_\varepsilon^\zeta(x'')}{\dot{x}(x'')}dx''\right)2\frac{\gamma_\varepsilon^\zeta(x')}{\dot{x}(x')}\int_0^{x'} \phi_\varepsilon^\zeta(x'')P(x'',x')dx''dx'$$

with χ a constant that is obviously determined by the boundary condition $\phi_\varepsilon^\zeta(1-\zeta) = 0$:

$$\chi = \int_0^{1-\zeta}\exp\left(-\int_0^{x'}\frac{\tilde{\lambda}}{\dot{x}(x'')}dx'' - \int_0^{x'}\frac{\gamma_\varepsilon^\zeta(x'')}{\dot{x}(x'')}dx''\right)2\frac{\gamma_\varepsilon^\zeta(x')}{\dot{x}(x')}\int_0^{x'}\phi_\varepsilon^\zeta(x'')P(x'',x')dx''dx'$$

yielding, having introduced the quantity H such as:

$$H : x \mapsto \exp\left(\frac{\tilde{\lambda}}{K'(1-\beta)}(1-x)^{1-\beta}\right):$$

$$\phi_\varepsilon^\zeta(x) = \frac{2C}{K'}\frac{E(x)}{H(x)}\int_x^{1-\zeta}\frac{H(x')}{E(x')}(1-x')^{\alpha-\beta-1}\int_0^{x'}\phi_\varepsilon^\zeta(x'')P(x'',x')dx''dx' \tag{2.27}$$

• ϕ_ε^ζ is equicontinuous: consider a real $\delta > 0$ and the ball $B(x,\delta) \subset [0, 1-\zeta]$. From (2.27), one remarks after adding/substracting the quantity $\frac{2C}{K'}\frac{E(x)}{H(x)}\int_{x+\delta}^1\frac{H(x')}{E(x')}(1-x')^{\alpha-\beta-1}\int_0^{x'}\phi_\varepsilon^\zeta(x'')P(x'',x')dx''dx'$ that the difference $\phi_\varepsilon^\zeta(x) - \phi_\varepsilon^\zeta(x+\delta)$ reads:

$$\begin{aligned}\phi_\varepsilon^\zeta(x) - \phi_\varepsilon^\zeta(x+\delta) &= \frac{2C}{K'}\frac{E(x)}{H(x)}\int_x^{x+\delta}\frac{H(x')}{E(x')}(1-x')^{\alpha-\beta-1}\int_0^{x'}\phi_\varepsilon^\zeta(x'')P(x'',x')dx''dx' \\ &+\frac{2C}{K'}\left(\frac{E(x)}{H(x)} - \frac{E(x+\delta)}{H(x+\delta)}\right)\int_{x+\delta}^{1-\zeta}\frac{H(x')}{E(x')}(1-x')^{\alpha-\beta-1}\int_0^{x'}\phi_\varepsilon^\zeta(x'')P(x'',x')dx''dx'\end{aligned}$$

From the triangle inequality, taking the absolute value on both sides yields:

$$\begin{aligned}|\phi_\varepsilon^\zeta(x) - \phi_\varepsilon^\zeta(x+\delta)| &\le \frac{2C}{K'}\frac{E(x)}{H(x)}\int_x^{x+\delta}\frac{H(x')}{E(x')}(1-x')^{\alpha-\beta-1}\int_0^{x'}\phi_\varepsilon^\zeta(x'')P(x'',x')dx''dx' \\ &+\frac{2C}{K'}\left|\frac{E(x)}{H(x)} - \frac{E(x+\delta)}{H(x+\delta)}\right|\int_{x+\delta}^{1-\zeta}\frac{H(x')}{E(x')}(1-x')^{\alpha-\beta-1}\int_0^{x'}\phi_\varepsilon^\zeta(x'')P(x'',x')dx''dx'\end{aligned}$$

The right–hand side's integrals with respect to x'' will be upper–bounded by $\|\phi_\varepsilon^\zeta\|_{L^\infty([0,1[)}$ since $\int_0^{x'}P(x'',x')dx'' = 1$. The first term will be bounded in δ to the extent that, for any $\epsilon > 0$, there exists a $\delta(\epsilon) > 0$ small enough to guarantee that:

$$\frac{2C}{K'}\|\phi_\varepsilon^\zeta\|_{L^\infty([0,1[)}\frac{E(x)}{H(x)}\int_x^{x+\delta}\frac{H(x')}{E(x')}(1-x')^{\alpha-\beta-1}dx' \le \frac{\epsilon}{2}\ \forall x' \in B(x,\delta(\epsilon)) \subset [0,1-\zeta]$$

In order to bound the second term, one gets once again from the (2.27) definition of ϕ_ε^ζ:

$$\int_{x+\delta}^{1-\zeta}\frac{H(x')}{E(x')}(1-x')^{\alpha-\beta-1}\int_0^{x'}\phi_\varepsilon^\zeta(x'')P(x'',x')dx''dx' = \frac{K'}{2C}\phi_\varepsilon^\zeta(x+\delta)\frac{H(x+\delta)}{E(x+\delta)}$$

and, using the facts that $\phi_\varepsilon^\zeta \in L^\infty$ and $|a||b| = |ab|$:

$$\begin{aligned}&\frac{2C}{K'}\left|\frac{E(x)}{H(x)} - \frac{E(x+\delta)}{H(x+\delta)}\right|\int_{x+\delta}^{1-\zeta}\frac{H(x')}{E(x')}(1-x')^{\alpha-\beta-1}\int_0^{x'}\phi_\varepsilon^\zeta(x'')P(x'',x')dx''dx' \\ &\le \|\phi_\varepsilon^\zeta\|_{L^\infty([0,1[)}\left|1 - \frac{E(x)}{H(x)}\frac{H(x+\delta)}{E(x+\delta)}\right|\end{aligned}$$

Once again, for any $\epsilon > 0$, the last factor on the right–hand side of the estimate can be bounded in $\delta(\epsilon)$, meaning that Ascoli's theorem can be applied to the $(\phi_\varepsilon^\zeta)_{\varepsilon\ge0,\zeta\ge0}$. As a consequence, a subsequence of $(\phi_\varepsilon^\zeta)_{\varepsilon\ge0,\zeta\ge0}$ converges to a solution ϕ to (2.15) on the interval $[0, 1-\zeta]$.

• One cannot yet pass to the limit though, since ϕ has to be estimated on $[1-\zeta, 1[$. To this end, following [20, 87, 109] amongst others, a supersolution to (2.15) has to be engineered before uniformly estimating ϕ. In this case, one remarks that : $x \mapsto (1-x)^{1+\beta-\alpha}$ is indeed a supersolution to (2.15):

$$\begin{aligned}\frac{\partial}{\partial x}\left[(1-x)^{1+\beta-\alpha}\right] - \frac{C}{K'}(1-x)^{\alpha-\beta-1}(1-x)^{1+\beta-\alpha} - \frac{\tilde{\lambda}}{K'}(1-x)^{-\beta}(1-x)^{1+\beta-\alpha}\\ +\frac{2C}{K'}(1-x)^{\alpha-\beta-1}\int_0^x (1-x')^{1+\beta-\alpha}P(x',x)dx'\\ \geq (1+\beta-\alpha)(1-x)^{\beta-\alpha} - \frac{\tilde{\lambda}}{K'}(1-x)^{1-\alpha} + \frac{C}{K'}\end{aligned}$$

having noticed that, $1+\beta-\alpha$ being positive:

$$x' \leq x \leftrightarrow (1-x')^{1+\beta-\alpha} \geq (1-x)^{1+\beta-\alpha}$$

The fact that $\beta < 1$ allows the existence of a real $x_{thr} < 1$ that satisfies:

$$(1+\beta-\alpha)(1-x)^{\beta-\alpha} \geq \frac{\tilde{\lambda}}{K'}(1-x)^{1-\alpha} \ \forall x \geq x_{thr}$$

allowing to conclude that $x \mapsto (1-x)^{1+\beta-\alpha}$ is a supersolution to (2.15). Considering that ϕ decreases faster than $(1-x)^{1+\beta-\alpha}$ for x close to 1, the L^∞ bounds to each of the terms of (2.15) follow instantly. Furthermore, none of these bounds depends on ε, meaning that taking the limit $\varepsilon \to 0 \leftrightarrow x_\varepsilon \to 0$ is straightforward, wrapping up the proof of the existence theorem.

Uniqueness of the solution to (2.16)-(2.17)

The uniqueness of the eigenelements is proven using a carbon copy of previously published reasonings by Perthame & Ryzhik [109] and Michel [87]. If (λ, N, ϕ), (λ', N', ϕ') are two solutions to (2.14)-(2.15), $N'\phi$ satisfies the differential equation:

$$\begin{aligned}\frac{\partial}{\partial x}\left[\dot{x}N'(x)\phi(x)\right] + (\lambda'-\lambda)N'(x)\phi(x) =& 2\phi(x)\int_x^1 \gamma(x')P(x,x')N'(x')dx'\\ &- 2\gamma(x)N'(x)\int_0^x \phi(x')P(x',x)dx'\end{aligned}$$

And an integration from 0 to 1 boils down to:

$$\begin{aligned}&(\lambda'-\lambda)\int_0^1 N'(x)\phi(x)dx\\ &= 2\left[\int_0^1 \phi(x)\int_x^1 \gamma(x')P(x,x')N(x')dx'dx - \int_0^1 \gamma(x)N'(x)\int_0^x \phi(x')P(x',x)dx'dx\right]\end{aligned}$$

A use of Fubini's theorem to treat the last term of the right–hand side entails:

$$\begin{aligned}(\lambda'-\lambda)\int_0^1 N'(x)\phi(x)dx =\\ 2\left[\int_0^1 \phi(x)\int_x^1 \gamma(x')P(x,x')N(x')dx'dx - \int_0^1 \phi(x')\int_{x'}^1 \gamma(x)N'(x)P(x',x)dxdx'\right] = 0\end{aligned}$$

meaning that $\lambda' = \lambda$.

To wrap the proof up, the uniqueness of the eigenvectors exploits the notion of general relative entropy (cf. Michel & al.'s [88, 89]), and has already been inked in Doumic & Gabriel's shot at a very similar PBE [20]. The backdrop is a study of the quantity $N'(x) - CN(x)$, with $C > 0$ a real number. Indeed, N is unique iff for any N' that satisfies (2.14): (1) $sgn(N'(x) - CN(x)) = 0 \ \forall \ x \in [0,1[$, and (2) $C = 1$.
Remark first that:

$$\int_0^1 \frac{\partial}{\partial x}\left(\left|\frac{N'(x)}{N(x)} - C\right|\dot{x}N(x)\phi(x)\right)dx = 0 = \int_0^1 \left|\frac{N'(x)}{N(x)} - C\right|\frac{\partial}{\partial x}(\dot{x}N(x)\phi(x))dx + \int_0^1 \frac{\partial}{\partial x}\left(\frac{N'(x)}{N(x)}\right)sgn\left(\frac{N'(x)}{N(x)} - C\right)\dot{x}N(x)\phi(x)dx$$

Then, since the product $\dot{x}N\phi$ satisfies the differential equation:

$$\frac{\partial}{\partial x}(\dot{x}(x)N(x)\phi(x)) = 2\phi(x)\int_x^1 \gamma(x')P(x,x')N(x')dx' - 2N(x)\gamma(x)\int_0^x P(x',x)\phi(x')dx'$$

injecting this into the penultimate equation entails:

$$\begin{aligned} 0 = &\int_0^1 \left|\frac{N'(x)}{N(x)} - C\right|2\phi(x)\int_x^1 \gamma(x')P(x,x')N(x')dx'dx \\ &- \int_0^1 \left|\frac{N'(x)}{N(x)} - C\right|2N(x)\gamma(x)\int_0^x P(x',x)\phi(x')dx'dx \\ &+ \int_0^1 \left(\frac{\partial N'(x)}{\partial x} - \frac{N'(x)}{N(x)}\frac{\partial N(x)}{\partial x}\right)sgn\left(\frac{N'(x)}{N(x)} - C\right)\dot{x}\phi(x)dx \end{aligned}$$

The next step is the use of the relationship $|x| = xsgn(x) \ \forall x$ to get rid of all absolute values:

$$\begin{aligned} 0 = &\int_0^1 \frac{N'(x)}{N(x)}sgn\left(\frac{N'(x)}{N(x)} - C\right)2\phi(x)\int_x^1 \gamma(x')P(x,x')N(x')dx'dx \\ &-C\int_0^1 sgn\left(\frac{N'(x)}{N(x)} - C\right)2\phi(x)\int_x^1 \gamma(x')P(x,x')N(x')dx'dx \\ &- \int_0^1 \frac{N'(x)}{N(x)}sgn\left(\frac{N'(x)}{N(x)} - C\right)2\gamma(x)N(x)\int_0^x P(x',x)\phi(x')dx'dx \\ &+C\int_0^1 sgn\left(\frac{N'(x)}{N(x)} - C\right)2\gamma(x)N(x)\int_0^x P(x',x)\phi(x')dx'dx \\ &+ \int_0^1 2\int_x^1 \gamma(x')P(x,x')\left(N'(x') - \frac{N'(x)}{N(x)}N(x')\right)dx'sgn\left(\frac{N'(x)}{N(x)} - C\right)\phi(x)dx \end{aligned}$$

The right–hand side can be simplified by gathering its first term with its last one:

$$
\begin{aligned}
0 = &-2C \int_0^1 sgn\left(\frac{N'(x)}{N(x)} - C\right)\phi(x) \int_x^1 \gamma(x')P(x,x')N(x')dx'dx \\
&-2\int_0^1 \frac{N'(x)}{N(x)} sgn\left(\frac{N'(x)}{N(x)} - C\right)\gamma(x)N(x) \int_0^x P(x',x)\phi(x')dx'dx \\
&+2C \int_0^1 sgn\left(\frac{N'(x)}{N(x)} - C\right)\gamma(x)N(x) \int_0^x P(x',x)\phi(x')dx'dx \\
&+\int_0^1 2\int_x^1 \gamma(x')P(x,x')N'(x')dx' sgn\left(\frac{N'(x)}{N(x)} - C\right)\phi(x)dx'dx
\end{aligned}
$$

A use of Fubini's theorem in the second and third terms follows, allowing to combine the first and third terms on the one hand, the second and fourth terms on the other hand:

$$
\begin{aligned}
0 = &-2C \int_0^1 \phi(x) \int_x^1 \gamma(x')P(x,x')N(x')\left(sgn\left(\frac{N'(x)}{N(x)} - C\right) - sgn\left(\frac{N'(x')}{N(x')} - C\right)\right)dx'dx \\
&+2\int_0^1 \phi(x) \int_x^1 \gamma(x')P(x,x')N'(x')\left(sgn\left(\frac{N'(x)}{N(x)} - C\right) - sgn\left(\frac{N'(x')}{N(x')} - C\right)\right)dx'dx
\end{aligned}
$$

Finally:

$$
\begin{aligned}
0 &= 2\int_0^1 \phi(x) \int_x^1 \gamma(x')P(x,x')N(x')\left(\frac{N'(x')}{N(x')} - C\right)\left(sgn\left(\frac{N'(x)}{N(x)} - C\right) - sgn\left(\frac{N'(x')}{N(x')} - C\right)\right)dx'dx \\
&= 2\int_0^1 \phi(x) \int_x^1 \gamma(x')P(x,x')N(x')\left|\frac{N'(x')}{N(x')} - C\right|\left(sgn\left(\frac{N'(x)}{N(x)} - C\right)sgn\left(\frac{N'(x')}{N(x')} - C\right) - 1\right)dx'dx
\end{aligned}
$$

Hence $\forall\ x \in [0,1[,\ x' > x,\ sgn\left(\frac{N'(x')}{N(x')} - C\right) = sgn\left(\frac{N'(x)}{N(x)} - C\right)$ and $N'(x) \geq CN(x)$. Both being positive and summing up to 1 by definition, C can only be equal to 1, that is $N' \equiv N$. The same reasoning evidences that $\phi' \equiv \phi$ and ends the proof of the uniqueness of the eigenelements. □

Discussion

This above–mentioned theorem holds only because of the generational redistribution operator's regularising property (cf. equation (2.21)), that was already mentioned when the existence of a solution to the PBE (1.22) was proven. Hence, the physics of the populational dynamics lies only in the right pick for γ and P. The case $l \in \mathbb{R}_+$ was extensively studied by Perthame & Ryzhik [109] and Doumic & Gabriel [20] and the requirements are γ and P polynomial. In the present case, γ being singular at a certain $\bar{l}$ provides a new insight at the dynamics, but $\bar{l}$ is an ad–hoc upper bound which physical meaning could be questioned when confronted with experimental data. Lengtheningwise, the sublinear assumption is not satisfied for instance in Yasuda's [148] experiment evidencing organelle–based enhancing when cells are individually cultivated on a chip. On the other hand, the latter case is observed in a seemingly ideal environment that is not encountered in an industrial reactor.

The $2\alpha - \beta > 0$ inequality is not necessary to ensure the application of Krein–Rutman theorem in the L^∞ space, but allows to conclude on the existence of a solution to (2.16) in the space of Lipschitz–continuous functions. This condition can easily be met, as will be shown in the modelling part of this work. The $\alpha < \beta$ inequality, on the other hand, is nothing but the translation of a fundamental assumption from Diekmann & al.'s [18] existence and uniqueness theorem for size–structured models. In fact, the proven result is a slight improvement from what Diekmann & al. have retrieved using the strong continuity property of the transport–fragmentation operator seen as a semigroup.

Chapter 3

Multiscale modelling of a biological reactor

The main goal of this chapter is the formulation of a biological model that can be run to simulate environmental perturbations of a fermenter which experimental measurements exist for.

The current state–of–the–art observations of the biological responses to fluctuating environments make a case for a breakdown into three main time scales. One of the quickest responses belongs to the substrate uptake process, operating at the cell–liquid border within a seconds–long time interval. A typically slow (i.e. counting in hours) time scale characterises the rate of anabolism's adaption: its prerequisites include gene induction, transcription, new enzyme synthesis and so on. In between these extrema, the cell cycle is in the order of a few dozens of minutes, meaning that the individuals' progress in the cell cycle counts in minutes. The corresponding modelling must consecrate some of or all these characteristic times. A naive approach could consist in focusing on long–term consequences of an ever–changing medium, that is treating the overall uptake as an algebraic function of S and distributing the cell length and lengthening rate only. A more demanding framework could lie in the distribution of all the aforementioned quantities, what requires a convenient formulation for the dynamics of all these coordinates. Also, as mentioned in the first chapter, the level of mixing can impact the substrate availability at the cell scale in real life industrial cultures, prompting the need to capture this hindering feature in an all–terrain model.

In this chapter, several versions of a multivariable structured model will be presented and a five–dimensional population balance model will emerge from the dismissal of more primitive formulations. An attempt at closing the substrates mass balance using the cells' metabolism will also be enforced.

3.1 Mass balances in the reactor

Here and below, the case study will consecrate a 1L glucose– and oxygen–fed Continuous Stirred Tank Reactor (CSTR) or chemostat. It will be assumed perfectly mixed to allow its treatment as one hydrodynamic compartment, without hindering the possibility of a sub–mesh refinement to take into account the interaction between the biological phase and its medium.

At the reactor scale, mass conservation equations dictate the time evolution of glucose, oxygen, and by–products in the liquid and/or gas phase. It is worth mentioning that when biological populations are addressed, these equations have to explicitly report the mass transfer between the medium and the biotic phase.

3.1.1 Preliminary: the treatment of the biotic phase

Whatever the appropriate coordinates, their distribution among a population of cells obey the (1.22) PBE:

$$\frac{\partial}{\partial t}n(t,\xi) + \nabla \cdot \left[\dot{\xi} n(t,\xi)\right] + \gamma(\xi)n(t,\xi) + Dn(t,\xi) = 2\int_{\Omega_\xi} \gamma(\xi')P(\xi,\xi')n(t,\xi')d\xi'$$

with appropriate (Dirichlet in the present case study) conditions at the boundary of Ω_ξ.

The cell division procedure is species–dependent and shall not be influenced by the proposed modelling of the organisms' uptake. From experimental grounds, the growth pattern is assumed sublinear to report the difficulty to maintain a growth rate constant as it is understood that the longer a cell gets, the more its biomechanics will be challenged by the effort to support its structure. Even though this view is not consensus, it is underpinned by recent raw data like Robert & al.'s [119] and seems more plausible than unrestrained exponential growth. It consequently allows a definition of the rupture function γ over $[0, \bar{l}[$ by:

$$\gamma : l \mapsto \frac{1}{T}\frac{\frac{1}{(\bar{l}-l)^v} - \frac{1}{(\bar{l}-l_{\text{inf}})^v}}{\frac{1}{(\bar{l}-l_c)^v} - \frac{1}{(\bar{l}-l_{\text{inf}})^v}} 1_{l_{inf} \leq l < \bar{l}} \tag{3.1}$$

as it was proven in section 2.2.3 that a singularity at $l = \bar{l}$ does not prevent the regularity of the eigenelements if v is chosen in agreement with the parameter dictating the sublinearity of the lengthening process.

Furthermore, considering that K has to be symmetric and the PBE's variables are compactly supported, the redistribution kernel will be picked among $\beta(p,p)$ beta laws with $p > 1$. K will also be assumed a tensor product of one–dimensional redistribution kernels to avoid the formulation of hazardous couplings between the partition of different quantities, any experimental smoking gun being out of reach for the time being.

3.1.2 Liquid phase mass balances

The computation of the said liquid–mass transfer involves a population balance approach taking into account the variability in physiological state among a collection of cells which marker is their

own history in the system.

$$\frac{dS}{dt} = D(S_f - S) - \int n(\xi) q_S(\xi) d\xi \tag{3.2}$$

$$\frac{dO_2}{dt} = D(O_2^e - O_2) + K_L a(O_2^\star - O_2) - \int n(\xi) q_{O_2}(\xi) d\xi \tag{3.3}$$

$$\frac{dAc}{dt} = -DAc + \int n(\xi) q_{Ac}(\xi) d\xi \tag{3.4}$$

$$\frac{dCO_2}{dt} = D(CO_2^e - CO_2) + K_L a(CO_2^\star - CO_2) - \int n(\xi) q_{CO_2}(\xi) d\xi \tag{3.5}$$

In the first equation, DS_f $(g_S/L/h)$ stands for the user–set external supply, DS $(g_S/L/h)$ for the washout term, and q_S (g_S/h) for the cell–scale glucose uptake. In the second one, q_{O_2} (g_{O_2}/h) is the cell–scale oxygen uptake rate (respiratory capability), $K_L a$ (h^{-1}) the gas–liquid mass transfer rate and $O_2^\star$ (g/L) the oxygen concentration at equilibrium emerging from Henry's law. O_2^e (g/L) is the dissolved oxygen concentration in the feed, its carbon dioxide counterpart being called CO_2^e (g/L) in the fourth equation.

In (3.2)-(3.5), q_{O_2} is necessarily a positive quantity whereas q_{CO_2} (g_{CO_2}/h) is of negative sign, in line with the assumption that *E. Coli* feeds its backbone machinery from respiration. On the other hand, $q_{Acetate}$ $(g_{Acetate}/h)$ has no predefined sign due to the versatile role of acetate in the cells' metabolism. The molecule is understood to be excreted when glucose has been overuptaken and reconsumed in case of glucose shortage.

3.1.3 Gas phase mass balances

In (3.3) and (3.5), O_{2g} (g/L) and CO_{2g} (g/L) are the respective offgas oxygen and carbon dioxide concentrations, which conservation in the reactor reads:

$$\frac{d}{dt}O_{2g} = \frac{2}{V_g}\left[Q_g(O_{2g,in} - O_{2g}) - K_L a\left(O_{2,g}^\star - O_2\right)V\right] \tag{3.6}$$

$$\frac{d}{dt}CO_{2g} = \frac{2}{V_g}\left[Q_g(CO_{2g,in} - CO_{2g}) - K_L a\left(CO_{2,g}^\star - CO_2\right)V\right] \tag{3.7}$$

with V_g (L) the volume fraction of the gas phase in the fermenter, Q_g (L/h) the gas flow rate and $O_{2g,in}$ (g/L) and $CO_{2g,in}$ (g/L) its oxygen and carbon dioxide content. All these parameters are supposed constant throughout the simulations. The equilibrium osygen cocnentration in the liquid phase is obtained through in $O_2^\star = H_{e,O_2}\langle O_{2,g}\rangle$ and $CO_2^\star = H_{e,CO_2}\langle CO_{2,g}\rangle$. H_e is the Henry constant defining the thermodynamic equilibrium at the gas–liquid interface [1]. With satisfying accuracy (cf. Morchain's [92]), one can assume $\langle O_{2,g}\rangle = \frac{1}{2}(O_{2g,in} + O_{2g})$ and $\langle CO_{2,g}\rangle = \frac{1}{2}(CO_{2g,in} + CO_{2g})$. Indeed, the gas feed consists in bubbles which transfer soluble gases to the liquid phase at a surface–dependent rate as they are steered out. If the reactor is small enough to limit the bubbles' residence time to a few seconds, the exponential decay of the gas phase–oxygen concentration can be supposed affine justifying the above mentioned calculation of an average concentration in the gas phase.

1. no such thermodynamic constant exists at the liquid–cell interface leading to modelling issues addressed in this work

3.2 Three different strategies for the calculation of q_S

A flurry of formulations for q_S can be cooked up depending on the desired attention to the details. Due to the cost of integrating a PBE in one hydrodynamic compartment, a balance must be stricken between level of biological complexity and computational power at the disposal of the researchers.

3.2.1 A two-dimensional Population Balance Model (reference [116])

Presentation of the model

In this article, that is attached for the sake of clarity, the distributed variables are the cells' size l (m) and their growth rate v (h^{-1}), with dynamics:

$$\begin{aligned} \dot{l} &= a \min(v, v^\star) \text{ with } v^\star = \mu_{\max} \frac{S}{K_S + S} \frac{O_2}{K_{O_2} + O_2} \\ \dot{v} &= \left(\frac{1}{T} + v\right)(v^\star - v) \text{ if } v \leq v^\star \\ &\quad \left(\frac{1}{T}\right)(v^\star - v) \text{ if } v \geq v^\star \end{aligned} \tag{3.8}$$

In (3.8), $v^\star$ (h^{-1}) can be thought of as a maximal population–averaged lengthening potential given the (S, O_2) environmental offer, $\mu_{\max}$ (h^{-1}) coming from batch (i.e. exponentially growing) culture experimental measurements. Half–saturation constants K_S (g_S/L) and K_{O_2} (g_{O_2}/L) also come from experimental observations, in steady–state CSTR this time. a (h^{-1}) and T (h^{-1}) consecrate the difference in temporal response to perturbations of a steady–state population.

If neither q_P nor q_p are distributed, algebraic formulae for both contributions to the total glucose uptake rate have to be contrived from the variables at disposal:

$$q_S = q_P + q_p = q_P(l, v, S) + q_p(l, v, S)$$

with no loss of generality. q_P is assumed a combination of the individuals' rate of anabolism and the whole population's growth rate. q_p is linked to a fictitious membrane permeability α which modelling involves population–averaged values of rate of anabolism ($\tilde{v}$) and uptake $\tilde{q_S}$. Also, the substrate–to–mass ratio Y_{SX} (g_S/g_X) is assumed constant, meaning that all the uptaken glucose translates into biomass with equal efficiency, whatever the cells' history in the reactor or the environment offer. In other words, the metabolism is assumed fixed and identical for all the organisms. q_{O_2} is a decreasing algebraic function of the biomass which maximum is set to q_S.

Finally, the micromixing's influence on the substrate availability to the individuals is taken into account via the introduction of a S_c parameter obeying the algebraic equation:

$$S_c = Y_{SX} t_M \tilde{v} X \tag{3.9}$$

where t_M (h) stands for a micromixing time constant and X the biomass. Then, q_S is corrected using a $1 - \exp(-S/S_c)$ factor.

Numerical implementation

The response of the said uptake mechanisms to user–set medium conditions is depicted in figure 3.1. The respective PTS and permease contributions to the overall uptake rate as functions of S are plotted, the trend being in line with 1990s experimental observations by Ferenci and co–workers [24, 26] or Kovárová–Kovar & Egli [57]: the higher the residual concentration, the bigger the contribution of the PTS system to the individuals' overall uptake rate, the permease activity adapting via the inhibiting feature of its PTS counterpart. In figure 3.2 representing the bacterium–scale uptake features as functions of the enforced dilution rate, the permeability profile testifies to a permease induction (respectively inactivation) when the substrate proves scarce (respectively plentiful) at the cells' vicinity.

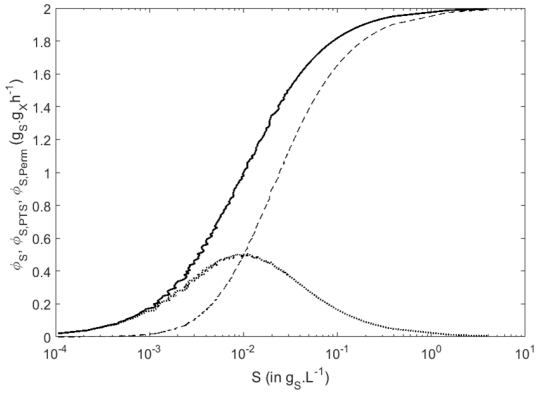

Figure 3.1 – Total uptake q_S (solid line) and respective PTS (dashed line) and permeases (dotted line) contributions as functions of the residual S (log scale). Borrowed from [116].

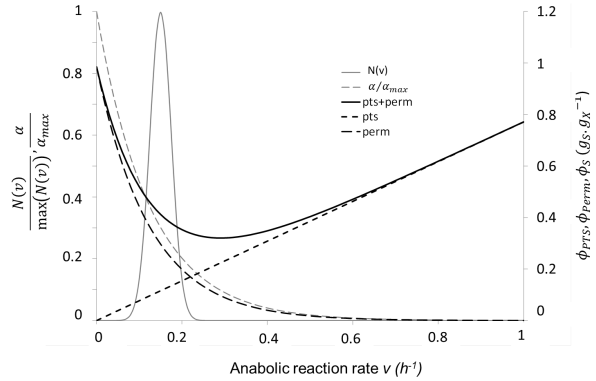

Figure 3.2 – q_S (full black line) and its PTS and permease contributions (black dashed lines) coming from a series of chemostat simulations at various dilution rates. The corresponding (normalised) NDF in rate of anabolism for the case $D = 0.15$ h^{-1} is superimposed to evaluate the consequential importance of q_P and q_p at steady state. The function $\alpha/\alpha_{\max}$ (light grey dashed line) controls the permease induction for slow growing cells. Borrowed from [116].

Limitations

Two observations have to be discussed to evaluate (3.8)'s performance:

- The assumptions formulated here have yet to allow the riddance to averaged quantities (like Y_{SX})
- q_S still obeys an algebraic Monod–shape law, despite the effort to make a step into the microscopical scale, at least regarding the cells' bio–mechanics.

From the former remark, one consequence is that no acetate consumption/production can be predicted due to the constant Y_{SX} (meaning that Sunya & al.'s [133] results are out of the reach of such a model). The latter remark leads to two unpleasantnesses when dynamical simulations are performed: the instantaneous adaption to the environment precludes any chance of over–uptake in the wake of a perturbation like a glucose pulse, unless the permeability is tampered with. Indeed, any permease regime/PTS regime transition would be instantaneous but the cells' uptake in profuse environmental conditions comes down to $q_S \approx q_P$ that can never overreach its maximum batch value by construction. Such an example is provided in figure 3.3 where the overall and marginal uptake profiles in the wake of a 1g glucose pulse in a steady–state reactor are computed using (3.8). Another inconvenience pertains to the influence of a low residual substrate concentration on the cell–scale features. It can happen that S is scarce because the feed perfectly suits the organisms' needs, in other words no waste is washed out of the reactor albeit the individuals are not short of substrate: in this case, Monod laws modelling q_S translate into a potentially misevaluated slender uptake capacity.

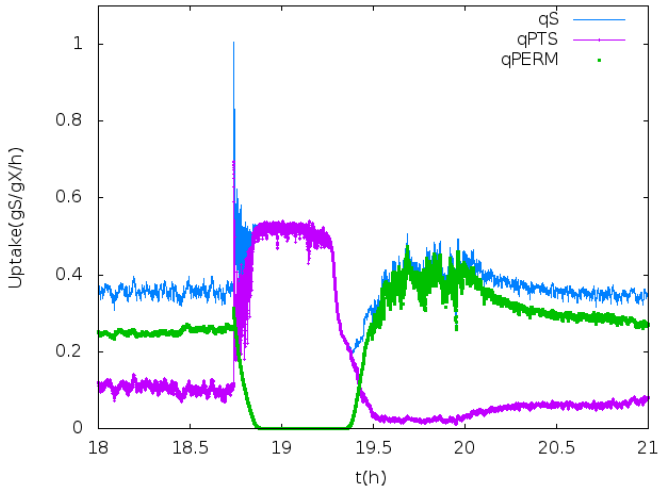

Figure 3.3 – Normalised substrate uptake shortly before and after a 1g glucose pulse in a steady–state CSTR, obtained from a Monte–Carlo simulation of the (3.8) model. Given that q_S instantaneously adapts to the environmental conditions, the peremase regime/PTS regime transition does not translate into an over–uptake in the wake of the pulse, contrary to Neubauer's 1990s [99] conclusions.

Thus, as light and computational time-wise affordable it is, this unsatisfying model cannot be implemented in a computational fluid dynamics code to couple the biological behaviour to micro and macro mixing.

3.2.2 A three-dimensional Population Balance Model

Presentation of the model

An improvement from the previous paragraph's model could consist in distributing both sources of glucose uptake over the whole population, giving birth to a 3-dimensional model involving l, q_P, q_p. The inner coordinates are computed via the differential equations:

$$\begin{aligned} \dot{l} &= \frac{q_S}{K_{conv} Y_{SX}} \left(1 - \frac{l}{\bar{l}}\right)^{\kappa} \\ \dot{q_P} &= \frac{1}{\tau_P} \left(q_{P_{\max}} \frac{S}{K_P + S} - q_P\right) \\ \dot{q_p} &= \frac{1}{\tau_p} \left[\psi_{\max} \frac{S}{K_p + S} \left(\frac{q_P}{q_{P_{\max}}} - 1\right)^2 - q_p\right] \end{aligned} \tag{3.10}$$

$Y_{SX}(g_S/g_X)$ standing once again for the substrate–to–biomass ratio. $q_{P_{\max}}$ (g_S/h) is the maximum uptake capability the PTS system is allowed, $\psi_{\max}$ (g_S/h) being its permease counterpart. K_{conv} (g_X/m_X) is a length–to–mass conversion ratio that allows $q_S/(K_{conv}Y_{SX})$ to be homogeneous to $\dot{l}$. κ is chosen in accordance with the value for (3.1)'s υ dictating the rate of rupture. In this framework, the dynamics of the cells' uptake is captured by the τ_P, τ_p constants but the metabolism is still fixed: Y_{SX} is a constant. (3.8)'s v variable is understood as $\dot{l}/l$ in this case, meaning that (3.10)'s longest characteristic time scale is missing from the population's dynamics. As a consequence, (3.10) is geared towards short–time transient responses to a change in the culture conditions and one of its weaknesses is its inability to quantitatively predict long–term variations in biomass or cell number.

In (3.10), K_P (g_S/h) is the PTS affinity constant to glucose, its permease counterpart being called K_p (g_S/h). q_P and q_p have a different affinity to glucose, the PTS allowing the organisms to cherry–pick their favorite source of organic carbon when it is in excess, whereas permeases allow both glucose and less desirable chemical species into the cytoplasm. One has thus to expect $K_P \ll K_p$. Also, the shutting term $(q_P/q_{P_{\max}} - 1)^2$ testifies that the PTS system inhibits its permease counterpart when sugar proves in excess. Without loss of generality, K_S can be set equal to the affinity constant of the standard uptake system K_P.

The micromixing's influence on the substrate availability to the biotic phase is once again corrected using a $(1 - \exp(-S/S_c))$ factor, the computation of S_c coming from the (3.9) definition of a limiting substrate concentration in physical regime.

Numerical implementation

The 3D model's adequacy with the established microbiological claims regarding *E. Coli* can be assessed similarly to the previous section's approach. Figure 3.4 compiles the numerical steady–state values for q_S from a collection of chemostat simulations using a mean individual version of (3.10). Once again, the numerics are sound with regards to Ferenci & al. [24] and Kovárová–Kovar & al. [57]'s conclusions: q_P is an increasing function of S and its inhibiting influence on q_p when S is large enough to allow the PTS system to thrive is conspicuous. It is worth a mention that even though $v \approx \dot{l}/l$ is a function of q_S, both quantities are significantly uncorrelated due to v's

dependence on l: at fixed q_S, the shorter cells will exhibit a much larger apparent growth rate than their longer counterparts. Hence, no equivalent of 3.2 is at the reach of (3.10).

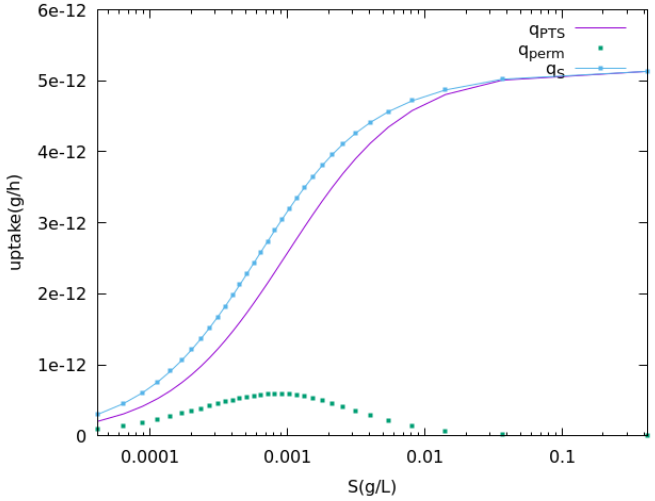

Figure 3.4 – Total uptake and its PTS and permease contributions for a collection of steady–state continuous cultures of varying dilution rate ($D \in [0.02\ \mathrm{h}^{-1}, 0.34\ \mathrm{h}^{-1}]$), as functions of S (log scale).

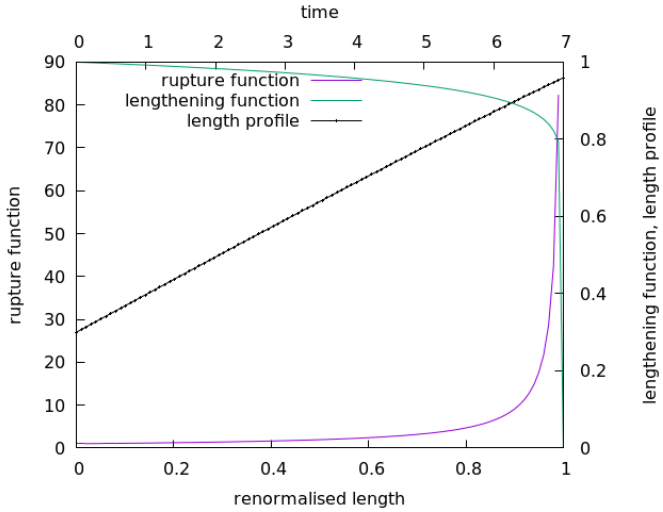

Figure 3.5 – Rupture function γ (left axis, purple line) and lengthening function $\dot{l}$ (right axis, green line) with respect to $l/\bar{l}$ (bottom axis). A typical length profile (black linespoints) is pictured on the right axis with respect to time (top axis). A cell is assumed born with length $l_0 = 6\mu$m and grows at constant rate ($1.8 \cdot 10^{-6}$ m/h) for 7h. The sublinear pattern starts materialising towards the end of the time window).

Limitations

When computational time is addressed, such an approach is still affordable whatever the numerical tool to solve the PBE. Furthermore, many of the transient behaviours will be reproduced by the implementation of this set. For instance, the latency in the PTS response to a glucose pulse would translate into a crossing of the $q_S = q_{P_{\max}}$ threshold over a time interval which measure would be dictated by τ_P and τ_p. From experimental grounds though, three conditions should be met in order to mesh the (3.10) set of ODEs with the biologists' claims:

- $\psi_{\max} \approx 5 q_{P_{\max}}$
- $K_p \approx 10 K_P$
- $q_P \approx q_p$ around $S \approx 0.001$ g/L.

At steady state, the analytical formulae are obvious:

$$q_P = q_{P_{\max}} \frac{S}{K_P + S}$$
$$q_p = \psi_{\max} \frac{S}{K_p + S} \frac{{K_P}^2}{(K_P + S)^2}$$

See if the model respects all these quantitative constraints by solving the equation $q_P = q_p$ in S:

$$q_{P_{\max}} \frac{S}{K_P + S} = 5 q_{P_{\max}} \frac{S}{10K_P + S} \frac{{K_P}^2}{(K_P + S)^2} \Leftrightarrow (K_P + S)(10K_P + S) = 5{K_P}^2$$

The roots of the polynomial $S^2 + 11K_P S + 5{K_P}^2$ are easily calculable ($S = \frac{K_P}{2}(-11 \pm \sqrt{101})$), both of these being strictly negative. For this reason, (3.10) shall not be put into practice due to its unreasonable simplicity.

Furthermore, the absence of v from the model and the ensuing loss of any long–term time scale lead to doubtful outputs when (3.10) is run to simulate tens of minutes to hours–long perturbations in a steady–state chemostat. In figure 3.6, a 1g glucose pulse in a steady–state chemostat is run using (3.10), the left–hand side focusing on the cell–scale state variables whereas the right–hand side concentrates on the reactor–scale biomass and substrate concentrations. A few inconsistencies between the microbiology theory and the numerics are highlighted: given the (3.10) framework, a hike in q_S translates into an instantaneous adaption of the cells' growth rate and the material balance makes no room for by–products (acetate, formate, ...) synthesis. The model is therefore at odds with the experimentalists' claim that acetate is synthetised in the wake of a glucose pulse, prompting the need to dissociate v from $\dot{l}$. A resulting refinement of the 3D model is then formulated in the next section.

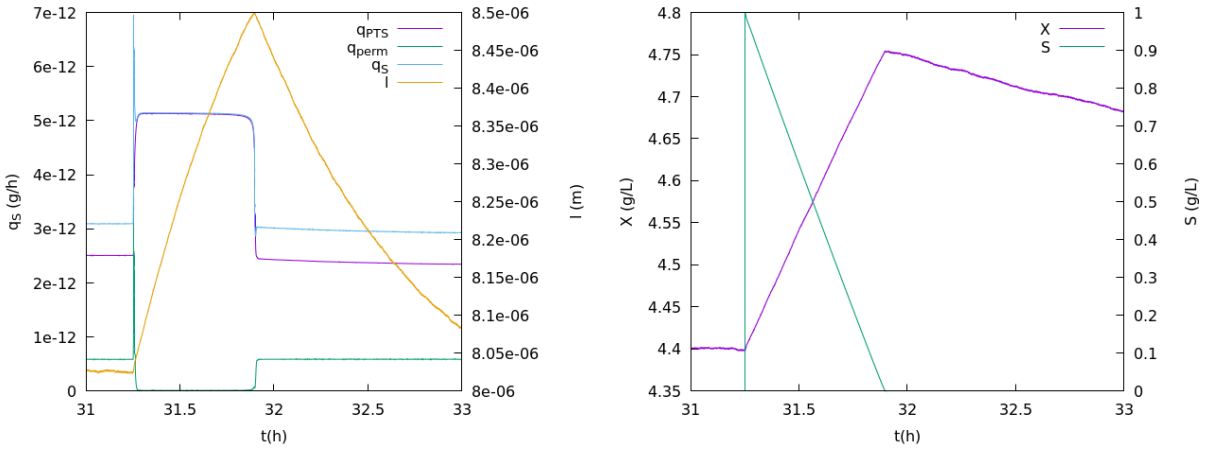

Figure 3.6 – Simulation of a 1g glucose pulse in a steady–state fermenter (D = 0.2 h^{-1}) running (3.10). Left: mean cell–scale uptake features and length shortly before and after the disruption. If $v = \dot{l}$, the cells' growth in mass is instantaneous and translates into a massive gain in biomass, at the expense of acetate production. Right: biomass and substrate concentrations throughout the perturbation. The glucose surplus translates into an overestimated proportional gain in mass.

3.2.3 A five-dimensional Population Balance Model

In both (3.8) and (3.10), two structural flaws were making the models unusuable: the immutability of Y_{SX} and the lack of degrees of freedom regarding q_S's dynamics. In response, a refinement of (3.10) will touch upon q_p and a reduced metabolic model to get rid of the untoward Y_{SX} parameter. Permeasewise, both their number N_p and activity A_p shall have their own dynamics to account for the fact that adjusting a degree of openness is a much easier task than fabricating / dismantling transmembrane sites.

Moreover, each individual is characterised by their own maximum lengthening capacity v^p (m/h), which serves as a markovian marker of their history in the reactor, and the variable's trend marks the environment offer over longer periods of time since it is understood that adapting a cell's growth rate to the resource at their neighbourhood requires deep structural transmutations which time scale is comparable to the organism's lifetime at least.

As a result, the following model is proposed to describe the influence of the uptake capability on the rate of anabolism:

$$\begin{aligned}
\dot{l} &= v^e \left(1 - \frac{l}{\bar{l}}\right)^{\kappa} \\
\dot{A_P} &= \frac{1}{\tau_P}\left(\frac{S}{K_S + S} - A_P\right) \\
\dot{N_p} &= \left(\frac{1}{\tau_i}\frac{K_S}{K_S + S} + \frac{1}{\tau_d}\frac{S}{K_S + S}\right)\left[N_{\max}\frac{K_S}{K_S + S} - N_p\right] \\
\dot{A_p} &= \left(\frac{1}{\tau_{A1}}\frac{q_P}{q_{P_{\max}}} + \frac{1}{\tau_{A2}}\left(1 - \frac{q_P}{q_{P_{\max}}}\right)\right)\left(1 - \frac{q_P}{q_{P_{\max}}} - A_p\right) \\
\dot{v^p} &= \frac{1}{\tau_{vp}}\left(\frac{q_S}{Y_{SX}} - v^p\right)
\end{aligned} \tag{3.11}$$

In (3.11), v^e (m/h) is the lengthening rate that emerges from a metabolic model, A_P stands for the PTS efficiency and τ_i, τ_d, τ_{A1}, τ_{A2} (h) the characteristic times of the upward / downward responses to external fluctuations. Since A_p defines a degree of opening, it is assumed ranging from 0 to 1. $N_{\max}$ is an arbitrary constant and is essentially linked to $\psi_{\max}$ (g_S/h), to the extent that the product $N_p\psi_{\max}$ must be understood as a cell–scale permease density.

Then, the total uptake rate q_S is simply the sum $q_P + q_{perm}$ and its different contributions are computed using:

$$\begin{aligned}
q_P &= A_P q_{P_{\max}} \frac{S}{K_P + S} \\
q_p &= N_p A_p \psi_{\max} \frac{S}{K_p + S}
\end{aligned} \tag{3.12}$$

In (3.11), all the inner variables but l obey a first–order dynamics of restoring force type. One can easily notice q_P's negative feedback on q_p via A_p and the complementary influence of S on the steady–state value for A_P and N_p, A_p.

With $\psi_{\max}$ still equal to 5 $q_{P_{\max}}$ and $K_p = 10\ K_P$, see which steady–state substrate concentration the equality $q_p = q_P$ is satisfied for:

$$\begin{aligned}
q_{P_{\max}} \frac{S}{K_P + S}\frac{S}{K_P + S} &= 5 q_{P_{\max}} \frac{K_S}{K_P + S}\left(1 - \frac{S}{K_P + S}\frac{S}{K_P + S}\right)\frac{S}{K_p + S} \\
\Leftrightarrow S &= \frac{5K_S}{K_p + S} K_S \frac{K_S + 2S}{K_S + S} \\
\Leftrightarrow S^3 + (K_P + K_p)S^2 - 5{K_P}^3 &= 0
\end{aligned}$$

With $K_P = 0.001$ g/L and $K_p = 0.01$ g/L, this third–order polynomial has only one positive root $S \approx 0.0006550$ g/L, that is a much more satisfying result than its (3.10) counterpart. Thus, this uptake model is the simplest formulation that can reproduce the experimental behaviours which were compiled in the introduction.

Remark: $\kappa = 0.05$ is consistent with the parameter $\upsilon = 0.04$ introduced in section 3.1.1. This allows the steady–state solution to the model a convenient Lipschitz regularity.

3.3 Calculation of the cell uptake features

3.3.1 Treatment of the glucose uptake rate: a sub–mesh hydrodynamic refinement

It is unreasonable to claim the substrate concentration in the neighbourhood of a thriving organism is exactly equal to the reactor–averaged value, even though the said reactor is assumed perfectly mixed. Indeed, the glucose uptake by ever–feeding cells has to translate into a slight concentration dip at the biotic phase/abiotic phase boundary and force a consequential gradient of concentration to stem the inhomogeneity. Hence, albeit the fermenter is treated as one hydrodynamic compartment, the glucose concentration at a particle's reach is refined using a sub–grid description at the cell's vicinity. This translates into solving the equation:

$$\frac{1}{t_m}(S - S_{@p}) - \int \left(q_S(S_{@p})n\right) d\xi = 0 \tag{3.13}$$

$S_{@p}(g_S/L)$ standing for the actual carbon concentration at the cell's neighbourhood and t_m (h) for a micromixing characteristic time, in line with Morchain's [92].

The idea behind the calculation of $S_{@p}$ consists in determining $q_S^e = q_S(S_{@p})$ (g_S/h) that is the actual glucose uptake taking into account a potential limitation by the external liquid–cell mass transfer. The procedure to solve (3.13) is detailed in section 3.3.3.

3.3.2 Treatment of the oxygen and acetate uptake rate

For the sake of completeness, q_{O_2} and q_{Ac} are compared with the environmental offer before the algorithms are run. Oxygenwise, the material transfer to the cells operates in two steps: the gas dissolves from bubbles into the liquid phase before its use by the bacteria as an electron transmitter in the glucose dissimilation. Integrating the gas–liquid transfer leads to an estimation of the oxygen quantity the bacteria can pick up from the liquid phase, refining the actual value of q_{O_2} via:

$$q_{O_2}^p = \min\left(q_{O_2}, \frac{O_2 + \delta t K_L a(O_2^\star - O_2) + D(O_2^e - O_2)}{\delta t N_{cell}}\right) \tag{3.14}$$

with $N_{cell} = \int_{\Omega_\xi} n d\xi$. $q_{O_2}^p$ (g_{O_2}/h) being a flux, it is interesting to remark that it does not necessarily tend to 0 as $O_2 \to 0$. Indeed, a small residual oxygen concentration in the liquid phase could indicate that the consumption is comparable with the influx with neither waste nor limitation to be reported.

Acetate is a light two–carbon molecule that is assumed diffusing without a hunch through the organisms' membrane. Indeed, as mentioned by Wolfe in [145], "Because acetate freely permeates the membrane in its undissociated form (...), assimilation does not require a dedicated transport system. However, under certain circumstances acetate uptake is saturable, suggesting that such a system exists". Hence, the acetate uptake rate has to be upper–bounded by a strain–dependent

constant $q_{Ac_{\max}}$ (g_{Ac}/h) within the limits of the medium offer, leading to the following modelling proposition:

$$q_{Ac}^{p} = \min\left(q_{Ac_{\max}}, \frac{Ac(1 - D\delta t)}{\delta t N_{cell}}\right) \tag{3.15}$$

.

3.3.3 Calculation of $S_{@p}$

(3.13) amounts to solve in x the equation:

$$\frac{1}{t_m}(S - x) - \int (q_S(x)n)\, d\xi = 0$$

The integral term is basically the cell number in the reactor multiplied by a population–averaged glucose uptake rate $\tilde{q_S}$. The latter is given by the following equation:

$$\begin{aligned}\tilde{q_S}(x) &= \frac{\int A_P(x)n(t, A_P)dA_P}{\int n(t, A_P)dA_P} q_{P_{\max}} \frac{x}{K_P + x} + \frac{\iint N_p(x)A_p(x)n(t, N_p, A_p)dN_p dA_p}{\iint n(t, N_p, A_p)dN_p dA_p} \psi_{\max} \frac{x}{K_p + x} \\ &:= \tilde{A_P} q_{P_{\max}} \frac{x}{K_P + x} + (\tilde{N_p A_p})\psi_{\max} \frac{x}{K_p + x}\end{aligned}$$

S_p is consequently one root of the third–order polynomial:

$$\begin{aligned}0 &= \frac{(S - x)(K_P + x)(K_p + S_p)}{t_m} - \tilde{A_P} q_{P_{\max}} x(K_p + x)N_{cell} - \psi_{\max}(\tilde{N_p A_p})x(K_P + x)N_{cell} \\ &= x^3 + x^2\left(t_m q_{P_{\max}} \tilde{A_P} N_{cell} + t_m \psi_{\max}(\tilde{N_p A_p}) \int n d\xi - S + K_P + K_p\right) \\ &+ x\left(t_m q_{P_{\max}} K_p \tilde{A_P} N_{cell} + t_m \psi_{\max}(\tilde{N_p A_p}) K_P N_{cell} - SK_P - SK_p + K_P K_p\right) - SK_P K_p \\ &\Leftrightarrow 0 = x^3 + \alpha x^2 + \beta x + \gamma\end{aligned} \tag{3.16}$$

with the obvious definitions:

- $\alpha = t_m q_{P_{\max}} \tilde{A_P} N_{cell} + t_m \psi_{\max}(\tilde{N_p A_p}) N_{cell} - S + K_P + K_p$
- $\beta = t_m q_{P_{\max}} K_p \tilde{A_P} N_{cell} + t_m \psi_{\max}(\tilde{N_p A_p}) N_{cell} K_P - SK_P - SK_p + K_P K_p$
- $\gamma = -SK_P K_p$

(3.3.3) will be solved using Cardan's method. To this end, the third–order polynomial in S_p is simplified using the change of variables $\Theta = S_p + \frac{\alpha}{3}$:

$$0 = S_p^3 + \alpha S_p^2 + \beta S_p + \gamma = \Theta^3 + \left(\beta - \frac{\alpha^2}{3}\right)\Theta + \left(\frac{2\alpha^3}{27} - \frac{\alpha\beta}{3} + \gamma\right) = \Theta^3 + a\Theta + b$$

having, still obviously, defined the quantities:

- $a = \beta - \alpha^2/3$
- $b = 2\alpha^3/27 - \alpha\beta/3 + \gamma$

Therefore, the solution will be picked in accordance with the value of $\Delta = b^2 + 4a^3/27$:

- If $\Delta > 0$: only one solution of the polynomial in Θ is real $\Rightarrow \Theta = 2^{-\frac{1}{3}}\left[\left(-b + \sqrt{\Delta}\right)^{\frac{1}{3}} + \left(-b - \sqrt{\Delta}\right)^{\frac{1}{3}}\right]$
- If $\Delta = 0$: the polynomial in Θ has two distinct roots, one of them only being positive $\Rightarrow$ $\Theta = \max((-b/2)^{\frac{1}{3}}, -1/2((-b/2)^{\frac{1}{3}})$

- If $\Delta < 0$: the polynomial in Θ has two distinct roots, and the calculation of S_p follows the short algorithm:

$$k = 0$$

$$\text{Step 1: } S_p = 2\sqrt{\frac{-a}{3}}\cos\left(\frac{2k\pi + \arccos\left(\frac{3b}{2a}\sqrt{\frac{-3}{a}}\right)}{3}\right) - \frac{\alpha}{3}$$

Step 2: If $S_p < 0$ or $S_p > S : k = k + 1$. Then restart Step 1

This refinement is implemented to treat the micromixing issues that disrupt the computation of $S(t)$. The difference between S and S_p is shown on figure 3.7

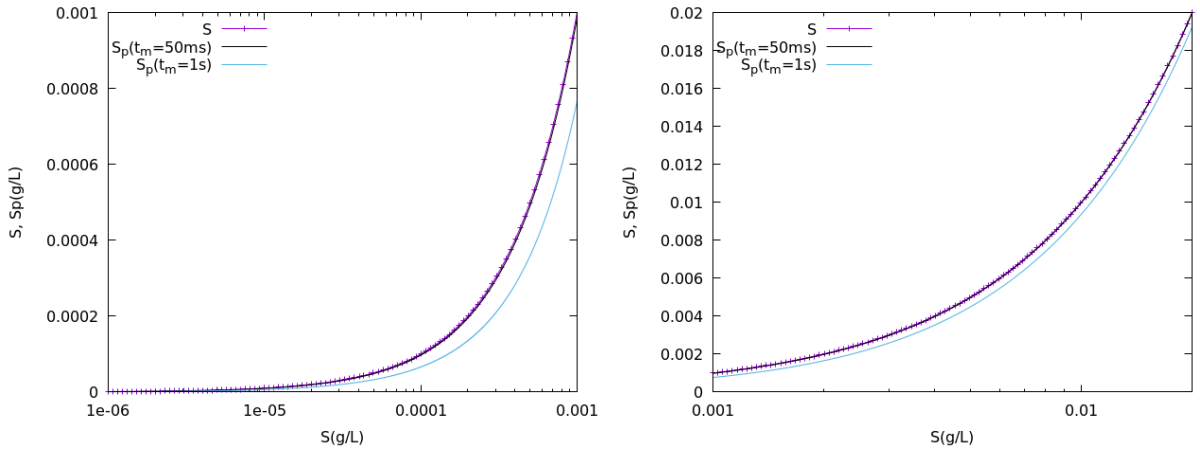

Figure 3.7 – $S_{@p}$ compared with S for small (between 10^{-6} g/L and 0.001 g/L) and larger (between 0.001 and 0.02 g/L) values of S. When S is scarce, so is S_p and the conclusion $S_p \underset{S\to 0}{\to} 0$ is obvious. On the other hand, when S is plentiful, the cell–scale consumption is ridiculous compared with the medium's supply, meaning that $S_{p@} \approx S$. The largest difference is recorded around 0.001 g/L, that is the standard residual substrate concentration in a chemostat. The level of mixing, which manifests itself in the modelling through the value of t_m, can significantly cut the substrate concentration at a cell's vicinity down to 64 % of S (when $t_m = 1s$, indicating a "bad" micromixing at the fermenter scale).

The need for such a refinement was conspicuous to update the permease–related quantities according to their extremely quick response time. It has to be enforced to integrate q_P's dynamics but this is not as crucial due the slower behaviour of this very mechanism, at least as long as S is not limiting. In an open industrial–scale fermenter, it is likely that the PTS system will not have as much time as its permease counterpart to adjust to the ever–changing substrate concentration at the cell's vicinity, prompting the assumption that q_P can equally be calculated as a function of $S_{@p}$ or S.

3.4 A metabolic model

To close the (3.11)–(3.12) set of equations, a metabolic model is formulated to summarise the main pathways which glucose is drawn to as a function of the cells physiological state:

$$S + \nu_{11} O_2 \overset{r_1}{\rightarrow} \nu_{12} X + \nu_{14} CO_2 \tag{1}$$

$$S \overset{r_2}{\rightarrow} \nu_{23} Ac + \nu_{22} X \tag{2}$$

$$S + \nu_{31} O_2 \overset{r_3}{\rightarrow} \nu_{33} Ac + \nu_{34} CO_2 \tag{3}$$

$$Ac + \nu_{41} O_2 \overset{r_4}{\rightarrow} \nu_{42} X + \nu_4 CO_2 \tag{4}$$

with ν_{ij} the yield coefficient for the j reagent/product involed in reaction i, which experimental data are known for (cf. [146]). The ν_{ij}s report the "stoichiometry in mass" at the pathway scale.

In essence, the fundamental way to consume glucose is its combustion into energy and growth-inducing proteins, that is the mode (1). Mode (2) depicts the glucose fermentation into biomass should oxygen prove limiting. Mode (3) is the overflow mechanism that degrades the sugar into acetate if the central metabolism's needs are satisfied. Finally, mode (4) occurs in case S is in default: acetate is consumed to offset the glucose scarcity, even though it proves much less energetic than glucose.

Due to the nature of the ν_{ij} coefficients, the (1)-(4) set of equations shall not be understood as chemical reactions. Indeed, the modes depict a substantial reduction of actual elementary reactions involving cofactors, ATP/ADP, coenzymes..., the main quantity of interest thereof being their resulting biomass production in the context of this work.

In this metabolic model, the known is made of the cells' inner properties v^p, q_S^e, $q_{O_2}^p$ and $q_{Ac_{\max}}^p$, its outputs being the rate of the reactions given the organisms' potential. This formulation consists in an original closure of the metabolic model. In fact, v^p plays the role of an objective function which environment–constrained maximisation is the main goal of the cells' biochemical strategy. This is not a cybernetic variable to the extent that it is an initial datum; instead, the optimisation procedure bears upon the r_i reaction rates, $i \in \{1, ..., 4\}$.

Finally, the CO_2 excretion has been assumed proportional to the oxygen consumption, 1 per 1 in terms of amount of substance (i.e. 44/32 in terms of mass), meaning that the ν_{i4} featuring on reactions (1), (3) and (4) are linearly correlated to the ν_{i1} standing for the required oxygen consumption. Both q_{Ac} and q_{O_2} are arbitrarily upper–bounded, $q_{Ac_{\max}}$ and $q_{O_{2\max}}$ coming from fits to experimental data like Sunya & al.'s [133].

3.4.1 Two numerical computations of the outputs of (1)-(4)

At this stage, the inputs of the metabolic model comprise the cell–scale uptake features q_S^e, $q_{O_2}^p$ and q_{Ac}^p along with the maximal lengthening capability v^p resulting from the bacteria's history in the fermenter. This dataset being given, the metabolites traffic is directed according to the cells' readiness to process the uptaken compounds. Figure 3.8 illustrates this procedure from a reagent–based standpoint, but the approach is extremely similar whatever the algorithm enforced.

A reaction–based scheme

The actual consumption rates over the course of an interval of Lebesgue measure δt can be calculated using the following algorithm:

Each particle is determined by its own uptake and growth capabilities: $\begin{pmatrix} q_S^e \\ q_{O_2}^p \\ q_{Ac}^p \\ v^p \end{pmatrix}$. From this inputs, the scheme's purpose is the calculation of the actual uptake and growth rates $q_{O_2}^e$, q_{Ac}^e and v^e as functions of the cells' potential capacities. Because q_S has already refined in section 3.3.1 through the calculation of an effective glucose concentration at the organisms' disposal, the computed q_S^e stands for the individuals' real glucose uptake over the considered time interval. It is worth a mention that q_{Ac}^e has no predefined sign because acetate can be either excreted (via mode (3), if S and O_2 are abundant enough to allow some overflow or mode (2) should the oxygen supply be the limiting contribution to the actual growth rate) or consumed in case reaction (4) is set forth (when S does not suffice to satisfy a cell's v^p through the sole reaction (1)).

$\diamond$ The glucose oxidation rate r_1, referring to the individuals' preferred source of energy, is determined by the limiting mechanism (glucose uptake, oxygen uptake, or anabolism rate):

$$r_1 = \min\left(q_S^e, \frac{q_{O_2^p}}{\nu_{11}}, \frac{v^p K_{conv}}{\nu_{12}}\right)$$

with K_{conv} (g_X/m_X) the mass–to–length ratio. Then: $\begin{pmatrix} q_S^e \\ q_{O_2}^p \\ q_{Ac}^p \\ v^p \end{pmatrix} \to \begin{pmatrix} q_S^e - r_1 \\ q_{O_2}^p - r_1\nu_{11} \\ q_{Ac}^p \\ v^p - \frac{r_1}{K_{conv}}\nu_{12} \end{pmatrix}$

$\diamond$ A real number r_2, standing for the fermentation rate, is computed using the equation:

$$r_2 = \min\left(q_S^e - r_1, \frac{1}{\nu_{22}}\left(v^p - \frac{r_1}{K_{conv}}\nu_{12}\right)K_{conv}\right)$$

Then: $\begin{pmatrix} q_S^e - r_1 \\ q_{O_2}^p - r_1\nu_{11} \\ q_{Ac}^p \\ v^p - \frac{r_1}{K_{conv}}\nu_{12} \end{pmatrix} \to \begin{pmatrix} q_S^e - r_1 - r_2 \\ q_{O_2}^p - r_1\nu_{11} \\ q_{Ac}^p \\ v^p - \frac{r_1}{K_{conv}}\nu_{12} - \frac{r_2}{K_{conv}}\nu_{22} \end{pmatrix}$

$\diamond$ The overflow rate r_3 is computed using the equation:

$$r_3 = \min\left(q_S^e - r_1 - r_2, \frac{1}{\nu_{31}}(q_{O_2}^p - r_1\nu_{11})\right)$$

Then: $\begin{pmatrix} q_S^e - r_1 - r_2 \\ q_{O_2}^p - r_1\nu_{11} \\ q_{Ac}^p \\ v^p - \frac{r_1}{K_{conv}}\nu_{12} - \frac{r_2}{K_{conv}}\nu_{22} \end{pmatrix} \to \begin{pmatrix} q_S^e - r_1 - r_2 - r_3 \\ q_{O_2}^p - r_1\nu_{11} - r_3\nu_{31} \\ q_{Ac}^p \\ v^p - \frac{r_1}{K_{conv}}\nu_{11} - \frac{r_2}{K_{conv}}\nu_{22} \end{pmatrix}$

$\diamond$ Finally, the acetate oxidation rate r_4 is computed using the equation:

$$r_4 = \min\left[q^p_{Ac}, \frac{1}{\nu_{41}}\left(q^p_{O_2} - r_1\nu_{11} - r_3\nu_{31}\right), \frac{1}{\nu_{42}}\left(v^p - \frac{r_1}{K_{conv}}\nu_{12} - \frac{r_2}{K_{conv}}\nu_{22}\right)K_{conv}\right]$$

Then: $\begin{pmatrix} q^e_S - r_1 - r_2 - r_3 \\ q^p_{O_2} - r_1\nu_{11} - r_3\nu_{31} \\ q^p_{Ac} \\ v^p - \frac{r_1}{K_{conv}}\nu_{12} - \frac{r_2}{K_{conv}}\nu_{22} \end{pmatrix} \rightarrow \begin{pmatrix} q^e_S - r_1 - r_2 - r_3 \\ q^p_{O_2} - r_1\nu_{11} - r_3\nu_{31} - r_4\nu_{41} \\ q^p_{Ac} - r_4 \\ v^p - \frac{r_1}{K_{conv}}\nu_{12} - \frac{r_2}{K_{conv}}\nu_{22} - \frac{r_4}{K_{conv}}\nu_{42} \end{pmatrix}$

Consequently, the outputs of the metabolic model are recorded in the table:

- $q^e_S = r_1 + r_2 + r_3$
- $q^e_{O_2} = r_1\nu_{11} + r_3\nu_{31} + r_4\nu_{41}$
- $q^e_{Ac_{consumption}} = r_4$
- $q^e_{Ac_{excretion}} = r_2\nu_{23} + r_3\nu_{33}$
- $v^e = r_1\nu_{12} + r_2\nu_{22} + r_4\nu_{42}$

with $q^e_{O_2}$ (g_{O_2}/h) the actual oxygen uptake rate, that is less than or equal to $q^p_{O_2}$, $q^e_{O_2} = q^p_{O_2}$ meaning that oxygen is a limiting strand of the cell's growth.

A reagent–based scheme

This formulation in modes is not the only way to compute the cell–specific metabolic yields from the S, O_2, Ac inputs, a formulation in reagents can also be considered, which consists in comparing the cells' own capabilities with the environmental offer first and evaluating the limiting strand before computing the desired outputs once the preliminary tests have allowed to pick out the appropriate configuration. An illustration of the algorithm is proposed for the sake of clarity and the implementation of the code is provided to the interested reader at the end of the chapter.

However, after a run of both algorithms, it has been observed that this logical formulation in interlocked tests is not as efficient as its algebraic mode–based counterpart and has consequently been ditched in the simulations.

The ν_{ij}s intervening in the algorithm are recorded in the following Table 3.1.

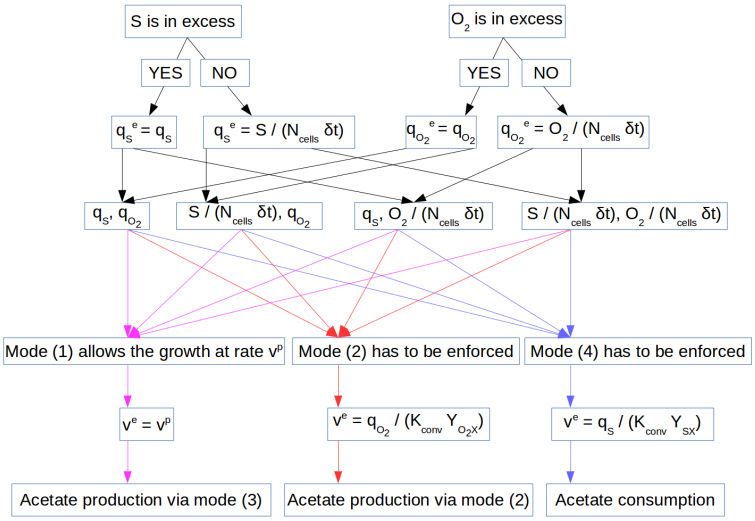

Figure 3.8 – An illustration of the reagent–based procedure to solve the metabolic model .

Table 3.1 – Metabolic ratios used in the simulations.

Parameter	value
ν_{11}	0.40104 g_{O_2}/g_S
ν_{12}	0.45091 g_X/g_S
ν_{23}	0.53378 g_{Ac}/g_S
ν_{22}	0.14404 g_X/g_S
ν_{31}	16/45 g_{O_2}/g_S
ν_{33}	2/3 g_{Ac}/g_S
ν_{41}	0.23767 g_{O_2}/g_{Ac}
ν_{42}	0.58600 g_X/g_{Ac}
$Y_{SX} = 1/\nu_{12}$	2.21774 g_S/g_X

3.5 Discussion

• Many impairments of this very simple model are questionable. For instance, temperature and pH were not taken into account among the dynamic variables, albeit having a well–known influence on the cells' metabolic capacity. Also, the CO_2's negative impact on the organisms' health was neglected in spite of its massive importance when experiments are carried out. This is partly due to the fact that the aim is a comparison with data collected in case–in–point experiments. Also, metabolic modelling is out of the scope of this work which mainly focuses on the use of structured models in biology. Metabolic models as they stand (given the limited knowledge at the disposal of mathematicians / engineers) involve culture medium–specific reaction rates and their coupling with an ever–fluctuating hydrodynamic environment is not an option.

• The reactor is assumed perfectly aerated, what is reasonable in a lab–scale fermenter. The gas–liquid transfer from a bunch of bubbles to the medium can be supposed linear in this context with a satisfying accuracy. This gross approximation originates in the reasoned choice to primarily address the liquid–solid transfer. A refinement of this model would consist in a computation of the gas–liquid transfer term from a population balance treatment of the gas phase, what would ratchet the computational time up to unreasonable durations.

• One of the features of such an individual-based approach is the definition of the cell growth rate from crossed moments of the PBE:

$$\frac{\partial}{\partial t}n(t,l,\xi_{-l}) + \frac{\partial}{\partial l}\left[\dot{l}n(t,l,\xi_{-l})\right] + \nabla\cdot\left[\dot{\xi_{-l}}n(t,l,\xi_{-l}\right] + \gamma(l)n(t,l,\xi_{-l}) + Dn(t,l,\xi_{-l})$$
$$= 2\int_{l}^{\bar{l}}\gamma(l')P(l,l')\int P(\xi_{-l},\xi'_{-l})n(t,l',\xi'_{-l})d\xi'_{-l}dl'$$

with ξ_{-l} standing for the relative complement of l in ξ. Given the (3.11) definition of $\dot{l} = \dot{l}(l, A_P, N_p, A_p, v^p)$, the cell–scale growth rate reads:

$$\int\int\frac{\dot{l}}{l}n(t,l,\xi_{-l})d\xi_{-l}dl = \int\int\frac{v^e(A_P,N_p,A_p,v^p)\left(1-\frac{l}{\bar{l}}\right)^{\kappa}}{l}n(t,l,\xi_{-l})d\xi_{-l}dl \neq \frac{\int\int\dot{l}n(t,l,\xi_{-l})d\xi_{-l}dl}{\int ln(t,l)dl}$$

• It could be understood that the maximum of the variable v^p is $\frac{q_{P_{\max}}+\psi_{\max}N_{\max}}{Y_{SX}K_{conv}}$. However, q_S overruns $q_{P_{\max}}$ in the wake of a perturbation only, and the uptake's return to normal is massively quicker than any change in v^p since $\max(\tau_{A1},\tau_{A2}) << \tau_{vp}$, imposing a $v_p^{\max}(q_{P_{\max}})$ upper bound to v^p that does not depend on $\psi_{\max}$. An illustration of this physical conclusion will feature in section 5.4.

• In (3.11), the permease system's characteristic time ranges between 1 and 10 s. In comparison, the mesomixing / micromixing characteristic time t_m is estimated at roughly 50 ms, that is two orders of magnitude faster. In this context, the bulk of the computation cost lies in the refinement (3.3.3).

• Neither q_P nor its permease conterpart are Monod-shaped functions of S. It is of very little importance though, as evidenced in a flurry of biotechnology publications such as Luong's [75] or Koch's [52]. More problematic is the fact that $q_{P_{\max}}$ and $\psi_{\max}$ are asymptotic values that have no reason to elude the hale and healthy bacteria. More generally, this remark raises the question of the relevance of saturation functions in enzymology, the shortcoming thereof being an assumption that optimal functioning is unattainable whatever the culture condition or the organisms' history.

3.6 First lessons from an unstructured model

The "unstructured" denomination comes from the field of mathematics and might confuse the biologists who interpret the notion of structured model as a cell–scale one, distributed or not. Without any semantic consensus, "unstructured" will be synonymous with "non–distributed" here and below.

A collection of continuous cultures simulated using (3.11) is shown in figure 3.9. Throughout this part, the metabolic scheme is first coupled with a non–segregated or average individual model that only keeps a record of the population–averaged state variables:

$$\begin{aligned}
\dot{A_P} &= \frac{1}{\tau_P}\left(\frac{S}{K_S+S} - A_P\right) \\
\dot{N_p} &= \left(\frac{1}{\tau_i}\frac{K_S}{K_S+S} + \frac{1}{\tau_d}\frac{S}{K_S+S}\right)\left[N_{\max}\frac{K_S}{K_S+S} - N_p\right] \\
\dot{A_p} &= \left(\frac{1}{\tau_{A1}}\frac{q_P}{q_{P_{\max}}} + \frac{1}{\tau_{A2}}\left(1-\frac{q_P}{q_{P_{\max}}}\right)\right)\left(1-\frac{q_P}{q_{P_{\max}}} - A_p\right) \\
\dot{v^p} &= \frac{1}{\tau_{vp}}\left(\frac{q_S}{Y_{SX}K_{conv}} - v^p\right)
\end{aligned} \tag{3.17}$$

At the population scale, the mass conservation translates into the balance:

$$\begin{aligned}
\frac{dX}{dt} &= v^e K_{conv} N - DX \\
N &= XR_{NX}
\end{aligned}$$

with N the cell number and R_{NX} a constant cell number per gram. This abuse of process is forced by the formulation of $\dot{l}$ that forbids the existence of an equilibrium point on $[0, \bar{l}[$, meaning that with no size distribution to report a population's variability in progress of the cell cycle, the total biomass has to be retrieved from a correlation with an assumed mean cell length.

In this context, the glucose concentration S obeys the subsequent differential equation:

$$\frac{dS}{dt} = D(S_f - S) - \left(A_P q_{P_{\max}}\frac{S}{K_P+S} + N_p A_p \psi_{\max}\frac{S}{K_p+S}\right)N$$

with a similar approximation concerning the other compounds.

Steady–state solution

It is understood that steady–state is characterised by the property $v^e = v^p$, given that acetate consumption or excretion signal an imbalance in the organism's metabolism. In other words, only mode (1) is operational. This way, the equilibrium values are analytically retrieved and depict a manifold indexed by D:

$$\begin{aligned}
&v^e = \frac{4D}{10^3\rho\pi d^2 R_{NX}} = v^p = \frac{4q_S}{10^3\rho\pi d^2 Y_{SX}} \Leftrightarrow q_S = \frac{DY_{SX}}{R_{NX}} = q_P + q_p \\
&A_P = \frac{S}{K_P+S}\ ,\ N_p = N_{\max}\frac{K_P(K_P+2S)}{(K_P+S)^2}\ ,\ A_p = \frac{K_P(K_P+2S)}{(K_P+S)^2} \\
&q_{P_{\max}}\left(\frac{S}{K_P+S}\right)^2 + \psi_{\max}N_{\max}\left(\frac{K_P(K_P+2S)}{(K_P+S)^2}\right)^2\frac{S}{K_p+S} = D\frac{S_f-S}{N_{cell}}
\end{aligned} \tag{3.18}$$

At the end of the day, the dilution rate–dependent steady–state solution reads:

- $q_P = q_{P_{\max}} \left(\frac{S_D}{K_P + S_D} \right)^2$
- $N_p = N_{\max} \frac{K_P(K_P + 2S_D)}{(K_P + S_D)^2}$
- $A_p = \frac{K_P(K_P + 2S_D)}{(K_P + S_D)^2}$
- $v^e = \frac{4}{10^3 \rho \pi d^2 Y_{SX}} \left[q_{P_{\max}} \left(\frac{S_D}{K_P + S_D} \right)^2 + N_{\max} \frac{(K_P(K_P + 2S_D))^2}{(K_P + S_D)^4} \frac{S_D}{K_p + S_D} \right]$

As discussed before, the set (3.17) is not closed because X's dynamics do not provide any information regarding X. At the end of the day, one single equation ($q_S = DY_{SX}/R_{NX}$) connects S to X.

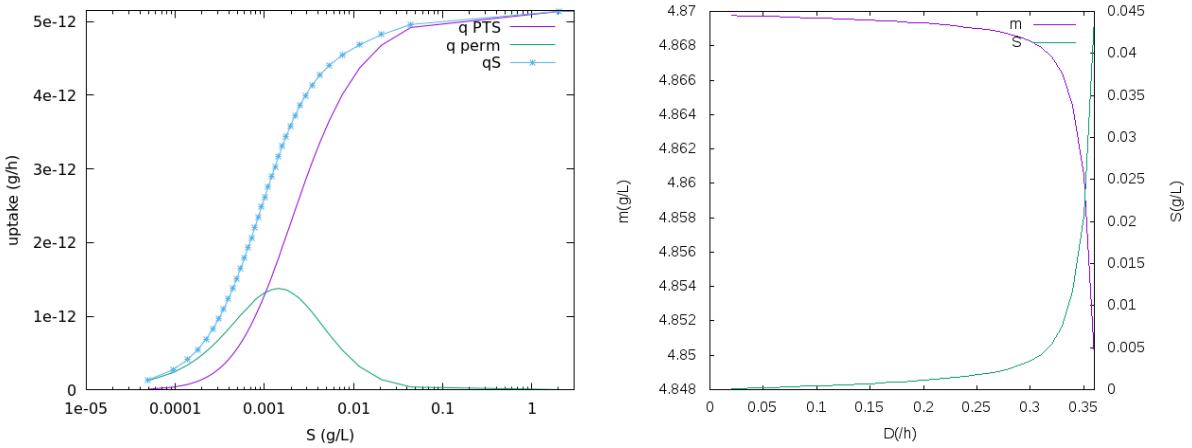

Figure 3.9 – Left : simulation of continuous cultures with variable dilution rate (from 0.01 h^{-1} to 0.4 h^{-1}). For the sake of clarity, the value of the different parameters is mentioned in Table 3.2. x axis: residual substrate concentration, y axis: steady–state values of q_P and q_{perm}. Right: steady–state biomass and sugar concentration associated for the same (3.11) set of equations, D ranging from 0.02 h^{-1} to 0.36 h^{-1}.

Some standard trends like those evidenced in [140] can be retrieved from the aforementioned unstructured model though. Figure 3.9 highlights two conclusions that have been discussed in the literature:

- The left–hand side shows that the permease mechanism allows starving cells some extra substrate (as was inferred by Neubauer & al. in [99]). It accounts for the majority of the total uptake when $S < 0.001$ g/L and is inhibited by the PTS mechanism when $S > 0.01$ g/L.
- The right–hand side illustrates the existence of a maximum growth rate beyond which any steady–state continuous culture will be abiotic only. It emerges from the assumption that $q_P(S)$ satisfies a Michaelis–Menten profile, to the point where the population is washed out as soon as $D > \frac{R_{NX} q_{P_{\max}}}{Y_{SX}}$. As a consequence, the chemostat's stability obeys an "all–or–none" logic: either the whole population subsists, should the mean individual be capable of so, or it is washed out at an exponential rate. Obviously, the continuous selection process is out of

the reach of the unstructured model that does not discriminate among the elements of the biotic phase.

Table 3.2 – Parameters used in the simulations.

Parameter	Value
K_P	0.001 g_S/L
K_p	0.01 g_S/L
$q_{P_{\max}}$	$5.145 \cdot 10^{-12}$ g_S/h
$\psi_{\max}$	$1.28625 \cdot 10^{-13}$ g_S/h
$N_{\max}$	200
t_m	0.05 s
τ_P	50 s
τ_i	100 s
τ_d	6 min
τ_{A_1}	5s
τ_{A_2}	15s
τ_{v^p}	1h

Appendix: A logical treatment of the (1)–(4) metabolic model

The reagent–based scheme to calculate $q^e_{O_2}$, q^e_{Ac} and v^e as functions of q^e_S, $q^p_{O_2}$, q^p_{Ac} and v^p breaks down into the following sequence of logical tests:

⋆ If $\int q^e_S n\delta t < S_p + D\delta t(S_f - S_p)$ (glucose is in excess)

◇ If $\int q^p_{O_2} n\delta t < O_2 + \delta t K_L a\,(O_2^\star - O_2) - \delta t D(O^e_2 - O_2)$ (oxygen is in excess)

• Then, at the cell scale, the following mechanism is implemented :

• If $\min\left(q^e_S, \frac{q^p_{O_2}}{\nu_{11}}\right) \geq \frac{1}{\nu_{12}} K_{conv} v^p$ (mode (1) is enough to ensure the cell's growth potential):

· $v^e = v^p$(the growth potential is ensured; the remaining nutrients will be involved in mode (3)).

· $q_{Ac_{excretion}} = \nu_{33} \min(q^e_S - \frac{1}{\nu_{12}} v^p K_{conv}, \frac{1}{\nu_{31}}\left(q^p_{O_2} - \frac{\nu_{11}}{\nu_{12}} v^p K_{conv}\right))$

• If $\min\left(q^e_S, \frac{q^p_{O_2}}{\nu_{11}}\right) < \frac{1}{\nu_{12}} K_{conv} v^p$ (mode (1) will not ensure the cell's growth potential):

· If $q^e_S > q^p_{O_2}/\nu_{11}$ (the oxygen uptake proves limiting ⇒ mode (2) will be enforced):

▪ $v^e \to v^e + \min(v^p - v^e, \frac{\nu_{22}}{K_{conv}}\left(q^e_S - \frac{q^p_{O_2}}{\nu_{11}}\right))$

▪ $Acetate \to Acetate + \delta t q_{Ac_{excretion}} = Acetate + \delta t \nu_{23} \min\left(q^e_S - \frac{q^p_{O_2}}{\nu_{11}}, \frac{1}{\nu_{22}} K_{conv}(v^p - v^e)\right)$

· If $q^e_S < q^p_{O_2}/\nu_{11}$ (the glucose uptake proves limiting ⇒ mode (4) will be enforced):

▪ $v^e \to v^e + \min(v^p - v^e, \frac{\nu_{42}}{K_{conv}} q^p_{Ac}, \frac{\nu_{42}}{\nu_{41} K_{conv}}(q^p_{O_2} - \nu_{11} q^e_S))$

▪ $Acetate \to Acetate - \delta t q_{Ac_{consumption}} = Acetate - \delta t \min\left(\frac{q^p_{O_2} - \nu_{11} q^e_S}{\nu_{41}}, q^p_{Ac}, \frac{1}{\nu_{42}} K_{conv}(v^p - v^e)\right)$

◇ If $\int q^p_{O_2} \delta t > O_2 + \delta t K_L a\,(O_2^\star - O_2) + D\delta t(O^e_2 - O_2)$ (oxygen is in default)

• The same algorithm repeats itself, albeit $q^p_{O_2}$ is potentially hindered by the environment availability: the actual oxygen uptake takes the shape:

$$\min(q^p_{O_2}, \frac{1}{N_{cell}}(O_2 + \delta t K_L a\,(O_2^\star - O_2) + D\delta t(O^e_2 - O_2)))$$

⋆ If $\int q_S^e \delta t > S_p + D\delta t(S_f - S_p)$ (glucose is in default):

⋄ The same algorithm repeats itself, albeit q_S^e is potentially hindered by the environment availability: the actual glucose uptake takes the shape:

$$\min(q_S^e, \frac{1}{N_{cell}}(S_p + \delta t D(S_f - S_p))$$

Chapter 4

Numerical methods

This chapter aims at breaking down the codes associated with the structured model.

4.1 A Monte–Carlo method

The numerical procedure

As mentioned in the introduction, the algorithm consists in a deterministic transport of a user–set amount of fictional particles which division or departure are stochastically treated. More precisely, an initial population of artificial cells which inner coordinates are sampled from normal laws (the procedure being repeated should the corresponding value be negative) are tracked with respect to time. In other words, for a given initial number of MC particles, a matrix of as many rows is filled with random inputs, the columns including l, q_P, N_p, A_p, v^p, along with the organisms' residence time in the reactor that obeys an exponential law:

$$x \sim \mathcal{E}(D^{-1}) \text{ for each cell}$$

The age is also tracked for the sake of comparison with the aforementioned sampled residence time.

In a second phase, the inner coordinates are updated at each time step using (3.11) for each cell and the chemical compounds' concentrations in the CSTR are calculated calling section 3.4's metabolic model. The (3.13) hydrodynamic refinement has to be enforced to guarantee that the bacteria's glucose uptake rate does not lead to a negative substrate concentration. From the oxygen and acetate availability, the individuals' potential uptake rates are corrected more coarsely if necessary, to ensure the said concentrations remain non–negative at the end of the operation. The formulation in reactions has been preferred from a computional time point of view, its reagents–based counterpart consisting in consecutive loops burdening the code with unnecessary logical tests. The new outputs are recorded as the basis for the next time step.

Then the mitosis and washout events are determined by the sampling of a random number u for each cell and the following comparisons:

Let $u \sim \mathcal{U}_{[0,1]}$: mitosis occures in case $1 - \exp(-\gamma(l)\delta t) < u$

Washout occurs as soon as the cell's age overbears x

When a cell divides, its inner properties are redistributed according to the kernel K (involving one random number defining the redistribution in length due to the assumed conservation of mass and four another ones pertaining to the daughter–cells' A_P and A_p, v^p being assumed equal to the mother–cell's for each newborn and N_p redistributed in proportion to l), and each new cell is given a residence time drawn from an $\mathcal{E}(D^{-1})$ distribution. The cell age is reset to zero for one of the daughter cells, making room for a new lineage in the fermenter, whereas the other daughter keeps the record of the mother–cell's lineage. In the following numerical simulations, the redistribution in A_P will be assumed following a $\beta(6,6)$ law supported over the whole $[0,1]$ (meaning that $P_{A_P}(A_P, A_P{}')$ does not in fact depend on $A_P{}'$), its length counterpart will be modelled with a $\beta(10,10)$, and the daughter–cells' A_p will be sampled from a $\mathcal{U}_{[0,1]}$ law. With no experimental data allowing to discriminate between the different admissible laws though, an endless number of models can be picked equivalently with little to no consequence to the marginal NDFs. Indeed, whatever the redistribution in A_P, N_p or A_p at a mother–cell's rupture, the characteristic time of their dynamics is significantly shorter than the daughter–cells' expected interdivision time, meaning that a cell's state at birth will be of little relevance during the majority of its lifetime. The most crucial modelling assumption bears upon $P_{v^p}(v^p, v^{p'})$ given the characteristic time of this variable's adaption to the organism's environment. Although it has been supposed in this model that a generational transition should have little influence on the daughter–cells' growth properties, experimental measurements of consecutive interdivision times amid one lineage (cf. Yasuda's [141,148]) hint at significant gaps in sister–cells length profile with respect to time. It is worth a mention that from biological grounds, there is no reason to suspect that the lengthening rate the daughter–cells are handed down is limited by an unsurpassable threshold such as the so–called maximum batch culture–growth rate, the latter standing for a population–averaged datum but shall not bind a cell–scale quantity.

Of course, one fictional Monte–Carlo particle carries the information of a certain N_{corr} actual cells in such a way that the product N_{corr} × "Number of MC particles" = $\int_{\Omega_\xi} n(\xi) d\xi$.

In the field of population dynamics, the massive advantage of lagrangian simulations is the lineage–tracking that makes it possible to extract some information that is out of the PBM's reach. Hence, this knowledge shall not be interpreted as an analytical result but as a method–dependent benefit. For example, considering a PBM which inner variables are the cells' mass and metabolism–related quantities, following the particles with respect to time makes it possible to extract their age and their (possibly consecutive) interdivision times until their residence time in the reactor is reached. In other words, $n+2$ cell–scale data are within the reach of a n-dimensional PBM.

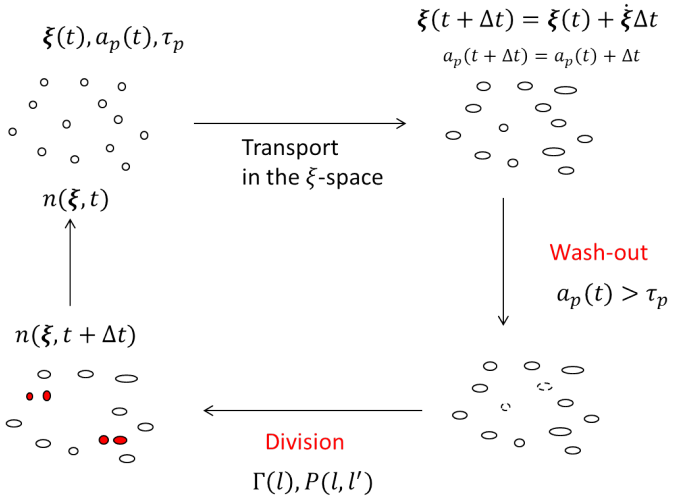

Figure 4.1 – An illustration of the Monte–Carlo procedure

The technical details

In order to draw real numbers from various probability laws, the sampling process follows a Box–Muller style algorithm which consists in algebraic transformations of uniform law–sampled numbers.
As an example, beta numbers are engineered using the lemmas:

Lemma: Let $(u_n)_{n\in\{1,\dots,N\}} \sim \mathcal{U}_[0,1]$. Then $-\sum_{n=1}^{N} \ln(u_n) \sim \Gamma(N)$

Proof. The proof will first be given for $N = 1$, the extension to $N > 1$ is a consequence of a following lemma. Let f be a bounded continuous function. Then the transfer theorem yields:

$$\int_0^1 f(u)du \underset{v=-\ln(u)}{=} \int_0^\infty f(v)e^{-v}dv$$

that is well an exponential law shape. □

Lemma: Let $u \sim \Gamma(N)$, $v \sim \Gamma(M)$. Then $\frac{u}{u+v} \sim \beta(N, M)$

Proof. The proof is once again a consequence of the transfer theorem. Let f be a bounded continuous function. The following integral will be transformed:

$$\frac{1}{\Gamma(N)\Gamma(M)} \iint_{\mathbb{R}_+\times\mathbb{R}_+} f(u,v)e^{-u}e^{-v}u^{N-1}v^{M-1}dudv$$

via a change of variables $(x, y) = (u + v, \frac{u}{u+v})$, which jacobian reads:

$$\left|\begin{matrix} 1 & 1 \\ \frac{1}{u+v} - \frac{u}{(u+v)^2} & -\frac{u}{(u+v)^2} \end{matrix}\right| = \frac{1}{u+v} \Rightarrow dxdy = \frac{1}{u+v}dudv \Leftrightarrow dudv = xdxdy$$

So:

$$\frac{1}{\Gamma(N)\Gamma(M)} \iint_{\mathbb{R}_+\times\mathbb{R}_+} f(u,v)e^{-u}e^{-v}u^{N-1}v^{M-1}dudv =$$
$$\frac{1}{\Gamma(N)\Gamma(M)} \iint_{\mathbb{R}_+\times\mathbb{R}_+} f(x,y)e^{-x}x^{N-1}y^{N-1}x^{M-1}(1-y)^{M-1}xdxdy =$$
$$\iint_{\mathbb{R}_+\times\mathbb{R}_+} f(x,y)x^{N+M-1}e^{-x}\frac{1}{\Gamma(N)\Gamma(M)}y^{N-1}(1-y)^{M-1}dxdy \tag{4.1}$$

At the end of the day, $u+v \sim \Gamma(N+M)$ (what in earnest completes the proof of the first lemma) and $\frac{u}{u+v} \sim \beta(N,M)$. □

Remark: Given uniform–sampled random numbers, the transfer theorem helps generating random numbers distributed according to many classic probability laws which emerge from the field of modelling. A non–exhaustive review of the algebraic transformations that are supported by these algorithms features on table 4.1.

Table 4.1 – Uniform–generated numbers transformation.

Law	Uniform numbers required	Algebraic transformation
Gamma(N)	N: $(u_n)_{1\leq n\leq N}$	$-\sum_{n=1}^{N}\ln(u_n)$
Beta(N, M)	$N+M$: $(u_n)_{1\leq n\leq N}$, $(v_m)_{1\leq m\leq M}$	$\frac{\sum_{n=1}^{N}\ln(u_n)}{\sum_{n=1}^{N}\ln(u_n)+\sum_{m=1}^{M}\ln(v_m)}$
$\mathcal{N}(0,1)$	2: u, v	$\sqrt{-2\ln(u)}\cos(2\pi\ln(v))$ or $\sqrt{-2\ln(u)}\sin(2\pi\ln(v))$
Weibull(n, α)	1: u	$-\frac{1}{\alpha}\ln(u)^{\frac{1}{n}}$
Log-normal(μ, σ^2)	2: u, v	$\exp\left(\mu+\sigma(\sqrt{-2\ln(u)}\cos(2\pi\ln(v)))\right)$ or $\exp\left(\mu+\sigma(\sqrt{-2\ln(u)}\sin(2\pi\ln(v)))\right)$
Chi-2(N)	$2N$: $(u_n)_{1\leq n\leq N}$, $(v_n)_{1\leq n\leq N}$	$\sum_{n=1}^{N}\sqrt{-2\ln(u_n)}\cos(2\pi\ln(v_n))$ or $\sum_{n=1}^{N}\sqrt{-2\ln(u_n)}\sin(2\pi\ln(v_n))$
Pareto(α, n)	1: u	$\left(\frac{\alpha}{u}\right)^{\frac{1}{n}}$

4.2 A Finite Volume method

Without loss of generality, the PBE reads:

$$\frac{\partial}{\partial t}n(t,\xi)+\nabla_\xi\left[\dot{\xi}n(t,\xi)\right]+\gamma(\xi)n(t,\xi)+Dn(t,\xi)=2\int_{\Omega_{\xi'}}\gamma(\xi')K(\xi,\xi')n(t,\xi')d\xi' \tag{4.2}$$

with the obvious definition of $\xi=(l,A_P,N_p,A_p,v^p)$ and $\Omega_{\xi'}=[l,\bar{l}[\times[0,1]\times[0,N_{\max}]\times[0,1]\times[0,v_p^{\max}]$.

The time order is defined by the discretisation of the numerical scheme with respect to t. A standard explicit Euler scheme would be consistent with a first–order method in time, whereas a Runge–Kutta method to go forward in time would ratchet the time order of the scheme at will. In this section, a first–order explicit Euler scheme is proposed, meaning that the evolution of n between t^m and t^{m+1} in each volume O of the phase space obeys:

$$n_O^{m+1} = n_O^m + \int_{t^m}^{t^{m+1}} \frac{\partial}{\partial t} n_O(t) dt$$

and the spotlight will once and for all be set on the right–hand side's integral.

The transport term

In this context, n is not known at each point of the phase space; instead, only the mean value $n_O(t) = \frac{1}{|O|}\int_O n(t,\xi)d\xi$ ($|\cdot|$ standing for the standard Lebesgue measure) is forthcoming at time t at the volume scale. An introduction of (4.2) into the integral leads to:

$$|O| \int_{t^m}^{t^{m+1}} \frac{\partial}{\partial t} n_O(t) dt \approx (t^{m+1} - t^m)$$
$$\left[- \int_{\partial O} \left[\dot{\xi} n(t,\xi)\right] d\sigma - \int_O \gamma(\xi) n(t,\xi) d\xi + 2 \int_O \sum_{O' | l' \geq l} |O'| \gamma(l') K(\xi,\xi') n(t,\xi') d\xi' \right]$$

The boundary integral over ∂O is understood as a sum over all the OO' interfaces where O' stands for a volume neighbouring O in any direction. If $\nu_{OO'}$ is the outward unit normal at the OO' interface, then:

$$\int_{\partial O} \left[\dot{\xi} n(t,\xi)\right] = \sum_{O'} \frac{|O|}{|OO'|} \left[\left(\dot{\xi} n\right)_O^+ (t) + \left(\dot{\xi} n\right)_{O'}^- (t) \right] \nu_{OO'} \tag{4.3}$$

Indeed, since the value of n is not known at any OO' interface, the desired quantity has to be interpolated from the known values of the NDF at each cell centre. The signs are consistent with the observation that an outward flux ($(\dot{\xi} n)_O^+ \nu_{OO'} > 0$) reports a transfer of matter from the cell under consideration towards its neighbour, the converse assessment translating into $(\dot{\xi} n)_{O'}^- \nu_{OO'} < 0$.

Set a meshing of the $[0,\bar{l}]\times[0,1]\times[0,N_{\max}]\times[0,1]\times[0,v_p^{\max}]$ domain in $N_l \times N_{A_P} \times N_{N_p} \times N_{A_p} \times N_{v^p}$ cells and five integers $0 \leq i < N_l$, $0 \leq j < N_{A_P}$, $0 \leq k < N_{N_p}$, $0 \leq o < N_{A_p}$, $0 \leq r \leq N_{v^p}$ (for the sake of simplicity, the mesh can be uniformly drawn in each dimension): the $\mathcal{F}^{x,m}_{\cdot,\cdot,\cdot,\cdot,\cdot}$ numerical fluxes $\dot{x}n_{\cdot,\cdot,\cdot,\cdot,\cdot}$ at time t^m will be calculated from (4.3) using a similar scheme as Nessyahu & Tadmor's [98]. Treating only one arbitrary dimension x, it will explicitly read:

$$\begin{aligned}
&\frac{1}{\Delta x}\Bigg[\max(\dot{x_{\iota+1}}^m, 0) \left(n(x_\iota)^m - \frac{t^{m+1}-t^m}{2\Delta x} \Delta\mathcal{F}^{x,m}|_{x\in[x_\iota, x_{\iota+1}]} \right) \\
&+ \min(\dot{x_{\iota+1}}^m, 0) \left(n(x_{\iota+1})^m - \frac{t^{m+1}-t^m}{2\Delta x} \Delta\mathcal{F}^{x,m}|_{x\in[x_{\iota+1}, x_{\iota+2}]} \right) \\
&- \max(\dot{x_\iota}^m, 0) \left(n(x_{\iota-1})^m - \frac{t^{m+1}-t^m}{2\Delta x} \Delta\mathcal{F}^{x,m}|_{x\in[x_{\iota-1}, x_\iota]} \right) \\
&- \min(\dot{x_\iota}^m, 0) \left(n(x_\iota)^m - \frac{t^{m+1}-t^m}{2\Delta x} \Delta\mathcal{F}^{x,m}|_{x\in[x_\iota, x_{\iota+1}]} \right) \Bigg]
\end{aligned} \tag{4.4}$$

for any ι, with the quantity $\Delta\mathcal{F}$ being computed via the superbee flux limiter:

$$\Delta\mathcal{F}^{x_\iota,m} = \text{Superbee}(\dot{x_{\iota+1}}^m n(x_{\iota+1}) - \dot{x_\iota}^m n(x_\iota), \dot{x_\iota}^m n(x_\iota) - \dot{x_{\iota-1}}^m n(x_{\iota-1}))$$

the Superbee (a, b) function being defined by the formula:

$$\text{Superbee}(a, b) = \max[0, \max(\min(2a, b), \min(a, 2b))]$$

Obviously, given the containment condition, no outward normal points to the exterior of the finite volume domain: in other words, the Superbee function is equal to 0 if $\nu_{OO'}$ is directed towards Ω_ξ^ε.

It is worth mentioning that the mesh has been built regularly in all the dimensions (downgrading the algorithm to a mere finite difference scheme), the rationale behind this choice being the massive dependence of the NDF on the dilution rate. Any refinement of the mesh would have to be concomitant with the update of the NDF (D being given once and for all), what would ratchet the computation time of an already time–consuming method.

The dilution and fragmentation terms

Once n has been transported using the (4.2) numerical scheme, the loss in each $\mathcal{M}_{i,j,k,o,r}$ mesh cell due to the fragmetation process is computed using a mid–point rule:

$$\int_{\mathcal{M}_{i,j,k,o,r}} \gamma(l) n(t,\xi) d\xi \approx \gamma\left(\frac{1}{2}(l_i + l_{i+1})\right) \int_{\mathcal{M}_{i,j,k,o,r}} n(t,\xi) d\xi \tag{4.5}$$

(4.5) is further used to compute the redistribution integral on the $\Omega_{\xi'}$ domain:

$$\int_{\Omega_{\xi'}} \gamma(l') K(\xi,\xi') n(t,\xi') d\xi' = \sum_{(i',j,k,o,r)|l_{i'}>l_i} \int_{\mathcal{M}_{i',j,k,o,r}} \gamma(l') K(\xi,\xi') n(t,\xi') d\xi'$$

A few simplifications can speed the calculation of the right–hand side up. For instance, v^p is not redistributed at the division because a division event is unlikely to improve or alter the daughter–cells' health. A bonanza emerging from the fixed mesh framework lies in the observation that if $P_{A_P}(A_P, A_P{}')$ is not contingent upon $A_P{}'$, the fraction of the newly–forrmed orgainsms landing in each $\mathcal{M}_{i,j,k,o,r}$ for $0 \leq j < N_{A_P}$ can be calculated before running the Finite Volume code and only needs to be called at the time the redistribution integral is computed. A_p being uniformly picked and also independent on $A_p{}'$, only the estimation of P_l and P_{N_p} need further consideration. These kernels can be viewed as $N_l \times N_l$ and $N_{N_p} \times N_{N_p}$ matrices, which respective entries are:

$$\int_{l_i}^{l_{i+1}} \int_{l_{i'}}^{l_{i'+1}} P_l(l,l') dl' dl \text{ and } \int_{N_{pk}}^{N_{pk+1}} \int_{N_{pk'}}^{N_{pk'+1}} P_{N_p}(N_p, N_p{}') dN_p{}' dN_p$$

$P_l \sim \beta(10,10)$, $P_{N_p} \sim \beta(10,10)$ being known beforehand (they are symmetric and close to gaussian pdfs) and the integrals being computed using a second–order trapezoidal method involving ≈ 2000 points each.

For the sake of exhaustiveness, the dilution term is implicited at the tail end of the code, meaning that the cell number in each mesh cell is multiplied by $1 - D(t^{m+1} - t^m)$. The rest of the algorithm follows the Monte–Carlo method, consisting in enforcing the (3.13) hydrodynamic refinement before calling the reaction–based metabolic model to update the chemical compounds' concentration. It is worth mentioning that given that $\dot{\xi}$ depends (directly or indirectly, through q_S for instance) on S, the time step ends with a revision of the numerical fluxes which will be employed on the $[t^{m+1}, t^{m+2}]$ time interval.

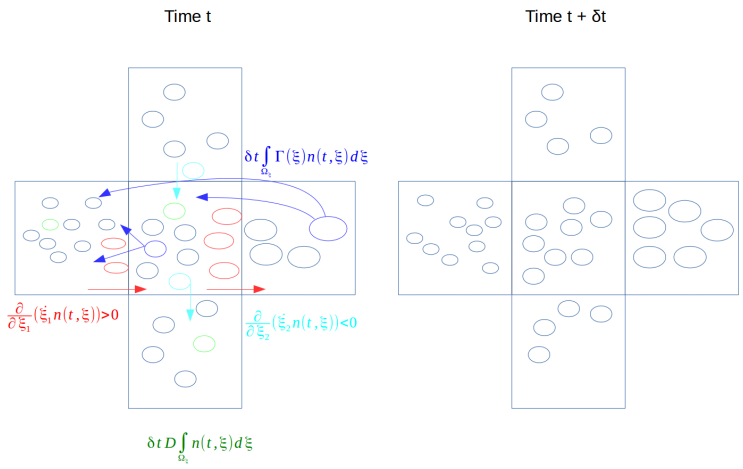

Figure 4.2 – An illustration of the Finite Volume procedure in two dimensions, the horizontal swathe pertaining to the length variable whereas the vertical one refers to any other of the model's dimensions, say v^p. At a given time t, the flux in the l direction transfers (in red) some organisms from a $\mathcal{M}_{i,j}$ cell to its $\mathcal{M}_{i+1,j}$ neighbour, some individuals coming up from the $\mathcal{M}_{i-1,j}$ volume in the meantime. The same procedure involving the $\mathcal{M}_{i,j-1}$, $\mathcal{M}_{i,j}$ and $\mathcal{M}_{i,j+1}$ in the v^p direction is pictured in cyan. A few bacteria are washed out between time t and time $t + \delta t$, the process being depicted in green. Finally, the cell division affecting the balance in each of the mesh's cell is represented in blue, the population's initial state at time $t + \delta t$ featuring on the right–hand side of the figure.

4.3 A hybrid Finite Volume - Method of Moments

In order to solve the PBE (4.2), a numerical method could consist in a call to two different schemes: a Finite Volume algorithm to solve the transport equation:

$$\frac{\partial}{\partial t} n(t,\xi) + \frac{\partial}{\partial l}\left[\dot{l} n(t,\xi)\right] + \frac{\partial}{\partial A_P}\left[\dot{A_P} n(t,\xi)\right] + \frac{\partial}{\partial A_p}\left[\dot{A_p} n(t,\xi)\right] + \frac{\partial}{\partial v^p}\left[\dot{v^p} n(t,\xi)\right] = 0 \tag{4.6}$$

and a Quadrature Method of Moments to treat the fragmentation equation:

$$\frac{\partial}{\partial t} n(t,\xi) + \frac{\partial}{\partial N_p}\left[\dot{N_p} n(t,\xi)\right] + D n(t,\xi) = 2 \int_{\Omega_{\xi'}} \gamma(l') K(\xi,\xi') n(t,\xi') d\xi' \tag{4.7}$$

The election of N_p as the variable to enforce QMOM on is not insignificant: indeed its dynamics does not depend on any other of the model's inner coordinates, alleviating the transport of the moments that comes from the treatment of (4.7)'s second term as will be seen below.

The solution of (4.6) follows from the previous section, and is transported over the time interval $[t^m, t^{m+1}]$ to give birth to $\hat{n}(t^{m+1})$. A potential use of $\hat{n}$ as an initial condition to (4.7) has to be gingerly considered because each cell–to–cell flux modifies the moments inside the receiving mesh element. More precisely, when the FV method is run, transfers between a $\mathcal{M}_{i,j,o,r}$ mesh cell and its immediate neighbours $\mathcal{M}_{i\pm1,j,o,r}$, $\mathcal{M}_{i,j\pm1,o,r}$, $\mathcal{M}_{i,j,o\pm1,r}$, $\mathcal{M}_{i,j,o,r\pm1}$ do not only affect the cell number in each volume; indeed, given that the moments associated to each organism inside a mesh cell are independently updated at each time step, the said transfers amount to mixing particles having a certain NDF in N_p with organisms emerging in the cell with their own distribution in N_p. As a consequence, before the Wheeler algorithm is called to address the fragmentation equation, one would have to recalculate the corresponding moments of the permease number NDF having taken into account the actual state of the population once processed by the FV scheme.

In the context of this particular method, one distribution in permease density is assumed in each cell of the mesh in l, A_P, A_p, v^p, and is approached by a sum of Dirac deltas:

$$n(t,v) = \sum_{i=1}^{I} \omega_i \delta_{N_p - N_{pi}}$$

The $(\omega_i)_{i\in\{1,\dots,I\}}$ weights and $(N_{pi})_{i\in\{1,\dots,I\}}$ nodes are computed from the $2I$ first moments of the PDE. The 0–th moment of equation (4.7) $\tilde{N_p}^0$ satisfies the differential equation:

$$\frac{\partial}{\partial t}\int_{\mathcal{M}_{i,j,o,r}}\int_0^{N_{\max}} n(t,\xi)d\xi + \int_{\mathcal{M}_{i,j,o,r}}\int_0^{N_{\max}} \gamma(l)n(t,\xi)d\xi + D\int_{\mathcal{M}_{i,j,o,r}}\int_0^{N_{\max}} n(t,\xi)d\xi$$
$$= 2\int_{\mathcal{M}_{i,j,o,r}}\int_0^{N_{\max}}\Bigg[\int_0^{l'} P_l(l,l')dl \int P_{A_P}(A_P,A_P{}')dA_P \int P_{N_p}(N_p,N_p{}')dN_p \int P_{A_p}(A_p,A_p{}')dA_p$$
$$\int P_{v^p}(v^p,v^{p'})dv^p\gamma(l')n(t,\xi')\Bigg]d\xi'$$

That is:
$$\frac{d}{dt}\tilde{N_p}^0 + \int_{\mathcal{M}_{i,j,o,r}}\int_0^{N_{\max}} \gamma(l)n(t,\xi)d\xi + D\tilde{N_p}^0 = 2\int_{\mathcal{M}_{i,j,o,r}}\int_0^{N_{\max}}\Bigg[\int_0^{l'} P_l(l,l')dl \tag{4.8}$$
$$\int P_{A_P}(A_P,A_P{}')dA_P \int P_{N_p}(N_p,N_p{}')dN_p \int P_{A_p}(A_p,A_p{}')dA_p \int P_{v^p}(v^p,v^{p'})dv^p\gamma(l')n(t,\xi')\Bigg]d\xi'$$

where in (4.8), for the sake of clarity, the mesh cell $[l_i, l_{i+1}]\times[A_{P_j}, A_{P_{j+1}}]\times[A_{p_o}, A_{p_{o+1}}]\times[v_r^p, v_r^{p+1}]$ has been christened $\mathcal{M}_{i,j,o,r}$.
The same reasoning is enforced to calculate the quantities $\tilde{N_p{}^u} = \int_{\mathcal{M}_{i,j,o,r}}\int_0^{N_{\max}} {N_p}^u n(t,\xi)d\xi$, $u \geq 1$. A two–node quadrature will require the computation of $\tilde{N_p}^1$, $\tilde{N_p}^2$ and $\tilde{N_p}^3$, with respective dynamics:

$$\frac{\partial}{\partial t}\tilde{N_p}^1 - \int_{\mathcal{M}_{i,j,o,r}}\int_0^{N_{\max}} \dot{N}_p n(t,\xi)d\xi + \int_{\mathcal{M}_{i,j,o,r}}\int_0^{N_{\max}} N_p\gamma(l)n(t,\xi)d\xi + D\tilde{N_p}^1 = 2\int_{\mathcal{M}_{i,j,o,r}}\int_0^{N_{\max}} \tag{4.9}$$
$$\Bigg[\int P_{A_P}(A_P,A_P{}')dA_P \int N_p P_{N_p}(N_p,N_p{}')dN_p \int P_{A_p}(A_p,A_p{}')dA_p \int P_{v^p}(v^p,v^{p'})dv^p\gamma(l')n(t,\xi')\Bigg]d\xi'$$

and, for $u \geq 2$:

$$\frac{\partial}{\partial t}\tilde{N_p}^u - u\int_{\mathcal{M}_{i,j,o,r}}\int_0^{N_{\max}} {N_p}^{u-1}\dot{N}_p n(t,\xi)d\xi + \int_{\mathcal{M}_{i,j,o,r}}\int_0^{N_{\max}} {N_p}^u \gamma(l) n(t,\xi)d\xi + D\tilde{N_p}^u \quad (4.10)$$

$$= 2\int_{\mathcal{M}_{i,j,o,r}}\int_0^{N_{\max}}\left[\int_0^{l'} P_l(l,l')dl \int P_{A_P}(A_P,{A_P}')dA_P \int {N_p}^u P_{N_p}(N_p,{N_p}')dN_p \int P_{A_p}(A_p,{A_p}')dA_p \right.$$

$$\left. \int P_{v^p}(v^p,v^{p'})dv^p \gamma(l')n(t,\xi')\right]d\xi'$$

Of course, the raw moments $\langle {N_p}^u \rangle = \frac{\tilde{N_p}^u}{\int n}$, $u \geq 1$ obey the differential equation:

$$\frac{d}{dt}\langle {N_p}^u \rangle = \frac{\frac{\partial}{\partial t}\tilde{N_p}^u}{\int n d\xi} - \frac{\tilde{N_p}^u}{\int n d\xi}\frac{\partial}{\partial t}\int n d\xi \quad (4.11)$$

all these quantities having been defined as above. Also, given the formulation of $\dot{N}_p$, the expected steady–state distribution in N_p has to be close enough to a dirac (the adaption to the environment–induced value of N_p for the newborn cells being driven by the rather small τ_i, τ_d time constants) to approximate it by moments up to order 3.

It is worth a mention that, should the QMOM algorithm involve A_p instead of N_p for instance, the second term of (4.11) for $u = 2$ would read:

$$-2\int_{\mathcal{M}_{i,j,k,r}}\int_0^1 \left[\frac{1}{\tau_{A_1}}\frac{q_P(A_P)}{q_{P_{\max}}} + \frac{1}{\tau_{A_2}}\left(1 - \frac{q_P(A_P)}{q_{P_{\max}}}\right)\right]\left(A_p - \frac{q_P(A_P)A_p}{q_{P_{\max}}} - {A_p}^2\right)n(t,\xi)d\xi$$

making the calculation of $\tilde{A_p^2}$ significantly harder than its $\tilde{N_p^2}$ counterpart due to the non–linearities binding A_P with A_p. A similar case can be made against the use of v^p or A_P as the QMOM–processed dimension.

Regarding the QMOM procedure in itself, the computation of the weights and nodes comes first, starting from the number density function n which moments are defined at time t^m. Wheeler's algorithm is used for this purpose, and ensures that the corresponding output abscissas lie within the support of the length distribution. The main idea behind Wheeler's algorithm consists in exploiting the following theorem, as mentioned for instance in Marchisio & Fox's [84]:

Theorem 4.3.1.

The approximation: $\sum_{i=1}^{I}\omega_i f(x_i)$ *of the integral:* $\int_\Omega f(x)n(x)dl$

is gaussian if and only if its nodes x_i *coincide with the* I *roots of the polynomial* P_I *orthogonal in* Ω *with respect to* n.

Then, once Wheeler's algorithm is enforced in each mesh cell, all the integrals on the right–hand side of (4.9) and (4.10) are reduced to weighted sums and (4.11) is used to compute the NDF's moments at $t = t^{m+1}$.

The computation of the appropriate weights and abscissas obeys the following procedure. To get started, one has to find a L^2–orthogonal polynomial basis on the support of the distribution of interest (in this case $[0, N_{\max}[$) through a Gram–Schmidt method. The $\{1, x, x^2, x^3\}$ family (which yields a two–node quadrature) is orthonormalised for the scalar product $\langle \cdot, \cdot \rangle$ such that $\langle P, Q \rangle = \int_0^{N_{\max}} P(x)Q(x)dx$ into:

$$\begin{aligned}
P_0(x) &= 1 \\
P_1(x) &= \frac{x - N_{\max}/2}{\int_0^{N_{\max}} x - N_{\max}/2 \, dx} \\
P_2(x) &= \frac{x^2 - N_{\max}x + (N_{\max})^2/6}{\int_0^{N_{\max}} x^2 - N_{\max}x + (N_{\max})^2/6 \, dx} \\
P_3(x) &= \frac{x^3 - \frac{3}{2}N_{\max}x^2 + \frac{3}{5}(N_{\max})^2 x - (N_{\max})^3/20}{\int_0^{N_{\max}} x^3 - \frac{3}{2}N_{\max}x^2 + \frac{3}{5}(N_{\max})^2 x - (N_{\max})^3/20 \, dx}
\end{aligned} \tag{4.12}$$

In Wheeler's algorithm, unitary polynomials are of particular interest, meaning that all denominators will be ditched in order to establish the recursive relation between polynomials of consecutive order:

$$P_{k+1}(x) = (x - a_k)P_k(x) - b_k P_{k-1}(x) \ \forall k \geq 0 \tag{4.13}$$

considering that $P_{-1} \equiv 0$.

Secondly, all the quantities $\int P_j(x)P_k(x)n(t,x)dx$ are gathered in the $M = (M_{j,k})_{1 \leq j \leq I, 1 \leq k \leq 2I}$ matrix. In accordance with (4.13):

$$M_{j+1,k} = \int_0^{N_{\max}} (x - a_j)P_j(x)P_k(x)n(t,x)dx - \int_0^{N_{\max}} b_j P_{j-1}(x)P_k(x)n(t,x)dx$$

$$= \int_0^{N_{\max}} P_j(x) \left(P_{k+1} + a_k P_k(x) + b_k P_{k-1}(x)\right) n(t,x)dx - a_j M_{j,k} - b_j M_{j-1,k} = M_{j,k+1} + (a_k - a_j)M_{j,k}$$

$$+ b_k M_{j,k-1} - b_j M_{j-1,k}$$

and M is filled row after row. Subsequently, the intermediate Jacobi matrix is computed:

$$J = \begin{pmatrix}
J_{11} & J_{12} & 0 & 0 & 0 & 0 & 0 & \dots \\
J_{21} & J_{22} & J_{23} & 0 & 0 & 0 & 0 & \dots \\
0 & J_{32} & J_{33} & J_{34} & 0 & 0 & 0 & \dots \\
 & & & \dots & \dots & & & \\
 & \dots & 0 & J_{j-1,k} & J_{j,k} & J_{j+1,k} & 0 & \dots \\
 & & & \dots & \dots & & & \\
 & & & & & & J_{I-1,I} & J_{I,I}
\end{pmatrix}$$

with

$$J_{i,i} = a_i - \frac{M_{i-1,i}}{M_{i-1,i-1}} + \frac{M_{i,i+1}}{M_{i,i}} \ , \ i \in \{1, \dots, I\} \ , \ J_{i,i+1} = \sqrt{\frac{M_{i,i}}{M_{i-1,i-1}}} = J_{i+1,i} \ , \ i \in \{1, \dots, I-1\}$$

in line with [84]. J's eigenvalues are the actual nodes x_i, $i \in \{1, \dots, I\}$ of the gaussian quadratures, and the weights ω_i, $i \in \{1, \dots, I\}$ are computed using the first component of its eigenvectors v_i: $\omega_i = v_{i,1}^2$.

It is worth mentioning that in the case of a two-node quadrature, the Jacobi matrix simply reads: $J = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$ and its eigenelements can be analytically calculated, allowing the riddance of the highly time–consuming matrix inversion. First, the recursive coefficients read:

$$a_k = \frac{N_{\max}}{2} \ \forall k \in \{0, ..., 2\} \text{ and } b_1 = \frac{(N_{\max})^2}{12}, \ b_2 = \frac{(N_{\max})^2}{15}$$

Then, the outputs of the Jacobi matrix inversion read:

$$x_1 = \frac{1}{2}\left(a + c + \sqrt{(a-c)^2 + 4b^2}\right) \text{ and } x_2 = \frac{1}{2}\left(a + c - \sqrt{(a-c)^2 + 4b^2}\right)$$

$$v_1 \propto \begin{pmatrix} 1 \\ \frac{1}{2b}\left(c - a + \sqrt{(c-a)^2 + 4b^2}\right) \end{pmatrix} \text{ and } v_2 \propto \begin{pmatrix} 1 \\ \frac{1}{2b}\left(c - a - \sqrt{(c-a)^2 + 4b^2}\right) \end{pmatrix}$$

$$\Leftrightarrow \omega_1 = \frac{4b^2}{(4b^2 + (c-a) + \sqrt{(c-a)^2 + 4b^2}))} \text{ and } \omega_2 = \frac{4b^2}{(4b^2 + (c-a) - \sqrt{(c-a)^2 + 4b^2}))}$$

having enforced the condition $||v_i||_2 = 1$, $i \in \{1, 2\}$.

Discussion

• Obviously, a three-node quadrature involving the growth rate PDF's six first moments would involve a 3 x 3 Jacobi matrix, which (real) eigenvalues are explicitly calculable for with the help of Cardan's method. The corresponding eigenvectors would be obtained through straightforward calculations. In the following though, a two-node quadrature will be enough to retrieve some steady–state information to cross-validate the Monte–Carlo and finite volume algorithms.

• Other methods of moments could have been implemented instead of this quadrature based - Wheeler algorithm; for instance, the shape of the distribution could be a priori prescribed in each mesh cell, and considering that the growth rate PDF has a compact support, a shoo–in would be a beta law $\beta(p, q)$, which moments would be tracked with respect to time in accordance with (4.8)-(4.9). Then, the PDF's parameters would be algebraically calculated from its moments:

$$p = \frac{\langle N \rangle}{N_{\max}} \frac{\langle N \rangle N_{\max} - \langle N^2 \rangle}{\langle N^2 \rangle - \langle N \rangle^2} \text{ and } q = p\frac{N_{\max}}{\langle N \rangle} - p$$

In this case, the reconstruction of the PDF from its moments would be straightforward. However, when p and / or q is large enough (meaning the distribution is "relatively" close to a Dirac), $\Gamma(p+q)$ cannot be calculated numerically, a machine being unable to handle numbers approaching $\sim 10^{308}$. If p and q were picked around 80 for instance, a standard computer could not produce the corresponding beta law, although such a beta distribution would exhibit a significant variance (≈ 0.00155), what would be problematic if this method was to be used in the context.

4.4 A test case: a chemostat convergence to steady–state

4.4.1 A stability analysis

In this section, a simulation of (3.11) with the metabolic model is performed at a dilution rate $D = 0.2h^{-1}$. The aim of this paragraph is a cross–validation of the different numerical schemes.

The above–mentioned numerical methods are set to give a ± satisfying approximation of the (1.22)-(3.11) combination given a user–defined computational time. The stability criteria differ depending on the numerical method (a CFL condition must be satisfied to ensure exploitable FV numerics, whereas the MC rate of accuracy is significantly determined by the number of fictional particles).

The lagrangian method consisting in tracking fictional particles along characteristic curves, only the local truncation error ($\mathcal{O}(\delta t^2)$) coming from the use of Euler's method to update the scalars of interest is made in the process. The statistical error over the population's features is comprehensively determined by the number of fictional MC particles: the said population being reconstructed from ~ 350000 particles, the method–dependent error scales as $1/\sqrt{350000} \approx 0.17$ %. The MC code is consequently considered as a reference algorithm any other tool can be tested against.

The time step for all the simulations ($\delta t = 2.75 \cdot 10^{-6}$ h) is dictated by the micromixing time constant $t_m = 50$ ms $\approx 1.389 \cdot 10^{-5}$ h. In the FV code, gross approximations of the maximal rates of change of the variables yield:

- $\delta l = 9 \cdot 10^{-8}$ m, $\max(\dot{l}) \approx 3 \cdot 10^{-6}$ m
- $\delta A_P = 0.1$, $\max(\dot{A_P}) \approx \frac{1}{\tau_P} = 72$
- $\delta N_p = 20$, $\max(\dot{N_p}) \approx \max\left(\frac{1}{\tau_i}, \frac{1}{\tau_d}\right) N_{\max} = 7200$
- $\delta A_p = 0.1$, $\max(\dot{A_p}) \approx \max\left(\frac{1}{\tau_{A_1}}, \frac{1}{\tau_{A_2}}\right) = 720$
- $\delta v^p = 1.5 \cdot 10^{-7}$ m/h, $\max(\dot{v^p}) \approx \frac{1}{\tau_{v^p}} \frac{q_{P\max}}{Y_{SX} K_{conv}} \approx 2.954 \cdot 10^{-6}$ m/h

$$\Rightarrow \delta t \sum_{i=1}^{5} \frac{\dot{\xi_i}}{\delta \xi_i} \approx 5.096 \cdot 10^{-3} << 1$$

meaning that the FV scheme's stability is guaranteed. In particular, the outputs of the MC simulation can be compared with the FV–generated distributions using dimensionless quantities like:

$$\sigma_{x,y} = \frac{\sum_i (x_i - y_i)^2}{\sqrt{\sum_i x_i^2} \sqrt{\sum_i y_i^2}}$$

if $(x_i)_i$ and $(y_i)_i$ stand for the respective datasets. A FV–MC comparison of the length PDFs translates into $\sigma_{x,y} = 5.81 \cdot 10^{-3}$, meaning that the distributions are close enough to validate the MC scheme. $\sigma_{x,y}$ is even lower ($3.50 \cdot 10^{-3}$) when the MC numerics are compared with the FV–QMOM outputs, testifying to an acceptable implementation of the hybrid algorithm. All steady–state distributions are displayed in figure 4.3, evidencing the accordance between the results of the different methods. The width of the FV–retrieved NDFs, which are supported over 2 (in q_P and N_p) to 4 (in A_p and v^p) nodes, is presumably a result of the numerical diffusion in the phase space. Increasing the method's order to at least second–order would most likely narrow the NDFs down to similar dirac distributions as their MC counterparts.

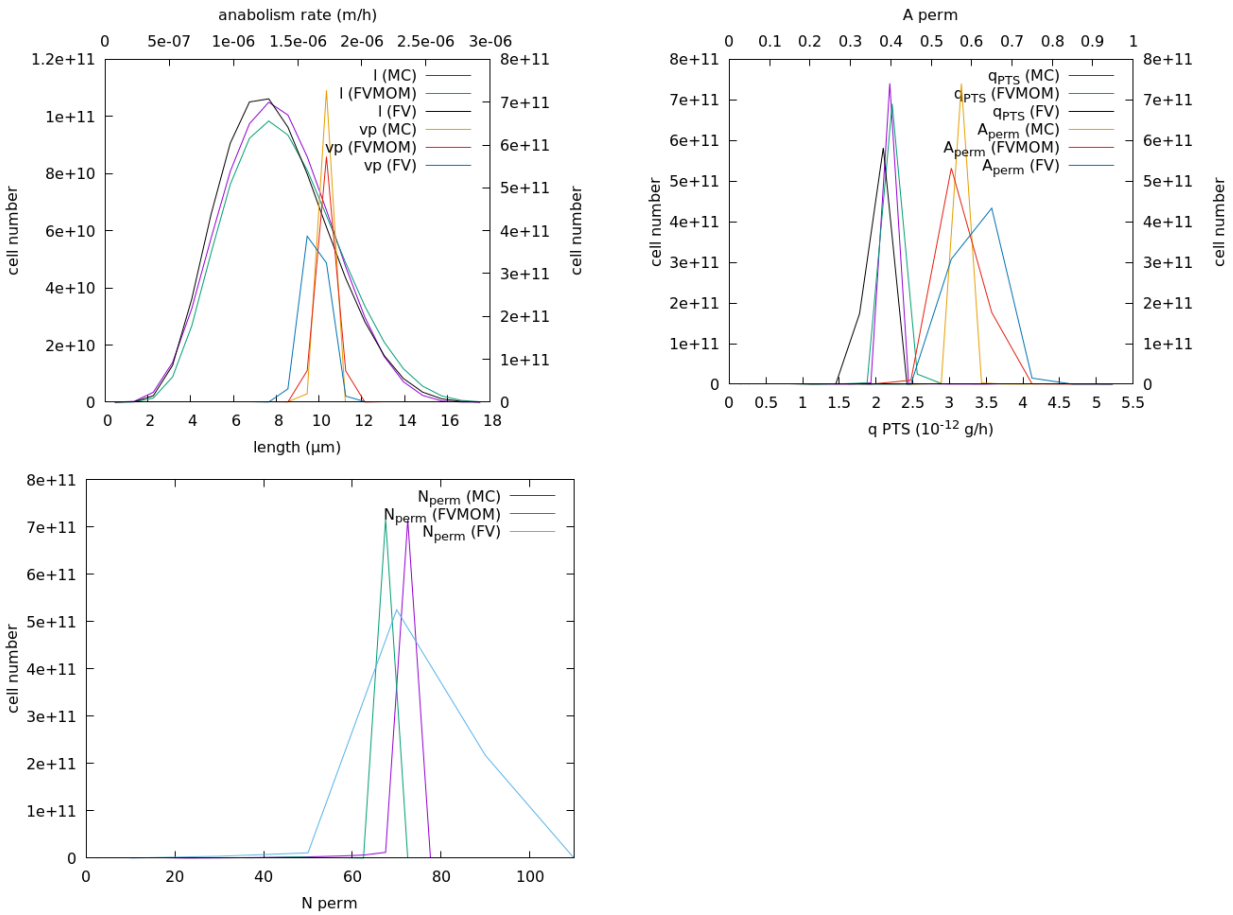

Figure 4.3 – Distributions retrieved from the Monte–Carlo, Finite Volume, and the Finite Volume–Quadrature of Moments algorithms. Top left: length and rate of anabolism. Top right: q_P and A_p. Bottom: N_p.

4.4.2 A comparison of the schemes' numerical efficiency

In this paragraph, the different numerical methods are run to sort them out time–consumingwise.

It shall be reminded that the formulation of (3.11) and its non-linear combination with the residual substrate concentration makes any exhaustive characterisation of a steady–state population a delusion. Indeed, the mean length has been proven to satisfy the equation:

$$\langle l \rangle = \frac{1}{D} \int \left(\dot{l} n \right) d\xi$$

At steady state, the cells are supposed in equilibrium with their immediate environment to the point where no acetate should be consumed or produced in the fermenter. All the uptaken glucose should be converted into biomass, entailing the conclusion $v^e = v^p$. In this context, the previous formula can be simplified:

$$\langle l \rangle = \frac{1}{D} \int \int v^p \left(1 - \frac{l}{\bar{l}} \right)^{\kappa} n(t, l, v^p) dv^p dl$$

but evidence an obvious non-linear dependence of $\langle l \rangle$ on l. Another non-linearity manifests itself if $\langle v^p \rangle$ is linked to the metabolic model:

$$\langle q_S \rangle(S) = \frac{\langle v^p \rangle}{Y_{SX}} K_{conv} = \langle q_P \rangle(S) + \langle q_p \rangle(S)$$

From the numerical data, an analytical value of S can only be estimated from the highly non-linear equation:

$$q_{P_{\max}} \langle A_P \rangle \frac{S}{K_{PTS} + S} + \psi_{\max} \langle N_p A_p \rangle \frac{S}{K_p + S} = D(S_f - S)$$

each of the population–averaged quantities depending on S. As a consequence, the validity of the different numerics is accessible from out-of-thin-air correlations only.

In the Monte–Carlo simulation, from the steady–state glucose concentration (≈ 0.00173 g/L), a guesstimate of the population–averaged inner coordinates follows:

- $\langle q_P \rangle \approx 2.067 \cdot 10^{-12}$ g/h
- $\langle A_p \rangle \approx 0.598$
- $\langle N_p \rangle \approx 74.03$

The other equilibrium data are $\langle l \rangle \approx 8.043 \cdot 10^{-6}$ m, $\langle v^e \rangle \approx 1.659 \cdot 10^{-6}$ m/h and $\langle v^p \rangle \approx 1.660 \cdot 10^{-6}$ m/h. Of course:

$$\langle l \rangle = \frac{\langle \dot{l} \rangle}{D} = \frac{\langle v^p \left(1 - \frac{l}{\bar{l}}\right)^\kappa \rangle}{D} \approx 8.043 \cdot 10^{-6} \text{m} \neq 8.056 \cdot 10^{-6} \text{m} \approx \frac{\langle v^p \rangle \left(1 - \frac{\langle l \rangle}{\bar{l}}\right)^\kappa}{D}$$

as expected. With these raw data, a calculation of the total cell mass and cell number is available. Considering that what has gone in the reactor without going out has been eaten, a trivial mass balance on a time interval of Lebesgue measure Δt would yield:

$$\text{“Cell number ”} \times \langle q_S \rangle \Delta t \approx D(S_f - S)\Delta t \Rightarrow \text{“Cell number ”} \approx 7.474 \cdot 10^{11}$$

and given their average length:

$$\text{“Cell mass ”} \approx 4.722 g$$

The same computations can be carried out for the Finite Volume and FV–QMOM steady–state numerics:

- For the FV code: $S \approx 0.00168325$ g/L $\Rightarrow \langle q_P \rangle \approx 2.025 \cdot 10^{-12}$ g/h, $\langle A_p \rangle \approx 0.606$, $\langle N_p \rangle \approx 74.54$ and “Cell number” $\approx 7.572 \cdot 10^{11}$. $\langle l \rangle$ being equal to $7.940 \cdot 10^{-6}$ m, “Cell mass” is expected to be roughly 4.722 g.
- For the FV–QMOM code: $S \approx 0.00196279$ g/L $\Rightarrow \langle q_P \rangle \approx 2.258 \cdot 10^{-12}$ g/h, $\langle A_p \rangle \approx 0.561$, $\langle N_p \rangle \approx 67.50$ and “Cell number” $\approx 7.131 \cdot 10^{11}$. $\langle l \rangle$ being equal to $8.314 \cdot 10^{-6}$ m, “Cell mass” is expected to be roughly 4.656 g.

These gross estimations are here to give the reader an order of magnitude of the different variables featuring in the simulations.

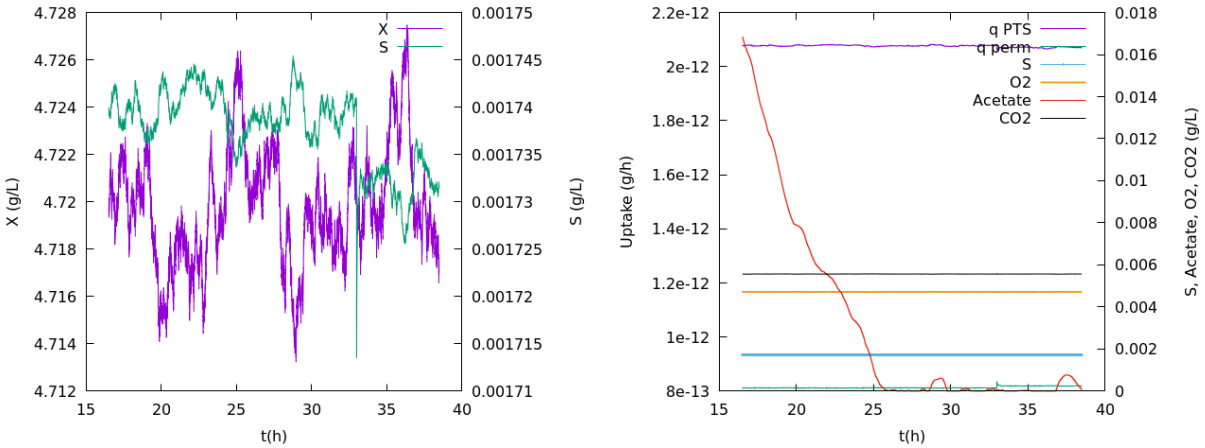

Figure 4.4 – Monte–Carlo simulation of a convergence to steady–state. Left: biomass and glucose. Right: uptake and by–products

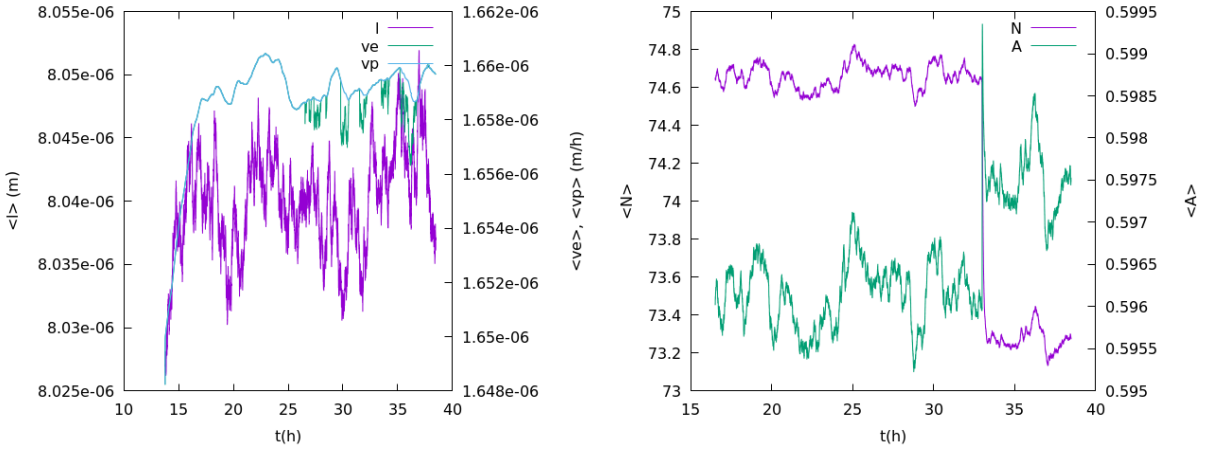

Figure 4.5 – Monte–Carlo simulation of a convergence to steady–state. Left: mean length and rate of anabolism. Right: mean permease number and permease activity.

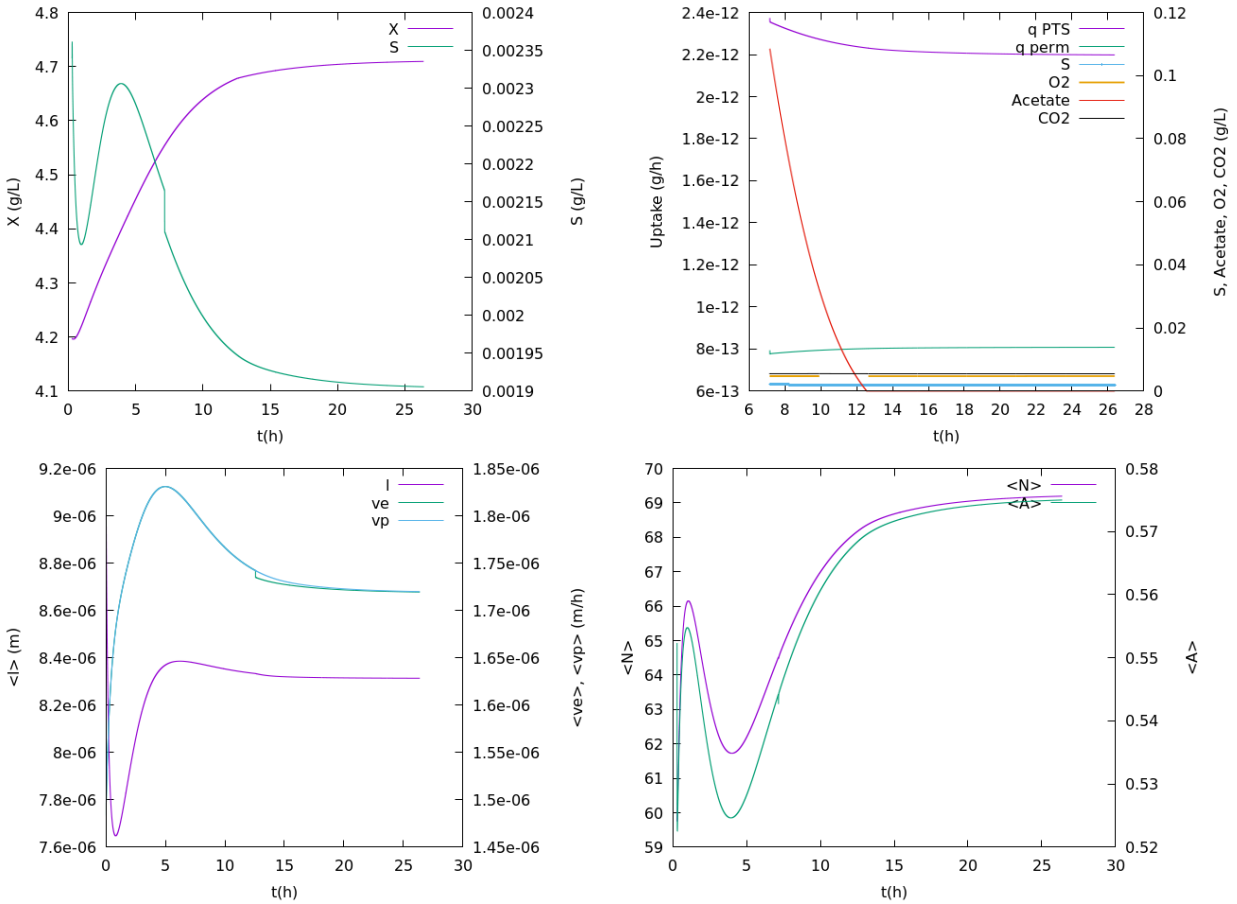

Figure 4.6 – Finite Volume–Quadrature Method of Moments simulation of a convergence to steady state. Top left: biomass and glucose. Top right: uptake and by–products. Bottom left: mean length and rate of anabolism. Bottom right: population–averaged permease number and permease activity

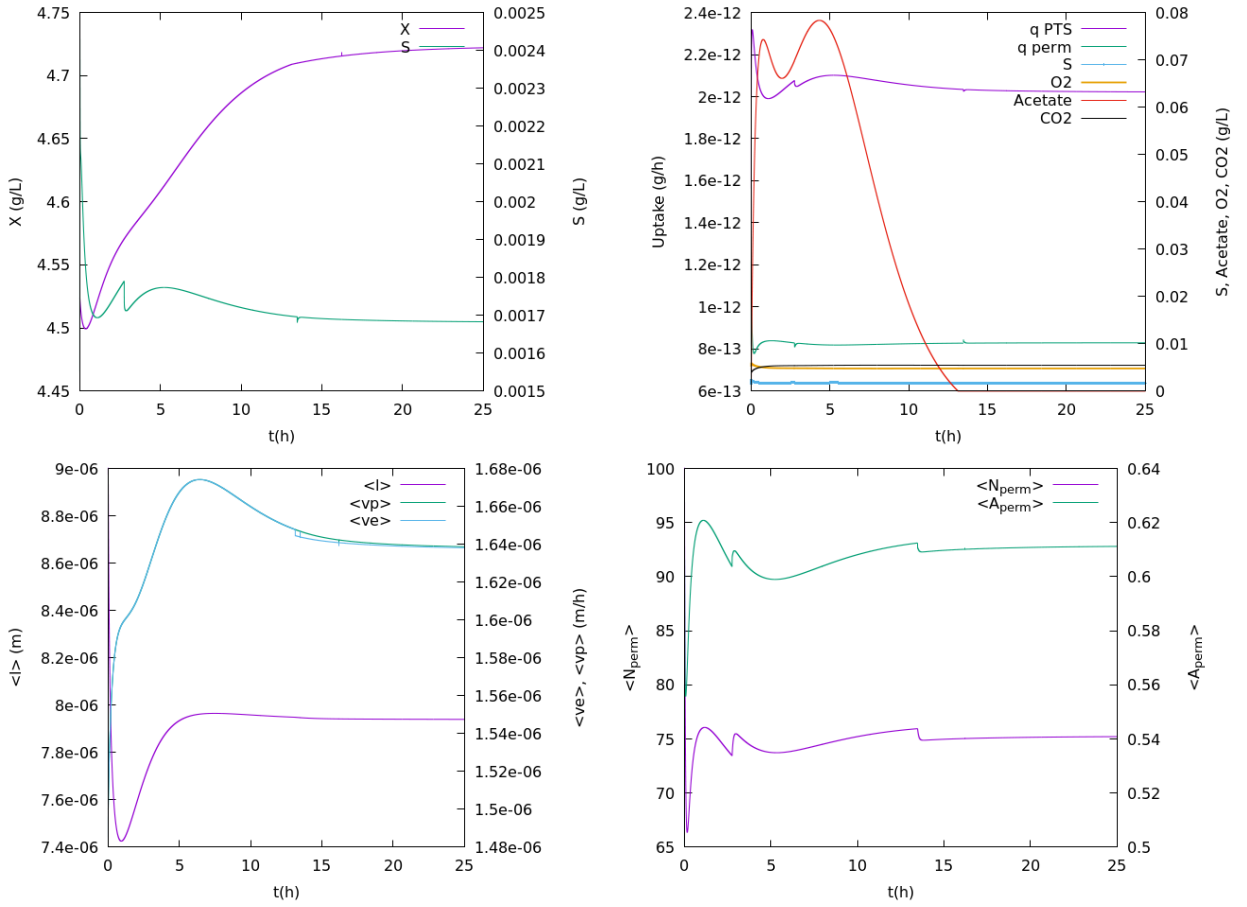

Figure 4.7 – Finite Volume simulation of a convergence to steady state. Top left: biomass and substrate as functions of time. Top right: uptake–related features and by–products. Bottom left: population–averaged length and rate of anabolism. Bottom right: mean permease number and permease activity.

The following table 4.2 sums up the settings of the different computations:

Table 4.2 – Operational set-ups.

Method	Number of cells	Number of computation nodes	Time consumption for 10^6 time steps
MC	$\approx$ 350000 fictional particles	17	2 days
FVMOM	40000 mesh cells	7	4 days
FV	400000 mesh cells	17	9 days

A comparison between figures 4.4 and 4.6 or 4.7 which depict the convergence to equilibrium at both cell– and reactor–scales points to seemingly noised Monte–Carlo data when compared with their FV or FV–QMOM counterparts. However, the oscillations in the steady–state values of residual substrate do not exceed $1.2 \cdot 10^{-5}$ g/L (roughly 0.7 % of S) over the last 5h of culture, whereas the

biomass does not stray from a $1.15 \cdot 10^{-2}$ g/L interval in the meantime, corresponding to a 0.2 % variation. It might happen that steady state has yet to be reached in the chemostat, but the amplitude of the oscillations is only emphasised by the time window the plot zeroes in. Similar remarks can be made regarding the statistic error on the population's mean length and lengthening rate for instance, although the trends are in agreement with the FV and FV–QMOM numerics. The top right plot of each figure highlight similar uptake and by–products profiles, with the exception of sporadic low–amplitude acetate excretion periods.

The conclusion of this performance comparison highlights the benefits of the Monte–Carlo method to solve multivariate PBEs. Indeed, it compares favourably with the Finite Volume solution while operating more than 4 times quicker than its eulerian counterpart. The method's main flaw is its intrinsic stochastic noise, but its range is controlled by the number of fictional MC particles (roughly 0.2 % if 350000 particles are tracked in the system). The hybrid FV–QMOM algorithm which operates ≈ 5 times as fast as the FV one at the expense tracking 5 times fewer abscissas in the N_p variable does not lead to the same steady–state numerics as the other aforementioned codes, probably because a two node–approximation of the permease number NDF does not provide enough information to retrieve the complete profile with satisfying accuracy. The mean permease number (≈ 73.29) is indeed 5.6 % below the equivalent Monte–Carlo datum and 7.9 % under its Finite Volume counterpart, whereas the respective differences in terms of permease activity, for example, amount to 3.8 % and 5.9 %. Making up an attempt to strike a balance between computational time and precision, this method should not be fastened to the point where it can be run on a standard 8 core–computer though, leaving wide open the problem of implementing light and accurate numerical methods to solve PBEs. In this context, the Monte–Carlo tool will be preferred in the next chapter dedicated to numerical simulations of different litmus test–experiments.

Chapter 5

Simulations and comparison with experimental data

5.1 Growth in mass vs. growth in number

Section 2.2.1's claim that a population's growth in number exhibits a latency in comparison with its growth in mass when disrupted is exemplified in figure 5.1. To illustrate this reasoning, the (3.10) model is run until steady state is reached, and the equilibrium is disrupted by a 1g glucose pulse in the reactor.

When a perturbation forces a population to stray from its equilibrium behaviour, growth in mass always precedes growth in number for obvious biological reasons. Indeed, if given the possibility, a cell will start lengthening before engaging in the division process. Hence, the gain in mass is dictated by the continuous evolution of the organisms' length, whereas the gain in number operates at the time scale of the cell cycle (the order of magnitude thereof being 1h).

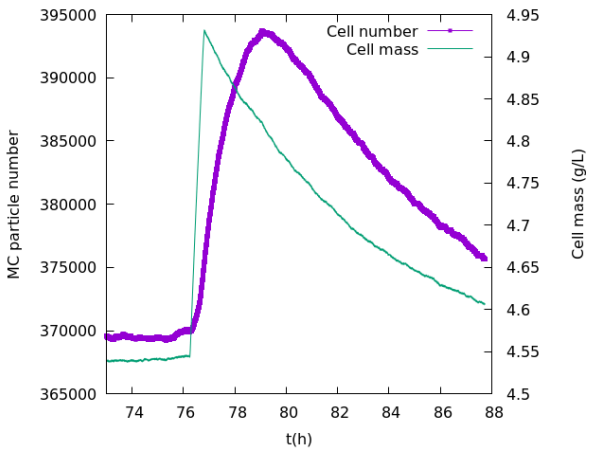

Figure 5.1 – Monte–Carlo simulation of a glucose pulse in a steady–state chemostat, evidencing a difference between the time scales of growth in mass and growth in number. Left: number of fictional Monte–Carlo particles, with constant N_{corr}. Right: biomass.

This uncoupling between the time scales is an integral part of any population balance model

featuring a size–related inner coordinate; by the way, this property has also been claimed by Doumic [19] amongst others.

5.2 Simulations of the metabolic model

In order to evidence the primary field of each biochemical equation of the (1)-(4) set page 111, the metabolic model is run using a non–distributed version of (3.11), the outputs being plotted on figures 5.2 and 5.3. More precisely:

- The cell number is calibrated to match the expected biomass at steady state once a mean length has been assumed for each bacterium.
- The environmental offer is defined before the run, S and O_2 being set to arbitrary non–limiting values.
- q_S (respectively q_{O_2}) varies between 0 and $q_{P_{\max}}$ (respectively $q_{O_{2\max}}$), scanning its domain in arbitrary steps
- v^p is fixed ($v^p = 1.6 \cdot 10^{-6}\ m/h$) in each case.
- Two cases are considered: a 10^{-4} g/L residual acetate concentration is available to the bacteria (the corresponding numerics featuring on figure 5.2) and no acetate in the fermenter (the pertaining data being plotted on figure 5.3)

Given this set–up, the function returns v^e which is compared with the predefined v^p and q_{Ac} which sign indicates acetate production ($q_{Ac} > 0$) or consumption ($q_{Ac} < 0$). Whether the $v^e = v^p$ equality is satisfied or not results from the strategy the cells are allowed to ambition, considering both their physiology and the substrate availabilty at their neighbourhood. The sign of q_{Ac} hints at the pathways the individuals have followed to make the best of the uptaken substrate.

On figures 5.2 and 5.3, $v^e = v^p$ (in red on both figures) is achieved from glucose oxidation (reaction (1)) if the organisms' q_S and q_{O_2} allow it (top right quarter of both plots). In this case, acetic fermentation (reaction (3)) can be enforced to the point where some acetate is excreted in the medium, in other words $q_{Ac} > 0$. The difference between the two simulated cases manifests through:

- the involvement of acetate oxidation (reaction (4), top left quarter of figure 5.2's plots) if q_S does not suffice to achieve $v^e = v^p$, at the expense of acetate consumption though. The consequence is $q_{Ac} < 0$ as soon as some acetate is available to the cells
- the impossibility to offset the glucose uptake deficiency in case of acetate shortage. On figure 5.3, $q_{Ac} = 0$ translates into $v^e < v^p$.

Anaerobic growth on glucose (reaction (2)) is set in motion when q_{O_2} is the element that prevents v^e from reaching v^p from reaction (1) only (bottom of both plots). In line with the involved compounds, some acetate is produced in proportion to the glucose assimilation, and $v^e > 0$; $v^e = v^p$ is assured when a little oxygen can be uptaken from the culture medium. It is worth a mention that acetate excretion is triggered by both over–uptake and contingency metabolic modes, but not by the (1) standard biochemical reaction, leading to a transitioning zone between the primary fields of the (2) and (4) peripheral pathways.

Finally, it is worth a mention that the position of the frontiers between the different primary fields depends significantly on the cells' needs through their v^p, as can be seen from figure 5.4 representing similar plots involving the same environmental offer with the exception of $v^p = 2.5 \cdot 10^{-6}$ m/h instead of $1.6 \cdot 10^{-6}$ m/h.

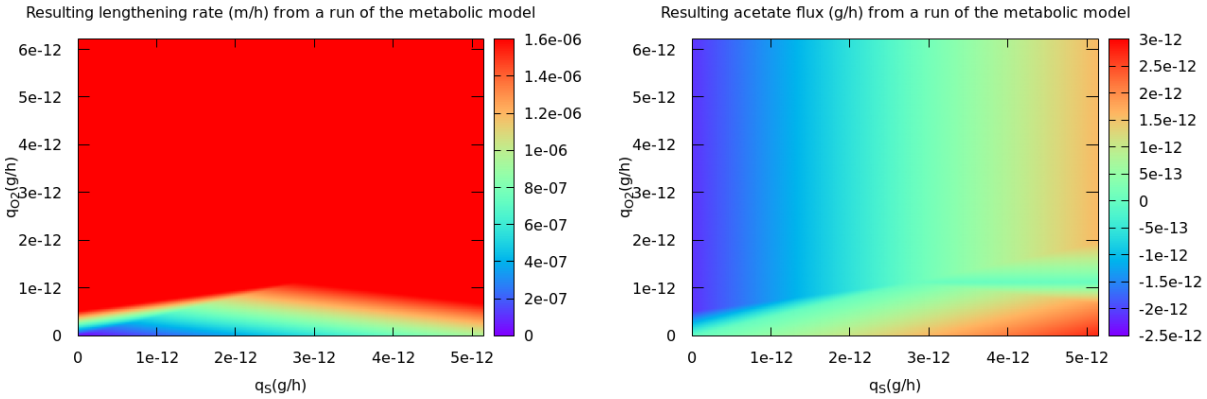

Figure 5.2 – Outputs of the metabolic model using a non–distributed version of (3.11). Left: v^e as a function of q_S and q_{O_2}. Right: q_{Ac} as a function of q_S and q_{O_2}. In this case, the environment consists in $S = 0.0001\ g/L$, $O_2 = 0.0001\ g/L$, and $Acetate = 0.0001\ g/L$.

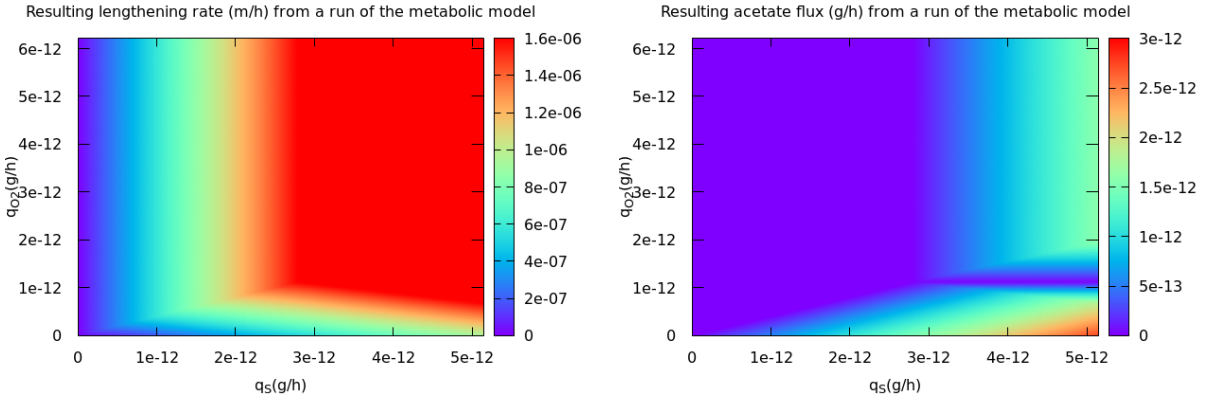

Figure 5.3 – Outputs of the metabolic model using a non–distributed version of (3.11). Left: v^e as a function of q_S and q_{O_2}. Right: q_{Ac} as a function of q_S and q_{O_2}. In this case, the environment consists in $S = 0.0001$ g/L, $O_2 = 0.0001$ g/L, and $Acetate = 0$ g/L.

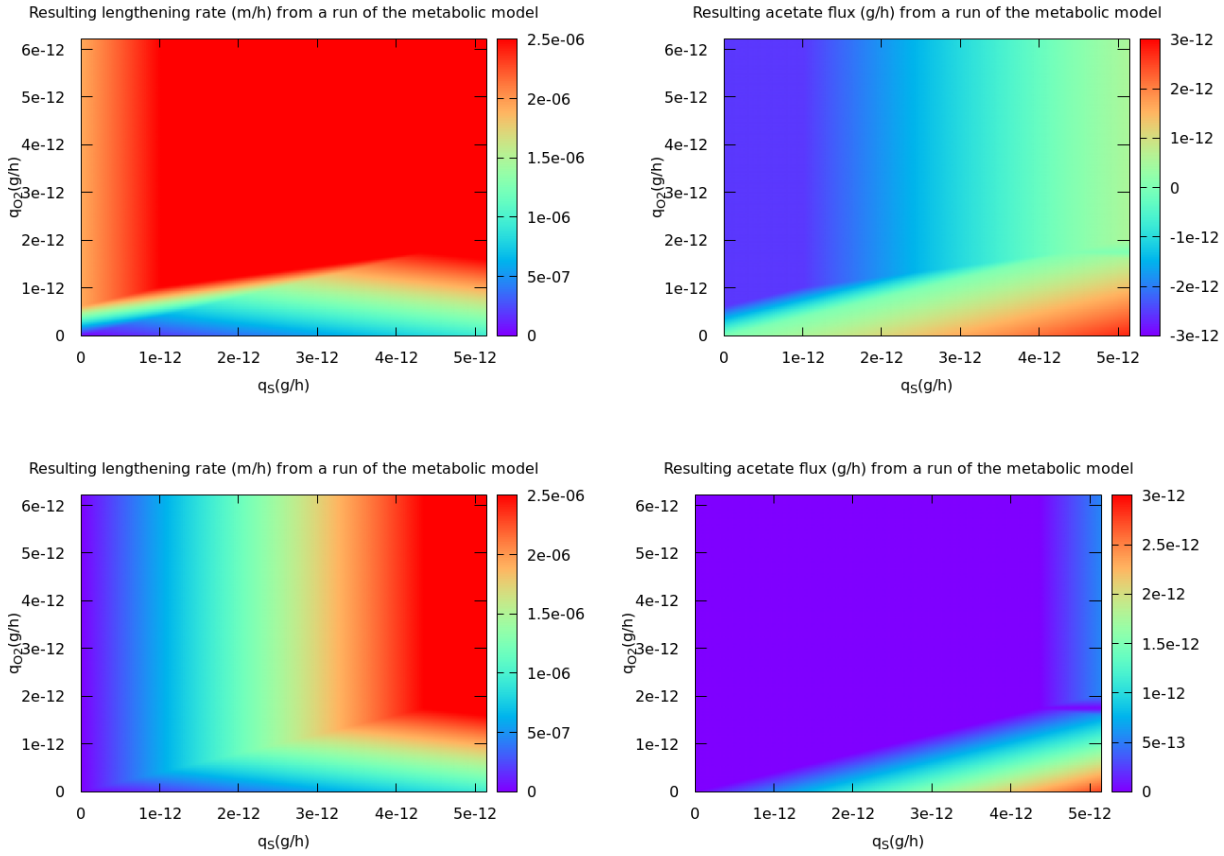

Figure 5.4 – Outputs of the metabolic model using a non–distributed version of (3.11). In this case, v^p is set equal to $2.5 \cdot 10^{-6}$ m/h. The environment consists in $S = 0.0001$ g/L, $O_2 = 0.0001$ g/L, and $Acetate = 0.0001$ g/L on the top swathe, 0 g/L on the bottom one. The left–hand side plots feature v^e (color bar) as a function of q_S and q_{O_2}, the right–hand side being made of $q_{Ac}(q_S, q_{O_2})$.

In the standard approach of biological population modelling, the substrate–to–biomass ratio is fixed whatever the organisms' physiological state, whereas the (1)–(4) metabolic model suggests that this population–averaged quantity is retrieved from the cells' q_S/v^e quotient, the latter exhibiting a massive dependence on the individuals' v^p. With these considerations in mind, the said model aims at evidencing the response to strong perturbations in terms of acetate production (reaction (3)) when the equilibrium $S \to X$ pathway is saturated, this raw loss of efficiency having been confirmed by a flurry of experiments such as Sunya & al.'s [133] glucose pulse. Moreover, it is illustrated that the concept of limiting concentration encompasses considerations of both heterogeneous catalysis and metabolism. On top of that, the proposed metabolic model is not geared towards a maximisation of the bacteria's lengthening rate, contrary to the customary hypothesis that the cells' strategy is the best return–on–investment growthwise (cf. [149]). The latter standpoint is tantamount to the claim that the living cells can instantaneously adapt to a sudden change in their environment, what would once again go against the observed latency in their response to strong perturbations.

5.3 The batch experiment

In this section, the Monte–Carlo algorithm has been run to simulate the consumption of a defined quantity glucose by a user–set initial biomass. Figure 5.5 illustrates the $X_0 \approx 0.05$ g/L, $S_0 = 8$ g/L case, whereas figure 5.6 refers to a $X_0 \approx 0.001$ g/L, $S_0 = 0.5$ g/L set–up. Special caution will bear upon the qualitative cell–scale features when $S \to 0$ as one conclusion from Notley–McRobb & al.'s [104] raw data featuring on page 25's figure 1.17 is the cells' ability to anticipate a substrate dearth by setting in motion their contingency uptake mechanism. Obviously, the well–known batch culture trends like exponential growth need to be looked at first.

Exponential growth is reached from the moment v^e reaches on average its maximum value. The latter is correlated to the population's maximum growth rate $\mu_{\max}$ via the equation:

$$\mu_{\max} = \int \left(\frac{1}{l} \frac{dl}{dt} n \right) d\xi = \int \left(\frac{v^e}{l} \left(1 - \frac{l}{\bar{l}} \right)^{\kappa} n \right) d\xi \tag{5.1}$$

$$\left(\text{Obviously: } \mu_{\max} \neq \frac{\int (v^e n)\, d\xi}{\int ln(l) dl} \left(1 - \frac{\int ln(l) dl}{\bar{l}} \right)^{\kappa} \right)$$

In the context of simplifying the formulae for $\mu_{\max}$ that evidences nonlinearities binding v^e to the state variables, a coarse but somehow informational approximation could consist in $v^e \approx v^p$, turning (5.1) into a double integral:

$$\mu_{\max} \approx \iint \frac{v^p}{l} \left(1 - \frac{l}{\bar{l}} \right)^{\kappa} n(l, v^p) dv^p dl$$

Both the exact and the approached values of $\mu_{\max}$ are at the Monte–Carlo code's reach and the values are:

$$\mu_{\max_{exact}} \approx 0.362 h^{-1} \text{ and } \mu_{\max_{approached}} \approx 0.362 h^{-1}$$

It must be emphasised that as low at it seems, this value is massively influenced by the cells' length and as a consequence the parameters $\bar{l}$, l_c and $l_{\inf}$ which dictate the rupture process. It is also worth a mention that $\mu_{\max}$ pertains to an exponential growth situation that can be viewed as an equilibrium; whether or not it can be momentarily overreached in the wake of a massive environmental perturbation is still an open question in the biology community.

In this context, a more interesting datum could consist in the lengthening rate in exponential growth, which is also available from the Monte–Carlo code ($\dot{l}_{\max} \approx 2.834 \cdot 10^{-6}$ m/h). In terms of order of magnitude, this value is consistent with Zheng & al.'s claim (cf. [152]) that a bacterium's volume doubles in a ~ 1 h time interval. Overall, lengthening rate seems to be the pivotal quantity for the sake of comparison between different experimental set–ups.

On the other hand, the above mentioned numerical values have no impact on the qualitative behaviour of the whole population, what is the topic of the present work.

The progressive substrate depletion by proactive organisms is illustrated on figure 5.5. Over the course of the batch experiment, exponential growth is achieved after 4 to $5h$ that is slightly less than 5 times the characteristic time of v^p's dynamics ($1h$). A negligible ($< 0.02g$) acetate excretion is associated with the transient regime, the surplus being available to the newly–born cells which q_{PTS} at birth is not enough to satisfy their v^p. The permeases are inactive as soon as the PTS system

supplies enough glucose for the cells' needs, with a slight anticipation of the looming shortage that characterises the individuals' proactive behaviour in a closed system, what is qualitatively in line with Notley–McRobb & al.'s [104] raw data (cf. figure 1.17).

Albeit identical processwise, two batch experiments involving a different terminal biomass concentration will lead to a different analysis. More precisely, if the trends are expected to be similar (q_P skyrocketing and q_p falling in the pre–exponential phase, all cell–scale features reaching their feast mode equilibrium as the exponential phase starts, the permease trying to take over as S is plunging, ...), the data are a point for discussion in themselves to the extent that at the end of the culture, the larger the cell mass, the shorter the growth deceleration, the shorter the time given to the permeases to compensate the nose dive in q_P. In order to illustrate this claim, similar data are plotted on figures 5.5 and 5.6, with $X(t = 0) \approx 0.055$ g and $S(t = 0) = 8$ g in the former case, whereas $X(t = 0) \approx 0.001$ g and $S(t = 0) = 0.5$ g in the latter. If X is around 3.5 g at the time of substrate depletion, N_p does not overreach $N_{\max}/8$ and A_p is roughly equal to $2/3$; on the other hand, a 0.45 g terminal biomass does not exhaust the reactor at the same pace, allowing time for N_p to reach 40 % $N_{\max}$ and for A_p to skyrocket to more than $3/4$.

Another look at the bacteria's behaviour in front of the substrate depletion is provided on figure 5.7, emphasising the influence of the culture's terminal conditions on the cell–scale features when $S \to 0$. The q_p uptick appears insignificant if $X(S \to 0) \approx 3.5$ g/L but can be observed if $X(S \to 0) \approx 0.45$ g/L. The values of both q_S and v^e need to be discussed too. At the tail end of the simulation, for a given substrate concentration in the fermenter (say 10^{-4} g/L), the more the population has grown, the more the substrate offer is limiting, the closer q_P is to 0. On the other hand, the so–called "lag" (that is pre–exponential) phase goes with a much larger acetate excretion when the initial census is more numerous, the consequence being the possibility to sustain for a longer period of time the crusing exponential growth rate.

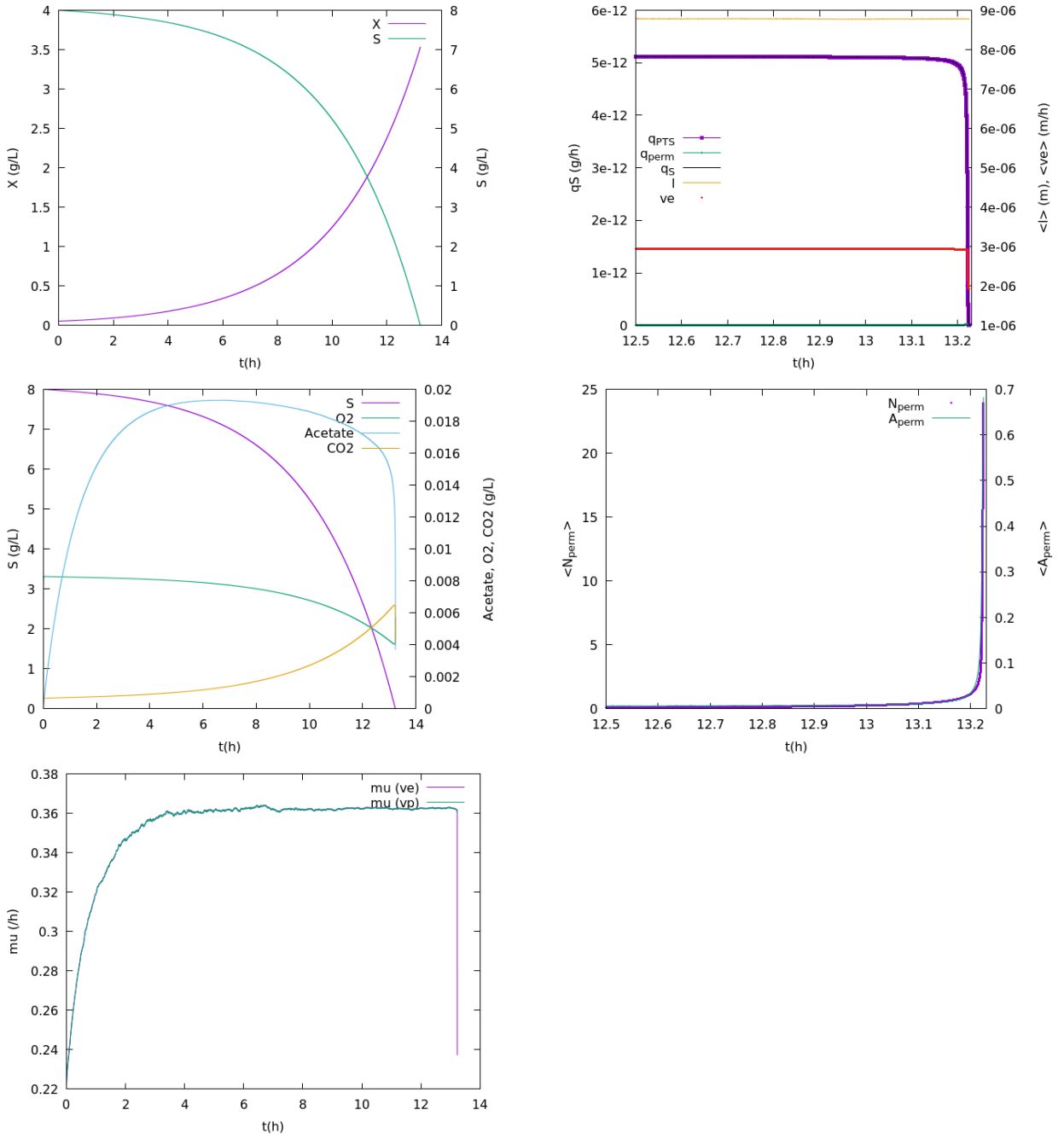

Figure 5.5 – Batch culture simulation, with initial biomass ≈ 0.055 g and substrate concentration 8 g. Top left: $X(t)$ and $S(t)$. Top right: q_S and its PTS/permease contributions as S is plunging. Centre left: substrates and byproducts with respect to time. Centre right: population–averaged N_p and A_p at the final moments of the glucose exhaustion. Bottom: exact (from (5.1)) growth rate and its approximation involving the organisms' v^p.

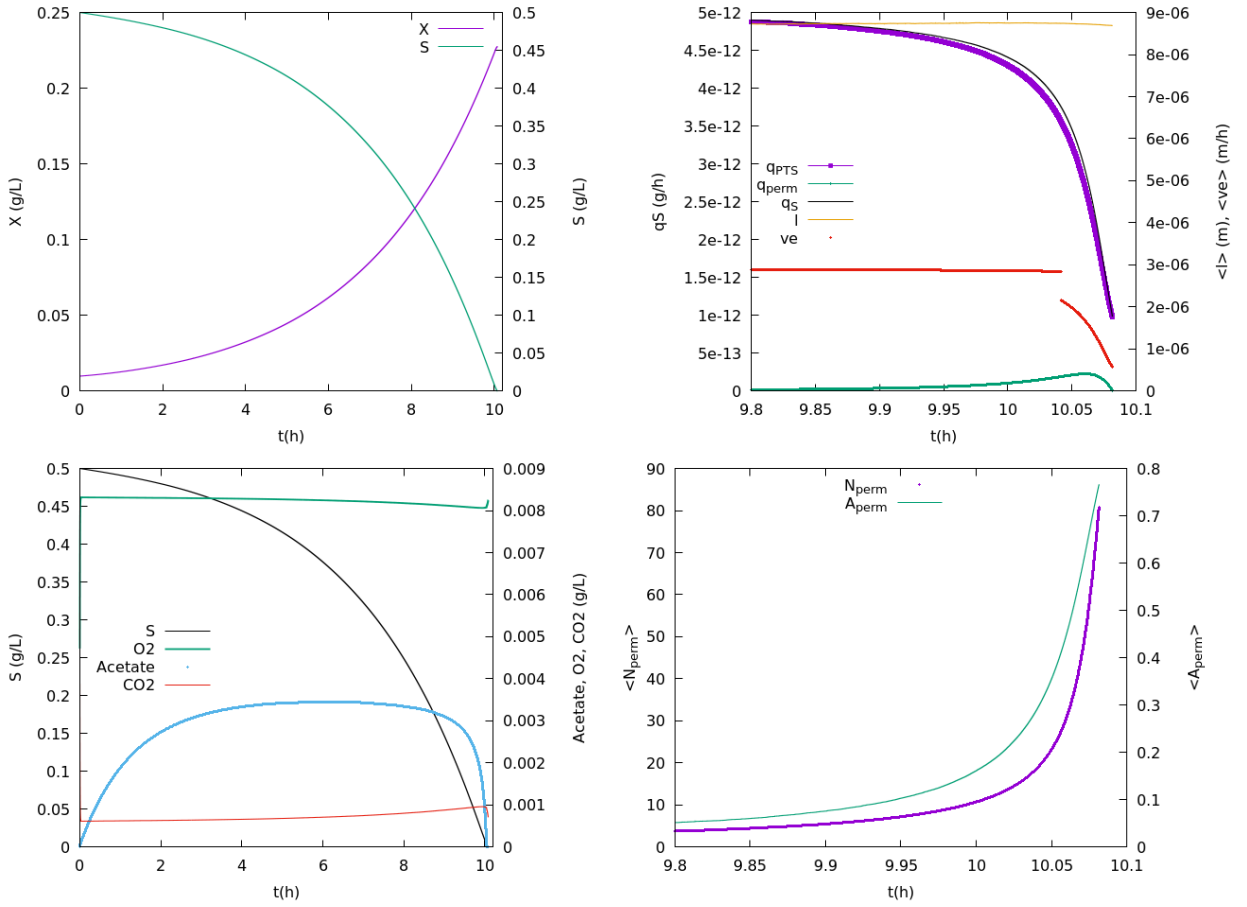

Figure 5.6 – Batch culture simulation, with initial biomass ≈ 0.001 g and substrate concentration 0.5 g. Top left: $X(t)$ and $S(t)$. Top right: q_S and its PTS/permease contributions as glucose shortage is looming. Bottom left: substrates and byproducts with respect to time. Bottom right: population–averaged N_p and A_p near substrate dearth.

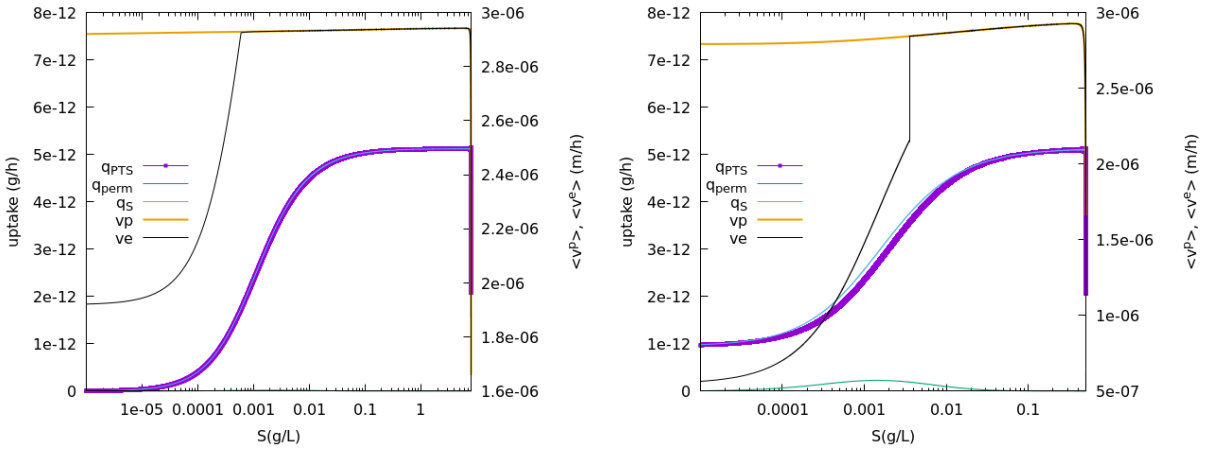

Figure 5.7 – Uptake and lengthening rate throughout a batch culture, represented as functions of S (log scale), involving two sets of initial conditions. Left: $X_0 \approx 0.05$ g/L, $S_0 = 8$ g/L. Right: $X_0 \approx 0.001$ g/L, $S_0 = 0.5$ g/L.

The takeaway is the massive dependence of the observed transient and terminal behaviours on the initial conditions: in particular, a comparison between two batch experiments for which $X(S = 0)$ varies 10–fold needs particular caution since the environmental offer will necessarily yield different responses from the affected organisms.

5.4 The glucose pulse experiment

In this section, the steady–state from section 4.4 is the starting point for a disruption consisting in the sudden injection of $1g$ of glucose in a steady–state continuous reactor. Sunya & al.'s [133] experiments have been carried out at a dilution rate of $0.15h^{-1}$ but the talk will mostly involve the trend in the response, quantitative fits being out of the scope of this present work. Uptakewise, this configuration should yield comparable qualitative responses with Lara & al.'s [65] BioScope set–up that is described in section 1.3.3.

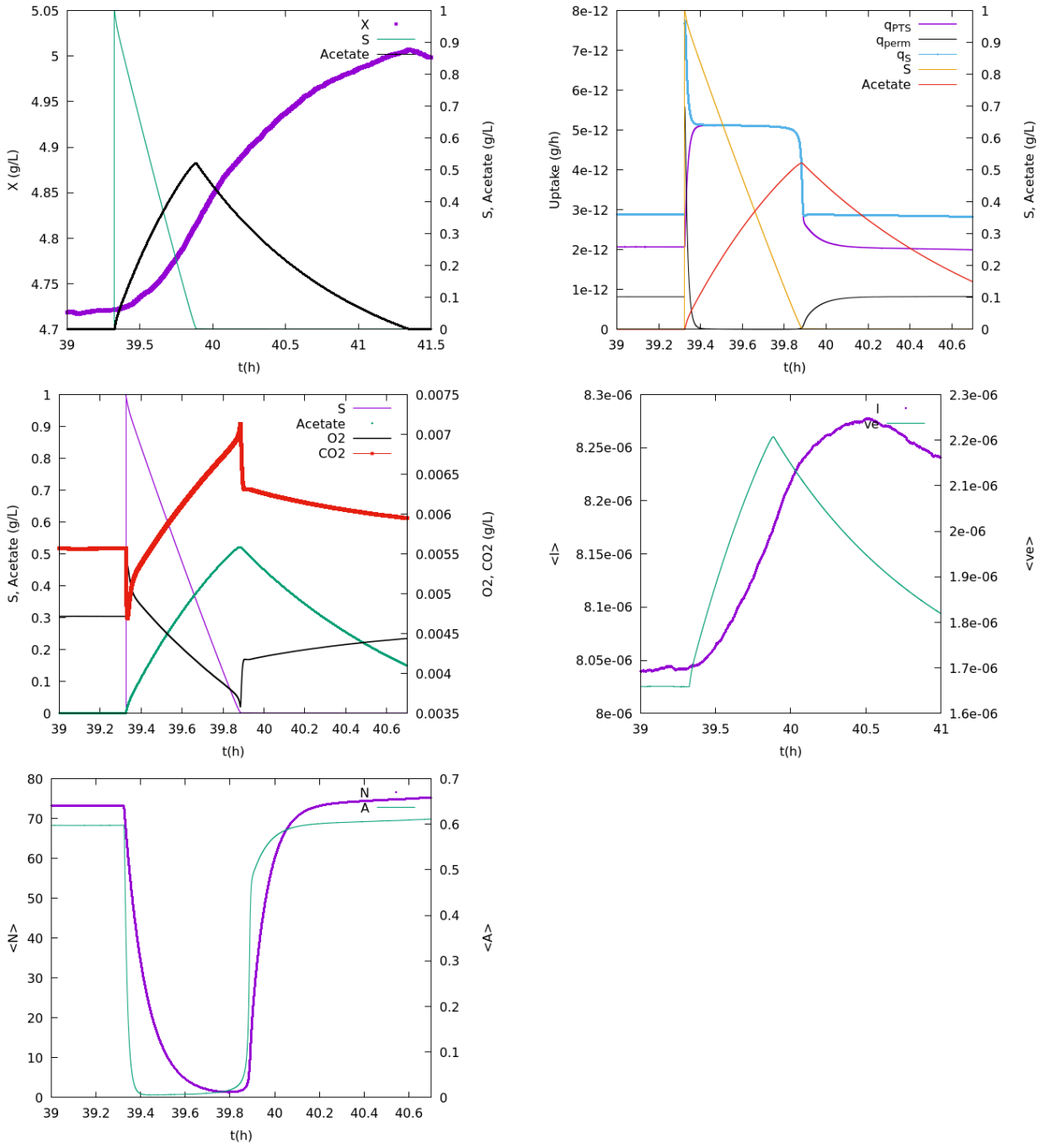

Figure 5.8 – Simulation of a glucose pulse in a steady–state chemostat ($D = 0.2h^{-1}$). Top left: X, S and *Acetate* shortly before and after the injection. Top right: the substrate uptake, looked at against the glucose / acetate concentration in the reactor. Centre left: the tracked chemical species over the course of the disruption. Centre right: the mean cell length and cell growth rate throughout the perturbation. Bottom left: mean N_p and A_p from steady–state to return to normal.

The sudden injection of substrate makes the cells' neighbourhood suddenly much richer and the

organisms which were not prepared for such a disruption burned some of the uptaken carbon into CO_2 while dissimilating a fraction of it into acetate.

The response breaks down into two parts. Quantitatively, the first phase of the perturbation (glucoe consumption and acetate production) lasts slightly less than 30 minutes, entailing the gross estimations:

- $1+0.5DS_f \approx 2.1$g of glucose are introduced in the system over the course of the perturbation.
- $D\int_t^{t+0.5h} S(s)ds \approx 0.01$g are washed out of the reactor.
- ≈ 0.75g are dissimilated into ≈ 0.5g acetate.
- ≈ 1.3g make 0.59g of new cells, while roughly $D\int_t^{t+0.5h} X(s)ds \approx 0.48$g of cells are washed out. X has hence increased by slightly more than 0.1g over the course of the experiment, what proves quite undetectable by the experimentalists given the measurement technique (Optical Density at 500 nm for instance).
- ≈ 0.41g of oxygen has been consumed to burn the uptaken carbon in total, translating into a 0.57g carbon dioxide production ($\approx$ 21 % of the total carbon).

As glucose becomes scarcer in the second phase of the response, the lately well–nourished bacteria are goaded to reconsume the available acetate in order to compensate the limiting glucose offer. Quantitatively, over the 1hour 30min–long period:

- $\int_t^{t+1.5} DAc(t)dt \approx 0.075$g of acetate are washed out of the fermenter. Hence, ≈ 0.43g are reconsumed to yield ≈ 0.25g of biotic output.
- $\int_t^{t+1.5} DX(t) \approx 1.5$g of biomass are eliminated from the reactor, meaning that roughly 1.7g of new cells have been produced in the meantime.
- $\int_t^{t+1.5} DS_f = 3.24$g of sugar are injected in the system, a negligible fraction of it being dragged out. Out of this raw glucose absorption, 1.46g of cells would result given the assumed ν_{12} metabolic ratio.

At the end of the day, the mass balance is in complete agreement with the above–mentioned set of biochemical reactions. Figure 5.8 sums up the time evolution of both cell–scale features (q_S, l, v^p, v^e, N_p and A_p) and reactor–scale quantities (X, S) which satisfy the mass conservation when $1g$ of glucose is injected in the system. A closer look at this dynamic evolution of both uptake and growth throughout the pulse is plotted on figure 5.9. q_p starts skyrocketing (clockwise) as S is set to $1g/L$, meaning that $S/(K_p + S) \approx 1$, what drives q_S accordingly. The apparent highest uptake rate overreaches the maximum batch value by roughly 50 %, in agreement with Lara & al.'s [65] experimental measurements. Then, as q_P starts hiking (counter–clockwise), A_p decreases under the weight of the $1 - q_P/q_{P_{\max}}$ factor contributing to its dynamics, and N_p follows suit due to the $K_S/(K_S + S) \approx 0$ contribution in its dynamics. As the perturbation proceeds, $q_S \approx q_P$ similarly to a batch culture until S plunges, what signals a necessary change in the cells' uptake strategy.

By the end of the disruption, the population's mean rate of anabolism has been raised by $\approx 1/3$, far from what is claimed in [133], partly because the authors probably mistook the uptake rate for the growth rate. Indeed, not all the assimilated carbon yields new cells, due to the necessary latency that comes with the adaption from a glucose–constrained extracellular medium to an unlimited environment. Uptakewise, the permease activation is not anticipated when S turns scarce because as the origin of the perturbations is this time external, the bacteria can only react to the fluctuations in substrate concentration which trigger is, before all, the draining circuit.

One shortcoming in the metabolic model manifests itself through a slight glitch in the oxygen and carbon dioxide concentration at the start of the perturbation, the lack of biological reactions being to blame. Indeed, being agreed that the equilibrium at the population scale has nothing to do with the equilibrium at the cell scale, the macroscopic marker of steady state (negligible residual acetate in a reactor) is in fact a consequence of the quantitative equivalence between acetate production from the elements which overuptake and consumption from the starving organisms. Whatever the cells' v^p, a substantial fraction of them are poised to burn some oxygen to set in motion either mode (3) or mode (4) of the so–configured biochemical network, both reactions requiring O_2. When S is ratcheted up, with little to no effect on v^p, a massive amount of sugar will be dissimilated into acetate (what should be expected) via the metabolic model's mode (2) ... which does not require O_2 at all. Hence, the big picture is an early slump in oxygen consumption in luxuriant culture conditions, what does not hinge upon any biological considerations. Albeit dubious, this biological oversimplification can be found in the pertaining literature (Pigou & al.'s [111] being an appropriate example) and has yet to be questioned to date.

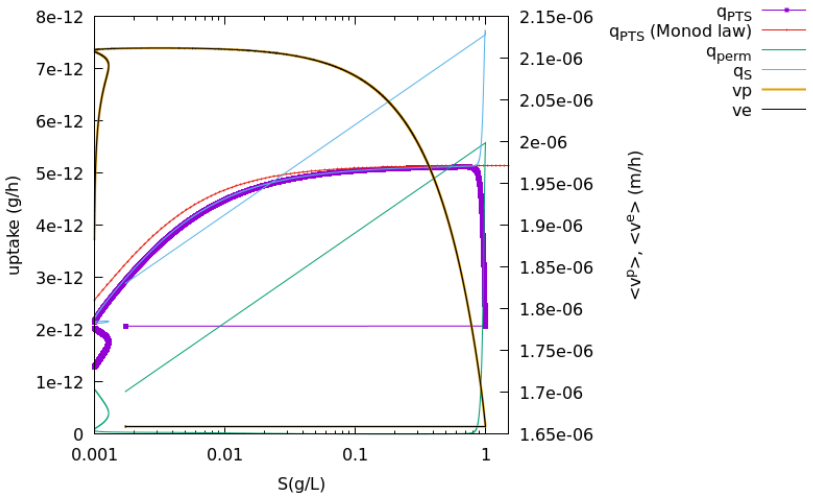

Figure 5.9 – q_S, v^e and v^p as the surplus S (log scale) is being consumed.

To conclude on the uptake profile, the amplitude of the permease response to the pulse is massively determined by the residual substate concentration at steady state, pointing to the crucial role of the dilution rate in setting the chemostat up: the lower the dilution rate, the measlier the residual glucose, the higher the permease induction, the more pronounced the response to a perturbation like a 1g glucose pulse. Indeed, the permease activation in a steady–state continuous reactor is the result of the individuals' strategy in the face of the external substrate offer, the latter being impacted by the user–set D and S_f parameters: in this context, the 50 % $q_{P_{\max}}$ over–uptake could be massively overreached should the reactor's dilution rate be feeble enough to make for a small S in comparison with the PTS range of application.

5.5 Influence of the micromixing time on the numerics

In section 3.3.1, a characteristic mixing time was introduced in the modelling to account for possible mass transfer limitations, what led to a distinction between physical and biochemical regime.

In this section, the consequence of an imperfect mixing on the cell uptake dynamics and overall metabolism is discussed. It is recalled that solving in $S_{@p}$ the (3.13) equation yields an expected glucose concentration at the cells' vicinity given their uptake capacity and the degree of mixing inside the fermenter. The t_m time constant models the efficiency of the micromixing inside one hydrodynamic compartment: the smoother the monomer supply to the bacteria, the smaller t_m. As a consequence, $S_{@p}$ is expected to be extremely close to S when t_m ranges between 10 to 100 ms, indicating a satisfying micromixing inside the system under consideration (only the organisms' uptake is playing a role in this case), whereas $S_{@p}$ can be massively hampered by the hydrodynamics should t_m approach 1s. In the latter case, the cells can be considered in permanent imbalance, having to compensate their subpar glucose uptake with acetate absorption at certain times, the consequential slump in q_S triggering a sudden excess in S they take advantage of by excreting acetate until the next dearth.

Figure 5.10 reports these claims in a test case consisting in setting a chemostat to equilibrium (D = 0.2h^{-1}) with t_m = 1s. It is visible from the top left plot that the variation in biomass is around 0.5 % of the overall value over the $[6.5 \times D^{-1}, 10 \times D^{-1}]$ time interval, what exceeds the method's intrinsic noise ($\approx$ 0.2 %) by a too significant margin to attribute these oscillations to the MC scheme's accuracy. The mean uptake rate (bottom plot) exhibits a 0.65 % oscillation around its time average over the same window, these variations playing a conspicuous role in the acetate concentration in the reactor, the over–uptake going with acetate excretion by virtue of the metabolic model's equation (3). On the top right plot that makes for a close–up of the time evolution of both q_P and q_p with respect to time, the respective fluctuations in comparison with the time averages (1 % and 0.46 %) illustrate the imbalance that characterises a population which individuals face a continual glucose gradient.

Figure 5.11 testifies to the influence of this never–ending imbalance on the cells' lengthening rate. The left–hand side plot, which depicts the substantially fluctuating acetate concentration (up to 0.002 g/L) in the system for t_m = 1s, the organisms' average v^p is shown varying in a 0.5 % range around its equilibrium value, the oscillations in v^e being even larger in proportion to the time average (0.6 %). On the right–hand side plot that features two acetate upticks only (the amplitude thereof remaining below the 10^{-4} g/L level), the respective variations amount to 0.01 % and 0.08 % and could be attributed to the MC algorithm's precision. Both the acetate profile and the lengthening rate look much less chaotic, what heads to the conclusion that an imperfect degree of mixing in a reactor is perceived as a constant perturbation by the suspended organisms.

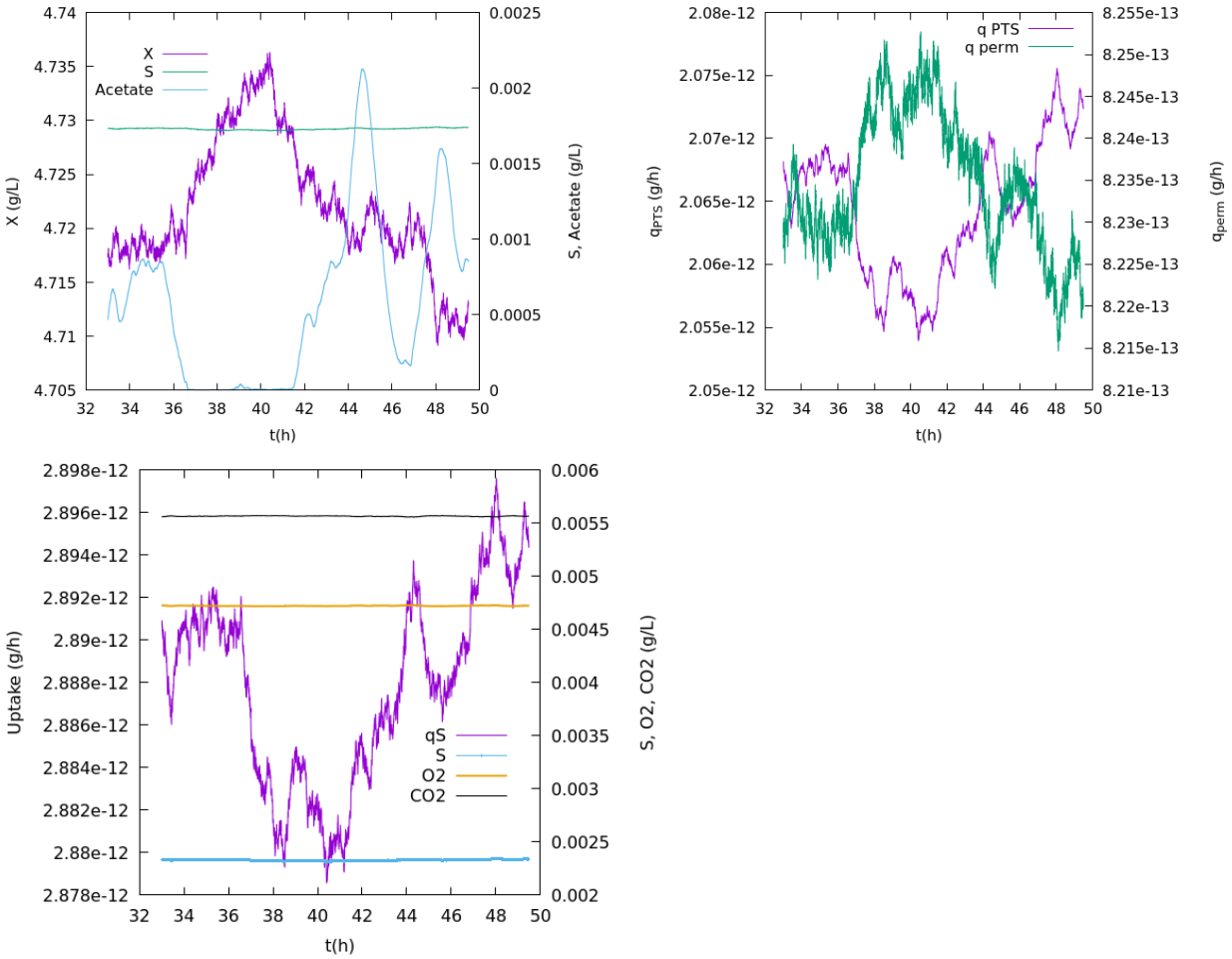

Figure 5.10 – Monte–Carlo simulation of a chemostat convergence to steady–state ($D = 0.2\text{h}^{-1}$), in case $t_m = 1s$. Top left: biomass, glucose and acetate with respect to time. Top right: close–up of the population–averaged q_P and q_p. Bottom: mean glucose uptake rate and liquid phase sugar, oxygen and carbon dioxide concentrations.

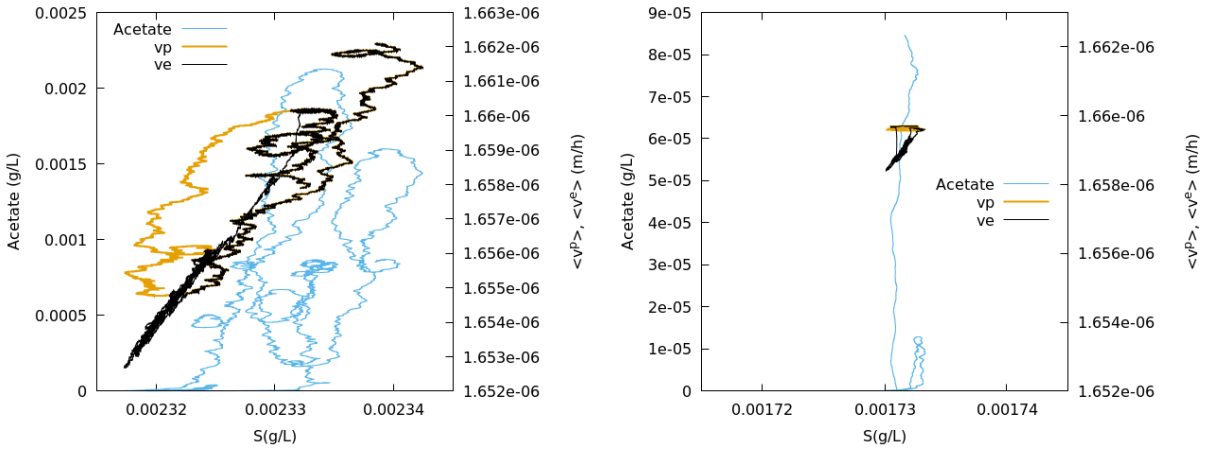

Figure 5.11 – Monte–Carlo simulation of a chemostat convergence to steady–state ($D = 0.2\text{h}^{-1}$), in case $t_m = 1s$ (left plot) and $t_m = 50ms$ (right graph). Evolution of the acetate concentration in the fermenter (blue lines), and the population–averaged potential (yellow curves) and actual (black lines) lengthening rate as functions of S.

All in all, section 3.3.1's modelling of the imperfect degree of mixing in one compartment provides a first overview of the hydrodynamics' influence on the bacteria's response to ever–fluctuating concentrations when cultivated in 10^5L–industrial fermenters. Considering that the amplitude of the substrate gradients is much more significant in the latter case than in a 1L–laboratory reactor, incorporating such a hydrodynamic refinement has to be considered a mandatory step towards the simulation of biological population dynamics in industrial–scale tanks.

5.6 The dilution rate step–up experiment

5.6.1 $D = 0.2\ \mathbf{h}^{-1} \rightarrow 0.3\ \mathbf{h}^{-1}$

From the steady state that made up the talk of section 4.4, a hike in the dilution rate from $0.2h^{-1}$ to $0.3h^{-1}$ is enforced. In this case, the reader has to cautiously consider one of the pillars of the Monte–Carlo algorithm. A cell's residence time in the reactor is sampled at its inception from a $\mathcal{E}(D)$ probability law, meaning it will be determined by the value of D at this very moment. It is also recalled that the removal of a cell from the reactor results from the comparison between its age and its sampled–at–birth residence time. If D was suddenly ratcheted up to a certain D', signalling a brutal washout in the culture, the residence time of a group of extant cells would be on average larger than the lifetimes the new culture condition would offer. As a consequence, many organisms would stay in the system longer than they should and the big picture would be a non–physical initial gain in biomass at transient state. In order to correlate the organisms' residence times to the newly effective dilution rate, a scalar must be redrawn for each MC particle from the new $\mathcal{E}(D')$ exponential law.

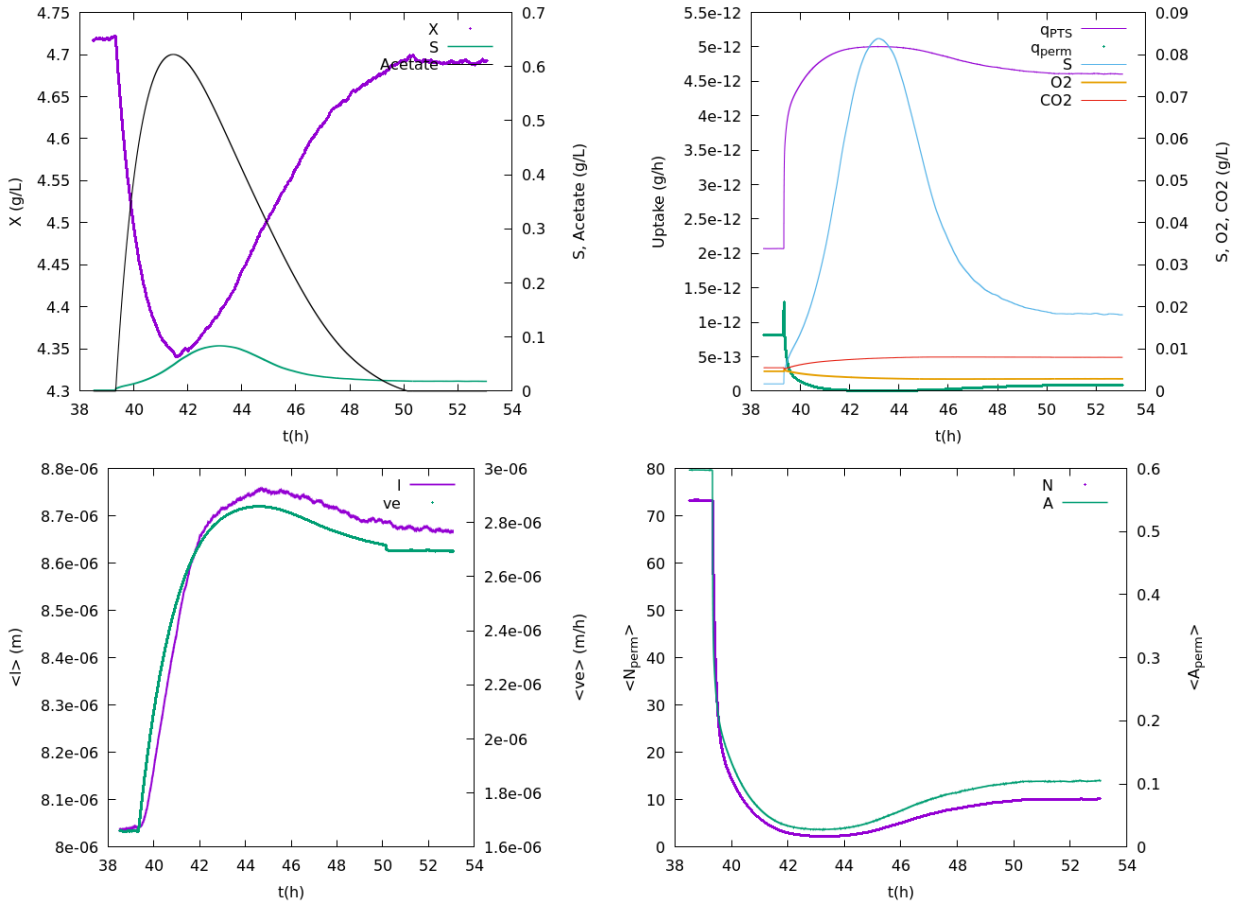

Figure 5.12 – Simulation of an instantaneous hike in dilution rate ($D = 0.2h^{-1} \rightarrow 0.3h^{-1}$). Top left: X, S and *Acetate* as functions of time. Top right: uptake compared with the time evolution of S, O_2 and CO_2 throught the perturbation. Bottom left: mean length and rate of anabolism towards the new steady state. Bottom right: mean permease number and permease activity with respect to time.

The population is tracked over the course of 13.75 h following the perturbation, that is slightly more than 4 times the longest characteristic time featuring in the modelling ($1/D$). In other words a new equilibrium is manifesting itself and can be investigated. The evolution towards the new equilibrium features in figure 5.12. Given the residual substrate concentration in the reactor (0.01814 g/L > K_P), the PTS system has to account for the overwhelming majority of the glucose uptake and reduce its permease counterpart to almost nothing. This is indeed predicted by the model and the numerical values could also be grossly be evaluated as in section 4.4. A massive acetate excretion goes with the disruption while the bacteria's lengthening rate adapts to the suddenly glucose–enriched culture medium.

As the reactor's renewal time has consequentially dropped, the fraction of the population with the smaller lifetime in the system is washed out of the fermenter, leaving the cells with the larger sampled residence time only. These extant cells have been able to raise their v^p at will, prompting a

shift to the right in the rate of anabolism NDF, in line with the observed uptick in the distribution's first moment featuring on figure 5.12's bottom left plot.

For the sake of clarity, a comparison between the $D = 0.2$ h^{-1} and $D = 0.3$ h^{-1} number density functions in the 5 inner coordinates is provided in figure 5.13. As was visible from the mean values, the NDF in q_P is shifted to the right as D is increased, to the extent that S turns abundant. The shift to the left of the N_p and A_p NDFs is not surprising either, considering the little interest from the cells' perspective in sustaining a scavenging uptake mechanism when S is plentiful in the CSTR.

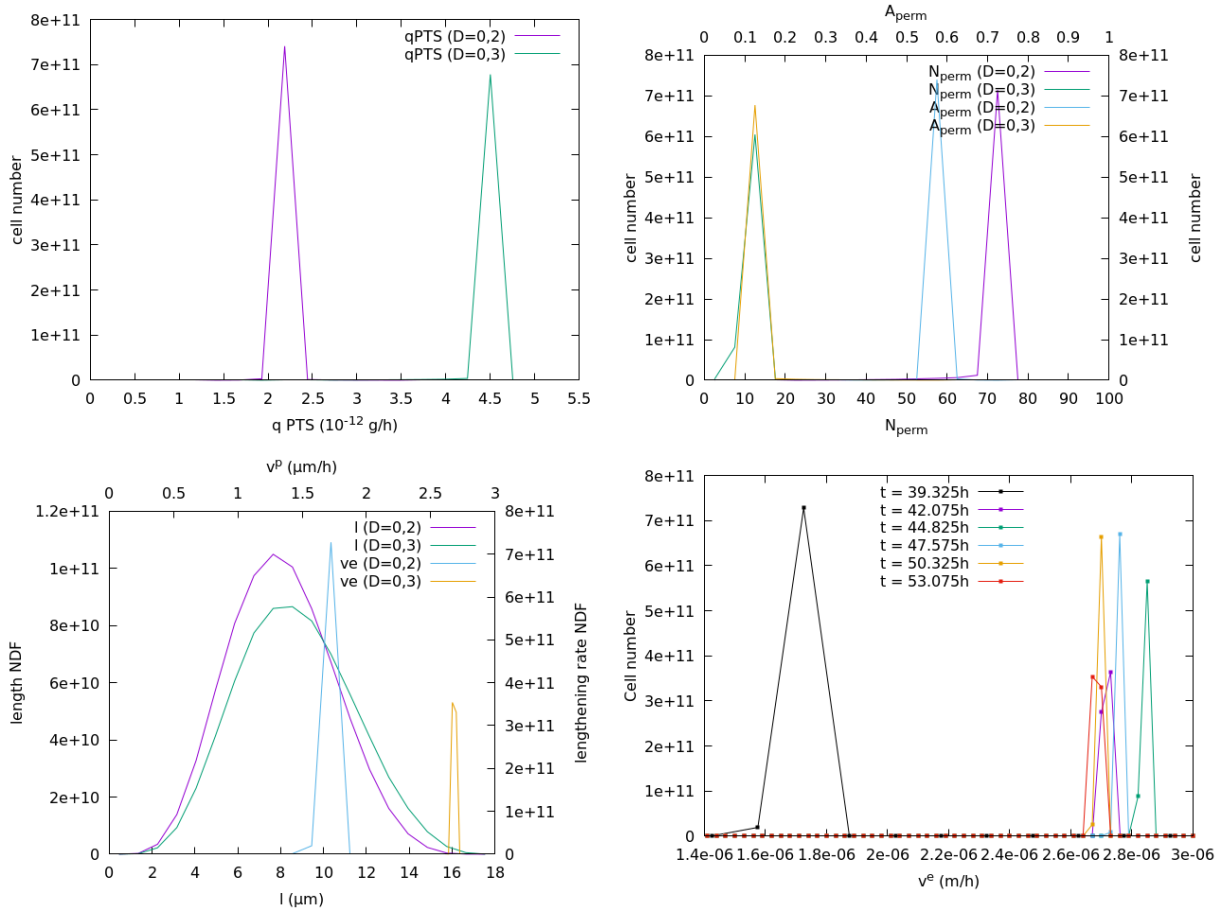

Figure 5.13 – Equilibrium NDFs in all 5 inner coordinates of (3.11) for $D = 0.2$ h^{-1} and $D = 0.3$ h^{-1}, in: top left: q_P; top right: N_p and A_p; bottom left: l and v^p. Bottom right: distribution in v^e towards the steady–state associated to $D = 0.3$ h^{-1}, as the disruption starts at 39.325 h.

From figure 5.13's bottom right plot, it is visible that the population's effective anabolic rate does not follow the increased draining in the immediate aftermath of the dilution rate upshift: it remains insufficient to offset the washout until roughly $t = 42$h, what is consistent with the $X(t)$ profile that features on figure 5.12's top left plot. The plotted lengthening rate NDFs suggest that in comparison with the D' induced–equilibrium (red linespoints, at approximately $t = 53$h), the population catches up this biomass loss over the next 8h due to a luxuriant external supply. Interestingly, its meeting with the v^e steady–state profile does not proceed due to the introduction

of an averaged population growth rate (contrary to equation (1.32) involving a $\mu^\star$ real number, page 48), but the $n(v^e)$ distribution is attracted to its newly–induced steady–state through the slump in substrate uptake that goes with the continuous decrease of S following the said 8h–long period of over–uptake.

The transient $X(t)$ pattern has been experimentally evidenced in the literature, such data having been provided for instance by Yun & al. in [150]. The biologists' measurements testify to a return to normal after roughly 10 h following a D = 0.2 h^{-1} → 0.6 h^{-1} upshift, the apparent "specific growth rate" countering the challenging draining rate in less than 5 hours as can be seen from figure 5.14. This means that as long as the bacteria are capable of withstanding the user–defined dilution rate, the rate of virtually all the biological processes occurring in the reactor, uptake included, is completely dictated by D^{-1}.

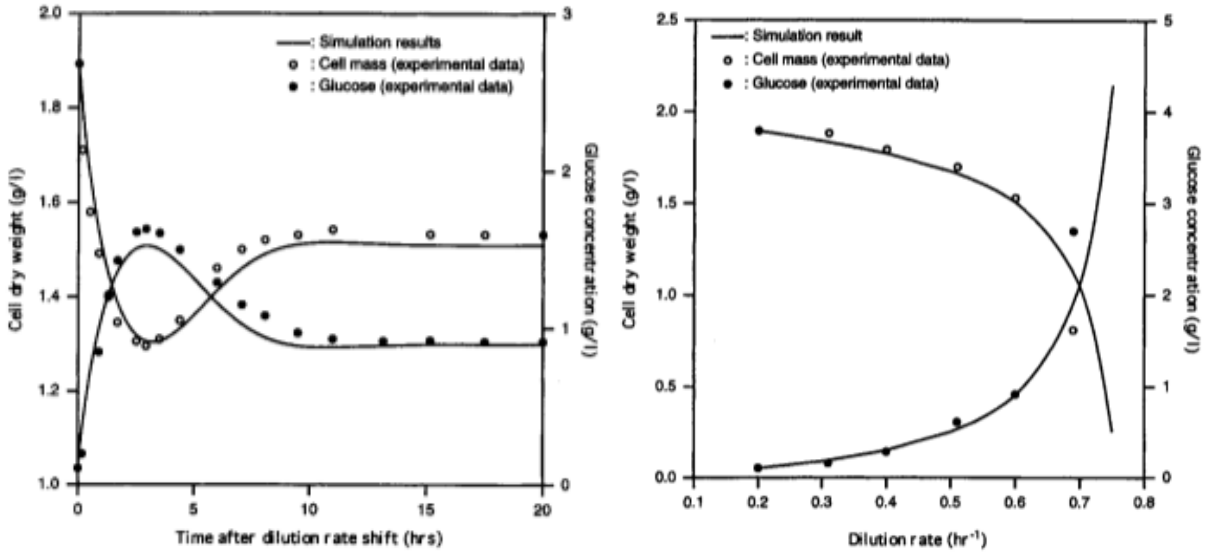

Figure 5.14 – Left: experimental measurements of a D = 0.2 h^{-1} → 0.6 h^{-1} dilution rate increase in an *E. Coli* K–12 continuous culture, which maximal batch growth rate is estimated around 0.803 h^{-1}. Right: determination of the *E. Coli* population's highest sustainable dilution rate in an open fermenter. Borrowed from Yun & al.'s [150].

One last remark from Yun & al.'s experimental data relates to the measured specific growth rate as the population is heading to the steady state following the dilution rate step–up. It is assumed that X linearly grows from 1.3 g/L at $t = 3h$ to 1.45 g/L at $t = 6h$, the washout draining roughly $3 \times 0.6 \times 1.375 = 2.475$ g of cells. If ≈ 2.625 g of cells have erupted in the meantime, the population's specific growth rate (0.875 h^{-1}) overreaches any batch culture estimation for the so–called maximal growth rate $\mu_{\max}$ (≈ 0.7 h^{-1} in Taymaz–Nikerel & al.'s [136] but estimated around 0.803 h^{-1} in Yun & al.'s [150]).

The latter claim could be explained by the very definition of a maximal growth rate: as a population–averaged quantity, it could allow for organisms which instantaneous anabolic rate v^e is higher than the supposedly impassable maximal batch value. The lengthening rate being distributed among the individuals comprising the observed population, it has no reason to be bounded

by a population–scale datum. In this context, following a strong perturbation in an open fermenter, the fraction of the organisms exhibiting the $v^e > v_p^{\max}$ property would dominate in the system at the expense of their less active counterparts. The $v^e > v_p^{\max}$ feature could not be maintained over time, otherwise an exponentially growing population would stray from its dilution rate–imposed equilibrium, what is not observed in the experiments. Presumably, a cell which v^e overreaches $v_p^{\max}$ would exhaust most of its resources to do so and be likely to give birth to one (or two) less active daughter–cell(s) when it divides. This would be tantamount to setting a more asymmetric $P_{v^p}(v^p, v^{p'})$ redistribution kernel, without hampering the rest of the modelling. This is in agreement with experimental measurements provided by Yasuda [148] pointing to the interdivision time distribution of a constantly–fed population in a perfused reactor: even though some of the extant organisms are thriving, they can give birth to much less active daughter–cells which interdivision time is twice the mother–cell's. The data are provided on figure 5.15 for the sake of clarity.

Uptakewise, that scenario would make it necessary to release the $q_S \leq q_{P_{\max}}$ long–term constraint, for instance through the mobilisation of another glucose assimilation system which preferred field of action overlaps the PTS's. Indeed, Ferenci [24] noted in his 1990s research works the existence of more than two dedicated membrane transporters, the so–called symporter playing a role in the overall substrate uptake should S be high enough at the cell's vicinity.

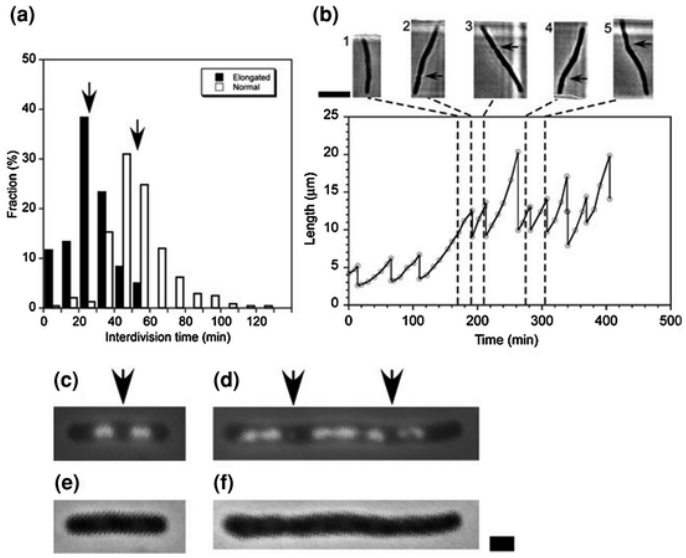

Figure 5.15 – Interdivision time distribution of an "elongated cell" population, which individuals are constantly fed with glucose. The mean interdivision time is roughly 25 minutes, but the measurements span a wide value range (from 0 to 50 minutes), hinting at a heterogeneous distribution of the cells' fitness from generation n to $n+1$. Borrowed from [148].

One last remark from Yasuda's experimental data pertains to the difference in size allocation among normal and elongated cells. It seems that the size redistribution is more asymmetric among the longer mother–cells which happen to be the fitter in the population, what could be explained by the little they are given to settle down and consolidate their lengthening poles before ceasing to exist. This observation hints at the conclusion that not only is the daugther–cells' growth potential redistributed over a wide range if their mother is thriving, but the fraction of the dividing organism they are handed down is also massively impacted by the mother–cell's fitness: in other words, it is

reasonable to conjecture that $P_l(l,l')$ is also a function of $v^{p'}$.

5.6.2 $D = 0.2\ \mathbf{h}^{-1} \rightarrow 0.32\ \mathbf{h}^{-1}$

It would seem that this perturbation is quite similar to the change in dilution rate from 0.2 h^{-1} to 0.3 h^{-1}, entailing extremely resembling transient states. If the opening washout phase is taking place as expected, draining a larger (11 %) chunk of the total biomass though, the return to equilibrium obeys a different dynamics from the previous case to the extent that limiting the glucose uptake to a constant $q_{P_{\max}}$ prevents the maximal growth potential from crossing the $q_{P_{\max}}$ / (Y_{SX} K_{conv}) barrier. More precisely, if the washout still operates at an exponential rate, the gain in biomass reads:

$$K_{conv} \int \dot{l} n(t,\xi) d\xi = K_{conv} \iint v^e \left(1 - \frac{l}{\bar{l}}\right)^{\kappa} n(t,l,v^e) dl dv^e$$

Even if v^e was equal to its maximum value ($q_{P_{\max}}$ / (Y_{SX} K_{conv})) for each cell, the biomass balance in the CSTR would read:

$$\frac{dX}{dt} = -DX + \frac{q_{P_{\max}}}{Y_{SX}} \int_0^{\bar{l}} \left(1 - \frac{l}{\bar{l}}\right)^{\kappa} n(t,l) dl$$

By virtue of the relationship between cell number and biomass ($X = \int_0^l n(t,l) dl K_{conv} \langle l \rangle$):

$$\frac{dX}{dt} = -DX + \frac{q_{P_{\max}}}{Y_{SX}} \int_0^{l} \left(1 - \frac{l}{\bar{l}}\right)^{\kappa} \frac{n(t,l)}{\int_0^{\bar{l}} n(t,l) dl} \frac{X}{K_{conv} \langle l \rangle} dl$$

and finally, all it would take to conclude on the rate of exponential growth would lie in the calculation of the right–hand side's integral. Once the length profile has become self–similar, this computation would yield:

$$\int_0^{l} \left(1 - \frac{l}{\bar{l}}\right)^{\kappa} \frac{n(t,l)}{\int_0^{\bar{l}} n(t,l) dl} dl \approx 0.965 \text{ and } \langle l \rangle \approx 8.791 \cdot 10^{-6} m$$

and the right–hand side's last term would amount to, roughly, 0.324 X. As a consequence, due to the introduction of $q_{P_{\max}}$ in the model, skyrocketing the reactor's dilution rate to demanding growth conditions leads to a somewhat unsatisfying observation: after a $5D^{-1}$ time interval, the biotic phase is miles from the expected steady state that would emerge from a real–life experiment. Also, from a quantitative point of view, the highest sustainable dilution rate is less than the maximal batch growth rate due to the fitness variability observed in a population and experimentally quantified (cf. Yasuda's measurements featuring on figure 5.15): a significant chunk of little active cells are poised to be washed out at each generation whatever the culture conditions, making the more active cells' task to stabilise the open reactor impossible.

The hindering feature of $q_{P_{\max}}$ impacts the transient NDF in v^e too, its shape being a $\delta_{q_{P_{\max}}/(Y_{SX}K_{conv})}$ distribution and self–similar over the course of the return to normal. Also, $n(t,v^e)$ being dirac–shaped tends to stabilise the transient size profile to the extent that $\dot{l}$ is a function of l only, hence identical for organisms of equal length. This claim is illustrated on figure 5.17.

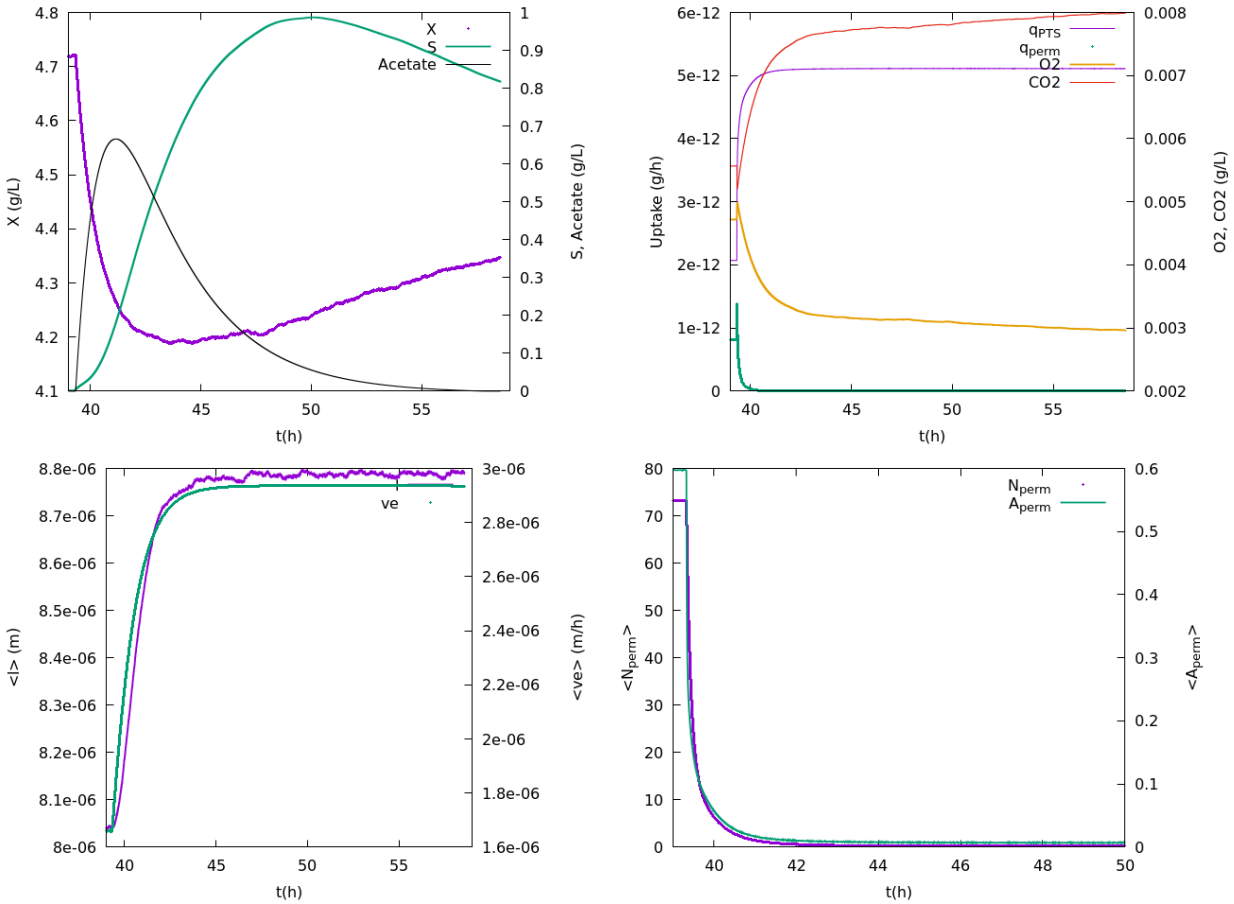

Figure 5.16 – Transient states associated to a $D = 0.2\ \mathrm{h}^{-1} \rightarrow 0.32\ \mathrm{h}^{-1}$ disruption. Top left: biomass, glucose and acetate shortly before and after the perturbation. Top right: population–averaged uptake and gas with respect to time. Bottom left: mean length and lengthening capacity as functions of time. Bottom right: time evolution of the mean permease number and permease activity.

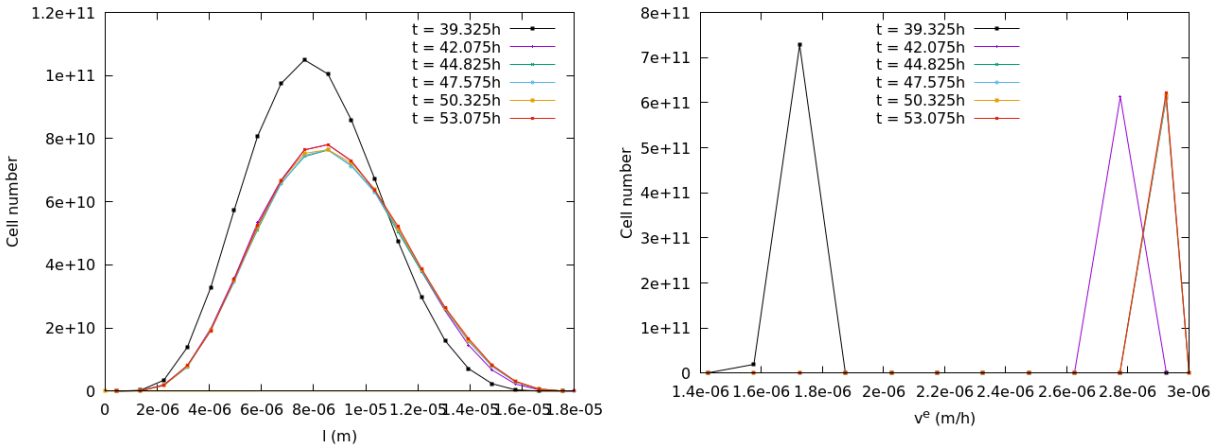

Figure 5.17 – Comparison in length (left) and rate of anabolism (right) NDFs following a $D = 0.2$ $h^{-1} \rightarrow 0.32$ h^{-1} dilution rate increase.

It is worth a mention that in equation (3.12), the bacteria's maximal uptake rate $q_{P_{\max}}$ has been picked in order to fit the order of magnitude of the 1g glucose pulse's assimilation with Sunya & al.'s experiment operating at a similar dilution rate. It is possible though that the extant cells do not uptake the neighbouring glucose at the $q_{P_{\max}}$ rate to the extent that this liquid–cell transfer is energy– hence oxygen–consuming and the dissolved oxygen dwindles to almost 0 throughout the pulse according to [133]. This refinement missing in the (1)–(4) metabolic model, the value of $q_{P_{\max}}$ may be underestimated, explaining the seemingly measly maximal growth rate. This has no impact on the qualitative trends the (3.11) model is reproducing though.

5.7 The dilution rate step–down experiment

Starting once again from section 4.4's data, a slump in the dilution rate from $0.2h^{-1}$ to $0.05h^{-1}$ is performed in this paragraph, the numerics featuring on figure 5.18. As S becomes scarce, the permease system attempts to offset the lack of efficiency of the substrate-specific PTS mechanism. Nevertheless, it does not prevent a 5–time decrease in the actual growth rate in starving conditions.

A closer look at figure 5.18 might deceive the reader who would hastily interpret from the numerics that acetate is transiently produced in an "glucose–empty" chemostat. This production is an statistical artifact that comes along with a proportional decrease in biomass due to the randomness in the cells' residence time, the Monte–Carlo algorithm being the prime culprit in this glitch. More precisely the removal of, say, 0.1 % of the extant cells within a short time interval makes the same substrate offer slightly less restrictive, prompting some of the individuals to process the suddenly more profuse glucose into acetate, the converse phenomenon occuring when X increases due to the statistical noise in the Monte–Carlo treatment of the biotic phase.

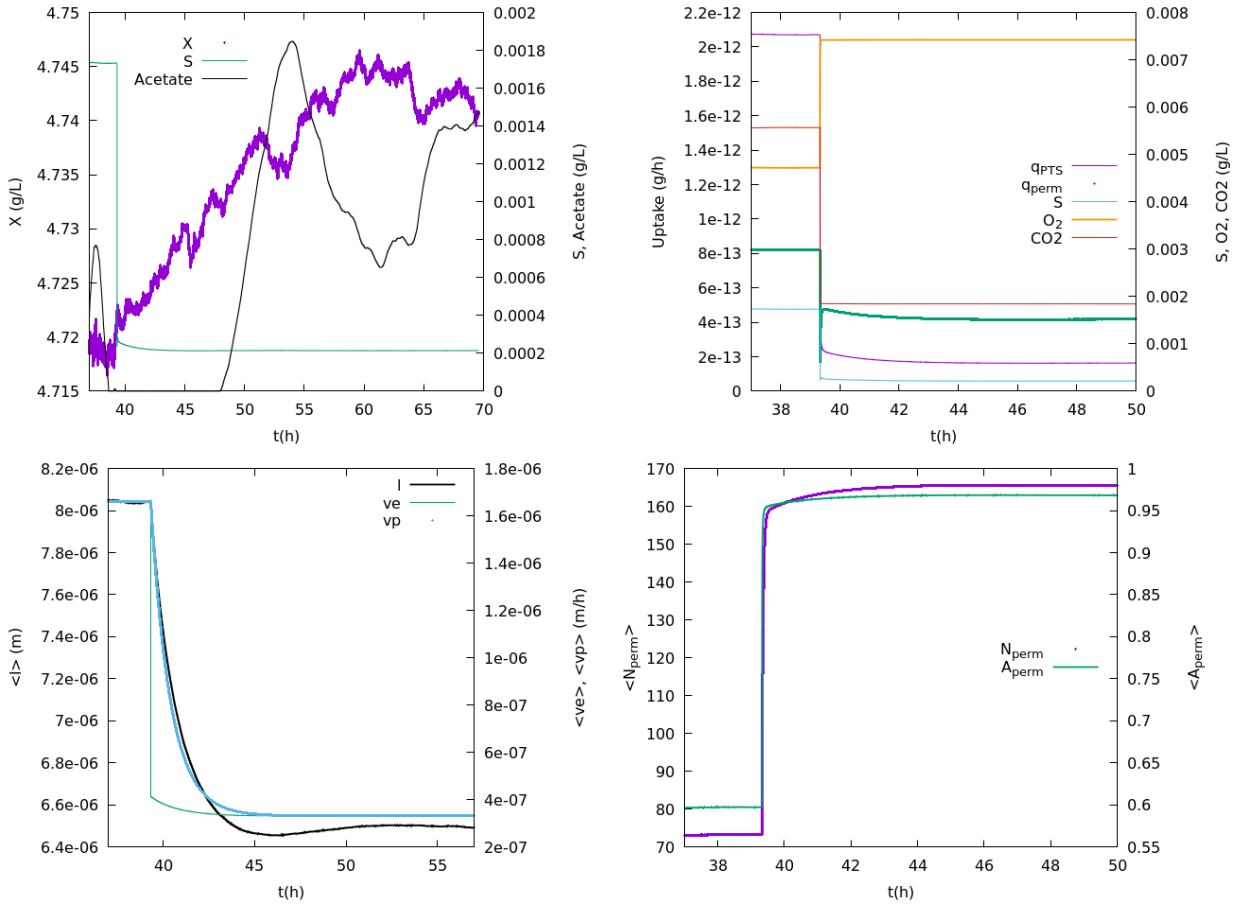

Figure 5.18 – Simulation of a sudden decrease in dilution rate $D = 0.2h^{-1} \rightarrow 0.05h^{-1}$. Top left: X, S, and *Acetate* in terms of t. Top right: time evolution of the mean uptake and the by–products. Bottom left: mean length and lengthening rate as the feed is reduced. Bottom right: mean permease number and permease activity while the perturbation proceeds.

A similar comparison as section 5.7's between $D = 0.2$ h^{-1} and $D = 0.3$ h^{-1} steady–state NDFs illustrates the seizure of power by the permease system when the residual substrate is too scarce to allow the PTS a noticeable contribution to the organisms' glucose uptake. A_P drops to almost 0 for all cells and both NDFs in N_p and A_p exhibit a substantial shift to the right that is explained by the necessary need for every organism to scavenge the reactor for the elusive substrate.

A consequence of these assessments is the size distribution's shift to the left that can be understood by the treatment of the division procedure in the Monte–Carlo code. To illustrate this claim, consider a cell of length $l = l_c$ $(= 11\mu m)$ at a certain time t_0, which $\gamma(l)$ is equal to 1 for. Suppose it is at equilibrium with its environment. If, in the Monte–Carlo code, the rupture is triggered by the $1 - \exp(-\gamma(l)\delta t) < u$ inequality for a randomly picked u, it is statistically likely that the bacterium will divide over a 1h time window. Its v^e being constant, the cell's length at

$t_0 + 1$ will be given by the equation:

$$\left.\begin{array}{c} \frac{dl}{dt} = v^e \left(1 - \frac{l}{\bar{l}}\right)^{0.05} \\ l(t_0) = l_c \end{array}\right\} \Leftrightarrow \left(1 - \frac{l_c}{\bar{l}}\right)^{0.95} - \left(1 - \frac{l(t_0+1)}{\bar{l}}\right)^{0.95} = \frac{0.95 v^e}{\bar{l}}$$

and a numerical application for $v^e \approx 3.334 \cdot 10^{-7} m \cdot \mathrm{h}^{-1}$ leads to $l(t_0+1) \approx 11.34 \mu m$. In other words, v^e is so small in a substrateless chemostat that one cannot expect the longer extant organisms to substantially overreach l_c at division. A blatant illustration of this claim features on figure 5.19.

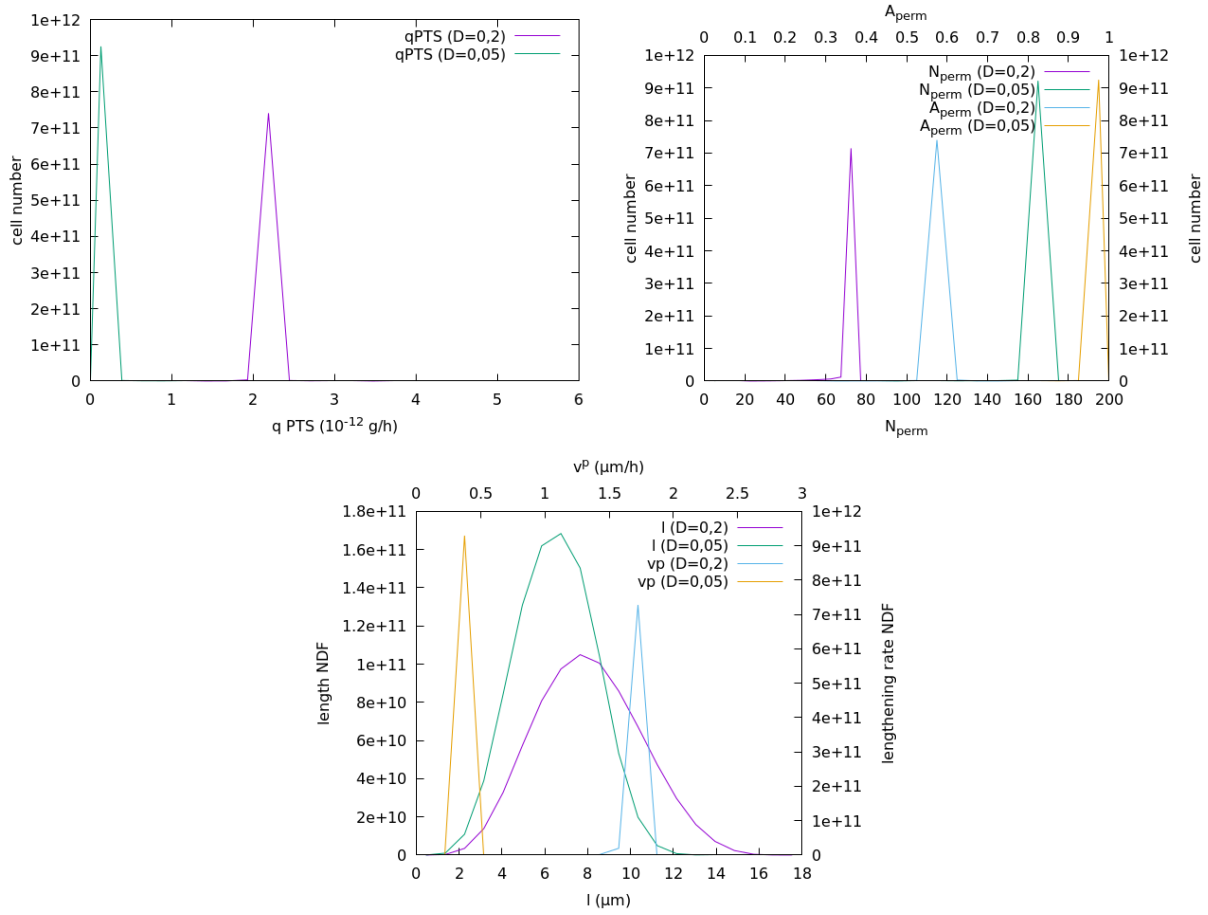

Figure 5.19 – Comparison between D = 0.2 h^{-1} and D = 0.05 h^{-1} steady–state NDFs in all state variables of (3.11). Top left: q_P. Top right: N_p and A_p. Bottom: l and v^p.

The shape of the length distribution finds its root in the very small probability that a cell reaches a significantly larger length than l_c before dividing if D is responsible for a growth–limiting residual glucose concentration. The lower S in the reactor, the lower the individuals' lengthening rate, crippling the bacteria's growth beyond the presumed standard length at rupture l_c.

5.8 The dilution rate slope experiment

In this paragraph, a continuous change in dilution rate aims at mimicking Nakhu & al.'s experiment [97] in a so–called accelerostat. This constant increase in D makes the use of the Monte–Carlo algorithm highly unpractical considering the residence time uptake procedure. Indeed, if D is incremented at each time step, a new residence time would have to be sampled for each organism and only the elements which residence time is less than the said time step would be washed out of the reactor. To avoid this hiccup, the hybrid FV–QMOM code has to be run for this purpose, with no impact on the section's conclusions.

A continuous increase in dilution rate from 0.2 h^{-1} to roughly 0.3 h^{-1} over the course of 8.25 h is performed, the numerics being reported on figure 5.21. As the feed is ramped up, the extant cells are able to adapt their lengthening rate at will until they are dragged by the flow. Acetate reaches a maximum despite the continuous overflow of substrate that turns the metabolic model's reaction (3) on. Indeed, as v^p adjusts to the ever more profuse environment, the uptake surplus wanes while the washout waxes until the acetate outflow ends up overtopping the production by the biological phase.

Because of the first–order dynamics ushering the uptake–related variables, the medium is always one step ahead of the biotic phase, contradicting the experiment presented by Nahku & al. [97], the data thereof hinting at an instantaneous adaption of the growth rate to the dilution rate until the latter turns unsustainable to the organisms, in other words a biomass maintenance for hours of dilution acceleration.

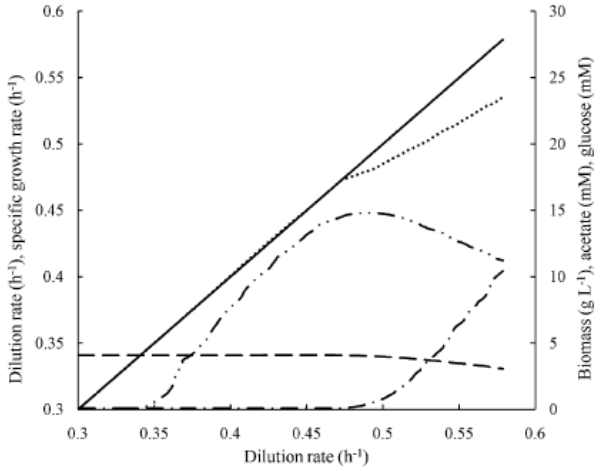

Figure 5.20 – Experimental measurements of biomass (dashed lines), glucose (dashed, single–dotted lines) and acetate (dashed, double–dotted lines) in an accelerostat. D increases from $0.3h^{-1}$ to $0.6h^{-1}$ at a $0.01h^{-2}$ rate. The population's growth rate (dots) starts following the dilution rate (lines) until the latter becomes unsustainable to allow the biotic phase to maintain in the reactor. Borrowed from [97]

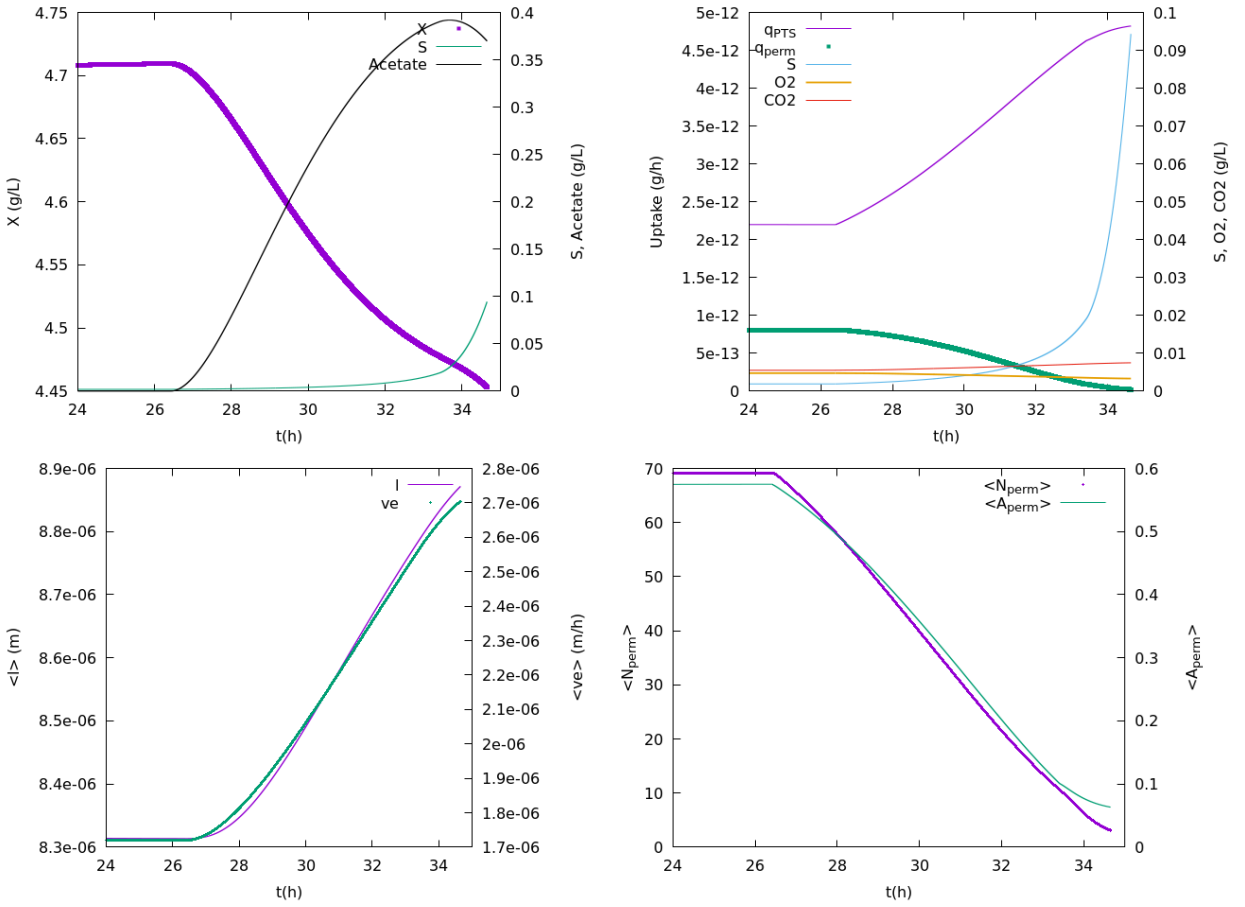

Figure 5.21 – FV-QMOM simulation of a slope in dilution rate from 0.2 h^{-1} at $t = 27$ h, at a constant 0.0121 h^{-2} rate. Top left: biomass and glucose concentrations. Top right: uptake and by–products. Bottom left: length and lengthening rate. Bottom right: permease number and permease activity.

Such a behaviour is beyond the reach of (3.11), mostly because it is geared towards strong perturbations the biotic phase responds to with an observable latency. Indeed, if a stable chemostat is disrupted by a massive glucose pulse, the translation in terms of growth rate is measured in hours. However the progressive acceleration of the draining circuit is viewed by the organisms as a small perturbation they are most likely capable to deal with at first through a mobilisation of some cytoplasmic resources. This feature could be included in the model via the addition of a stock variable with the aim of providing a contingency supply for a short period of time if required, the stock's feeding coming for instance from the diversion of some over–uptaken glucose in case of abundant supply.

Chapter 6

Conclusion

Understanding bacteria population dynamics requires complementary standpoints with the aim of formulating evolution laws which numerical simulation allows for qualitative and quantitative comparisons with experimental measurements. In this context, this work has delved into different paths geared towards a finer understanding of biological PBEs.

Mathematical advances

From an analysis point of view, most of the results published in the literature were consecrating a bounded rupture function with little care to the fictitious weight that lies on non–physical particle sizes, what has been corrected in part 2. The solution to a size–structured model has been proven continuous with respect to both time and the inner coordinate, and it has been demonstrated that the Malthus eigenfunction associated to this integro–differential equation is $\mathcal{C}^1$ under reasonable modelling assumptions, testifying to the well–posedness of the problem. A protracted debate concerning the relationship between the mean interdivision time in a reactor and the macroscopic time scale of the renewal has also been wrapped up and the following analytical formulae can be tested against microbiology experiments:

$$\langle \tau \rangle \leq \frac{\ln(2)}{D}$$
$$\langle a \rangle + \langle \tau \rangle = \frac{1}{D}$$
$$\langle a^2 \rangle + \langle \tau^2 \rangle = \frac{2}{D}\langle a \rangle$$

It must be put to the reader's attention that no expression for the cell length PDF is accessible because the integral $\int \gamma(l')P(l,l')n(t,l')$ has no specific shape. More precisely, the introduction of the kernel P comes from experimental data fitting but does not capture the biological triggers of the division phenomenon. Age–structured models are much better suited to this end, being agreed that one dividing cell gives birth to two null–age cells.

Modelling claims

Modellingwise, a 5D formulation has been proposed in chapter 3 with the idea of uncoupling growth and uptake by delving into the cells' inner machinery. Its inner coordinates are made of the

bacteria's size, their lengthening rate and the main features of two membrane transport mechanisms: activity of the PTS system, permease number and degree of induction. A metabolic model reduction, essentially consisting in removing the intermediary metabolites (which balance is thought to be neutral at any time) from the operating biochemical pathways, was required for the computational time remain reasonable in view of the state–of–the–art routine processors. Albeit affordable, this formulation is the source of two conclusions which experimental measurements make for supporting evidence. Firstly, due to their own markovian history in a reactor, all cells will respond to the same environmental offer according to their particular fitness. Secondly, the organisms will not look for a maximisation of their growth rate but are geared towards the most energetic metabolic mode allowed at each time by their physiological state, the latter being highly unlikely to instantaneously adapt to the environmental offer. These original conclusions give rise to a cell–centred interpretation of the observed by–product excretions in the wake of a strong environmental perturbation.

All the inner coordinates are described with a first–order ODE, meaning that their adaption to the culture condition exhibits a necessary latency. This very delay is essential to capture the transient surges in substrate uptake in the seconds following a glucose pulse for instance, in line with [133]. It also allows for a removal of the standard algebraic relationships between residual substrate concentration and cell–scale uptake and lengthening rate. Such algebraic formulae are retrieved at steady state though, but have to be viewed as a consequence of the environmental regulation of the individuals' features instead of intrinsic properties consecrated by the model.

The idea behind the said model reduction is a coupling with CFD codes to simulate "real life" experiments and predict yield coefficients from industrial–scale cultures. In this context, massive substrate concentration gradients have to be taken into account due to their influence on the cells' fitness. This consideration is addressed in section 3.3.1 in the context of one perfectly mixed hydrodynamic compartment. In a 10^5 L industrial fermenter, the physical regimes exhibit a much larger spectrum because some zones are hardly fed and the turbulence can affect the monomer transport at the mesoscopical scale too.

Numerical aspects

Three numerical methods (eulerian, lagrangian and statistical) have been implemented to solve the five–dimensional PBE discussed in the third part of the document and cross–validated in a test case consisting in a chemostat convergence to steady state. Both the numerical accuracy and the compuational time have been evaluated, entailing a preference for the Monte–Carlo code throughout most of the fifth chapter. The calculation of the cells' residence time has also been discussed in case a dilution rate shift is treated by the Monte–Carlo algorithm.

Perspectives

Mathematical advances would mostly consist in attempting to find similar relations pertaining to size–structured models as chapter 2's formulae emerging from age–structured models. To this end, the formulation would have to be headed towards the removal of the data–fitting P_l redistribution kernel, what probably requires the assessment of a fundamental biology principle that has yet to be understood.

Future works can evolve in many directions with the aim of striking a balance between computational time to solve the population dynamics and level of detail of the cells' metabolism.

Modellingwise, although the latency in the response to environmental perturbations is compulsory to explain the uncoupling between growth and uptake, experiments carried out in an accelerostat (cf. [97]) tend to hint at an instantaneous increase in growth rate when a *E. Coli* strain is submitted to a progressive speed–up. This gears future research works towards the inclusion of a stock variable to leave some slack in the modelling. In parallel, the metabolic model could be enriched by considering the inclusion of the energetic cost associated with the $S_{ext} \rightarrow S_{stock}$ reaction, what would close the energy balance that goes with the material balance driven by the liquid–cell transfer of chemical compounds.

The organisms' uptake can also be refined by taking into account the existence of 3 or more transport systems (cf. [24]), which preferred field of action overlap, to release a maximal growth rate constraint that is thought to be challenged when a steady–state is disrupted by strong perturbations (cf. [150]).

The generational redistribution in growth potential could be refined from experimental evidences bearing upon so–called "elongated cells". It is thought that a thriving bacterium reaches its septation size too quickly to settle its lengthening poles before dividing, triggering an asymmetric size allocation in comparison with the "normal" case that involves a $\approx$ 1h time interval between two consecutive division events. Including some information pertaining to the mother–cell's fitness in the length redistribution and proposing a (potentially bimodal) $P_{v^p}(v^p, v^{p'})$ kernel that heeds the microbiologists' conclusions are two examples of model improvements from the framework that constituted the foundation to chapter 5's numerical simulations.

If the biological model is to be incorporated in a CFD code to simulate a population's behaviour in an industrial tank, solving Navier–Stokes's equation in coupling with the PBE through a splitting method is a too time–consuming approach as it stands, prompting the need to focus on a simplification of the hydrodynamics via algebraic methods such as Proper Orthogonal Decomposition to approach the vector field by a few of its most revelating modes. The biology could also be simplified by assuming for instance that the uptake machinery adapts instantly to the substrate offer the bacteria are confronted to throughout their trajectory in the system, what is a convenient assumption if the time step of the solution to the PBE is much larger than the characteristic time of the transport. In the latter framework, both the cells' uptake and lengthening rate would be calculated from the summation of a function of the time–averaged substrate concentration and a deviation from this integral quantity, emulating the so–called Reynolds–Averaged (RANS) numerical treatment of the Navier–Stokes equation. These are long–term perspectives, but the roadmap towards this ultimate goal appears more clearly and the steps to follow are already underway.

Acknowledgements

The author gratefully acknowledges the financial support from the French National Research Agency through the 2015 "Attractivity Chair" program, Idex UNITI-"Biological, Reacting, Multiphase Flows (BIREM)" conv-ANR-11-Idex-0002-02.
This work was granted access to the HPC resources of CALMIP supercomputing center under the allocation 2016-p19023.

Bibliography

[1] Luis M. Abia, Óscar Angulo, Juan Carlos López-Marcos, and Miguel Ángel López-Marcos. A numerical study on the estimation of the stable size distribution for a cell population balance model. *Mathematical Methods in the Applied Sciences*, 41(8):2894–2905, May 2018.

[2] J. P. Barford. A general model for aerobic yeast growth: Batch growth. *Biotechnology and Bioengineering*, 35(9):907–920, April 1990.

[3] Agnieszka Bartlomiejczyk, Henryk Leszczyński, and Piotr Zwierkowski. Existence and uniqueness of solutions for single-population McKendrick-von Foerster models with renewal. *Rocky Mountain Journal of Mathematics*, 45(2):401–426, February 2015.

[4] George I. Bell and Ernest C. Anderson. Cell Growth and Division I. A Mathematical Model with Applications to Cell Volume Distributions in Mammalian Suspension Cultures. *Biophys. Journal*, 7:329–351, 1967.

[5] W. A. Beyer. Solution to a Mathematical Model of Cell Growth, Division, and Death. *Mathematical Biosciences*, 6:431–436, 1970.

[6] Frédérique Billy, Jean Clairambault, Franck Delaunay, Céline Feillet, and Natalia Robert. Age-structured cell population model to study the influence of growth factors on cell cycle dynamics. *Mathematical Biosciences and Engineering*, 10(1):1–17, February 2013.

[7] E. Boye and K. Nordstrom. Coupling the cell cycle to cell growth. *EMBO Rep*, 4:757 – 760, 2003.

[8] Onn Brandman, James E. Ferrell, Rong Li, and Tobias Meyer. Interlinked Fast and Slow Positive Feedback Loops Drive Reliable Cell Decisions. *Science*, 310(5747):496, October 2005.

[9] Manuel Campos, Ivan V. Surovtsev, Setsu Kato, Ahmad Paintdakhi, Bruno Beltran, Sarah E. Ebmeier, and Christine Jacobs-Wagner. A Constant Size Extension Drives Bacterial Cell Size Homeostasis. *Cell*, 159(6):1433–1446, December 2014.

[10] Christophe Chassagnole, Naruemol Noisommit-Rizzi, Joachim W. Schmid, Klaus Mauch, and Matthias Reuss. Dynamic modeling of the central carbon metabolism of Escherichia coli. *Biotechnology and Bioengineering*, 79(1):53–73, 2002.

[11] M. Chipot. On the equations of age-dependent population dynamics. *Archive for Rational Mechanics and Analysis*, 82(1):13–25, March 1983.

[12] Panagiotis Chrysinas, Michail Kavousanakis, and Andreas Boudouvis. Effect of cell heterogeneity on isogenic populations with the synthetic toggle switch network: bifurcation analysis of two-dimensional Cell Population Balance Models. *Computers and Chemical Engineering*, 112:27–36, April 2018.

[13] Jean Clairambault, Béatrice Laroche, Stéphane Mischler, and Benoît Perthame. A mathematical model of the cell cycle and its control. Technical report, INRIA, 2003.

[14] Antoni Leon Dawidowicz and Krzysztof Loskot. Existence And Uniqueness of Solution of Some Integro-differential Equation. *Annales Polonici Mathematici*, 47:79–87, 1986.

[15] Frank Delvigne, Alison Brognaux, Nathalie Gorret, Peter Neubauer, Angélique Delafosse, Marie-Laure Collignon, Dominique Toye, Michel Crine, Mathieu Boxus, and Philippe Thonart. Characterization of the response of GFP microbial biosensors sensitive to substrate limitation in scale-down bioreactors. *Biochemical Engineering Journal*, 55(2):131–139, 2011.

[16] Mario Diaz, Monica Herrero, Luis A. Garcia, and Covadonga Quiros. Application of flow cytometry to industrial microbial bioprocesses. *Biochemical Engineering Journal*, 48(3):385–407, 2010.

[17] O. Diekmann, M. Gyllenberg, J. A. J. Metz, and A. M. de Roos. *The Dynamics of Physiologically Structured Populations: A Mathematical Framework and Modelling Explorations*, October 2012.

[18] O. Diekmann, H. J. A. M. Heijmans, and H. R. Thieme. On the stability of the cell size distribution. *Journal of Mathematical Biology*, 19:227–248, 1984.

[19] Marie Doumic. Analysis of a Population Model Structured by the Cells Molecular Content. *Math. Model. Nat. Phenom.*, 2(3):121–152, 2007.

[20] Marie Doumic and Pierre Gabriel. Eigenelements of a General Aggregation-Fragmentation Model. *Mathematical Models and Methods in Applied Sciences*, 20(5):757–783, 2010.

[21] François Dubois. Equivalent partial differential equations of a lattice Boltzmann scheme. *Computers & Mathematics with Applications*, 55:1441–1449, February 2008.

[22] Sarah Fadda, Alberto Cincotti, and Giacomo Cao. A Novel Population Balance Model to Investigate the Kinetics of In Vitro Cell Proliferation: Part I. Model Development. *Biotechnology and Bioengineering*, 109(3):772–781, March 2012.

[23] T. Ferenci. Adaptation to life at micromolar nutrient levels: the regulation of Escherichia coli glucose transport by endoinduction and cAMP. *FEMS Microbiology Reviews*, 18(4):301–317, 1996.

[24] T. Ferenci. Adaptation to life at micromolar nutrient levels, the regulation of Escherichia coli glucose transport by endoinduction and cAMP. *FEMS Microbiology Reviews*, 18(4):301 – 317, 1996.

[25] Thomas Ferenci. 'Growth of bacterial cultures' 50 years on: towards an uncertainty principle instead of constants in bacterial growth kinetics. *Research in Microbiology*, 150(7):431–438, 1999.

[26] Thomas Ferenci. Regulation by nutrient limitation. *Current Opinion in Microbiology*, 2(2):208–213, 1999.

[27] Thomas Ferenci and K. Poole Robert. Bacterial Physiology, Regulation and Mutational Adaptation in a Chemostat Environment. In *Advances in Microbial Physiology*, volume Volume 53, pages 169–229. Academic Press, 2007.

[28] Francis Filbet and Philippe Laurençot. Numerical simulation of the Smoluchowski coagulation equation. *SIAM Journal on Scientific Computing*, 25(6):2004–2028, 2004.

[29] A. G. Fredrickson. Population balance equations for cell and microbial cultures revisited. *AIChE Journal*, 49(4):1050–1059, 2003.

[30] A. G. Fredrickson and Nikolaos V. Mantzaris. A new set of population balance equations for microbial and cell cultures. *Chemical Engineering Science*, 57(12):2265–2278, 2002.

[31] A. G. Fredrickson, D. Ramkrishna, and H. M. Tsuchiya. Statistics and Dynamics of Procaryotic Cell Populations. *Mathematical Biosciences*, 103:327–374, 1967.

[32] A. G. Fredrickson and H. M. Tsuchiya. Continuous Propagation of Microorganisms. *AIChE Journal*, 9(4):459–468, July 1963.

[33] Pierre Gabriel. *Équations de transport-fragmentation et applications aux maladies à prions.* PhD thesis, Université Pierre et Marie Curie, 2011.

[34] Sashimukaar Ganesan. An Operator-Splitting Galerkin/SUPG Finite Element Method for Population Balance Equations: Stability and Convergence. *ESAIM M2AN*, 46:1447–1465, 2012.

[35] S. George, G. Larsson, and S.-O. Enfors. A scale-down two-compartment reactor with controlled substrate oscillations: Metabolic response of Saccharomyces cerevisiae. *Bioprocess Engineering*, 9(6):249–257, November 1993.

[36] Michel Godin, Francisco Feijó Delgado, Sungmin Son, William H. Grover, Andrea K. Bryan, Amit Tzur, Paul Jorgensen, Kris Payer, Alan D. Grossman, Marc W. Kirschner, and Scott R. Manalis. Using buoyant mass to measure the growth of single cells. *Nature methods*, 7(5):387–390, may 2010.

[37] Léa Guérin, Carole Coufort-Saudejaud, Alain Liné, and Christine Frances. Dynamics of aggregate size and shape properties under sequenced flocculation in a turbulent Taylor-Couette reactor. *Journal of Colloid and Interface Science*, 491:467–478, 2017.

[38] Vincent Guillou, Lucile Plourde-Owobi, Jean-Luc Parrou, Gerard Goma, and Jean François. Role of reserve carbohydrates in the growth dynamics of Saccharomyces cerevisiae. *FEMS Yeast Research*, 4(8):773–787, September 2004.

[39] Christos Hatzis, Friedrich Srienc, and A. G. Fredrickson. Multistaged corpuscular models of microbial growth: Monte Carlo simulations. *BioSystems*, 36:19–35, 1995.

[40] Anna-Lena Heins, Rita Lencastre Fernandes, Krist V. Gernaey, and Anna Eliasson Lantz. Experimental and in silico investigation of population heterogeneity in continuous Saccharomyces cerevisiae scale-down fermentation in a two-compartment setup. *J. Chem Technol Biotechnol*, 90:324–340, 2015.

[41] Michael A. Henson. Dynamic modeling and control of yeast cell populations in continuous biochemical reactors. *Computers and Chemical Engineering*, 27:1185–1199, 2003.

[42] Michael A. Henson. Cell ensemble modeling of metabolic oscillations in continuous yeast cultures. *Computers and Chemical Engineering*, 29:645–661, 2005.

[43] Christopher J. Hewitt, Gerhard Nebe-Von Caron, Britta Axelsson, Caroline M. McFarlane, and Alvin W. Nienow. Studies related to the scale-up of high-cell-density E. coli fed-batch fermentations using multiparameter flow cytometry: Effect of a changing microenvironment with respect to glucose and dissolved oxygen concentration. *Biotechnology and Bioengineering*, 70(4):381–390, 2000.

[44] Martin A. Hjortso and Jens Nielsen. Population balance models of autonomous microbial oscillations. *Journal of Biotechnology*, 42:255–269, 1995.

[45] H.M. Hulburt and S. Katz. Some Problems in Particle Technology. *Chemical Engineering Science*, 19:555–574, 1964.

[46] Farshid Jafarpour, Charles S. Wright, Herman Gudjonson, Jedidiah Riebling, Emma Dawson, Klevin Lo, Aretha Fiebig, Sean Crosson, Aaron R. Dinner, and Srividya Iyer-Biswas. Bridging the Timescales of Single-Cell and Population Dynamics. *Physical Review X*, 8, 2018.

[47] James Kacmar, Abdelqader Zamamiri, Ross Carlson, Nicholas R. Abu-Absi, and Friedrich Srienc. Single-cell variability in growing Saccharomyces cerevisiae cell populations measured with automated flow cytometry. *Journal of Biotechnology*, 109(3):239–254, 2004.

[48] L. Kätterer, H. Allemann, O. Käppeli, and A. Fiechter. Transient responses of continuously growing yeast cultures to dilution rate shifts: A sensitive means to analyze biology and the performance of equipment. *Biotechnology and Bioengineering*, 28(1):146–150, 1986.

[49] A. Kayser, J. Weber, V. Hecht, and U. Rinas. Metabolic flux analysis of Escherichia coli in glucose-limited continuous culture. *I. Growth-rate-dependent metabolic efficiency at steady state. Microbiology*, 151, 2005.

[50] David G. Kendall. An artificial realization of a simple "birth-and-death" process. *Journal of Royal Statistical Society*, 12(1):116–119, 1950.

[51] A.L. Koch and M. Schaechter. A Model for Statistics of the Cell Division Process. *Journal of General Microbiology*, 29:435–454, 1962.

[52] Arthur L. Koch. Multistep Kinetics: Choice of Models for the Growth of Bacteria. *Journal of Theoretical Biology*, 98:401–417, 1982.

[53] Dhinakar S. Kompala, Doraiswami Ramkrishna, Norman B. Jansen, and George T. Tsao. Investigation of Bacterial Growth on Mixed Substrates: Experimental Evaluation of Cybernetic Models. *Biotechnology and Bioengineering*, 28:1044–1055, 1986.

[54] Dhinakar S. Kompala, Doraiswami Ramkrishna, and George T. Tsao. Cybernetic Modeling of Microbial Growth on Multiple Substrates. *Biotechnology and Bioengineering*, 266:1272–1281, 1984.

[55] Tanya Kostova. Numerical Solutions to Equations Modelling Nonlinearity Interacting Age-Dependent Populations. *Computers Math. Applic.*, 19(8/9):95–103, 1990.

[56] Oliver Kotte, Judith B. Zaugg, and Matthias Heinemann. Bacterial adaptation through distributed sensing of metabolic fluxes. *Molecular systems biology*, 6:355, 2010.

[57] K. Kovàrovà-Kovar and T. Egli. Growth Kinetics of Suspended Microbial Cells: From Single-Substrate-Controlled Growth to Mixed-Substrate Kinetics. *Microbiol. Mol. Biol. Rev.*, 62:646–666, 1998.

[58] Erik E. Krommenhoek, Michiel van Leeuwen, Han Gardeniers Walter, M. van Gulik, Albert van den Berg, Xiaonan Li, Marcel Ottens, Luuk A. M. van der Wielen, and Joseph J. Heijnen. Lab-scale fermentation tests of microchip with integrated electrochemical sensors for pH, temperature, dissolved oxygen and viable biomass concentration. *Biotechnology and Bioengineering*, 99(4):884–892, 2008.

[59] Herbert E. Kubitschek. Growth during the bacterial cell cycle: analysis of cell size distribution. *Biophys. Journal*, 9:792–809, 1969.

[60] Herbert E. Kubitschek and J.A. Friske. Determination of Bacterial Cell Volume with the Coulter Counter. *Journal of Bacteriology*, 168(3):1466–1467, December 1986.

[61] Michael Kuhn, Christoph Kirse, and Heiko Briesen. Population Balance Modeling and Opinion Dynamics—A Mutually Beneficial Liaison? *Processes*, 6(9):164–183, 2018.

[62] Sanjeev Kumar and Doraiswami Ramkrishna. On the solution of population balance equations by discretization–I. A fixed pivot technique. *Chemical Engineering Science*, 51(8):1311–1332, 1996.

[63] Sanjeev Kumar and Doraiswami Ramkrishna. On the solution of population balance equations by discretization–II. A moving pivot technique. *Chemical Engineering Science*, 51(8):1333–1342, 1996.

[64] Guillaume Lambert and Edo Kussell. Memory and Fitness Optimization of Bacteria under Fluctuating Environments. *PLOS Genetics*, 10(9):1–10, September 2014.

[65] Alvaro R. Lara, Hilal Taymaz-Nikerel, Mlawule R. Mashego, Walter M. van Gulik, Joseph J. Heijnen, Octavio T. Ramírez, and Wouter A. van Winden. Fast dynamic response of the fermentative metabolism of Escherichia coli to aerobic and anaerobic glucose pulses. *Biotechnology and Bioengineering*, 104(6):1153–1161, 2009.

[66] Alvaro R. Lara, Hilal Taymaz-Nikerel, Mlawule R. Mashego, Walter M. van Gulik, Joseph J. Heijnen, Octavio T. Ramírez, and Wouter A. van Winden. Fast dynamic response of the fermentative metabolism of Escherichia coli to aerobic and anaerobic glucose pulses. *Biotechnology and Bioengineering*, 104(6):1153–1161, December 2009.

[67] David Laupsien. *Hydrodynamics, Mass Transfer and Mixing induced by Bubble Plumes in Viscous Fluids*. PhD thesis, INSA de Toulouse, 2017.

[68] Noureddine Lebaz, Arnaud Cockx, Mathieu Spérandio, and Jérôme Morchain. Population Balance Approach For The Modelling Of Enzymatic Hydrolysis Of Cellulose. *The Canadian Journal of Chemical Engineering*, 93:276–284, February 2015.

[69] J.L. Lebowitz and S.I. Rubinow. A Theory for the Age and Generation Time Distribution of a Microbial Population. *Journal of Mathematical Biology*, 103:17–36, 1974.

[70] Urs Lendenmann and Thomas Egli. Is Escherichia coli growing in glucose-limited chemostat culture able to utilize other sugars without lag ? *Microbioogy*, 141:71–78, 1995.

[71] Randall J. Leveque. *Finite Volume Methods for Hyperbolic Problems*. Cambridge University Press, 2002.

[72] Mengqiu Li, Jian Xu, Maria Romero-Gonzalez, Steve A Banwart, and Wei E Huang. Single cell Raman spectroscopy for cell sorting and imaging. *Current Opinion in Biotechnology*, 23(1):56–63, February 2012.

[73] Francisco Llaneras and Jesús Picó. Stoichiometric modelling of cell metabolism. *Journal of Bioscience and Bioengineering*, 105(1):1–11, 2008.

[74] V. Looser, F. Hammes, M. Keller, M. Berney, K. Kovar, and Thomas Egli. Flow-cytometric detection of changes in the physiological state of E. coli expressing a heterologous membrane protein during carbon-limited fedbatch cultivation. *Biotechnology and Bioengineering*, 92(1):69–78, 2005.

[75] J. H. T. Luong. Generalization of Monod Kinetics for Analysis of Growth Data with Substrate Inhibition . *Biotechnology and Bioengineering*, 29:242–248, 1987.

[76] Aniruddha Majumder, Vinay Kariwala, Santosh Ansumali, and Arvind Rajendran. Lattice Boltzmann method for multi-dimensional population balance models in crystallization. *Chemical Engineering Science*, 70:121–134, 2012.

[77] Nikolaos V. Mantzaris, Prodromos Daoutidis, and Friedrich Srienc. Numerical solution of multi-variable cell population balance models: I. Finite difference methods. *Computers and Chemical Engineering*, 25:1411–1440, 2001.

[78] Nikolaos V. Mantzaris, Prodromos Daoutidis, and Friedrich Srienc. Numerical solution of multi-variable cell population balance models: II. Spectral methods. *Computers and Chemical Engineering*, 25:1441–1462, 2001.

[79] Nikolaos V. Mantzaris, Prodromos Daoutidis, and Friedrich Srienc. Numerical solution of multi-variable cell population balance models: III. Finite Element Methods. *Computers and Chemical Engineering*, 25:1463–1481, 2001.

[80] Nikolaos V. Mantzaris, Jia-Jer Liou, Prodromos Daoutidis, and Friedrich Srienc. Numerical solution of a mass structured cell population balance model in an environment of changing substrate concentration. *Journal of Biotechnology*, 71:157–174, 1999.

[81] Nikos V. Mantzaris. A cell population balance model describing positive feedback loop expression dynamics. *Computers and Chemical Engineering*, 29:897–909, 2005.

[82] Nikos V. Mantzaris. From Single-Cell Genetic Architecture to Cell Population Dynamics: Quantitatively Decomposing the Effects of Different Population Heterogeneity Sources for a Genetic Network with Positive Feedback Architecture. *Biophys. Journal*, 92:4271–4288, June 2007.

[83] Nikos V. Mantzatis. Stochastic and deterministic simulation of heterogeneous cell population dynamics. *Journal of Theoretical Biology*, 241:690–706, 2006.

[84] Daniele L. Marchisio and Rodney O. Fox. *Computational Models for Polydisperse Particulate and Multiphase Systems*. Cambridge University Press, The Edinburgh Building, Cambridge CB2 8RU, UK, 2013.

[85] Aline Marguet, Marc Lavielle, and Eugenio Cinquemani. Inheritance and variability of kinetic gene expression parameters in microbial cells: modeling and inference from lineage tree data. *Bioinformatics*, 35:i586–i595, 2019.

[86] Z. A. Melzak. A Scalar Transport Equation. *Am. Math. Society*, July 1957.

[87] Philippe Michel. Existence of a solution to the cell division eigenproblem. *Mathematical Models and Methods in Applied Sciences*, 168(7S):1125–1153, 2006.

[88] Philippe Michel, Stéphane Mischler, and Benoît Perthame. General entropy equations for struc- tured population models and scattering. *Comptes Rendus de l'Académie des Sciences de Paris*, 338(9):697–702, 2004.

[89] Philippe Michel, Stéphane Mischler, and Benoît Perthame. General relative entropy inequality: an illustration on growth models. *Journal de Maths Pures et Appliquées*, 84(9):1235–1260, 2005.

[90] Stéphane Mischler, Benoît Perthame, and Lenya Ryzhik. Stability in a Nonlinear Population Maturation Model. *Math. Models Methods Aplpl. Sci.*, 12(12):1751, 2004.

[91] J Monod. The Growth of Bacterial Cultures. *Annual Review of Microbiology*, 3(1):371–394, 1949.

[92] Jérôme Morchain. *Modélisation des bioréacteurs*. iSTE, 2018.

[93] Jérôme Morchain and Christian Fonade. A structured model for the simulation of bioreactors under transient conditions. *AIChE Journal*, 55(11):2973–2984, 2009.

[94] Jérôme Morchain, Jean-Christophe Gabelle, and Arnaud Cockx. Coupling of biokinetic and population balance models to account for biological heterogeneity in bioreactors. *AIChE journal*, 59(2):369–379, feb 2013.

[95] Jérôme Morchain, Maxime Pigou, and Noureddine Lebaz. A population balance model for bioreactors combining interdivision time distributions and micromixing concepts. *Biochemical Engineering Journal*, 126:135–145, October 2017.

[96] M.P.M Leegwater, O. M. Neijssel, and D. W. Tempest. Apects of microbial physiology in relation to process control. *J. Chem. Tech. Biotechnol.*, 32:92–99, 1982.

[97] Ranno Nahku, Kaspar Valgepea, Petri-Jaan Lahtvee, Sten Erm, Kristo Abner, Kaarel Adamberg, and Raivo Vilu. Specific growth rate dependent transcriptome profiling of Escherichia coli K12MG1655 in accelerostat cultures. *Journal of Biotechnology*, 145:60–65, 2010.

[98] Haim Nessyahu and Eitan Tadmor. Non-oscillatory Central Differencing for Hyperbolic Conservation Laws. *Journal of Computational Physics*, 87(2):408–463, April 1990.

[99] P. Neubauer, L. Häggström, and S.-O. Enfors. Influence of substrate oscillations on acetate formation and growth yield in Escherichia coli glucose limited fed-batch cultivations. *Biotech. Bioeng.*, 47(2):139–146, July 1995.

[100] Tan Trung Nguyen, Frédérique Laurent, Rodney O. Fox, and Marc Massot. Solution of population balance equations in applications with fine particles: mathematical modeling and numerical schemes. *Journal of Computational Physics*, 325:129–156, 2016.

[101] Jens Nielsen and John Villadsen. Modelling of Microbial Kinetics. *Chemical Engineering Science*, 47(17/18):4225–4270, 1992.

[102] Jean-Bernard Nobs and Sebastian J. Maerkl. Long-Term Single Cell Analysis of S. pombe on a Microfluidic Microchemostat Array. *PLOS ONE*, 9(4):1–11, April 2014.

[103] Elad Noor, Eran Eden, Ron Milo, and Uri Alon. Central Carbon Metabolism as a Minimal Biochemical Walk between Precursors for Biomass and Energy. *Molecular Cell*, 39(5):809–820, September 2010.

[104] Lucinda Notley-McRobb, Alison Death, and Thomas Ferenci. The relationship between external glucose concentration and cAMP levels inside Escherichia coli: implications for models of phosphotransferase-mediated regulation of adenylate cyclase. *Microbiology*, 143:1909–1918, 1997.

[105] Page R. Painter and Allen G. Marr. Inequality of Mean Interdivision Time and Doubling Time. *Journal of General Microbiology*, 48:155–159, 1967.

[106] Page R. Painter and Allen G. Marr. Mathematics of microbial populations. *Annual Reviews Microbiology*, 22:519–548, 1968.

[107] J. Perret. A new kinetic model of growing bacteria population. *Journal of General Microbiology*, 2:589–617, 1960.

[108] Benoît Perthame. *Transport Equations in Biology*. Springer Verlag, Birkhäuser, 2007.

[109] Benoît Perthame and Lenya Ryzhik. Exponential decay for the fragmentation or cell-division equation. *Journal of Differential Equations*, 210(1):155–177, March 2005.

[110] Maxime Pigou. *Modélisation du comportement cinétique, des phénomènes de mélange, de transfert locaux et des effets d'hétérogénéité de population dans les fermenteurs industriels.* Génie des Procédés et de l'Environnement, Université de Toulouse, Institut National des Sciences Appliquées, Toulouse, France, October 2018.

[111] Maxime Pigou and Jérôme Morchain. Investigating the interactions between physical and biological heterogeneities in bioreactors using compartment, population balance and metabolic models. *Chemical Engineering Science*, 126:267–282, 2015.

[112] E. O. Powell. Growth Rate and Generation Time of Bacteria, with Special Reference to Continuous Culture. *Journal of General Microbiology*, 15:492–511, 1956.

[113] E. O. Powell. A Note on Koch & Schaechter's Hypothesis about Growth and Fission of Bacteria. *Journal of General Microbiology*, 37:231–249, 1964.

[114] Shamsul Qamar and Gerald Warnecke. Solving population balance equations for two-component aggregation by a finite volume scheme. *Chemical Engineering Science*, 62:679–693, 2007.

[115] Vincent Quedeville, Jérôme Morchain, Philippe Villedieu, and Rodney O. Fox. A critical analysis of Powell's results on the interdivision time distribution. *Scientific Reports*, 9, 2019.

[116] Vincent Quedeville, Hicham Ouazaite, Bastien Polizzi, Rodney O. Fox, Philippe Villedieu, Pascal Fede, Fabien Létisse, and Jérôme Morchain. A two-dimensional population balance model for cell growth including multiple uptake systems. *Chemical Engineering Research and Design*, 132:966–981, 2018.

[117] Doraiswami Ramkrishna. Statistical Models of Cell Populations. *Adv. Biochem. Eng.*, 11:1–47, 1979.

[118] Galina Reshes, Sharon Vanounou, Itzhak Fishov, and Mario Feingold. Cell shape dynamics in Escherichia coli. *Biophysical journal*, 94(1):251–264, January 2008.

[119] Lydia Robert, Marc Hoffmann, Nathalie Krell, Stéphane Aymerich, Jérôme Robert, and Marie Doumic. Division in Escherichia Coli is triggered by a size-sensing rather than a timing mechanism. *BMC Biology*, 12(17), 2014.

[120] Manuel Rotenberg. Selective Synchrony of Cells of Differing Cycle Times. *Journal of Theoretical Biology*, 66:389–398, 1977.

[121] Manuel Rotenberg. Transport Theory for Growing Cell Populations. *Journal of Theoretical Biology*, 103:181–199, 1983.

[122] S.I. Rubinow. A Maturity-Time Representation For Cell Population. *Biophys. Journal*, 8:1055–1073, 1968.

[123] Jitraj Saha, Jitendra Kumar, Andreas Bück, and Evangelos Tsotsas. Finite volume approximations of breakage population balance equation. *Chemical Engineering Research and Design*, 110:114–122, 2016.

[124] M. Schaechter, Joan P. Williamson, J.R. Hood, and A.L. Koch. Growth, Cell and Nuclear Divisions in some Bacteria. *Journal of Bacteriology*, 29:421–434, 1962.

[125] William T. Scott. Analytical studies of Cloud Droplet Coalescence. *Journal of The Atmospheric Sciences*, 25:54–65, 1968.

[126] B. H. Shah, J. D. Borwanker, and D. Ramkrishna. Monte Carlo Simulation of Microbial Population Growth. *Math. Biosc.*, 31:1–23, 1976.

[127] Marian Smoluchowski. Drei Vorträge über Diffusion, Brownsche Molekularbewegung und Koagulation von Kolloidteilchen. *Physik. Zeit.*, 17:557–571, 585–599, 1916.

[128] Friedrich Srienc. Cytometric data as the basis for rigorous models of cell population dynamics. *Journal of Biotechnology*, 71(1-3):233–238, 1999.

[129] Michail Stamatakis. Cell Population Balance, Ensemble And Continuum Modeling Frameworks: Conditional Equivalence And Hybrid Approaches. *Chemical Engineering Science*, 65:1008–1015, 2010.

[130] Michail Stamatakis and Kyriacos Zygourakis. A mathematical and computational approach for integrating the major sources of cell population heterogeneity. *Journal of Theoretical Biology*, 266:41–61, 2010.

[131] Jeffrey V. Straight and Doraiswami Ramkrishna. Cybernetic Modeling and Regulation of Metabolic Pathways. Gowth on Complementary Nutrients. *Biotechnology Progress*, 10(6):574–587, 1994.

[132] G. Subramanian, D. Ramkrishna, A. G. Fredrickson, and H. M. Tsuchiya. On the mass distribution model for microbial cell populations. *Bull. Math. Biophys.*, 32:521–537, 1970.

[133] Sirichai Sunya, Frank Delvigne, Jean-Louis Uribelarrea, Carole Molina-Jouve, and Nathalie Gorret. Comparison of the transient responses of Escherichia coli to a glucose pulse of various intensities. *Applied Microbiol Biotechnol*, 95:1021–1034, 2012.

[134] Sattar Taheri-Araghi, Serena Bradde, John T. Sauls, Norbert S. Hill, Petra Anne Levin, Johan Paulsson, Massimo Vergassola, and Suckjoon Jun. Cell-Size Control and Homeostasis in Bacteria. *Current Biology*, 25:385–391, February 2015.

[135] Yu Tanouchi, Anand Pai, Heungwon Park, Shuqiang Huang, Rumen Stamatov, Nicolas E. Buchler, and Lingchong You. A noisy linear map underlies oscillations in cell size and gene expression in bacteria. *Nature*, 523:357–360, jul 2015.

[136] Hilal Taymaz-Nikerel, Walter M. van Gulik, and Joseph J. Heijnen. Escherichia coli responds with a rapid and large change in growth rate upon a shift from glucose-limited to glucose-excess conditions. *Metabolic Engineering*, 13:307–318, March 2011.

[137] E. Trucco. Mathematical Models for Cellular Systems. The Von Foerster equation. *Bull. of Math. Biophys.*, 27:285–304, 1965.

[138] H. M. Tsuchiya, A. G. Fredrickson, and R. Aris. Dynamics of Microbial Cell Populations. *Advances in Chemical Engineering*, 6:125–206, 1966.

[139] John J. Tyson and Kenneth B. Hannsgen. Cell growth and division: a deterministic/probabilistic model of the cell cycle. *Journal of Mathematical Biology*, 23:231–246, 1986.

[140] Kaspar Valgepea, Kaarel Adamberg, Ranno Nahku, Petri-Jaan Lahtvee, Liisa Arike, and Raivo Vilu. Systems biology approach reveals that overflow metabolism of acetate in Escherichia coli is triggered by carbon catabolite repression of acetyl-CoA synthetase. *BMC Systems Biology*, 4:166–178, 2010.

[141] Johan van Heerden, Mannus Kempe, Anne Doerr, Timo Maarleveld, Niclas Nordholt, and Frank Johannes Bruggeman. Statistics and simulation of growth of single bacterial cells: Illustrations with B. subtilis and E. coli. *Scientific Reports*, 7(1):1–11, December 2017.

[142] Johan H. van Heerden, Meike T. Wortel, Frank J. Bruggeman, Joseph J. Heijnen, Yves J. M. Bollen, Robert Planqué, Josephus Hulshof, Tom G. O'Toole, S. Aljoscha Wahl, and Bas

Teusink. Lost in Transition: Start-Up of Glycolysis Yields Subpopulations of Nongrowing Cells. *Science*, 343(6174):1245114, February 2014.

[143] J. Varner and D. Ramkrishna. Application of cybernetic models to metabolic engineering: Investigation of storage pathways. *Biotechnology and Bioengineering*, 58(2-3):282–291, 1998.

[144] John C. Wheeler. Modified Moments and Gaussian Quadratures. *Rocky Mountain Journal of Mathematics*, 4(2), 1974.

[145] Alan J. Wolfe. The Acetate Switch. *Microbiol. Mol. Biol. Rev.*, 69(1):12–50, March 2005.

[146] Bo Xu, Mehmedalija Jahic, and Sven-Olof Enfors. Modeling of Overflow Metabolism in Batch and Fed-Batch Cultures of Escherichia coli. *Biotechnology Progress*, 15(1):81–90, 1999.

[147] Jiangang Yang, Zhiwen Wang, Nianqing Zhu, Baiyun Wang, Tao Chen, and Xueming Zhao. Metabolic engineering of Escherichia coli and in silico comparing of carboxylation pathways for high succinate productivity under aerobic conditions. *Microbiological Research*, 169(5):432–440, May 2014.

[148] Kenji Yasuda. Algebraic and geometric understanding of cells: Epigenetic inheritance of phenotypes between generations. *High Resolution Microbial Single Cell Analytics*, 124:55–81, 2011.

[149] Jamey D. Young, Kristene L. Henne, John A. Morgan, Allan E. Konopka, and Doraiswami Ramkrishna. Integrating cybernetic modeling with pathway analysis provides a dynamic, systems-level description of metabolic control. *Biotechnology and Bioengineering*, 100(3):542–559, June 2008.

[150] Hyun Shik Yun, Juan Hong, and Henry C. Lim. Regulation of Ribosome Synthesis in Escherichia coli Effects of Temperature and Dilution Rate Changes. *Biotechnology and Bioengineering*, 523(5):615–624, December 1996.

[151] Yongchun Zhang, Michael A. Henson, and Yannis G. Kevrekidis. Nonlinear model reduction for dynamic analysis of cell population models. *Chemical Engineering Science*, 58:429–445, 2003.

[152] Hai Zheng, Po-Yi Ho, Meiling Jiang, Bin Tang, Weirong Liu, Dengjin Li, Xuefeng Yu, Nancy E. Kleckner, Ariel Amir, and Chenli Liu. Interrogating the Escherichia coli cell cycle by cell dimension perturbations. *PNAS*, 113(52):15000–15005, December 2016.

Contents lists available at ScienceDirect

Chemical Engineering Research and Design

journal homepage: www.elsevier.com/locate/cherd

A two-dimensional population balance model for cell growth including multiple uptake systems

V. Quedeville[b,a], H. Ouazaite[a], B. Polizzi[b], R.O. Fox[c], P. Villedieu[d,a], P. Fede[b], F. Létisse[a], J. Morchain[a,*]

[a] Laboratoire d'Ingénierie des Systèmes biologiques et des Procédés, INSA de Toulouse, 135 Avenue de Rangueil, 31400 Toulouse, France
[b] Institut de Mécanique des Fluides de Toulouse, 2 Allée du professeur Camille Soula, 31000 Toulouse, France
[c] Department of Chemical and Biological Engineering, Iowa State University, 618 Bissell Road, Ames, IA 50011-1098, United States
[d] ONERA de Toulouse, 2 Avenue Edouard Belin, 31055 Toulouse, France

ARTICLE INFO

Article history:
Received 18 October 2017
Received in revised form 14 February 2018
Accepted 14 February 2018
Available online 2 March 2018

Keywords:
Cell growth
Population balance model
Uptake rate
Bioreactor
Dynamic simulation

ABSTRACT

Cell growth in a chemostat is a well-documented research topic. How cells uptake the available substrate to gain weight and engage cell division is not generally taken into account in the modelling bioreactors. In fact, the growth rate is related to a population doubling time whereas the microorganisms' growth in mass is due to the mass transfer of substrates from the liquid phase to the biotic phase. Clearly, growth in mass precedes growth in number. Similarly, the transport of substrates down to the cell scale precedes the mass transfer. This article's main feature is a two-dimensional population balance model that allows to uncouple growth in mass and growth in number when the equilibrium between a cell population and its environment is disrupted. The cell length and the rate of anabolism are chosen as internal variables. It is proved that the hypothesis "growth in number = growth in mass" is valid at steady-state or in exponential growth only. The glucose uptake is assumed driven by two transport systems with a different affinity constant for the substrate. This combination of two regulated uptake systems operating in parallel explains a 3-fold increase in the uptake following a glucose pulse, but can also predict substrate uptake rates higher than the maximal batch value as observed in some experiments. These features are obtained by considering carbon fluxes in the formulation of regulation principles for uptake dynamics. The population balance's implementation in a multi-compartment reactor is a natural prospective work and allows extensions to industrial processes.

1. Introduction

From a chemical-engineering perspective, aerated bioreactors have to be regarded as three-phase reactors, and the prediction of mass transfer between phases is a central issue. Given the abundance of literature pertaining to the gas–liquid aspects of the problem, this topic will be put aside here. One point specific to liquid–cell mass transfer is that there is no thermodynamic law to prescribe the relationship between the concentrations at the cell interface. Thus, in living systems, the mass-transfer intensity through the cell membrane is dynamically adjusted in order to fit the cell's needs (Ferenci, 1996). The latter can correspond to a maximum growth rate in a non-limiting environment, or be dictated by the environmental conditions such as the imposed dilution rate in a chemostat. As a consequence, in exponentially growing cultures (balanced-growth phase) and in chemostat cultures, a strict proportionality is observed between the mass-transfer rate (or uptake rate in the field of biochemical engineering) and the growth rate. Moreover the latter is correlated to the

* Corresponding author.
E-mail address: morchain@insa-toulouse.fr (J. Morchain).
https://doi.org/10.1016/j.cherd.2018.02.025

residual concentrations of nutrients in the liquid phase. These well-known observations led to the formulation of specific growth (μ) and uptake rates (q_S) as algebraic functions of the substrate concentration in the liquid phase:

$$\mu = f(S), \tag{1.1}$$

$$q_S = Y_{SX}(\mu)\ \mu, \tag{1.2}$$

where Y_{SX} is the average mass of substrate required to yield 1 g of cells. Note that these models are relevant to fit experimental data (they are in fact empirical correlations), and thus are limited to the situation in which they are fitted. In particular, the steady-state or balanced-growth assumption implies that they are time-averaged laws.

In light of these remarks, we could preferably use a more explicit notation for the specific growth ($\bar{\mu}$) and uptake ($\bar{q}_S$) rates:

$$\bar{\mu} = f(S) \tag{1.3}$$

$$\bar{q}_S = \frac{1}{\tau}\int_t^{t+\tau} q_S(t)\ dt \approx Y_{SX}(\bar{\mu})\ \bar{\mu}, \tag{1.4}$$

where the overbar indicates a time average. Clearly, the value of the time scale τ must be large enough so that the mean uptake rate becomes constant and proportional to the mean specific growth rate. These time-averaged quantities are relevant to describe the pseudo-steady-state dynamics, however they are not applicable to the transient response (over time scales shorter than τ) because they assume an instantaneous adaptation of the living system (Silveston et al., 2008). In view of improving the dynamical modelling of bioreactors, it is important to be able to prescribe a substrate uptake model valid on the shortest time scale possible. As far as we know, there are only a few studies dedicated to this point. Chassagnole and co-workers derived a dynamic model for the glucose uptake through the PhosphoTransferase System (PTS) (Chassagnole et al., 2002) based on a detailed description of the glycolysis and pentose-phosphate pathways. Even in its reduced form, this model involves five internal concentrations and requires a large number of parameters to be fitted. Moreover, Ferenci has identified the existence of multiple transport systems whose activity depends on the substrate concentration (Ferenci, 1996) and proposed that the uptake rate is computed as the sum of the contribution of each system (Ferenci, 1999a).

The experimental measurement of the substrate uptake rate has received much attention in the last decades. Neubauer's 1990s experimental work (Neubauer et al., 1995) revealed that the instantaneous uptake rate of *Escherichia coli* cells, cultivated in a two-compartment (Continuous Stirred Reactor + Plug Flow Reactor) bioreactor operated in fed-batch mode and subject to repeated exposures to high glucose concentrations, could largely exceed the maximum uptake rate observed in a batch reactor. The experimental device was such that the first 120 s after the exposure to a glucose pulse could be observed with a temporal resolution of 30 s. More recently, Lara et al. (2009) using a bioscope measured the instantaneous uptake rate of *E. coli* cells sampled from a continuous stirred bioreactor. Their results confirmed in both aerobic and anoxic conditions that the uptake rate in the few seconds following the addition of glucose largely exceeds the maximum uptake rate measured in batch culture (based on the definition of a substrate to biomass yield and a maximum specific growth rate). The temporal resolution here is raised up to ≈3 s and the duration of the observation limited to 90 s. These are experimental evidence that the correlation between growth, uptake and the substrate concentration established in non-limiting or steady-state conditions is not valid on very short time scales ($\tau < 10$ s) when the transient response of the uptake system is involved. The results obtained by Sunya et al. (2012) who examined the dynamic response of *E. coli* cells to glucose pulses in chemostat cultures, with a temporal resolution of ≈25 s over longer periods of time offer an opportunity to establish a closure model in the situation where $\tau \approx 5$–30 min. Natarajan and Srienc (2000) examined the uptake of a glucose analogue at the cell level using cytometry. Their results revealed that the substrate uptake rate (after 5 min following a pulse addition) is distributed in the population of cells.

The use of the population balance concept to deal with the population dynamics has been identified as the most natural way to proceed, for some time (Fredrickson and Tsuchiya, 1963). PBMs were first introduced by Smoluchowski (1916) to model the size of particles undergoing coalescence and rupture.

In biology, PBMs are rife to describe the dynamics of a cell property (age, size, mass, intracellular concentration of an enzyme representative of the cell's state) among a population of individuals. Such modelling of biological systems was introduced by Von Forster (1959) to take into account the influence of mortality over the age of a population. The cell-cycle effect was then emphasised by Bell and Anderson (1967) under the assumption that one cell gives birth to two identical daughters. In the earliest works, 1-D PBM have been derived. Most of proposed models (Eakman et al., 1966; Subramanian et al., 1970; Shah et al., 1976), and the many papers these references have inspired, relate a cell's state to its mass or volume, which requires a formulation of other properties (such as age, growth and reaction rates, or substrate consumption for instance) as functions of the mass, which has proved insufficient and rather inconclusive (for all processes that are not related to mass in the cell functioning). Many other variables may turn out to be relevant depending on the biological behaviour of the cells under consideration. One of them is maturity, highlighted by Trucco (1965), and understood (Lebowitz and Rubinow, 1974) as the cytological age, $a_c \in [0, 1]$. This formulation yields a boundary condition that connects the number density at $a_c = 0$ and $a_c = 1$: the production of new born cells (with $a_c = 0$) equals the flux of cells reaching $a_c = 1$. A unique solution in $\mathcal{C}^0(\mathbb{R}_+, L^1(\Omega))$ (Ω once again standing for the internal variable's domain) is inferred from the initial condition.

In general, 1-D PBMs fail to provide a comprehensive perspective across different time scales. Models aimed at depicting the cell cycle (for which the internal variable may be mass, length or volume) are ill-adapted for explaining the variations in a population's total mass. On the other hand, when it comes to maturity (consequently the doubling time), all age-related information is filtered and only phenomena driven by a characteristic time equating to cell growth are reachable.

An alternative is to develop a multi-dimensional PBM that includes many cell properties like DNA concentration (Hatzis et al., 1995; Stamatakis and Zygourakis, 2010) or enzymes expression levels (Mantzaris, 2005), which has turned out to be a significant step forward regarding mathematical modelling in biology. However, the completeness of such a model is always subject to doubt, and due to the large number of internal variables they can quickly become computationally

intractable. Rotenberg (1977, 1983) gave a complete 2-D model for the cell cycle, which includes age and growth rate as internal variables. Mischler et al. (2004) extracted an eigenvector that geometrically shapes the steady-state solution and an eigenvalue (the so-called Malthus parameter) that drives the exponential steady-state growth in time. The existence and uniqueness of the solution in $\mathcal{C}^0(\mathbb{R}_+, L^1(\Omega))$ is guaranteed provided the initial condition lies in $L^1(\Omega)$ and the fragmentation function in L^∞. In other words the existence of a division phenomena at the cell level ensures that the number of cells will eventually grow exponentially with time after a transition period. This result is independent from the growth rate law rate prescribed at the cell level. Perthame's seminal work (Perthame, 2007) was enlightening regarding the L^1 exponential decay of the solution to a transport-fragmentation equation such as the cell cycle dynamics, along with bounded variation regularity, provided the breakage function lies in L^∞ on its domain.

At the end of the day, in order to have a description of both the cell-scale and population-scale behaviours, a PBM needs a minimum of two degrees of freedom. These observations argues for the development of a PBM describing the cell cycle as the result of the following steps: (i) transport of nutrients down to the cell membrane, (ii) substrate uptake, (iii) transformation into new cell constituents leading to cell elongation, and (iv) cell division.

The principal objective of this work is therefore to introduce a 2-D PBM for cell growth that allows to distinguish between growth in mass and growth in number. The present model can be regarded as an extension of a previously published 1-D model whose characteristic time scale is the inverse of the population maximum specific growth rate (Morchain et al., 2017). Introducing a second dimension allows the uncoupling of the growth in mass (related to substrate uptake) and the growth in number (related to cell division). In the first part of this paper, the 2-D model is presented along with the hypothesis and assumptions. Then some properties of the model are examined. In particular, we discuss the situations leading to the equivalence between growth in mass and growth in number. We also propose an integration of the 2-D PBM leading to a population-averaged model and we enlighten the consequences of such a simplification on the predictive capacity of the integrated model. In the third part, the issue of formulating an instantaneous uptake law is addressed. Finally, the proposed model is subjected to validation through the simulation of a pulse experiment in a chemostat for which data are available in the literature.

2. Modelling framework

2.1. Statement of assumptions regarding the biological system

(i) We consider in this work the case of rod-shaped cells having a constant diameter d and a varying length l. This assumption corresponds to various, widespread microorganisms such as *E. coli* (Subramanian et al., 1970) and *Schizosaccharomyces pombe* (Nobs and Maerkl, 2014). Note that the cell volume $\pi\frac{d^2}{4}l$ and surface $\pi d\,l$ are only functions of l. In this particular case, the ratio of the cell surface to the cell mass is constant $A_e = \frac{4}{\rho_c d}$ where ρ is the cell mass density.

(ii) Besides its length, each cell in the population is characterized by its elongation rate v. This rate is related to the physiological state of the microorganism and more specifically to the rate of anabolism. For a given strain, Nobs and Maerkl (2014) found that the elongation rate is cell-specific and constant throughout the cell cycle.

(iii) The cell division is driven by a size mechanism, thus the probability that a cell divides is related to its length (Robert et al., 2014). When a cell of length l divides, two daughters of approximate size $l/2$ are formed. The sum of the daughters' lengths equals that of the mother cell, which can also be regarded as a conservation of the total cell mass through cell division.

(iv) The elongation rate is redistributed at cell division, which means the two daughters may not be able to perform anabolic reactions at the same rate as their mother. If both newborn cells can thrive with an arbitrary rate of anabolism, the latter is reportedly distributed around that of the mother's. This is consistent with the fact that the cell content is not evenly distributed among the two daughters (for instance Stamatakis and Zygourakis (2010) assumed the redistribution follows a hypergeometric law).

(v) In order to grow, cells uptake a carbon source (typically glucose) and oxygen. At steady state in a chemostat or in the balanced-growth phase of a batch culture, the specific growth rate of the entire population is correlated to the concentration of the substrate in the culture medium. This correlation takes the usual Monod form:

$$\mu^\star = \mu_{\max}\frac{S}{K_S + S}\frac{O_2}{K_{O_2} + O_2}. \tag{2.1}$$

where μ_{max} is the maximum specific growth rate, K_S the affinity constant for substrate and K_{O2} the affinity constant for oxygen. Note that these *constants* are empirically determined and correspond to population averaged values (Ferenci, 1999a).

(vi) At the cell scale, the uptake rates differ from one cell to another as revealed by Natarajan and Srienc (2000). However, as stated in the Introduction, the overall uptake rate is algebraically linked to the population growth rate at steady state.

(vii) During the transition period, the population specific growth rate relaxes toward the equilibrium growth rate defined by Eq. (2.1). The shape of this adaptation in a biological-systems context is discussed in Morchain et al. (2013) and Morchain (2017).

2.2. 2-D population balance model

Let $\xi \in \mathbb{R}^2$ be a set of internal properties that fully characterize a cell's state. For the sake of completeness, the cell growth is explored in a continuous and perfectly mixed bioreactor, characterized by its dilution rate D (1/h). The PBM for such a population reads

$$\frac{\partial}{\partial t}\mathcal{N}(t,\xi) + \nabla_\xi\cdot[\dot{\xi}\mathcal{N}(t,\xi)] + D\mathcal{N}(t,\xi) + \gamma(\xi)\mathcal{N}(t,\xi) = 2\int \gamma(\xi')K(\xi,\xi')\mathcal{N}(t,\xi')\cdot d\xi' \tag{2.2}$$

where

$\xi = (l, v)^T$ is the vector of internal properties,

$\mathcal{N}$ is a number density function, and $\mathcal{N}(t, l, v)dldv$ is the number of cells with a rate of anabolism v and length l at time $t>0$,
$\dot{\xi}$ is the vector of velocities in the space of internal variables, namely $\dot{l} = \partial l/\partial t$ and $\dot{v} = \partial v/\partial t$,
$\gamma(\xi)$ is the rate of cell division, and
$K(\xi, \xi')$ is a redistribution kernel that defines the probability that a cell in state ξ' gives birth to a cell in state ξ.

The factor 2 on the right-hand side of the equation indicates that one mother cell produces to two daughter cells. The boundary condition assumes a regularity condition (i.e. $\int_{\partial\Omega} \dot{\xi}_k \mathcal{N}(t, \xi) = 0 \quad \forall k \in \{1, \ldots, n\}$), and the initial condition $\mathcal{N}(0, l, v)$ belongs to $L^1(\mathbb{R}_+, [0, v_{max}])$.

In order to get the full set of equations for the dynamic simulation of a continuous bioreactor, the PBM is complemented with two mass balances for the carbon substrate, S, and the dissolved oxygen, O_2:

$$\frac{dS}{dt} = D(S_f - S) - \frac{\rho_c \pi d^2}{4} \iint q_S l \mathcal{N}(t, l, v) \cdot dldv \tag{2.3}$$

$$\frac{dO_2}{dt} = K_L a(O_2^\star - O_2) - DO_2 - \frac{\rho_c \pi d^2}{4} \iint q_{O_2} l \mathcal{N}(t, l, v) \cdot dldv \tag{2.4}$$

where ρ_c is the density of cells (assumed constant and equal to 1000 kg/m^3), q_S and q_{O_2} are the substrate and oxygen specific uptake rates, respectively, S_f is the substrate concentration in the feed (g/L), and $O_2^\star$ is assumed constant and given by Henry's law. The reader should note that all mass densities are given in kg/m^3, but are converted into g/L in the simulations. These mass balances are coupled to the PBM through the integral terms on the right-hand sides, which represent the contribution of the entire population. We may recall here that $\rho_c \pi d^2 l/4$ is a cell mass.

2.3. Modelling the cell division

Many breakage laws have been implemented in the literature (Hatzis et al., 1995; Fadda et al., 2012; Subramanian et al., 1970). Here, following Mantzatis (2006), we take

$$\gamma(l) = \left(\frac{l}{l_c}\right)^\kappa \tag{2.5}$$

where l_c is a constant characteristic length and κ, a shape parameter, is set to 5 as proposed by Mantzatis (2006). This smooth function allows cells to grow up to a length comparable to l_c, and cell division is almost guaranteed at $l = 2l_c$. However, it does not theoretically preclude the possibility that a cell never stops growing. The value of the parameter l_c is set to 10^{-5} m according to Nobs and Maerkl (2014).

2.4. Modelling velocities in the internal phase space

The length change is taken to be proportional to the rate of anabolism:

$$\dot{l} = \frac{\partial l}{\partial t} = a \min(v, \mu^\star). \tag{2.6}$$

This formulation ensures that a cell elongates at a rate proportional to the rate of anabolism under the condition that the medium is not depleted in carbon substrate and/or oxygen. The time scale associated with this elongation is the interdivision time. Since the elongation rate v is a distributed property, it results that the combination of (2.5) and (2.6) will produce the experimentally observed interdivision time distribution (Yasuda, 2011; Nobs and Maerkl, 2014).

The parameter a is a conversion constant that connects the rate of anabolism to the rate of elongation. This value is clearly strain dependent and, in the present work, it is adjusted to the total cell mass measured in the experiments simulated (see Table 1 for parameter values).

The relationships (1.3) and (1.4) reflect the well-established fact that the uptake rate is proportional to the specific growth rate at steady state. The latter is correlated to the residual substrate concentration. However there is much experimental evidence that the specific growth rate (and hence the elongation rate) is not correlated to the substrate concentration in the transient regime (see Perret, 1960; Abulesz and Lyberatos, 1989; Patarinska et al., 2000; Kätterer et al., 1986; Guillou et al., 2004 for the response to a step-up in the feed concentration or the dilution rate, and Adamberg et al. (2009) for the response to a gradual increase of the dilution rate). In fact, cells adapt their rate of anabolism in response to changes in environmental concentrations. The dynamics of such an adaptation has been investigated in previous works and the following expression was proposed and validated against experimental data (Morchain and Fonade, 2009; Morchain et al., 2017):

$$\dot{v} = \frac{\partial v}{\partial t} = \begin{cases} \left(\frac{1}{T} + v\right)(\mu^\star - v) & \text{if} \quad v \le \mu^\star \\ \frac{1}{T}(\mu^\star - v) & \text{if} \quad v \ge \mu^\star \end{cases} \tag{2.7}$$

where $\mu^\star$ is given by Eq. (2.1). The value of the parameter T was found to be around $1.25/\mu_{max}$.

2.5. Modelling the redistribution kernels

The redistribution kernel $K(\xi, \xi')$ accounts for the probability that a mother cell with internal variables ξ' gives birth to a daughter with internal variables ξ. Without precise empirical knowledge of the redistribution process of internal variables, we assume that the cell length and the rate of anabolism are independently redistributed at cell division. Thus, the redistribution kernel is a tensor product of two independent kernels, each one involving one internal variable. This leads to $K(\xi, \xi') = P(l, l')Q(v, v')$ with P being the length redistribution kernel, and Q the rate of anabolism redistribution kernel.

In general, $P(l, l')$ and $Q(v, v')$ satisfy the following properties:

$$\begin{aligned} &P(l, l') = P(l' - l, l'), \\ &\int_0^{l'} P(l, l') \cdot dl = 1, \\ &\int_0^{v_{max}} Q(v, v') \cdot dv = 1. \end{aligned} \tag{2.8}$$

The first states that a cell gives birth to two daughters cells, and the two others are normalization constraints. In this work, the redistribution kernel in length is

$$P(l, l') = \left(\frac{l}{l'}\right)^{m-1} \left(1 - \frac{l}{l'}\right)^{m-1} \frac{(2m-1)!}{(m-1)!^2} \tag{2.9}$$

with $m = 10$, so that P is a beta distribution centred around $l'/2$. We recall here that the cell division occurs around the

Table 1 – The parameter values used in the simulations.

Name	Value	Units	Description	Ref.
d	10^{-6}	m	Cell diameter	Assumed
l_c	10^{-5}	m	Critical length for the cell division rate	(Nobs and Maerkl, 2014)
υ_{max}	1	h^{-1}	Maximal rate of anabolism	Assumed
κ	5	None	Stiffness in the cell division rate	(Mantzatis, 2006)
a	6.51×10^{-6}	m	Cell lengthening	Assumed
σ	10^{-2}	None	Variance in the redistribution kernel for υ	Assumed
m	10	None	Parameter in the length redistribution kernel	Assumed
T	$\frac{1.25}{\mu_{max}}$	h	Characteristic time of adaptation	(Morchain, 2017)
t_M	50×10^{-3}	s	Micromixing time	(Morchain et al., 2017)
k_{cat}	6.15×10^{-6}	g_S permease^{-1} h^{-1}		Assumed
α_{max}	100	permease/m^2	Maximum permease surface density	Assumed
D	0.15	h^{-1}	Dilution rate	Assumed
k_{PTS}	0.01	g/L	PTS affinity constant	(Ferenci, 1996)
k_{Perm}	10^{-3}	g/L	Permease affinity constant	(Ferenci, 1996)
μ_{max}	0.46	h^{-1}	Maximum specific growth rate	(Sunya et al., 2012)
K_S	0.01	g/L	Macroscopic affinity constant for glucose	(Sunya et al., 2012)
K_{O_2}	10^{-4}	g/L	Macroscopic affinity constant for oxygen	Assumed
Y_{SX}	1/0.42 = 2.38	g_S/g_X	Substrate into biomass yield	(Sunya et al., 2012)
K_La	300	h^{-1}	Global gas–liquid transfer rate	Identified (Sunya et al., 2012)
O_2^*	8×10^{-3}	g/L	O_2 partial pressure in the gas phase	Henry's law

characteristic length l_c appearing in the definition of $\gamma(l)$. The kernel for the rate of anabolism is such that the daughter's rate is distributed around that of the mother cell according to a truncated Gaussian distribution:

$$Q(\upsilon, \upsilon') = \frac{\exp\left[-\frac{1}{2\sigma^2}(\upsilon' - \upsilon)^2\right] \mathbf{1}_{0\le\upsilon\le1}}{\int_0^{\upsilon_{max}} \exp\left[-\frac{1}{2\sigma^2}(\upsilon' - \upsilon)^2\right] d\upsilon}. \tag{2.10}$$

The variance is set arbitrarily to $\sigma = 10^{-2}$ in order to limit the randomizing effect of cell division on the rate of anabolism. With few experimental data regarding this parameter, it was assumed here that υ is evenly distributed by the daughter cells, which therefore inherit a state rather similar to their mother's.

2.6. *Uptake rate models*

The formulation of an suitable uptake model expressed as a function of the cell properties and the available substrate is a key issue. Indeed, the uptake model actually defines the rate of mass transfer between the abiotic and the biotic phases; it consequently plays the role of a closure law for the set of Eqs. (2.2) and (2.3). In this paper, following Ferenci (1996), we consider here two mechanisms contributing to the glucose uptake at a cell's membrane:

(i) PTS are responsible for the substrate uptake at high S,
(ii) Permeases are instead characterized by a smaller affinity constant, which makes this system more efficient at low S.

These processes ease the cell's adaptation to the local environment and take into account the anticipatory effect that allows cells to ratchet up/down their backbone metabolism. When glucose is omnipresent, cells favour the sugar specific PTS system, whereas starvation triggers the non-selective option that consists in increasing the membrane permeability to allow the uptake of various carbon sources into the cell.

Accordingly, in this work, it is proposed that the uptake capacity ϕ_S of the cells takes the form

$$\phi_S = Y_{SX}\left[r_{PTS}\frac{S}{k_{PTS}+S} + \alpha(\upsilon, S)\frac{4}{\rho d}k_{cat}\frac{S}{k_{Perm}+S}\right]$$
$$= \phi_{PTS} + \phi_{perm}. \tag{2.11}$$

The first term accounts for the PTS contribution to the overall substrate absorption. The second one models the permease transport system. Both correspond to active transport and are mediated by an enzymatic process with its own affinity constant. Therefore, k_{Perm} and k_{PTS} are defined at the cell scale. The ratio $4/(\rho_c d)$ corresponds to the cell surface-to-mass ratio, valid for rod-shaped cells. Thus α indicates the permease density (number of permeases per surface unit), bounded above by a maximum value α_{max}. k_{cat} is a rate constant for the substrate uptake through one permease (g_S permease^{-1} h^{-1}). The value of these parameters have been set in order that both contributions are comparable in the interval $S \in [k_{Perm}, k_{PTS}]$ (see Table 1).

In (2.11), the PTS uptake rate is defined by

$$r_{PTS} = \frac{\mu^\star + \min(\upsilon, \mu^\star)}{2} \tag{2.12}$$

This formulation is inspired by the work of Chassagnole et al. (2002) in which the uptake through PTS was correlated to both the external glucose concentration and the ratio Phosphoenolpyruvate (PEP)/Pyruvate(PYR). The transformation of PEP into PYR is the node connecting the glycolysis pathway to the TCA Cycle. The carbon flux that can enter this cycle is dependent on the availability of an electron acceptor (namely oxygen, if present). In case of oxygen depletion, the assimilation capacity through the PTS system is thus hindered by the accumulation of PEP whose conversion is slower because of insufficient energy. However, on short time scales, the pool of PYR may allow a transient boost in the carbon uptake through the PTS system in response to a glucose pulse. Rather than using a complex dynamical model for this, we propose the formulation (2.12) in which the absence of oxygen will limit uptake whilst allowing a transient over-assimilation if oxygen is present. The dependence on υ ensures that the uptake rate through the PTS system, r_{PTS}, reaches an equilibrium value at steady state. At equilibrium, the population average specific growth rate $\tilde{\upsilon}$ equals the optimal growth rate $\mu^\star$. Thus one obtains that the equilibrium value of r_{PTS} equals $\mu^\star$, reflect-

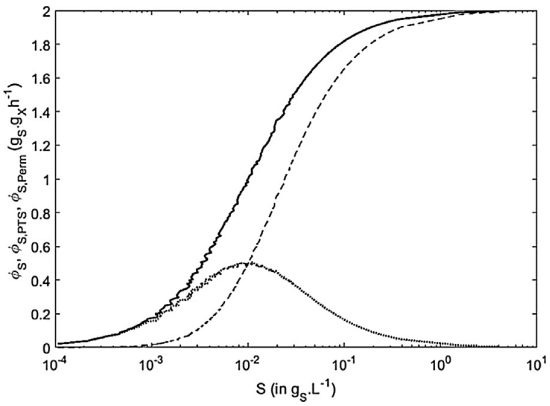

Fig. 1 – Total uptake q_S (solid) and respective contributions of the PTS (dashed) and permeases (dotted) as a function of residual substrate concentration S.

ing that the PTS system adjusts itself to allow for the optimal growth rate under the given environmental conditions.

In (2.11), the assumed permease density distribution is

$$\alpha(\upsilon, S) = \alpha_{\max} \exp\left(-\frac{\alpha_{\max}}{u}\upsilon\right) \tag{2.13}$$

where α_{max} is a constant parameter that reflects the cell's membrane permeability, and u obeys the dynamical equation

$$\frac{du}{dt} = -\frac{1}{\tau}(\tilde{q_S} - Y_{SX}\tilde{\upsilon}) \tag{2.14}$$

with a constant characteristic time τ that controls the time scale of change in the membrane permeability. The driving force of this adaptation is the difference between the overall substrate uptake rate and the overall population need. Here the notation $\tilde{\cdot}$ stands for the average over the whole population.

Fig. 1 shows the cells' more efficient glucose uptake mechanism as a function of the residual substrate concentration. This figure is found from a simulation of chemostat cultures at various dilution rates leading to different residual substrate concentrations. At low S, the majority of the total capacity ϕ_S is the contribution of the permeases, whereas the latter are inhibited by the PTS at high S, entailing a decrease in the ϕ_{perm} during the ϕ_{PTS} uptick. The rationale behind the above assumptions is that at the single-cell level, the available substrate that is likely to be absorbed is determined by phenomena operating at different scales: (a) meso/micromixing, (b) the effect of neighbouring cells, and (c) the cell's growth history (the so-called memory effect). The biological meaning of α is to represent a multiscale quantity. Thus, it is a function of the cells internal properties, but is heavily influenced by the hydrodynamics that affect the available amount of substrate at the cell membrane. It consequently blends mesoscopic and microscopic phenomena and operates upon a mechanical and physiological boundary. The permease density distribution is implemented such that at small D (entailing a small $\tilde{\upsilon}$), the substrate uptake is controlled by the permease (i.e. $\alpha(\tilde{\upsilon}, \tilde{S})$ is substantial), whereas at high D, $\alpha(\tilde{\upsilon}, \tilde{S})$ will be negligible in comparison to the PTS system that then yields the majority of the overall assimilation. These features fit the observations by reported by Ferenci (1996), Ferenci (1999b) and Kovàrovà-Kovar and Egli (1998).

Fig. 2 illustrates the contributions of PTS and permeases to the total uptake as a function of the rate of anabolism. Here, for demonstration purposes, it is assumed that the $\mathcal{N}(t, l, \upsilon)$ distribution is Gaussian and this distribution is normalized by its maximum value. The situation depicted is that of a population at steady state in a chemostat at $D = 0.15\,h^{-1}$. Thus the distribution is centred around $\tilde{\upsilon} = 0.15\,h^{-1}$. The normalized function $\alpha(S, \upsilon)/\alpha_{\max}$ decreases and consequently limits the role of permeases in the substrate uptake for those cells with a high rate of anabolism. On the contrary, for those cells with a lower rate of anabolism ($\upsilon = 0.15\,h^{-1}$), the permeases contribute to two thirds of the total uptake capacity. This feature of the model allows an overshoot in the uptake rate when starving cells are exposed to high substrate concentrations as has been observed experimentally (Lara et al., 2009; Neubauer et al., 1995).

2.7. *Interphase mass-transfer limitations*

A salient feature of the substrate uptake model proposed here is that we distinguish between the uptake capacity of the cell, ϕ_S, and the actual uptake, q_S, which can be limited by the transport of the substrate down to the cell scale. This limitation is due to the meso/micromixing that can hamper the uptake regardless the cells' capability to consume the substrate. This modelling approach, which has already been presented in previous work (Morchain et al., 2017), is based on the ratio of uptake and micromixing times. It states that the uptake rate is defined by

$$q_S = \phi_S\left[1 - \exp\left(-\frac{S}{S_c}\right)\right] \tag{2.15}$$

where the model for ϕ_S is ((2.11). The characteristic substrate concentration S_c is defined by (Morchain et al., 2017)

$$S_c = t_M Y_{SX}\tilde{\upsilon}X \tag{2.16}$$

with X the total biomass (calculated as an integral over the entire cell population), t_M the micromixing time (set to 50 ms in this study), and $\tilde{\upsilon}$ is the mean rate of anabolism of the population.

Following the same logic the actual oxygen uptake rate definition includes the possible limitation by the gas–liquid mass transfer:

$$q_{O_2} = \phi_{O_2}\left[1 - \exp\left(-\frac{K_L a O_2^*}{X\ \phi_{O_2}}\right)\right] \tag{2.17}$$

where $K_L a$ is a constant that takes into account the O_2 transfer from the gas phase to the liquid phase. Since O_2 is uptaken when glucose is absorbed by a cell and is not stored in the cytoplasm, ϕ_{O_2} is assumed equal to q_S, the eventual glucose uptake (a ratio of 1 g of oxygen per gram of glucose is assumed).

3. Numerical methods

The population balance (2.2) is solved with both a first-order finite-volume (FV) method and a Monte-Carlo (MC) algorithm. Throughout this section, the redistribution kernels $P(l, l')$ and $Q(\upsilon, \upsilon')$ are chosen beta and Gaussian, respectively. All codes are implemented in C++14.

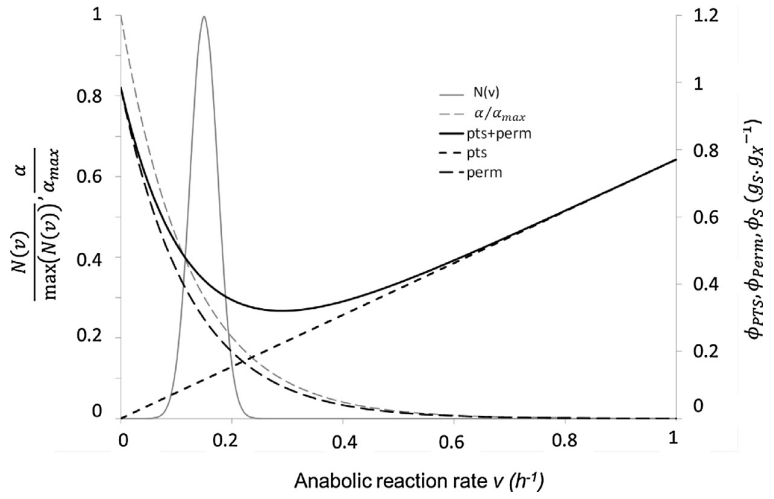

Fig. 2 – The function $\alpha(S, v)/\alpha_{max}$ (light grey dashed line) controls the induction of permeases for slow growing cells.

3.1. *The FV method*

The domain on which the PBM is solved is assumed to be a rectangle $[l_{min}, l_{max}] \times [0, v_{max}]$ and the PDF is computed on $(N+1) \times (M+1)$ nodes. In the following, $\mathcal{N}^n_{i,j}$ stands for the discrete approximation of $\mathcal{N}$ at time t^n and nodes $(l_i, v_j) \in [0, N+1] \times [0, M+1]$, and S^n the substrate concentration computed at time t^n. For the number density function, a classic local Lax–Friedrichs scheme was implemented:

$$\begin{aligned}\mathcal{N}^{n+1}_{i,j} = \; & \mathcal{N}^n_{i,j} - \frac{\Delta t}{\Delta l}\left(\mathcal{F}^n_{i+\frac{1}{2},j} - \mathcal{F}^n_{i-\frac{1}{2},j}\right) \\ & - \frac{\Delta t}{\Delta v}\left(\mathcal{G}^n_{i,j+\frac{1}{2}} - \mathcal{G}^n_{i,j-\frac{1}{2}}\right) - D\mathcal{N}^n_{i,j} - \gamma(l_i)\mathcal{N}^n_{i,j} \\ & + \Delta l \Delta v \sum_{i'>i}\sum_{j'\geq 0}\gamma(l_{i'})P(l_i, l_{i'})Q(v_j, v_{j'})\mathcal{N}^n_{i',j'}\end{aligned}$$

where $\mathcal{F}^n_{\cdot,\cdot}$ and $\mathcal{G}^n_{\cdot,\cdot}$ are the discrete fluxes associated to the respective advection contributions $\mathcal{F}(\dot{l}, \mathcal{N}) = \dot{l}\mathcal{N}(t, l, v)$ and $\mathcal{G}(\dot{v}, \mathcal{N}) = \dot{v}\mathcal{N}(t, l, v)$.

According to the Rusanov scheme, the fluxes explicitly read

$$\begin{aligned}\mathcal{F}^n_{i\pm\frac{1}{2},j} &= \frac{1}{2}\left(\mathcal{F}(\dot{l}^n_{i\pm1}, \mathcal{N}^n_{i\pm1,j}) + \mathcal{F}(\dot{l}^n_i, \mathcal{N}^n_{i,j})\right) + \frac{1}{2}\max(\dot{l}^n_{i\pm1}, \dot{l}^n_i)\left(\mathcal{N}^n_{i+1,j} - \mathcal{N}^n_{i,j}\right) \\ \mathcal{G}^n_{i,j\pm\frac{1}{2}} &= \frac{1}{2}\left(\mathcal{G}(\dot{v}^n_{j\pm1}, \mathcal{N}^n_{i,j\pm1}) + \mathcal{G}(\dot{v}^n_j, \mathcal{N}^n_{i,j})\right) + \frac{1}{2}\max(\dot{v}^n_{j\pm1}, \dot{v}^n_j)\left(\mathcal{N}^n_{i,j+1} - \mathcal{N}^n_{i,j}\right)\end{aligned}$$

The substrate concentration $S(t)$ is computed using a Riemann sum, with first-order integration at the domain boundary and second-order inside:

$$\begin{aligned}S^{n+1} = \; & S^n + \Delta t\Big(D(S_f - S^n) \\ & - \rho\pi\frac{d^2}{4}\Sigma\Sigma(l_i + l_{i+1}\mathbf{1}_{i<N})\left(\frac{v_j + \min(v_j, \mu^\star)}{2}\right) \\ & \left(\frac{v_{j+1} + \min(v_{j+1}, \mu^\star)}{2}\mathbf{1}_{j<M}\right)\frac{1}{(1+\mathbf{1}_{i<N})(1+\mathbf{1}_{j<M})} \\ & \left[(\mathcal{N}^n_{i,j} + \mathcal{N}^n_{i+1,j})\mathbf{1}_{i<N} + \mathcal{N}^n_{i,j+1}\mathbf{1}_{j<M} + \mathcal{N}^n_{i+1,j+1}\mathbf{1}_{\substack{i<N \\ j<M}}\right]\Big)\end{aligned}$$

When $O_2(t)$ reached small values, O_2^{n+1} was computed with the help of a semi-implicit scheme that removed the positivity challenge due to the term $K_L a(O_2^\star - O_2)$:

$$O_2^{n+1} = \frac{O_2^n + \Delta t K_L a O_2^\star - \Delta t \rho \sum_i\sum_j \pi\frac{d^2}{4} l q_{O_2}\mathcal{N}^n_{i,j}\cdot dl dv}{1 + \Delta t(K_L a + D)}$$

3.2. *The MC method*

This Lagrangian tool was used in a test case for comparison with the results given by the FV code. Given an initial number of cells N_0 at $t = t_0$ whose values were Gaussian distributed, the procedure is the following:

(1) Set $t \to t + \Delta t$. An integer A is set equal to 0. A cell's residence time τ_i is given by the value of D and a random number v_i: $\tau_i = -\frac{\log(v_i)}{D}$.
(2) Random numbers u_i, $i \in \{0, \ldots, N_0 - 1\}$ are drawn from a uniform distribution. A is then set equal to $\sum_{0\leq i\leq N_0-1}\mathbf{1}_{u_i<1-\exp(-\Delta t\gamma_i)}$. If this sum is equal to 0, go back to 1).
(3) If $A > 0$, all magnitudes are computed using the explicit Euler method on the interval $[t_0, t]$. All ages a_i are also updated. Considering the set $B = \{i | u_i < 1 - \exp(-\Delta t\gamma_i)\}$, all cells with subscript in B will give birth to a daughter cell whose length, rate of anabolism and residence time are computed using respectively $P(l, l')$, $Q(v, v')$ and D. N_0 is also updated: $N_0 \to N_0 + |B|$.
(4) Considering the set $C = \{i | a_i > \tau_i\}$, all cells whose subscript lies in C are withdrawn from the reactor and $N_0 \to N_0 - |C|$.
(5) The last step is the update $t_0 \to t$ and the conservative resize with respect to the system's cardinal.

Fig. 3 illustrates the convergence of the two methods toward the same solution. The number of cells in the MC approach was limited to 20,000 which may explain the minor discrepancies.

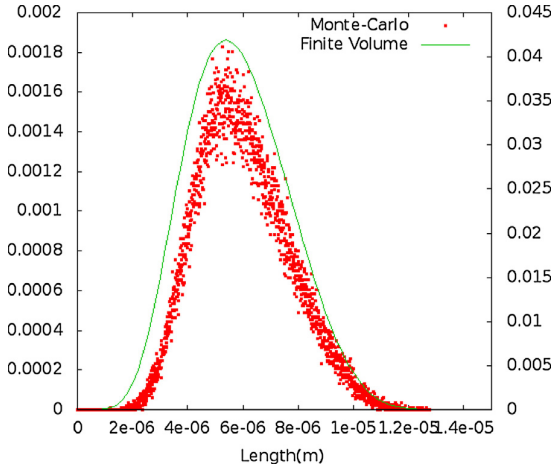

Fig. 3 – Comparison of the length distribution in an unlimited environment using different numerical simulation approaches.

4. Moments of the population balance

In this section, the moments of the population balance are introduced and used to demonstrate some properties of the PBM.

4.1. *Equivalence between growth in number and growth in mass*

Solving Eq. (2.2) is equivalent to solving the infinite set of equations on its moments. The moments of the $\mathcal{N}(t, l, v)$ distribution are defined as follows:

$$m_{p,q}(t) = \iint l^p v^q \mathcal{N}(t, l, v) \cdot dl dv. \tag{4.1}$$

Thus, the moment $m_{0,0}$ refers to the total number of cells, and the moment $m_{1,0}$ corresponds to the total mass, since the cell mass is proportional to the cell length. Let us establish the equation for these two moments. Starting from the general PBM (2.2), we obtain after some mathematical manipulation described in detail in Appendix A the following relationships:

$$\frac{d}{dt}\iint \mathcal{N}(t, l, v) \cdot dl dv = -D \iint \mathcal{N}(t, l, v) \cdot dl dv + \iint \gamma(l)\mathcal{N}(t, l, v) \cdot dl dv, \tag{4.2}$$

$$\frac{d}{dt}\iint l\, \mathcal{N}(t, l, v) \cdot dl dv = -D \iint l\, \mathcal{N}(t, l, v) \cdot dl dv + \iint \dot{l}\, \mathcal{N}(t, l, v) \cdot dl dv. \tag{4.3}$$

In other words, the notation $\langle \cdot \rangle$ standing for the double integral over the entire population, (4.2) and (4.3) read

$$\frac{dm_{0,0}}{dt} = -Dm_{00} + \langle \gamma \mathcal{N} \rangle = -Dm_{00} + \frac{\langle \gamma \mathcal{N} \rangle}{m_{0,0}} m_{0,0} \tag{4.4}$$

$$\frac{dm_{1,0}}{dt} = -Dm_{1,0} + \langle \dot{l} \mathcal{N} \rangle = -Dm_{1,0} + \frac{\langle \dot{l} \mathcal{N} \rangle}{m_{1,0}} m_{1,0} \tag{4.5}$$

These two equations degenerate into a single one on condition that

$$\frac{\langle \gamma \mathcal{N} \rangle}{m_{0,0}} = \frac{\langle \dot{l} \mathcal{N} \rangle}{m_{1,0}} \tag{4.6}$$

If the above condition holds, reasoning in terms of total cell mass or total cell number is equivalent.

At steady state ($\mathcal{N}(t, l, v) = \bar{\mathcal{N}}$), multiplying Eqs. (4.4) and (4.5) by the steady-state moments $\bar{m}_{1,0}$ and $\bar{m}_{0,0}$, respectively, and equating the two right-hand sides leads to

$$\bar{m}_{1,0}\left(-D\bar{m}_{0,0} + \langle \gamma \bar{\mathcal{N}} \rangle\right) = \bar{m}_{0,0}\left(-D\bar{m}_{1,0} + \langle \dot{l} \bar{\mathcal{N}} \rangle\right) \tag{4.7}$$

This equation can be rewritten in the same form as Eq. (4.6), so one condition for the growth in mass and in number to be equivalent is that the continuous culture is at steady state.

In a batch culture ($D = 0$), Eqs. (4.4) and (4.5) reduce to

$$\frac{1}{m_{00}}\frac{dm_{00}}{dt} = \frac{\langle \gamma \mathcal{N} \rangle}{m_{0,0}}, \tag{4.8}$$

$$\frac{1}{m_{10}}\frac{dm_{1,0}}{dt} = \frac{\langle \dot{l} \mathcal{N} \rangle}{m_{1,0}}. \tag{4.9}$$

It was well documented in Perthame (2007) and references therein that the number density function converges to a distribution whose geometry is shaped by the eigenvector associated to the population balance equation's largest eigenvalue (the so-called Malthus parameter). Then, for unlimited growth, the distribution becomes self-similar (Subramanian et al., 1970), meaning that

$$\frac{1}{m_{00}}\frac{dm_{00}}{dt} = 0.$$

The length distribution also remains self-similar, what yields

$$\frac{1}{m_{10}}\frac{dm_{10}}{dt} = 0.$$

Combining (4.8) and (4.9) then leads to the relationship

$$\frac{1}{m_{00}}\frac{dm_{00}}{dt} = \frac{1}{m_{10}}\frac{dm_{10}}{dt} \Rightarrow \frac{\langle \gamma \mathcal{N} \rangle}{m_{0,0}} = \frac{\langle \dot{l} \mathcal{N} \rangle}{m_{1,0}}. \tag{4.10}$$

Therefore, we have demonstrated that the population specific growth rate in number equals the population specific growth rate in mass if the culture is at steady state or if the population is growing exponentially.

4.2. *On the relationship between the 2-D and population-averaged model*

The usual unstructured model (or population-averaged model) takes the following form:

$$\frac{dX}{dt} = (\mu - D)\,X,$$

$$\frac{dS}{dt} = D(S_f - S) - Y_{SX}\mu X, \tag{4.11}$$

$$\mu = \mu_{\max}\frac{S}{k_{PTS} + S}.$$

Dividing Eq. (4.3) by the volume of liquid produces an equation for the cell concentration X (in g_X/L) very similar to the corresponding equation of the standard unstructured model

(4.11). However, these two equations are not equivalent and the unstructured model equation results from an approximation used to simplify the last term of Eq. (4.3):

$$\iint \dot{l}\mathcal{N}(t,l,v)\cdot dl dv = \iint \left(\frac{\dot{l}}{\bar{l}}\right) l\mathcal{N}(t,l,v)\cdot dl dv$$
$$\approx \left(\frac{\dot{l}}{\bar{l}}\right) \iint l\mathcal{N}(t,l,v)\cdot dl dv = \mu X. \quad (4.12)$$

In other words, the integral of the cell growth rate in mass over the population is roughly expressed as the product of the total mass multiplied by an average specific growth rate. This is only a rough approximation which, however, is justified if $\dot{l}/l$ is constant, meaning that the growth is exponential.

This observation clarifies the definition of μ as it appears in the standard unstructured model. For that model to be exact, μ should always be equal to $\langle \dot{l}\mathcal{N}\rangle / m_{1,0}$ (see Eq. (4.5)), whilst there is no information on the distribution in the unstructured modelling approach. It is therefore of no surprise that the unstructured model is not suitable to predict the transient behaviour of cell populations. One can further observe that cell division modifies the distribution $\mathcal{N}$, but leaves the total mass unchanged. Because μ in Eq. (4.11) is defined on a mass basis, it is not possible to investigate the consequences of any process that would impact the cell number and the cell mass on separate time scales. Clearly, the mass transfer from the liquid to the cell (uptake) is one such phenomena since uptake obviously precedes cell division. As a matter of fact, it is a paramount interest to recall that the relationship between growth and uptake rates is made on a mass basis, whereas the exact definition of specific growth rate is made on a number basis, i.e.,$\bar{\mu} = \ln 2/t_d$

The discussion above shows that (4.11) comes down to approximating an integral over the population by the product of the averaged quantities, the latter being only first-order accurate. This approximation is good for $t \to \infty$, but it can be highly inaccurate if the system is disturbed from the outside, for instance if a pulse of substrate is injected into the reactor. The first-order approximation in this case is misleading for it states that the specific growth rate immediately ratchets up from an equilibrium value to an algebraic $\mu(S(t))$. Because no distinction is made between cell mass and cell number, any gain in mass (uptake) is translated immediately into a higher specific growth rate which turns into a higher cell concentration which is simply inaccurate based on experimental observations.

5. Results and discussion

In this part, we first perform a series of experiments to test our model behaviour in response to a pulse under fully aerobic conditions. Then we assess our model against experimental data obtained by Sunya et al. (2012). A population of *E. coli* cells is set to equilibrium in a chemostat. Pulsed addition of substrate with various intensities (0.08, 0.4 and 1 g/L) are imposed on the cell population. The macroscopic properties of cell growth were given by the authors ($\mu_{max} = 0.46\,h^{-1}$, apparent substrate affinity $K_S = 0.01$ g/L, substrate to biomass yield at steady state $Y_{XS} = 0.42\,g_X/g_S$). The oxygen concentration in the liquid phase was measured with a fast responding probe, the oxygen and CO_2 concentrations in the outlet gas flow were obtained from a gas analyser. Glucose, acetate and formate were also analyzed at high temporal resolution using a mass spectrometer.

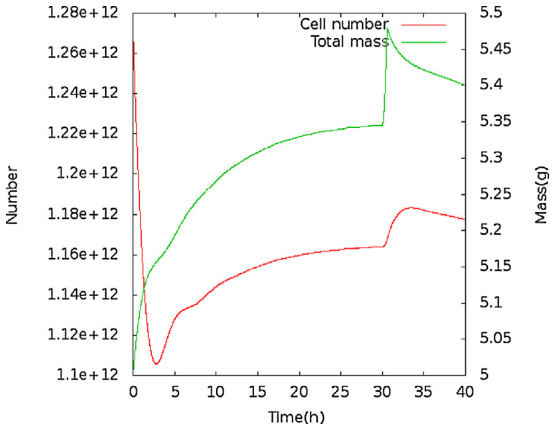

Fig. 4 – Evolution of the total mass and total number of cell during the continuous culture.

The parameter values used in simulations are reported in Table 1.

5.1. *Pulse addition without oxygen limitation*

In this part, oxygen is set constant, equal to its saturation value so that it has no impact on the results. A continuous culture at $D = 0.15\,h^{-1}$ is simulated and a pulse of substrate of 1 g/L is imposed at $t = 30$ h. We present the evolution of the integral properties of the cell population (mass, number) as well as the distributions in length and rate of anabolism before the establishment of a steady-state.

Interestingly, one can observe in Fig. 4 that the total mass and total number evolve separately and eventually tend to become proportional when the culture approaches steady state. From that point onward, examining the population growth on a number or on a mass basis becomes equivalent since the average cell mass is now time independent. Fig. 5 shows that the distributions in length and rate of anabolism become self-similar at the end of the preliminary phase that precedes the pulse addition.

The pulse addition of substrate results in a instantaneous increase in the cell mass because all cells can now elongate at their potential rate v instead of being limited by the environment (see Eq. (2.6)). However the evolution in terms of total cell number is not so sudden since cells have to elongate before they can divide into two daughter cells. Fig. 6 presents a closer view of this decoupling between growth in mass and growth in number.

These numerical results were theoretically predicted in Section 4. Note also that the slight differences in terms of number and mass between Figs. 4 and 6 are due to a change in the parameter l_c which indirectly controls the average cell size and hence the total cell mass at steady-state.

5.2. *Pulse addition with oxygen limitation: comparison to experimental data*

In this part, glucose, oxygen, cell mass and cell number are calculated. A constant $K_L a = 300\,h^{-1}$ was deduced from the experimental data at steady state prior to the glucose pulse. Integrating over the entire population leads to the total uptake rates for glucose and oxygen as well as the specific growth

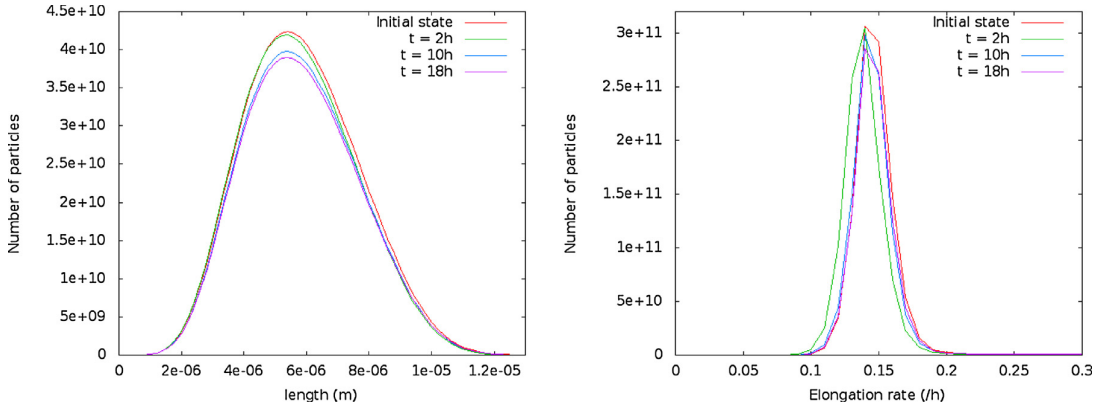

Fig. 5 – Distribution in length (left) and anabolism rate (right) toward steady-state. A population is cultivated in a chemostat ($D = 0.15$). For the sake of simplicity, the distribution were inferred before the simulation so that the code was only run for three times the residence time $1/D \approx 6.6$ h. Red: initial pdf; green: at $t = 2$ h; blue: at $t = 10$ h; purple: at $t = 18$ h. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of the article.)

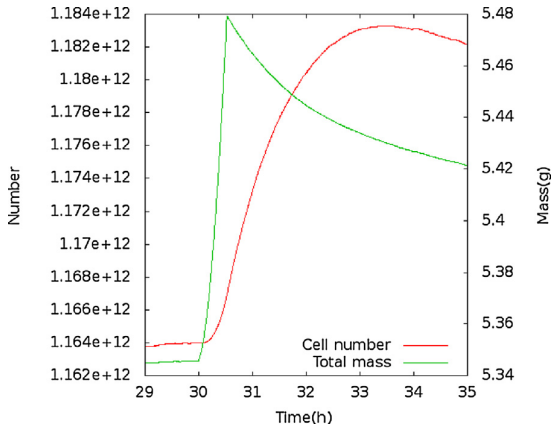

Fig. 6 – Evolution of the total mass and total number in response to a glucose pulse in a continuous culture.

rate in number and the rate of mass change due to cell growth (elongation).

Fig. 7 presents the evolution of the key variables of the dynamic response to a glucose pulse for the 0.4 g/L experiment. Just before the pulse, the PTS system contributes to approximately one third of the total glucose uptake (0.16 out of 0.37). Immediately after the pulse ($t < 0.1$ min), the total uptake rate jumps because of the PTS contribution. At that moment, oxygen is still present and this allows r_{PTS} to step-up and to overtake the contribution of permeases to the total transport. Because the glucose uptake rate has increased but the cell had hardly no time to change its rate of anabolism, the total glucose uptake rate now exceeds the glucose consumption due to anabolism. The permeases start shutting down, because of that excess carbon flux, according to Eq. (2.14). Oxygen depletion takes place at $t = 0.1$ min. As explained in the model presentation, the lack of oxygen slows down the catabolic activity and r_{PTS} falls down. Nevertheless, the still active permease system compensate the diminution of r_{PTS} so that eventually the total uptake rate remains stable. The plateau following the initial overshoot is clearly visible in the experimental data but remained unexplained until now.

From 1 to 10 min, the glucose concentration linearly decreases which progressively reduces the demand for oxygen. Thus the dissolved oxygen concentration slightly increases, resulting in an increased catabolic activity. Hence r_{PTS} goes up and permeases keep closing for the same reason as before (uptake exceeds the cell needs for anabolism). At substrate exhaustion, the rapid decrease of the sugar concentration creates a situation where q_S becomes smaller than $Y_{SX}.\upsilon$, so the cell receive a signal that the substrate flux into the cell becomes insufficient with respect to the cell needs (note that the concentration does not have to be extremely low for that message to be recorded by the cell (see Ferenci, 1996). This insufficient carbon flux signal triggers the increase in the permease activity. It is very interesting to observe that the dynamic model predicts that, as the sugar gets exhausted, the permease activity increases which explains how cells can anticipate the glucose exhaustion and manifest an apparent "anticipation capacity". This important characteristic of our model is clearly a benefit from a flux formulation in Eq. (2.14).

Fig. 8 shows the prediction of our numerical model compared to the experimental data for the three pulse intensities. In the simulation, the glucose uptake rate is slightly underestimated. The agreement is rather satisfying and we were particularly interested in understanding the origin of such similar uptake rates, irrespective of the pulse intensity. We came to the conclusion that the glucose uptake rate was in fact limited by the oxygen mass transfer during the experiments. Indeed, the maximum oxygen transfer rate is given by $K_L a O_2^*$. Considering the $K_L a$ value identified from the steady state concentrations, we conclude that the oxygen uptake rate is limited to 2.4 $g_{O_2}\,L^{-1}\,h^{-1}$ which closely matches the observed glucose consumption rate in all experiments (0.4 g in 10 min or 1 g in 25 min). This observation supports the assumption made in the model that 1 g of oxygen is consumed per gram of substrate uptaken.

In the simulations, the glucose uptake is limited by the low concentration of oxygen (see the role of O_2 in r_{PTS}). Therefore the whole dynamics is controlled by this residual concentration of oxygen whose value is highly dependent on the constant K_{O2}. Fig. 9 shows the evolution of the dissolved oxygen concentration as predicted by the model. It is put to the reader's attention that O_2 approaching zero hampers the glucose uptake. This underestimation of the glucose uptake rate is therefore related to the low oxygen residual concentration. A lower K_{O2} would cause the oxygen concentration to be much closer to zero and this would also adversely reduce the r_{PTS}. It appears that, the Monod kinetics states that the reaction stops if oxygen is zero whereas in fact the reaction

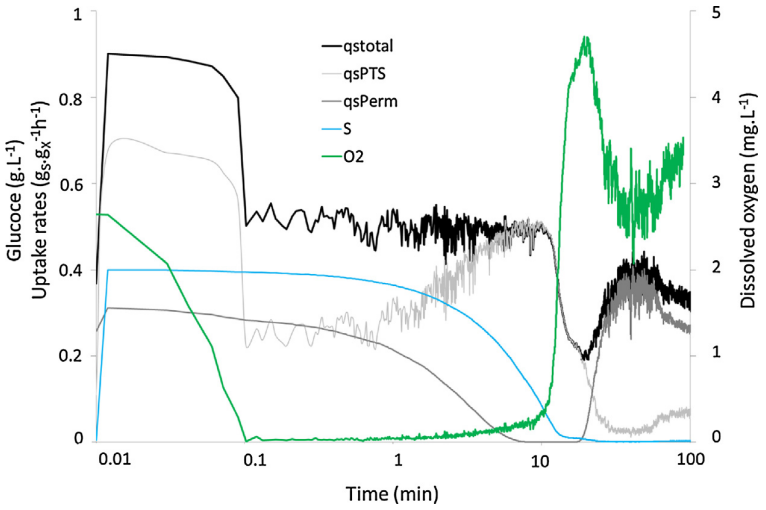

Fig. 7 – Dynamics of substrate uptake during the response to a pulse addition of 0.4 g/L of glucose in a chemostat at $D = 0.16h^{-1}$. A log scale is used for the abscissae to emphasize the first instants after the pulse. *S* and O_2 are represented in blue and green respectively, uptake rates are in light grey (PTS), dark grey (Permeases) and black (total).

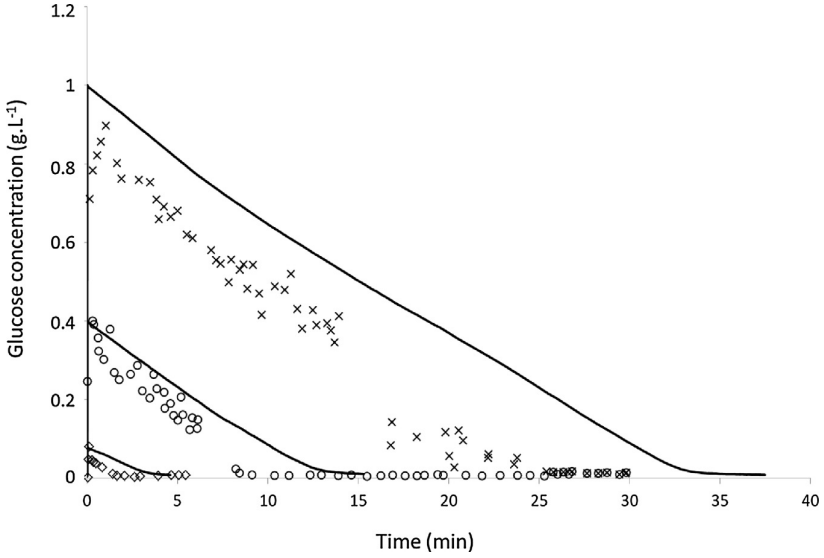

Fig. 8 – Response to a pulse of substrate of various intensities in a chemostat. Comparison of the model prediction to the experimental data of Sunya et al. (2012).

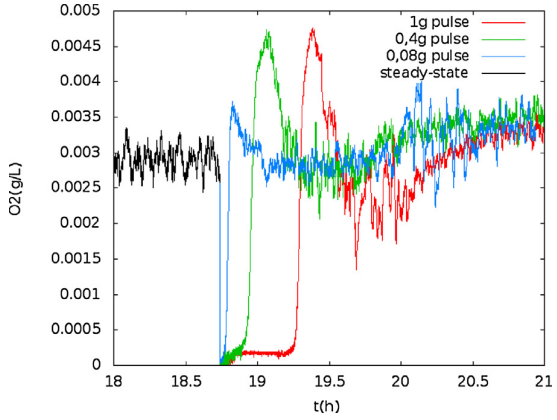

Fig. 9 – Simulated oxygen concentrations in the liquid phase as a function of time, shortly before and after the pulses.

rate would become limited by the gas-liquid mass transfer rate. This entails a caution related to the use of Monod laws to describe local phenomena in biology in general. The use of concentration in the kinetic law makes them unadapted to the situation where interphase mass transfer is the limiting phenomena. Actually, it would be preferable to avoid completely the use of Monod kinetics and to include a limitation of r_{PTS} by the maximum oxygen transfer rate without any reference to the residual concentration of oxygen. Furthermore, the use of concentration based kinetic rates creates some numerical noise that could not completely avoided despite the semi-implicit strategy for the resolution. The behaviour after glucose exhaustion is not representative of the experiments because the oxygen uptake associated to the reconsumption of acetate is not included in our model.

Fig. 10 gives an insight at which transport phenomenon drives the glucose uptake in function of the substrate flux entering the cells. At steady-state, $S \approx 5 \times 10^{-3}$ g/L, i.e. 5 times the permease affinity constant k_{perm} but only one half of k_{PTS}.

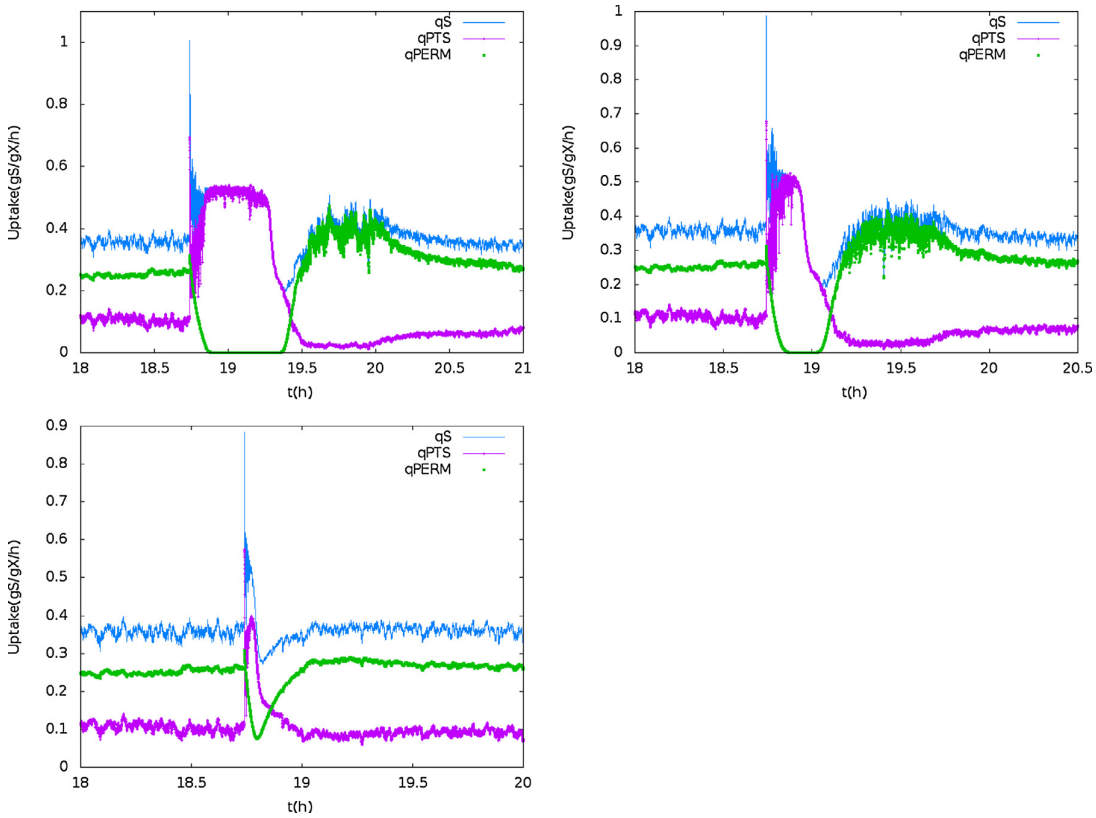

Fig. 10 – Total substrate uptake rate, PTS and Permeases contributions, shortly before and after the pulse. Top right: 1 g/L, top left: 0.4 g/L, bottom left: 0.08 g/L.

In this case, the permease contribution to the total q_S is 2–3 times over its PTS counterpart. The permease system in that case is already almost saturated. At the beginning of the pulse, the PTS are immediately requested and their full-on functioning ratchets up the q_S while the permeases' contribution does increase a little. In all cases, the instantaneous total uptake rate is multiplied by a factor 3 as reported in the experiments. Shortly after, the permeases' shutting leads to a plateau situation characterized by a PTS-only work. When $S \to 0$, the PTS becomes less and less effective and the permeases relay the PTS until steady-state is reached once again.

Up to now, it was shown that the glucose uptake was correctly predicted, it was explained why it was slightly underestimated and it was also shown that the gain in the total cell mass following a pulse addition of a subsequent amount of carbon is in contrast very limited (see Fig. 6). So there is an apparent contradiction between the quantitative uptake of substrate and the small increase in the total cell mass. This is not surprising in our model since there is no carbon balance over the cell. Such a balance could be implemented through a minimal metabolic modelling and although it is possible (see Pigou and Morchain, 2015), it was not the central objective in this paper. Nevertheless, the question of the fate of the uptaken carbon remains and we will see now how the experimental data provide a way to confirm our model prediction. For this purpose, a carbon mass balance was established from the experimental measurements of the carbon dioxide in the gas outflow and the concentrations of by-products excreted during the pulse. Fig. 11 presents the difference between the concentration in the gas phase during the pulse and the concentration measured at steady state. Since the duration of the glucose exhaustion is relatively small compared to the residence time, it can be assumed that the totality of the carbon transformed into CO_2 finally exits the reactor in the gas phase. We performed the integration of these curves and multiplied by the gas flow rate to quantify the amount of CO_2 produced during the pulse. Acetate and formate are also produced during the glucose excess period. However acetate is reconsumed and therefore also contribute to CO_2 production. This was taken into account in the mass balance. The results presented in Fig. 12 show that 65% of the total carbon uptaken is transformed into CO_2, acetate and formate during the 1 g/L pulse experiment. This fraction goes up to 93% for the 0.08 g/L pulse. The simultaneous consumption of oxygen during the period of high glucose uptake confirms the fact that the first response of the biological system is to transform the glucose uptaken in excess into CO_2. We can conclude that there is actually a significant uptake of carbon but a simultaneous release of CO_2 in the liquid phase. The carbon uptaken is not mainly metabolized to form new cells. Note that our model quantifies this excess through the difference $q_S - Y_{SX}\tilde{\nu}$.

In the 1 g/L pulse experiment, it can be calculated from the growth yield on glucose that the amount of carbon available for growth would correspond to a maximum of 0.11 g of new cells (assuming that Y_{SX} remains constant during the pulse). This value is obtained from the carbon mass balance excluding any storage, so it constitutes an upper limit. In our simulation, the boost in the uptake rate of substrate is not directly interpreted in terms of increase growth rate but we proposed that the dynamics of the population would be dictated by its rate of anabolism prior to the pulse. This hypothesis actually leads to a net production of biomass equal to 0.07 g after 25 min (glucose exhaustion). This results is therefore consistent with the carbon mass balance and confirms the fact that the substrate

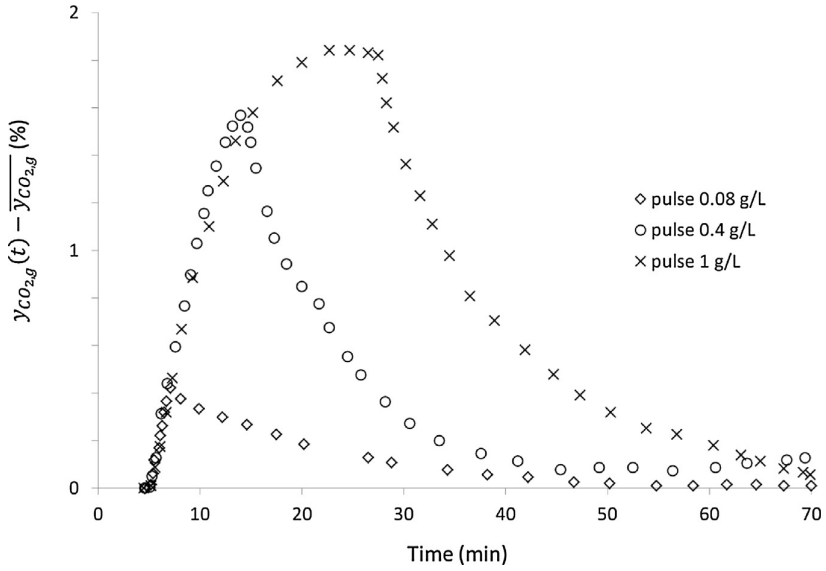

Fig. 11 – Relative change in the concentration of CO_2 in the gas phase during the pulse experiments. Data from Sunya et al. (2012).

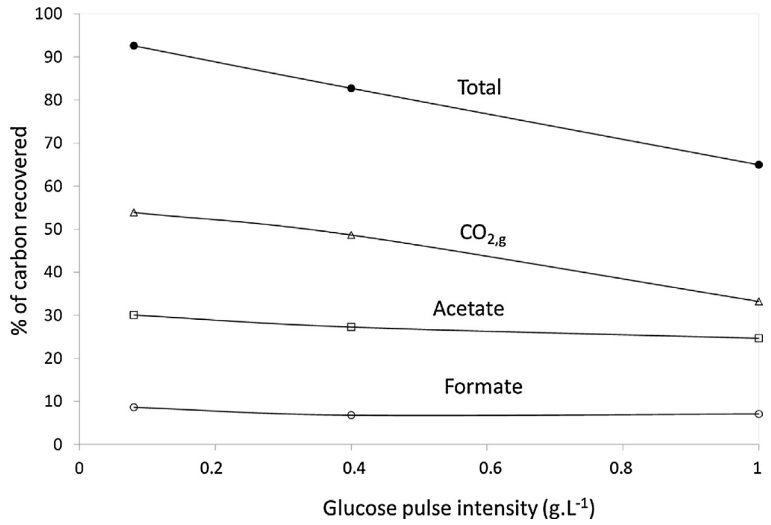

Fig. 12 – Carbon mass balance.

Data from Sunya et al. (2012).

uptake rate was significant whilst the population gain in mass was in practice undetectable.

5.3. *The instantaneous uptake rate can exceed its maximum value observed in batch*

In order to model the permease contribution to the total q_S, a value for k_{cat} is inferred and entails an equilibrium value for α. Indeed, αk_{cat} is a constant factor that emerges from the steady-state equation:

$$\left(D\frac{\bar{S}}{\bar{S}+k_{PTS}}+\alpha k_{cat}\frac{4}{\rho d}\frac{\bar{S}}{\bar{S}+k_{Perm}}\right)\left(1-\exp\left(-\frac{\bar{S}}{S_c}\right)\right)=D$$

where $\bar{S}$ stands for the residual substrate concentration at steady-state. Therefore (D being given and assumed constant), the surge in q_{perm} in the wake of a glucose pulse is totally determined by k_{cat}: the higher it is, the lower α is at steady-state, the less is the permease overactivity following the disruption.

It is highlighted in Joseph (2005) that "genetic interventions usually lead to very large changes in enzyme activity". In other terms, the cells' membrane permeability can be altered in order that the total q_S overtakes the maximum value encountered in batch culture (here, $\upsilon_{max}/Y_{SX} \approx 1.095$).

The model's equivalent would be a decrease in k_{cat} that would lead to a higher value for α before the disruption. Consequently, in the seconds following a glucose pulse, the q_{perm} contribution to q_S would be high enough to allow q_S exceeding the maximum value in batch conditions.

This is indeed predicted by the model. Fig. 13 is testament to the response of a population cultivated at low (0.05) D (meaning that the permease transport accounts for the majority of the total uptake) to a step in S. k_{cat} was divided by 2 so that the permease are significantly open before the glucose step. The glucose uptake rate q_S ratchets up in the wake of the injection and momentarily surpasses υ_{max}/Y_{SX}.

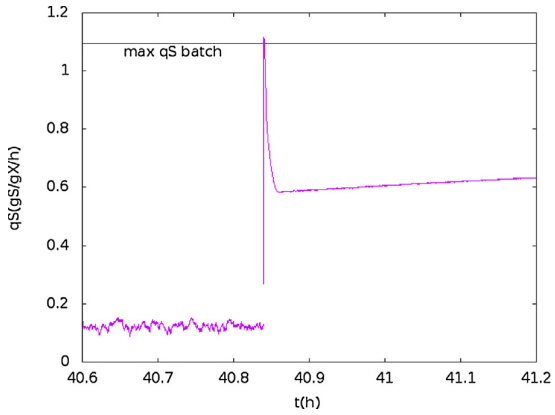

Fig. 13 – Instantaneous total substrate uptake rate following a 1 g/L pulse at $D = 0.05\ h^{-1}$.

6. Conclusion

This article highlights the relevance of a 2-D PBM formulated in terms of length (equivalent to mass for the rod-shaped cells) and rate of anabolism to investigate the population dynamics under transient conditions. The moments of the PBM were used to demonstrate that growth in mass and growth in number are only equivalent at steady-state in a continuous bioreactor and during the unlimited growth phase in a batch culture. In general, growth in mass is a matter of cell elongation (controlled by the rate of anabolic reactions) whereas growth in number is a matter of cell division (controlled by the cell division kernel). The proposed model allows for the description of transient behaviours of a cell population when growth in mass and growth in number are no longer equivalent. As an illustration, it is shown that the pulse addition of the carbon source in a substrate limited culture first causes an increase in the cell mass followed by an increase in the cell number. The second originality of this work is the formulation of an uptake law as the sum of two contributions standing for as many transport systems (PTS and permeases). Taking into account multiple transport systems is certainly necessary when modelling bacteria populations. In the proposed model, the contribution of each system to the total uptake evolves because of the difference between the actual substrate uptake and the cell needs for growth (deduced from the rate of anabolism). The dynamics of each system obey different time scales. This model compares favourably with experimental results. It is also consistent with experimental observation such as the apparent capacity of cells to anticipate the substrate exhaustion (Ferenci, 1996) or the fact that the instantaneous substrate uptake rate may exceed the maximum uptake rate observed in batch culture (Lara et al., 2009). As it is, the proposed model already appears as a valuable tool to understand and analyze one experimental data set existing in the literature. It could be observed that the dilution rate and pulse intensity are not sufficient to fully characterize the experiments. However, it is clear to the authors that some parts of the model need further improvements. A detailed sensitivity analysis has still to be conducted. It is also reasonable to include minimal energy and mass balances at the cell level. The ongoing works now concern the assessment of this model against various experimental data regarding substrate limited continuous culture at different dilution rate, different S/X ratio, inlet feed concentration and global mass transfer coefficients.

This model is now set to be implemented in a computational fluid dynamics (CFD) code to couple the biological behaviour to both micro and macro mixing. It will then be used as a tool to numerical simulations on a $\sim 10^5$ L-reactor routinely put in place in many industrial processes.

Acknowledgment

The authors gratefully acknowledge the financial support provided by the University of Toulouse via the IDEX program "Chaire d'attractivité", funded by the ANR.

Appendix A. Integral expression of the 2-D PBE

We provide below some details about the mathematical calculation of the integral expression obtained from the PBE (2.2). The passage from Eq. (2.2) to Eq. (4.3) is detailed first.

Regarding the cell number, an integration of (2.2) yields:

$$\frac{d}{dt}\iint \mathcal{N}(t,l,v)\cdot\cdot\, dl dv + D\iint \mathcal{N}(t,l,v)\cdot\cdot\, dl dv + \iint \gamma(l)\mathcal{N}(t,l,v)\cdot\cdot\, dl dv = 2\int\int\int\int_{l'>l} \gamma(l')P(l,l')Q(v,v')\mathcal{N}(t,l',v')\cdot dl' dv'\cdot dl dv \tag{A.1}$$

Indeed, the regularity boundary condition imposed on v leads to

$$\iint \frac{\partial}{\partial v}[\dot{v}\mathcal{N}(t,l,v)]\cdot dl dv = \int [\dot{v}\mathcal{N}(t,l,v)]_0^1 dl = 0$$

and its counterpart

$$\iint \frac{\partial}{\partial l}[\dot{l}\mathcal{N}(t,l,v)]\cdot dl dv$$

vanishes the same way.

Using the Fubini theorem to transform the right-hand side of (A.1) yields:

$$2\int\int\int\int_l^\infty \gamma(l')P(l,l')Q(v,v')\mathcal{N}(t,l',v')dl'dv'\cdot dl dv = 2\iint \gamma(l')\mathcal{N}(t,l',v')\{\int_0^{l'} P(l,l')dl\int Q(v,v')dv\}dl'dv'$$

Recalling the properties (2.8), the right-hand side takes the shape of:

$$2\int\int\int\int_l^\infty \gamma(l')P(l,l')Q(v,v')\mathcal{N}(t,l',v')dl'dv'\cdot dl dv = 2\iint \gamma(l')\mathcal{N}(t,l',v')dl'dv'$$

One can then rewrite (A.1):

$$\frac{d}{dt}\iint \mathcal{N}(t,l,v)\cdot dl dv = -D\iint \mathcal{N}(t,l,v)\cdot dl dv + \iint \gamma(l)\mathcal{N}(t,l,v)\cdot dl dv \tag{A.2}$$

Since it was assumed that a cell's mass is proportional to their length, an integration of the first moment on l of (2.2) gives the balance on the biomass. It reads:

$$\frac{d}{dt}\int\int l\mathcal{N}(t,l,v)\cdot dldv - \int\int \dot{l}\mathcal{N}(t,l,v)\cdot dldv + D\int\int ln(t,l,v)\cdot dldv + \int\int l\gamma(l)\mathcal{N}(t,l,v)\cdot dldv = 2\int\int\int\int_{l'>l}\gamma(l')lP(l,l')Q(v,v')\mathcal{N}(t,l',v')dl'dv'\cdot dldv \quad (A.3)$$

Indeed, an integration of the advection terms gives:

$$\int\int l\frac{\partial}{\partial l}[\dot{l}\mathcal{N}(t,l,v)]\cdot dldv = -\int\int \dot{l}\mathcal{N}(t,l,v)\cdot dldv + \int[l\dot{l}\mathcal{N}(t,l,v)]_0^\infty dv = -\int\int \dot{l}\mathcal{N}(t,l,v)\cdot dldv$$

having integrated by parts and used the regular boundary conditions, and

$$\int\int l\frac{\partial}{\partial v}[\dot{v}\mathcal{N}(t,l,v)]\cdot dldv = \int l[\dot{v}\mathcal{N}(t,l,v)]_{v=0}^{v=1}dl = 0$$

Regarding the right-hand side of (A.3), another use of the Fubini theorem gives:

$$2\int\int\int\int_{l'>l}\gamma(l')lP(l,l')Q(v,v')\mathcal{N}(t,l',v')dl'dv'\cdot dldv = 2\int\int\gamma(l')\mathcal{N}(t,l',v')\left\{\int_0^{l'} lP(l,l')dl\int Q(v,v')dv\right\}dl'dv'$$

To compute the integral over l, one uses the change of variables $l \to l' - l$:

$$\int_0^{l'} lP(l,l')dl = \int_0^{l'}(l'-l)P(l'-l,l')dl = l'\int_0^{l'}P(l,l')dl - l\int_0^{l'}P(l,l')dl$$

using the hypothesis $P(l,l') = P(l'-l,l')$.

At the end of the day,

$$\int_0^{l'} lP(l,l') = \frac{1}{2}l'$$

and:

$$2\int\int\int\int_{l'>l}\gamma(l')lP(l,l')Q(v,v')\mathcal{N}(t,l',v')dl'dv'\cdot dldv = \int\int\gamma(l')\mathcal{N}(t,l',v')dl'dv'$$

References

Abulesz, E.M., Lyberatos, G., 1989. Periodic operation of a continuous culture of Baker's yeast. Biotechnol. Bioeng. 34 (September (6)), 741–749.

Adamberg, K., Lahtvee, P.-J., Valgepea, K., Abner, K., Vilu, R., 2009. Quasi steady state growth of *Lactococcus lactis* in glucose-limited acceleration stat (A-stat) cultures. Antonie van Leeuwenhoek 95 (3), 219–226.

Bell, G.I., Anderson, E.C., 1967. Cell growth and division I. A mathematical model with applications to cell volume distributions in mammalian suspension cultures. Biophys. J. 7, 329–351.

Chassagnole, C., Noisommit-Rizzi, N., Schmid, J.W., Mauch, K., Reuss, M., 2002. Dynamic modeling of the central carbon metabolism of *Escherichia coli*. Biotechnol. Bioeng. 79 (12), 53–73.

Eakman, J.M., Fredrickson, A.G., Tsuchiya, H.M., 1966. Statistics and dynamics of microbial cell populations. Chem. Eng. Prog. Symp. Ser. 69, 37–49.

Fadda, S., Cincotti, A., Cao, G., 2012. A novel population balance model to investigate the kinetics of in vitro cell proliferation: Part I. Model development. Biotechnol. Bioeng. 109 (3), 772–781.

Ferenci, T., 1996 July. Adaptation to life at micromolar nutrient levels: the regulation of *Escherichia coli* glucose transport by endoinduction and cAMP. FEMS. Microbiol. Rev. 18, 301–317.

Ferenci, T., 1999a. Growth of bacterial cultures' 50 years on: towards an uncertainty principle instead of constants in bacterial growth kinetics. Res. Microbiol. 150 (7), 431–438.

Ferenci, T., 1999b. Regulation by nutrient limitation. Curr. Opin. Microbiol. 22 (2), 208–213.

Von Forster, H., 1959. Some Remarks on Changing Populations. Grune & Stratton, pp. 382.

Fredrickson, A.G., Tsuchiya, H.M., 1963. Continuous propagation of microorganisms. AIChE J. 9 (4), 459–468.

Guillou, V., Plourde-Owobi, L., Parrou, J.-L., Goma, G., François, J., 2004. Role of reserve carbohydrates in the growth dynamics of *Saccharomyces cerevisiae*. FEMS Yeast Res. 4 (September (8)), 773–787.

Hatzis, C., Srienc, F., Fredrickson, A.G., 1995. Multistaged corpuscular models of microbial growth: Monte Carlo simulations. BioSystems 36, 19–35.

Joseph, H., 2005. Approximative kinetic formats used in metabolic network modeling. Biotechnol. Bioeng. 91 (September (5)), 534–545.

Kätterer, L., Allemann, H., Käppeli, O., Fiechter, A., 1986. Transient responses of continuously growing yeast cultures to dilution rate shifts: a sensitive means to analyze biology and the performance of equipment. Biotechnol. Bioeng. 28 (1), 146–150.

Kovàrovà-Kovar, K., Egli, T., 1998. Growth kinetics of suspended microbial cells: from single-substrate-controlled growth to mixed-substrate kinetics. Microbiol. Mol. Biol. Rev. 62, 646–666.

Lara, A.R., Taymaz-Nikerel, H., Mashego, M.R., van Gulik, W.M., Heijnen, J.J., Ramírez, O.T., van Winden, W.A., 2009. Fast dynamic response of the fermentative metabolism of *Escherichia coli* to aerobic and anaerobic glucose pulses. Biotechnol. Bioeng. 104 (6), 1153–1161.

Lebowitz, J.L., Rubinow, S.I., 1974. A theory for the age and generation time distribution of a microbial population. J. Math. Biol. 103, 17–36.

Mantzaris, N.V., 2005. A cell population balance model describing positive feedback loop expression dynamics. Comput. Chem. Eng. 29, 897–909.

Mantzatis, N.V., 2006. Stochastic and deterministic simulation of heterogeneous cell population dynamics. J. Theor. Biol. 241, 690–706.

Mischler, S., Perthame, B., Ryzhik, L., 2004. Stability in a nonlinear population maturation model. Math. Models Methods Appl. Sci. 12 (12), 1751.

Morchain, J., Fonade, C., 2009. A structured model for the simulation of bioreactors under transient conditions. AIChE J. 55 (11), 2973–2984.

Morchain, J., Gabelle, J.-C., Cockx, A., 2013. Coupling of biokinetic and population balance models to account for biological heterogeneity in bioreactors. AIChE J. 59 (2), 369–379.
Morchain, J., Pigou, M., Lebaz, N., 2017. A population balance model for bioreactors combining interdivision time distributions and micromixing concepts. Biochem. Eng. J. 126 (October), 135–145.
Morchain, J., 2017. Bioreactor Modeling. Interactions between Hydrodynamics and Biology. ISTE Press Ltd/Elsevier Ltd, London/Oxford.
Natarajan, A., Srienc, F., 2000. Glucose uptake rates of single *E. coli* cells grown in glucose-limited chemostat cultures. J. Microbiol. Methods 42 (1), 87–96.
Neubauer, P., Häggström, L., Enfors, S.-O., 1995. Influence of substrate oscillations on acetate formation and growth yield in *Escherichia coli* glucose limited fed-batch cultivations. Biotechnol. Bioeng. 47 (2), 139–146.
Nobs, J.-B., Maerkl, S.J., 2014. Long-term single cell analysis of *S. pombe* on a microfluidic microchemostat array. PLOS ONE 9 (4), 1–11.
Patarinska, T., Dochain, D., Agathos, S.N., Ganovski, L., 2000. Modelling of continuous microbial cultivation taking into account the memory effects. Bioprocess Biosyst. Eng. 22 (6), 517–527.
Perret, J., 1960. A New kinetic model of growing bacteria population. J. Gen. Microbiol. 241, 589–617.
Perthame, B., 2007. Transport Equations in Biology. Springer, birkhäuser verlag edition.
Pigou, M., Morchain, J., 2015. Investigating the interactions between physical and biological heterogeneities in bioreactors using compartment, population balance and metabolic models. Chem. Eng. Sci. 126 (0), 267–282.
Robert, L., Hoffmann, M., Krell, N., Aymerich, S., Robert, J., Doumic, M., 2014. Division in *Escherichia coli* is triggered by a size-sensing rather than a timing mechanism. BMC Biol. 12 (17).
Rotenberg, M., 1977. Selective synchrony of cells of differing cycle times. J. Theor. Biol. 66, 389–398.
Rotenberg, M., 1983. Transport theory for growing cell populations. J. Theor. Biol. 103, 181–199.
Shah, B.H., Borwanker, J.D., Ramkrishna, D., 1976. Monte Carlo simulation of microbial population growth. Math. Biosci. 31, 1–23.
Silveston, P.L., Budman, H., Jervis, E., 2008. Forced modulation of biological processes: a review. Chem. Eng. Sci. 63 (20), 5089–5105.
Smoluchowski, M., 1916. Drei Vorträge über Diffusion, Brownsche Molekularbewegung und Koagulation von Kolloidteilchen. Phys. Z. 17, 557–571, 585-599.
Stamatakis, M., Zygourakis, K., 2010. A mathematical and computational approach for integrating the major sources of cell population heterogeneity. J. Theor. Biol. 266, 41–61.
Subramanian, G., Ramkrishna, D., Fredrickson, A.G., Tsuchiya, H.M., 1970. On the mass distribution model for microbial cell populations. Bull. Math. Biophys. 32, 521–537.
Sunya, S., Delvigne, F., Uribelarrea, J.-L., Molina-Jouve, C., Gorret, N., 2012. Comparison of the transient responses of *Escherichia coli* to a glucose pulse of various intensities. Appl. Microbiol. Biotechnol. 95, 1021–1034.
Trucco, E., 1965. Mathematical models for cellular systems. The Von Foerster equation. Bull. Math. Biophys. 27, 285–304.
Yasuda, K., 2011. Algebraic and geometric understanding of cells: epigenetic inheritance of phenotypes between generations. High Resolut. Microb. Single Cell Anal. 124, 55–81.