Studying the variability ofbacterial growth in microfluidicdroplets

Antoine Barizien

▶ **To cite this version:**

Antoine Barizien. Studying the variability ofbacterial growth in microfluidicdroplets. Bioinformatics [q-bio.QM]. Université Paris Saclay (COmUE), 2019. English. NNT : 2019SACLX020 . tel-02168120

HAL Id: tel-02168120

https://theses.hal.science/tel-02168120v1

Submitted on 28 Jun 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

université PARIS-SACLAY

ÉCOLE POLYTECHNIQUE
UNIVERSITÉ PARIS-SACLAY

NNT : 2019SACLX020

Thèse de doctorat

Etude de la variabilité de la croissance de bactéries en gouttes microfluidiques

Thèse de doctorat de l'Université Paris-Saclay
préparée à l'Ecole Polytechnique

Ecole doctorale n°573 Interfaces : approches interdisciplinaires, fondements, applications et innovation (Interfaces)
Spécialité de doctorat : Biologie

Thèse présentée et soutenue à Paris, le 28 Mai 2019, par

ANTOINE BARIZIEN

Composition du Jury :

Didier Chatenay Directeur de recherche, UMPC (Laboratoire Jean Perrin)	Président
Rosalind Allen Professeur, University of Edinburgh (School of Physics and Astronomy)	Rapporteur
Gregory Batt Chercheur, Institut Pasteur (InBio)	Rapporteur
Aleksandra Walczak Professeur, Ecole Normale Supérieure (Laboratoire de Physique Théorique)	Examinateur
Vincent Bansaye Professeur, Ecole Polytechnique (CMAP)	Examinateur
Charles Baroud Professeur, Ecole Polytechnique (Ladhyx)	Directeur de thèse

ECOLE POLYTECHNIQUE

DOCTORAL THESIS

Studying the variability of bacterial growth in microfluidic droplets

Author:
Antoine BARIZIEN

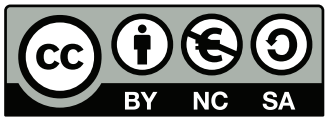
Supervisor:
Dr. Charles BAROUD

A thesis submitted in fulfillment of the requirements for the degree of Doctor of Philosophy

Physical Microfluidics and Bio-Engineering
LADHYX

March 28, 2019

Cette œuvre, création, site ou texte est sous licence Creative Commons Attribution - Pas d'Utilisation Commerciale - Partage dans les Mêmes Conditions 4.0 International. Pour accéder à une copie de cette licence, merci de vous rendre à l'adresse suivante http://creativecommons.org/licenses/by-nc-sa/4.0/ ou envoyez un courrier à Creative Commons, 444 Castro Street, Suite 900, Mountain View, California, 94041, USA.



"À la science, et en avant ! crie l'Ecclésiaste moderne, c'est-à-dire Tout le monde"

A.Rimbaud

Remerciements

Beaucoup de personnes m'ont aidé au cours de cette thèse, et je tiens à les remercier. Il y en a d'ailleurs tellement que je vais sûrement en oublier, alors merci à elleux aussi.

Tout d'abord, je tiens à remercier vraiment tout.e.s celles et ceux qui ont rendu ce travail possible, sans qu'on les remercie assez au quotidien : celles et ceux qui font le ménage derrière nous, assurent l'entretien des locaux, la livraison des colis, préparent à manger à la cantine... Mais aussi tout.e.s celles et ceux, un peu plus proches de notre travail, qui s'occupent des questions techniques, en particulier au Ladhyx Avin pour la bio, et Caroline pour la micro-fabrication. Merci aussi à Dani et Toaï pour m'avoir sorti de multiples trous noirs issus des mystères insolubles de l'informatique. Et merci à tout le personnel administratif, en particulier Marie à Pasteur, pour avoir tout fait pour que notre installation se passe au mieux.

Je voudrais aussi, bien sûr, remercier toute l'équipe de microfluidique. Déjà, Charles, pour m'avoir encadré pendant ces trois années, en me laissant une liberté à laquelle je suis attaché, tout en me donnant de bons conseils. Ensuite, merci Gabriel, pour t'être intéressé à absolument tout ce que je faisais avec bonne humeur et optimisme (ce qui contrastait pas mal avec moi), et aussi merci pour tes relectures de chapitres de cette thèse. Merci aussi à Cyprien pour m'avoir passé le flambeau en douceur, et avoir fait un super code Matlab bien plus propre que les miens, que je ne serai pas le dernier à utiliser. Merci à Benoît pour m'avoir enseigné les ficelles de la microbiologie. Merci beaucoup Adrien, pour avoir partagé mon bureau, pour tous les carreaux de chocolat, et parce qu'on sous-estime beaucoup l'impact positif sur le moral d'entendre un accent chantant du sud tous les matins. Merci aussi à Raphaël pour avoir été toujours disponible, Sebastien pour les signaling pathways expliqués avec clarté. Une pensée aussi pour Micaela avec qui ce fut super sympa de travailler, comme avec tout.e.s les autres d'ailleurs : Tiphaine, Roméo, Irma, et puis les dernier.e.s arrivé.e.s : Salomé, Christelle, Gustave et Aimee, bon courage à vous.

Vient maintenant le tour du reste du labo, avec un grand merci en particulier à toutes les filles de l'équipe d'Abdul (et Carlo), avec qui on a passé de supers moments dans le 65 (et en dehors, aussi). En particulier merci à Olga pour les shots de Tequila, merci à Thevy pour ta bonne humeur, et à Anna pour avoir commencé la thèse en même temps que moi et donné envie de finir six mois avant. Alex, merci beaucoup pour avoir partagé mon bureau, redonné vie à la plante verte et pour toutes les discussions qu'on a eu. Et merci à tout le Ladhyx, en particulier l'équipe de foot !

Je voudrais aussi remercier les gens du CMAP qui m'ont beaucoup aidé sur la partie théorique, et avec qui ce fut vraiment sympa de collaborer. Je pense à Vincent en particulier, qui a donné pas mal de son temps avec une super bonne humeur, et aussi à Tristan, Aline qui m'a initié aux mystères de Bellman-Harris, et Sylvie. Un grand merci aussi à l'équipe de Didier Mazel: à

Zeynep, pour avoir accepté de suivre ma thèse, à Julia pour sa disponibilité, et à Sébastien et Didier.

Enfin, je voudrais remercier ma famille et mes ami.e.s, qui me soutiennent tout le temps. Merci Maman, je ne crois pas que j'en serais là sans $pV = nRT$, ni que j'aurais tenu trois ans sans la mozzarella et les arancinis, t'es la meilleure, à égalité avec ma petite sœur qui est peut être ma petite sœur mais qui m'ouvre les yeux sur plein de sujets. J'adore parler avec toi Cam-cam. Merci aussi à mon petit frère qui est devenu vachement grand, au point de me rendre le plus heureux des tontons bientôt (avec l'aide non-négligeable de Marine, qu'on remercie de tolérer les soupes de kiwi). D'ailleurs en parlant de Tonton j'en ai un à remercier aussi, pour avoir toujours été là pour moi. Merci à mes grands-parents, Jean-Pi pour toutes ces soirées qui changent les idées, Minette pour toutes les petites attentions, Mémé pour toutes les tartes qui font chaud au ventre et au cœur et Pépé pour m'avoir initié à la haute œnologie et au patois savoyard.

Au tour des michel.le.s en tout genre. Merci à mon ours en chocolat, c'est fou qu'on se connaisse depuis aussi longtemps. Merci à la DTDG pour m'avoir fait gagner plus au poker qu'en faisant ma thèse. Merci à Louis de me faire déculpabiliser sur ma propre consommation de thé. Merci à Gianlu pour les soirées au Chinois. Merci à Nico de nous avoir invité au château. Merci à Severino pour m'avoir appris à recoudre mes pantalons avec un point infaillible. Merci Doudou pour m'avoir supporté ces derniers mois (dans tous les sens du terme), et m'avoir donné envie de venir au labo. T'es la meilleure des *caullaigueuh*. Et puis merci à l'apartheid, parce que si la vie était un pédalo à Lausanne, elle serait plus belle, je vous aime.

Quelques autres remerciements pour des choses qui me tiennent à cœur. Merci à Basket Pop' Paname d'ouvrir des brèches dans le monde sur-compétitif du sport contemporain. Merci à Lundi Matin d'avoir rendu attractifs mes lundis matins. Merci à PLI, les poto.es qui sont encore mieux que Lundi Matin, bien sûr. Merci à tout.e.s celles et ceux qui ont rendu magiques les manifs pendant ces trois années de thèse (et demi). Et pour finir merci à Cab (Calloway, j'aime bien le jazz), à Cab (ma sœur, quand je suis enrhumé), à CAB (Constructions Aéronautiques du Béarn, j'ai quand même un master d'aéronautique...).

Résumé de la thèse en français

Cette thèse porte sur l'étude de la variabilité de la croissance de bactéries en gouttes microfluidiques.

Dans un premier temps, la puce microfluidique utilisée pour mener des expériences tout au long de la thèse est présentée. Elle a été développée au sein de l'équipe de Charles Baroud, au Ladhyx, et permet la création d'un tableau de 1500 gouttes d'un volume de l'ordre du nanolitre. Les gouttes sont constituées de milieu de croissance aqueux, dans de l'huile fluorée, et des bactéries peuvent y être encapsulées. La croissance des bactéries dans chacune des gouttes peut alors être suivie au cours du temps par mesure de fluorescence, celle-ci étant proportionnelle au nombre de bactéries dans la goutte. Cette relation de proportionnalité est d'ailleurs démontrée expérimentalement et discutée, la valeur de la constante de proportionnalité étant difficile à obtenir dans les conditions de l'expérience de croissance. Les autres limitations de la puce microfluidique sont également discutées dans ce premier chapitre, en particulier la variabilité de la taille des gouttes. L'hétérogénéité de la fluorescence individuelle des bactéries est aussi mesurée, et son impact pour la mesure de croissance en goutte modélisé.

Dans un deuxième temps, nous revenons sur le résultat de la mesure de croissance en gouttes pour le modéliser. En effet, nous constatons une grande variation entre les gouttes en terme de fluorescence mesurée, et donc de taille de population de bactéries. A partir d'un modèle stochastique de croissance connu, le modèle de Bellman-Harris, qui prend en compte la variabilité des temps de division entre les bactéries individuelles, nous pouvons obtenir une estimation de la variabilité de taille de population. Nous montrons comment ajouter à ce modèle connu les contraintes expérimentales dues à notre puce, en particulier la distribution initiale du nombre de bactéries par goutte, qui est une distribution de Poisson. Cette variabilité initiale est en fait dominante par rapport à la variabilité inter-cellulaire dans le calcul de la variabilité de taille de population entre les gouttes. Les résultats expérimentaux en puce microfluidique sont en accord avec les prévisions du modèle, non seulement en ce qui concerne la variance mais également pour la forme de la distribution de taille de population entre les gouttes, qui peut être estimée numériquement.

Dans le troisième chapitre, nous nous attaquons au problème inverse, c'est à dire celui de retrouver la variabilité des temps de division des cellules individuelles à partir des mesures de croissance en goutte. Notre précédent modèle ne nous est ici d'aucune utilité, à cause de la domination de la variabilité initiale. Il nous faut donc élaborer un schéma d'inférence qui nous libère cette contrainte. Pour ce faire, nous suivons la trajectoire de chaque goutte au cours du temps, et quantifions la déviation stochastique de cette trajectoire par rapport à une pure croissance exponentielle. Cette déviation, qui dépend du temps, est appelée résidu de la trajectoire. Nous démontrons que la variance de ces résidus dépend fortement de la variabilité des temps

de divisions individuels. Ainsi, en mesurant cette variance, on peut accéder à ces temps de divisions. Ce schéma d'inférence fonctionne sur des simulations reproduisant les expériences en gouttes microfluidiques. Malheureusement, il ne peut pas être appliqué à ces expériences, en raison de la trop grande incertitude existante sur la constante de proportionnalité entre fluorescence et nombre de bactéries dans une goutte. Cependant, les résidus des trajectoires expérimentales se comportent en accord avec la théorie, ce qui démontre la validité de cette dernière. La méthode pourra être utilisée par d'autres groupes de recherche ayant des systèmes expérimentaux où la mesure est plus directe.

Enfin, dans le dernier chapitre de cette thèse, nous montrons comment la puce microfluidique peut être utilisée pour des mesures de susceptibilité de bactéries aux antibiotiques. L'emploi de la puce représente un gain de temps et de place par rapport aux techniques de mesure classiques, grâce en particulier à la possibilité, non utilisée dans les chapitres précédents, de contrôler l'environnement des bactéries au cours du temps. L'antibiotique peut donc être retiré, par exemple, pour mesurer l'influence du temps d'exposition sur la survie des bactéries. Le concept de mesure digitale est développé dans ce chapitre. Une mesure digitale se base sur le comptage du pourcentage de gouttes positives (la goutte est remplie de bactéries) ou négatives (aucune bactérie n'a poussé), pour en déduire mathématiquement combien de bactéries ont survécu au traitement. La puce est également utilisée dans ce dernier chapitre pour étudier le lien entre réponse SoS des bactéries individuelles et capacité de survivre à un traitement antibiotique, la réponse SoS étant une réponse bactérienne induite quand l'ADN de la bactérie est endommagé, en particulier sous l'effet de certains antibiotiques.

Contents

Remerciements **vi**

Résumé de la thèse en français **viii**

1 Introduction **1**
1.1 Cell to cell heterogeneity in bacteria 1
1.2 Microfluidics as a tool to study bacterial heterogeneity 6
1.3 Exponential growth 11
1.4 Plan of the thesis 14

2 Presentation of the experimental platform **17**
2.1 An original microfluidic chip for parallel measurements of bacterial growth 17
2.1.1 Design of the chip 17
2.1.2 A conceptual approach to microfluidics experiments . 21
2.2 Time-lapse measurements of bacterial growth 23
2.2.1 Principle of the experiment 24
2.2.2 Measuring the fluorescence 24
2.2.3 Analysis of the results 30
2.2.4 Volume and size of the colony 31
2.2.5 Fluorescence Heterogeneity 37
2.2.6 Summary of the Chapter 41

3 Distribution of the number of cells with time : the Bellman-Harris model **43**
3.1 Introduction 43
3.2 Theoretical and numerical results 44
3.2.1 Comparing the three models of cell division 44
3.2.2 Classical case studied by Bellman and Harris 47
3.2.3 Poisson distributed initial number of cells 54
3.2.4 Generation-dependent division time 58
3.2.5 Putting it all together: Three sources of stochasticity . . 60
3.3 Experimental results 62
3.3.1 Growth Rate, Mean and Standard Deviation 62
3.3.2 Individual division times 63
3.3.3 Full comparison 64
3.3.4 Towards inference ? 69
3.3.5 Experimental noise 73
3.4 Summary and conclusion 78

4 Following individual trajectories: the residuals **81**
4.1 Introduction . . . 81
4.2 Evolution of the residuals with time . . . 82
4.2.1 Idea and definition . . . 82
4.2.2 Mathematical Properties . . . 83
4.2.3 Choosing Δt . . . 87
4.3 Experimental residuals . . . 90
4.3.1 A slope problem . . . 90
4.3.2 Residuals and noise . . . 90
4.3.3 Experimental sampling . . . 92
4.4 Binning the residuals by the number of cells . . . 94
4.4.1 Definition, numerical results and inference method . . . 94
4.4.2 Back to the experiments . . . 101
4.5 Summary and Conclusion . . . 105

5 Antibiotics and SoS response **107**
5.1 Introduction . . . 107
5.2 Using the chip to test the response of bacteria to antibiotics . . . 108
5.2.1 Multi-Chip MIC testing and digital measurements . . . 108
5.2.2 Exposure time to antibiotics . . . 116
5.2.3 Growth under antibiotic stress . . . 119
5.3 Using microfluidics to study the SoS Response of bacteria . . . 123
5.3.1 A little introduction . . . 123
5.3.2 Preliminary results . . . 124
5.3.3 Pursuing experiments . . . 128
5.4 Summary and Conclusion . . . 136

6 Conclusion **139**

A Materials and Methods **145**
A.1 Bacterial strains and media . . . 145
A.2 Microscopy . . . 146

B Mathematical complements **147**
B.1 Equivalence . . . 147
B.2 Moments of a distribution . . . 147
B.3 Convergence in distribution : definition and lemma . . . 147
B.4 Sum of independent and identically distributed variables . . . 148
B.5 Gamma division times in a Bellman-Harris model . . . 149

C Approximated volume of a big droplet **151**

D Residuals and Noise **155**
D.1 Additive noise . . . 155
D.2 Multiplicative noise . . . 156
D.3 Heterogeneity of the fluorescence . . . 158

E URLs for Supplementary Movies **161**

Bibliography **163**

Chapter 1

Introduction

J'aime pas l'humain, j'préfère les espèces
DAMSO, Introduction à Lithopédion

Heterogeneity is ubiquitous within biological systems. Admiring the extreme diversity of human faces, although they are all representatives of the same species, is enough to be convinced of this. However, human beings are a quite complex and evolved organism, but even at smaller scales, or for simpler biological systems, heterogeneity plays a determinant role. For years, biology was mostly based on measurements of population averages, but recent discoveries have proven that taking heterogeneity into account was not superfluous, if not critical. Cancer, for instance, is known to be a highly heterogeneous disease [1], with huge intra-tumor and inter-tumor variablity, and this poses clinical problems for designing effective treatments from models that are based on population averages [2]. Moreover, cancer is not the only example in which the heterogeneity in a population of cells plays a key role. As we will discuss in the first part of this introduction, it is also crucial to understand the heterogeneity among bacteria. We will see thereafter that microfluidics has emerged as a powerful experimental tool to study this heterogeneity, and finish with a focus on the exponential growth of bacteria.

1.1 Cell to cell heterogeneity in bacteria

Let us start by defining what is meant by heterogeneity in a population of bacterial cells. It encompasses two different scenarios [3]: it can be phenotypic (genetically identical cells have different phenotypes, see Fig. 1.1) or genetic (genetically different cells in a population). Let us first examine the latter. If the emergence and selection of genetic mutants have been explained by Darwin in his classical theory of evolution [4], a strong selective pressure is required in this classical view to select for an individual phenotype corresponding to a genotypic change. But a lot of mechanisms are known to

increase the mutation rates of bacterial species, and favor genetic diversity. A well-known example is the SoS response, that will be described in depth in Chapter 5. It is a bacterial stress response induced when the DNA of the cell is damaged. It promotes, among other effects, the repair of DNA in an error-prone fashion [5], which paves the way for the emergence of resistant mutants [6].

Another mechanism that generates phenotypic diversity through genetic changes is phase variation. It encompasses several molecular mechanism in which some particular loci in the bacterial chromosome are subject to genetic changes at high frequency [7]. Those changes are heritable but reversible, in contrast with mutations selected by selective pressure that are fixed in the population. Some species can phase vary, for instance, to change their surface structure in order to escape the host immune system [8].

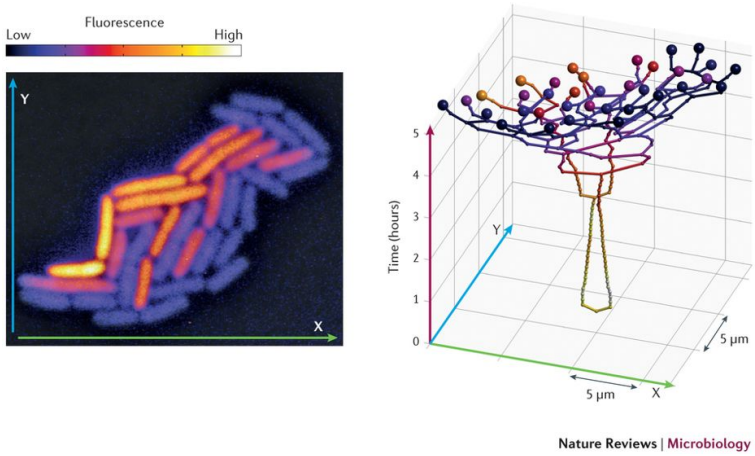


FIGURE 1.1: Taken from [9]. Emergence of phenotypic heterogeneity from a single-cell. Time-lapse analysis with a strain of *Salmonella Typhimurium* carrying a transcriptional gfp reporter for the promoter controlling the expression of the flagellin gene fliC [10]. The left panel shows a still image from a time-lapse microscopy experiment. The right panel shows a reconstruction of the lineage tree, based on the same time-lapse microscopy experiment. The root of the lineage tree is the single cell that founded the microcolony. Branching points correspond to cell division events, and the terminal branches correspond to the 35 cells that are shown in the left panel. The colour of the cells and of the branches represents the intensity of the fluorescence signal. X and Y refer to spatial dimensions, and vertical grid lines indicate 1 hour intervals. Analysis and images courtesy of S. van Vliet and M.Ackermann.

Let us come back now to phenotypic heterogeneity. The most famous example of such heterogeneity in bacterial populations is probably the existence of persister cells, which are individual dormant cells in an otherwise dividing population. It was shown that their existence is due to a phenotypic switch [11], and not to genetic divergence. These cells have the particularity to be more tolerant to antibiotics treatment because of their low

metabolism, this is how they were first discovered by Bigger, as early as 1944 [12]. These cells can switch back to a dividing state once the antibiotic treatment is stopped and may then be a cause of the failure of some antibiotic treatments.

Another example of phenotypic heterogeneity among bacteria can be found in biofilms, which are bacterial communities living on surfaces, created by the self-production of an exopolymer matrix [13]. Thanks to this matrix and the complex spatial organization of biofilms, the cells are protected from external stresses. In particular, biofilms are more tolerant to antibiotic treatments [14], or to host immune responses [15]. They have been shown to be involved in the infection mechanism of several important bacterial pathogens, such as *Vibrio Cholerae* [16]. The heterogeneity of the cells is a key in this case, some individual cells were shown to create the adhesive matrix components for the early stage of biofilm formation, with the rest of the population taking advantage of these early producers.

The question that arises then is to know how this phenotypic heterogeneity can emerge among cells that all have the same genetic background. It all comes down to the fact that a lot of important molecules, such as DNA or regulatory molecules are present in the cells in a very low copy number [17]. Hence biochemicals reactions involving these components are very noisy [18]. These reactions include all the reactions that are necessary for the production of a protein from a gene, such as transcription and translation. Thus single gene expression appears to be noisy, which in turn creates phenotypic variability among isoclonal cells [19]. Two sources of noise have to be distinguished to account for the stochasticity in gene expression. The extrinsic noise is due to fluctuations, from cell to cell, in concentration, location and conformal state of the molecules involved in gene expression (DNA polymerases, regulatory proteins..). But even if two cells had exactly the same molecular content, the random nature of the microscopic events leading to the expression of a gene would still create variation: this is the intrinsic noise. Both have to be taken into account [19].

Once this single-cell noise in gene expression is known, there are a lot of ways in witch it can be translated into phenotypic variability. The simplest is of course that the level of a certain protein gradually depends on the level of expression of a single gene that controls the production of this protein. But a small change in gene expression can also have dramatic phenotypic consequences [20], this is for instance the case for bistable switches, in which a gene expression can have two very different states, the "on" state in which it is activated and the "off" state in which it is almost silent (see Fig. 1.2). Thus, two phenotypically distinct sub-populations can be created. Those bistable motifs can be due to simple gene networks: if a single repressor has a very low on and off rates, for instance, a bistable switch is possible [23]. But they can also be due to more complicated gene networks, including in particular feedback loops. A double negative feedback loop, in which two repressors repress each other's gene, creates bistablity as long as there is cooperativity in between the two repressors [24]. A positive feedback loop can also create bistability in gene expression, as soon as a protein activates its

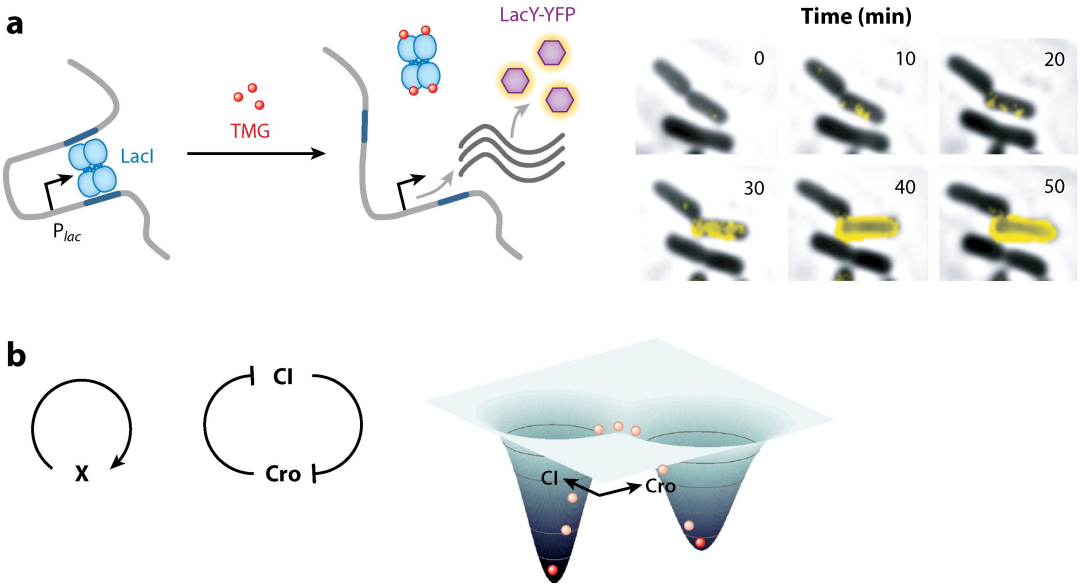


FIGURE 1.2: Taken and adapted from [20]. Mechanisms of stochastic phenotype switching. (a) When exposed to an inducer such as methyl-β-Dthiogalactoside (TMG), individual *E. coli* cells respond in an all-or-none fashion in which only some express the lac operon. This bifurcation in phenotype is mediated by rare complete dissociations of the Lac repressor (LacI) from its multiple operator sites. The resulting large bursts of transcription can be visualized directly by imaging single molecules of LacY-YFP (yellow fluorescent protein) (right; reprinted from [21]). (b) Bistable positive feedback loops in which a key component (X) reinforces its own production underlie many stochastic switches. A common architecture is the double negative loop composed of a pair of mutually repressing repressors such as CI and Cro from phage λ [22]. Once dominant, either repressor reinforces its own expression by repressing its counterpart, which leads to two stable steady states. These states are visualized here as wells in a potential landscape; the system's trajectory (red sphere) continues until it falls into one of the wells, where it is then stuck for long periods of time. Rare, large fluctuations can push the system back out of the well, which allows switching to occur.

own production with great sensibility [25]. Such bistable switches are widely spread among bacteria, and responsible for the expression of virulence factors, for instance in *Vibrio Cholerae* [26]. Another very famous class of bistable switches is the toxin/antitoxin systems, in which a stable toxin is inhibited by a unstable antitoxin, both encoded in closely linked genes [27]. Such systems are known to be involved in the formation of persisters cells [28, 29], and the regulation of a lot of bacterial stress responses [30].

Another characteristic of these bistable switches is that they can create heritable phenotypes without genetic changes [31]. Thanks to the regulary network, a stochastic gene expression can then be turned into a heritable

phenotypical difference. Indeed, one must not confuse phenotypic variability with inheritable variability. In *E. coli*, for instance, a biased partitioning of a multidrug efflux pump, called Acab-TolC, causes distinct and heritable drug tolerance phenotypes in a population [32]. This mechanism is here based on the fact that the efflux pumps are not transmitted equally between mother and daughter cells at division, but biased towards the mother. This is heritable, as the older cell will keep the phenotype.

Understanding the heterogeneity among seemingly similar bacteria is also of critical clinical importance: infections caused by bacterial pathogens often involve heterogeneous populations. A well-known example is the intestinal pathogen *Salmonella Typhimurium*. This bacteria is known to be involved in many food-borne illnesses, it was even found to be the leading cause of deaths due to food-borne diseases in the United States [33]. Its ability to infect hosts relies on the existence a small sub-population of cells that grow more slowly than the rest of the population, but express a virulence factor that is necessary for the early competition with the host's flora [34]. Once an infectious niche has been established thanks to this small sub-population, faster growing cells, that do not express the virulence factor, can take over. This is then an example of heterogeneity leading to cooperation in between the different cell types. Moreover, the bacteria expressing the virulence factor are known to be more tolerant to antibiotic treatment [35], which is another problem to efficiently cure patients. Understanding the heterogeneity in bacterial populations is therefore important for clinical reasons.

Last but not least, bacterial heterogeneity plays a key role in bacterial adaptation to new environments. A heterogeneous population is indeed more likely to survive if the environmental conditions change, as there may be a sub-population of cells that is well-suited for the new environment. It has been demonstrated for both phenotypic [36] and genetic heterogeneity [37]. This strategy is known as bet hedging, as the bacteria bet on future changes in the environment, maintaining less fit individuals in the population to have a better chance of survival if the environment changes.

So far, we have explained the origins of heterogeneity among bacterial cells and given examples of its importance for some microbiology problems, many of which have clinical implications, and also for bacterial evolution. This heterogeneity can create variability in the results of an experiment, which must not be disregarded. Indeed, if carefully analyzed, it can help answer fundamental biological questions. The most famous example of how observing variability was helpful for fundamental biological knowledge is probably the Luria-Delbruck experiment [38]. Luria has observed that the number of resistant mutants to a bacterial phage (a virus infecting bacteria) was highly variable from one experiment to another. Instead of just throwing away the data, he used this observation to answer the question of whether the resistance mutations were randomly acquired prior to the phage attack or would appear as a response to the aggression. Together with Delbruck, they built a mathematical model predicting the variance of the number of mutants in a culture in the two scenarios. It was shown that the observed variability

could only be explained by the random mutation scheme. Luria and Delbruck won the Nobel Prize together for solving this question. The concepts developed by Luria and Delbruck are used today, particularly the deviations from their original predictions. For instance, in a very nice paper, Kessler and Austin [39] showed that the variability of drug-response curves of tumors in patients could only be explained by adding a non-uniform death rate for sensitive cells to the Luria-Delbruck distribution of resistant mutants.

If Luria and Delbruck managed to extract information from the experimental variability with basic experimental tools (they just used plating of bacteria on agar plates), more advanced experimental tools are needed to be able to observe and quantify the heterogeneity at the single-cell level. This was made possible thanks to advances in microscopy techniques and also to the emergence of new experimental tools known as microfluidics, that has emerged during the last 20 years [40].

1.2 Microfluidics as a tool to study bacterial heterogeneity

What is microfluidics ? Let us quote directly one of the founders of the field, GM Whitesides: " It is the science and technology of systems that process or manipulate small (10^{-9} to 10^{-18} litres) amounts of fluids, using channels with dimensions of tens to hundreds of micrometres" [41]. Microfluidics is then at the same time a scientific and a technological field. We will focus on the latter here. The application of this technology to the study of biological systems was natural for one main reason: many biological systems can be studied at scales that are compatible with the length scales evoked by Whitesides. For the case that interests us, bacteria are usually a few μm long, so they can enter microfluidic channels and be studied in such systems. But microfluidic technologies have another decisive advantage over classical ones. Thanks to the miniaturization enabled by the small length scales, a lot of experiments can often be conducted in parallel in microfluidic devices: the throughput is often much higher than with classical experiments. And if a lot of experiments can be followed in parallel, the statistics on the variability of the experiments can be characterized with higher precision. This is why microfluidics is a natural tool to study the heterogeneity of biological systems.

Historically, the firsts examples of using microfluidics to study the cell to cell variability of bacteria were done using single-phase microfluidics technologies [42]. This means that there is only one fluid that is manipulated in those devices. The micro-scale in those systems is then mostly used as a way to geometrically constrain the bacteria to stay in a desired location in space, in order to be able to easily image them over time under a microscope. But the environment of the bacteria can also be finely controlled over time, and this turned out to be a great advantage over classical microbiology techniques. The first example of the use such a device to study cell to cell variability was the seminal study of persister cells by Balaban *et al.* [11]. It dates back to 2004. Straight trenches patterned in polymethylsiloxane (PDMS) were used

to force *E. coli* to grow in those trenches in one dimension, as the width of the trench was just enough for one bacteria to grow, see Fig. 1.3(a). The trenches were sealed with a permeable membrane to allow a precise control of the environment in which the bacteria grow. Thus, the differential sensitivity of individual cells to an antibiotic could be analyzed and persistence was proven to be due to a phenotypic switch. 2D chambers appeared soon after [43] (and Fig. 1.3(b)). Once again, it was made possible thanks to the precision of fabrication of microfluidic devices: the height of those chambers is comparable to the height of a single bacteria, forcing the bacteria to grow in a monolayer, which is perfect for microscopy. The advantage of the 2D device over the linear groove is that more generations can be traced thanks to the two dimensional aspect. Hence, they were first used to track the oscillations of gene expression over multiple generations [43]. However, because of the exponential growth of bacteria, even a two dimensional chamber is completely crowded after some time. The long-term monitoring of bacterial growth required therefore another technological advance.

This is where the most famous microfluidic device for the study of bacterial cell-to-cell heterogeneity comes into play. It is known as the "mother machine", see Fig. 1.3(c). It is once again based on constraining the bacteria to grow in one-dimension, but this time one end of the trench is closed, so that the mother cell is bound to stay in the trench, and the characteristics of the division events of this single-cell can be observed for hundreds of generations [44]. The other end of the trench is open to free flowing media, so that daughter cells get flushed away and the growth condition can be controlled over time. Hundreds of similar linear channels can be made on one microfluidic device, so that hundred of mother cells can be observed in parallel. This device proved to be useful for the deciphering of the cell division mechanism and cell-size control of *E. coli* [45], as we will see in more details in Chapter 2. But the mother machine has had numerous other applications, including for instance fast antibiotic susceptibility testing [46] or noise reduction in synthetic biology networks [47].

More recently, other single-phases microfluidic devices have been developed for the study of cell-to-cell heterogeneity. For instance, the MACS, which stands for Microfluidics Assisted Cell Screening [48]. In this very ingenious device, the cells flow continuously through the device, and are pumped from a reservoir into an observation chamber where the external pressure is very finely controlled. Thus, if the pressure is increase, the PDMS ceiling of the chamber collapses and bacteria get squeezed in between this ceiling and the bottom glass coverslip (Fig. 1.3(d)). They can be imaged with very high resolution (Fig. 1.3(e)): it was shown that single fluorescent proteins could be detected, keeping also a high throughput, as $10^5 - 10^6$ cells can be imaged per hour. The control mechanism of *RPoS*, an important promoter of the *E. coli* stress response could be quantified thanks to this device [49].

Single-phase microfluidic devices have then proven to be very useful in the study of bacterial heterogeneity. However, these devices have a few drawbacks. First of all, as their is only one fluid phase, all the bacteria are mixed up in this single fluid in the microfluidic device. As a consequence, if

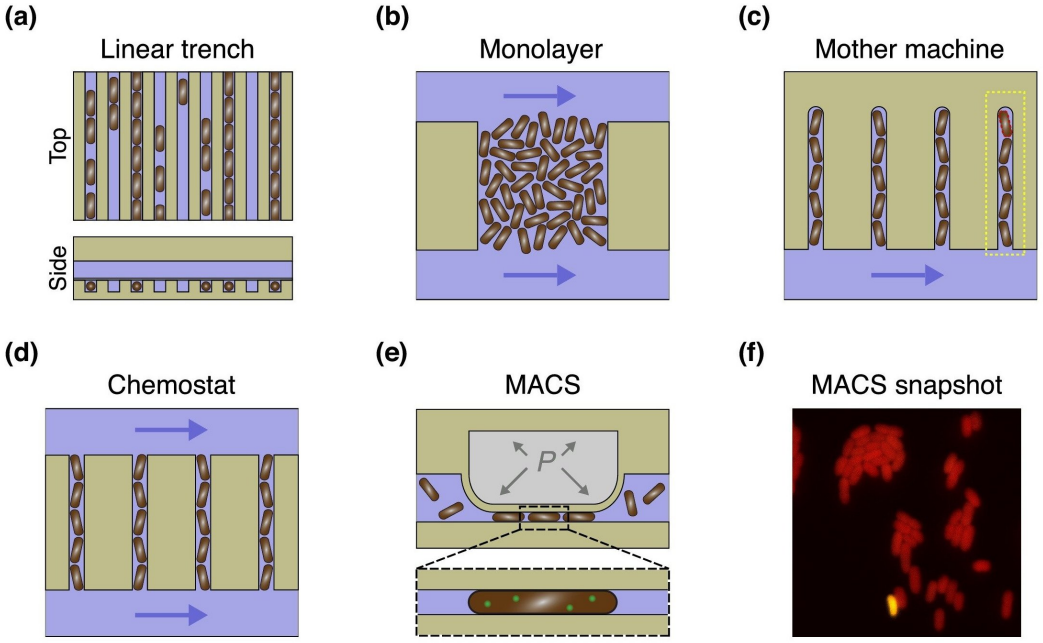


FIGURE 1.3: Taken from [42]. Microfluidic devices for single-cell analyses. (a) Diagram of the first device to support growth of bacteria in linear colonies with continuous time-lapse imaging. Cells grow and divide in strait grooves etched in PDMS (top panel, top-down view). Trenches are enclosed by a permeable membrane and fed diffusely from flowing media above (bottom panel, cross-sectional view). (b) Top-down depiction of monolayer growth device. Bacteria are confined to a two-dimensional monolayer, fed by flowing media on both sides of the colony. (c) Top-down schematic of the mother machine device. Cells are confined to dead-end linear trenches and fed by flowing media. Red dotted line represents segmentation of old-pole mother cells; yellow dotted line marks a single lineage with daughter cells sharing the same mother. (d) Top-down cartoon of the single-cell chemostat device. Bacteria are geometrically constrained to a single dimension, but no old pole accumulates due to the open-ended structure. Flowing media feeds the trapped bacteria and washes away daughter cells. (e) Cross-sectional view of the MACS platform. A soft ceiling is expanded by controlled pressure to trap single bacteria from a liquid culture. Squeezing the cells aids single-molecule detection (inset). (f) Single field-of-view from MACS. Bacteria with a constitutive segmentation reporter (red) are captured and imaged with MACS. A rare transient phenotype is observed in yellow.

one single bacteria is found to have special characteristics through imaging, and the experimenter wants to retrieve it from the device for further analysis (genome sequencing, for instance), this bacteria cannot be isolated from the others. A particular case is the one where this bacteria of interest can be

selected by selectively killing the others, which is the case for bacterial persistence. But even then, all the persisters are recovered together, lacking the ability to perform off-chip single cell analysis. Another drawback of these devices is that one way or another, descendants of the bacteria initially seeded into the observation chamber have to be flushed away. This is a consequence of the exponential growth of the bacteria: the observation chamber is rapidly filled. This implies that only a few generations can be completely tracked. Furthermore, descendants are mixed when they are flushed away, hence lineages are not kept separated, once again, for recovery and off-chip analysis. Finally, if the growth conditions are controlled in such devices, they are often the same for all bacteria, as they are all growing in the same medium. This can be a drawback for screening applications.

We see that this issues could be solved if the bacteria were separated from one another: this is where two-phases microfluidics comes into play. Indeed, in two-phases microfluidics, bacteria are encapsulated in droplets, more often droplets of aqueous medium in oil (although systems based on double emulsion also exist [50]). Hence single bacteria can be isolated from one another when encapsulated in the droplets. Because of the immiscibility of the two fluids, the lineages of these single bacteria are also kept separated. It was made possible thanks to recent chemical advances, such as the introduction of fluorinated oils as the carrier phase. They are biologically inert, limiting in particular the diffusion of biological molecule from one drop to another [51]. Biocompatible surfactants that stabilize the droplet interface while not interacting with the biological processes inside the droplet have also been developed [52].

The idea to grow bacteria in droplets is in fact not that new. Back in 1954, Lederberg already described a protocol to isolate and observe single bacteria in droplets placed on a glass slide covered with mineral oil [53]. Antoinette Gutman and André Lwoff also cultured bacteria in micro-droplets at Institut Pasteur in the 1950s [54]. What interested these precursors was the ability to encapsulate a single-cell and to observe the division tree emerging from this single cell. But they didn't have the multiple possibilities offered by modern droplet microfluidics.

What are these possibilities ? They can be summarized in three axes [55]. First, huge numbers (up to 10^6 [56]) of monodisperse droplets can be generated in parallel. The state of the art attains currently the generation of droplets at more than 10^4 Hz [57]. These droplets can be used as reactors for the parallel study of biological reactions. Second, these droplets can be monitored, sorted, recovered off the chip for further analysis. The fluorescence of up to 250 000 droplets per second can thus be monitored [58]. The sorting techniques are diverse, relying on electromagnetics, accoustics, mechanics... (see [59] for a recent review). Third, their chemical content can be controlled over time for complex study of bacterial responses. Of course, not all droplet microfluidic devices combine these three advantages. Let us examine though how some of them combine these possibilities for the study of cell to cell variability.

The high-throughput is of course a great advantage to study the cell-to-cell variability. Cottinet *et al.* [60] used for instance the parallel growth of hundreds of colonies from single *E. coli* cells in a millifluidic device to study the influence of starvation on the growth phenotype. Thanks to the high throughput, the variability of growth phenotypes could be proven and quantified. Furthermore, thanks to droplet-sorting, droplets with different growth phenotypes could be separated, and sequenced off-chip to detect changes in the genome of the bacteria. Using the same system, the encapsulation of single algae cells in droplets demonstrated the existence of a subpopulation of slow-growing cells, that would have been difficult to detect in bulk growth, as starting from a hundred cells in the droplet completely hides this phenotype [61]. It seems this time that this is due to a stochastic phenotypic switch rather than a different genotype.

Antibiotics susceptibility at the single cell level can also be explored using droplet microfluidics. The diversity of responses at the single-cell level can be quantified, and analyzed off-chip for the determination of the resistance mechanism [62]. Contrary to single-phase microfluidics, the morphology of the colonies growing from those single cell under antibiotic pressure can also be observed [63]. The control of the composition of the droplets over time can also be of great use for the study of antibiotics. The adaptation of a bacterial population to a changing concentration of antibiotics over time was for instance demonstrated [64].

Encapsulating bacteria in droplets also allows to separate one species from another in a sample of unknown composition [65]. It may allow the isolation of rare or very slow growing species from the environment, which is one of the most exciting challenges of microbiology, as most of the earth microbiome is for now "unculturable" [66]. The species variability is here investigated.

Last, droplet microfluidics is useful because it allows the screening of single cell contents in a very short time. The classical genomics techniques can nowadays be applied to single cells, which allows to explore the cell-to-cell variability in gene contents in a population. Real-time PCR was performed on different species of *Mycobacterium Tuberculosis*, that are of clinical importance, and identification of the different species was possible in a very short time [67]. Uniform amplification of the genome of single *E. coli* cells is also possible in droplets [68].

There are of course a few drawbacks of using droplet microfluidics. The one that matters the most for us is that compared to the single-phase devices that were evoked before, the growth of bacteria in droplets is not stringently spatially constrained. Bacteria can grow anywhere they want in the droplet, and thus it is much more difficult to track single cells, for instance. Usually, what is obtained with droplet-based microfluidics is then a measurement at the level of a small population growing in the droplet. So if we want to obtain some single-cell information from such experiments, they will have to be statistically inferred from population-level measurements. This approach will be developed in this thesis, in particular in Chapter 4.

To summarize, we have seen that both single-phase and droplet based microfluidics could be useful for the study of the heterogeneity of populations of bacteria. In the next part, we will show how this happy wedding could be used to study one case of cell-to-cell variability, which interests us in this thesis: the case of cell-to-cell variability in exponential growth.

1.3 Exponential growth

The question that arises at this point is why would we choose to focus on the exponential growth of bacteria ? Many answers can be given. The first one is a practical question. It has been known to micro-biologists for a long time that in order to prepare bacterial samples in a reproducible way, bacteria have to be in exponential phase. In order to build predictive models, this is then what has to be done. Another answer that we could give is that it has been shown that bacteria grow exponentially for several days in a host during an infection. Exponential growth then has a clinical importance.

But if bacterial physiologists have been obsessed with the growth curves of bacteria in exponential phase for more than 50 years, as ironically admitted by Neidhardt in a very nice commentary [69], it is for more fundamental reasons. It all began with Monod, who first put the bacterial growth in equation [70] (see also Fig. 1.4(D)), and declared: "The study of the growth of bacterial culture does not constitute a specialized subject or branch or research: it is the basic method of Microbiology". A lot of research in the physiology of bacterial cells has followed this intuition. In particular, the Copenhagen School, led by Maaloe, has shown that a lot of characteristics of the cells were directly dependent on the growth rate. The cell size, as well as its ribosomal content, are hence uniquely determined by the growth rate [71], which only depends on the quality of nutrients of the medium.

These growth laws, relating the ribosomal content and size of the cells to the growth rate, have recently been revisited as a analogy to Ohm's law in electronics [73] (see Fig. 1.4). In fact, these relations indicate that the protein content of a cell is determined by the growth rate [72]: gene expression and growth are thus intrinsically linked, and studying the growth of bacteria can thus help understand gene expression. But this is not the only application of these growth laws: they have been successfully used, for instance, to predict the effect of ribosome-targeting antibiotics on the growth of *E. coli* [74].

Nevertheless, these relations do not include any cell-to-cell heterogeneity. It is precisely by studying the single-cell deviations from these growth laws in the "mother machine" that the question mechanism of cell division in *E. coli* was recently deciphered [45]. The single-cell deviations from these growth laws are then of critical importance. It has been shown, also, that at the single-cell level, stochasticity of metabolism and growth are related [75]. Learning more on the variability of growth during exponential growth could hence help biologists to learn more on the variability of gene expression.

In parallel with these biological discoveries, exponential growth has received a lot of attention from mathematicians. Bellman and Harris, as early

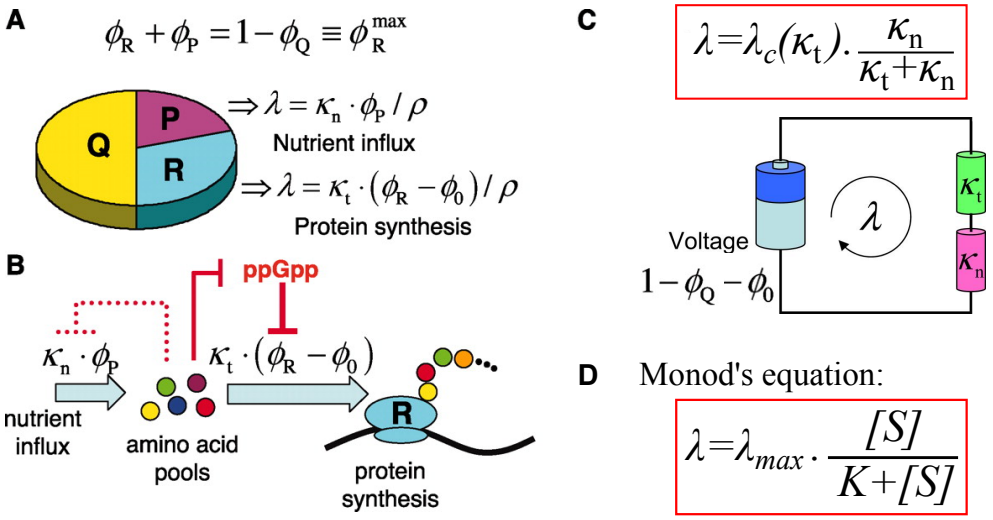


FIGURE 1.4: Adapted from [72, 73]: presentation of the modern growth law and comparison with Monod's classical equation. (A) The growth theory comprises three key ingredients: (i) a three-component partition of the protein content of the cell, consisting of a fixed core sector (Q) and two adjustable sectors (R and P) whose fractions (mass fractions: ϕ_R and ϕ_P) must add up to a constant (Embedded Image). (ii) A ribosomal fraction ϕ_R containing all the ribosomal proteins and their affiliates and exerting a positive effect on growth (with growth rate $\lambda \propto \phi_R - \phi_0$). ϕ_0 is an offset depending on the bacterial strain. κ_t is the "translational efficiency", it quantifies how efficient a microorganism is at translating DNA into proteins. ρ is the coefficient of proportionality between the mass fraction of ribosomal proteins ϕ_R and the ratio of the number of ribosomes to the total number of proteins. (iii) A remaining fraction ϕ_P exerting a similarly positive effect on growth (with growth rate $\lambda \propto \phi_P$) by providing an influx of nutrients. This fraction includes all the metabolic proteins. κ_n reflects the nutrient quality of the medium and is known as the "nutritional efficiency". (B) During steady-state exponential growth, efficient resource allocation requires that the nutrient influx ($\kappa_n \cdot \phi_P$) be flux-matched to the amino-acid outflux $\kappa_t(\phi_R - \phi_0)$. This can be coordinated by ppGpp, which up-regulates ribosome synthesis and hence amino acid outflux in response to increase in the amino acid pools, and has the opposite effect in response to decrease in the amino acid pools. (C) Analogy with Ohm's law: the value of the growth rate obtained by mixing the three previous equations, is mathematically identical to the description of electric current flow through a pair of resistors connected in series to a battery with fixed voltage $(1 - \phi_Q - \phi_0)$. In this analogy, the growth rate λ is the current through the resistors. $\lambda_C = (1 - \phi_Q - \phi_0) \cdot \kappa_t / \rho$. (D) Monod's classical equation relating the growth rate of a strain λ to the concentration of the limiting sugar $[S]$ and the maximal growth rate λ_{max}. K is the saturation constant. K and λ_{max} depend on the strain and the medium. We see that there is a formal analogy with equation (C). In fact, Monod's equation can be retrieved from (C) by supposing Micaelis-Menten kinetics for the transport of a sole sugar source inside the cell.

as 1952 [76], have set the mathematical basis of a description of exponential growth, taking into account the variability of division times of the cells. In this model, the cells are considered to be independent, and their division times are picked from a fixed distribution. In particular, the population growth rate and the mean growth could be computed for multiple distributions of the division times. This model was used soon after by Powell [77], who derived the age distribution of the bacteria in the culture and compared it to experimental data. He also compared the predicted growth rate of the population to the experimental one.

What is interesting about these early models of exponential growth is that they include a cell-to-cell variability. However, what was lacking is that this cell-to-cell stochasticity was not linked to any biological mechanism of cell division nor biochemical mechanism of noise production. Nevertheless, such phenomenological models are still used nowadays [78, 79, 80, 81], because they can be simulated easily. The link between the variability of individual division times and the variability in population size can thus be explored numerically [80]. The increase of the initial number of bacteria has also been shown to reduce the variability in terms of population size [81, 78]. It is also possible to also add to this simulation scheme a death rate [78] and study its effect on the population growth. We will see how such a model can be adapted to match our experimental constrains in Chapter 3, in order to be able to compare the prediction of this kind of model to experimental data, which is generally lacking in those simulation-oriented articles.

It is only very recently that the gap between population dynamics and biologically-relevant single-cell division mechanism has begun to be filled. In particular, Lin and Amir [82] have proposed a model that takes into account the recent biological findings on the cell division and size control mechanisms and include variability at the single-cell level. They have studied the effects of this model on the population growth rate. Another very interesting approach was developed by Iyer-Biswas in a series of articles [83, 84, 85, 86]. She describes of the cell cycle as a closed circle of stochastic autocatalytic reactions (reaction products catalyze the next reaction in the circle) [83]. Some scaling laws can be retrieved from this description [84], in particular the fluctuations in cell division times, that are found to scale with the mean division time. These scaling law were found to be in agreement with single cell observations on *Caulobacter Crescentus*. Similar scalings were also observed on *E. coli* [87]. To go further, the link between single cell variability and population dynamics was investigated in the case of an asymmetric cell division such as *Caulobacter Crescentus* [86]. The authors also evoke the possibility to use these relations to infer data on the single-cell from population measurements, but without any experimental confrontation.

We have already evoked this inference problem when discussing the drawbacks of droplet-based microfluidics. It is remarkable to note that it has been discussed for a long time in the mathematical literature, soon after the Bellman-Harris model was exposed. The estimation of the parameters of the single-cell distribution of division times from the mean population growth

was proven to be possible, in the case of a Gamma distribution of the division times [88], but not really applied to experimental cases: the point was more to prove the mathematical convergence of the estimators. Other approaches were developed, but for a continuous observation of the division process [89, 90], which is difficult to obtain experimentally. For discretely observed branching processes, maximum likelihood estimators were built, even in the case of multi-type processes [91], but they are usable only if the explicit distribution of the number of cells is available, which is not the case in a Bellman-Harris process. A pseudo-likelihood estimator, based on numerical simulations for the estimation of the time-dependent distribution of the number of cells, can then be used [92], and we will use it in Chapter 3.

This question of inference of single-cell parameters from the observation of dividing populations has regained a lot of attention in the last years [93], as the recent experimental developments allowed the production of more and more of these kinds of data. The fitness landscape of a resistance-conferring protein under drug pressure was for instance inferred from fluorescence data and observation of the division tree of *E. coli* [94]. In the context of gene expression, exploiting the structure of the lineage tree can lead to estimations of stochastic cell fate-switching parameters [95]. The observation of the division tree can also be used to get gene expression parameters for each single cell [96]. Notwithstanding, all these inference schemes rely on the the direct observation of the lineage tree, which is not what is usually obtained with droplet based measurements. We will see in chapter 4 how an inference problem on the parameters of the distribution of the division times of individual cells can be solved without any observation at the single-cell level.

1.4 Plan of the thesis

This thesis is composed of 4 chapters following this introduction.

In Chapter 2, we will present the microfluidic chip used throughout this thesis for microbiology experiments. We will expose the droplet generation technique that was developed in the lab and the way it was transferred into an easy-to-use microfluidic chip, where more than a thousand growth experiments can be followed in parallel. The rest of the chapter is dedicated to the technical questions related to the growth of bacteria on the chip, starting by explaining how a growth experiment is performed. We will then explore how bacterial growth can be quantified by measuring the fluorescence of the droplets, how the experimental noise has to be treated. We will also expose the difference between on-chip and batch growth, and try to quantify the volume heterogeneity of the droplets, as well as their shrinkage during an experiment.

The following chapter 3 is dedicated to the description of a stochastic model of the growth of the bacteria from a few cells to small colonies in the droplets that encompasses cell to cell variability. The exponential phase of the growth will be modeled, by extending the Bellman-Harris model to fit our experimental constrains. We will show that the variability in the early stages of growth, which is due to our experimental constrains, dominates

over the cell-to-cell variability to account for the observed variability in population size. Those theoretical predictions will be compared with the experimental results of growth of *E. coli* and *B. subtilis* on-chip. Besides, we will see that because of our experimental constrains, it is difficult to use the results of our extended Bellman-Harris model to learn information about the single cells from our populations based measurements.

This inference problem will be tackled in another way in Chapter 4: we will try to obtain the cell-to-cell variability of division times from the observation of the growth in droplets. Instead of using the distribution of the number of cells at the observation points, we will develop a theoretical and numerical inference approach based on following the trajectory of each droplet over time. The deviation from a pure exponential growth will be quantified and linked back to the cell-to-cell variability. This relation will be used for successful inference on simulations, and we will study how it could be used with experimental data.

The last chapter is dedicated to a slightly different subject: we will try to show how our microfluidic chip can be used to investigate how the bacteria respond to an antibiotic stress. We will show that the experimental possibilities offered by the chip can be of great use in this context. For instance, we will show how digital measurements can be used to study the survival rate to antibiotics varying the concentration and the exposure time, and compare the measurements that we get with our chip to more classical measurements of antibiotic susceptibility. Finally, we will demonstrate how our chip can be used for the study of a particular bacterial stress response to antibiotics, known as the SoS response.

Chapter 2

Presentation of the experimental platform

Jme présente, je m'apppelle Henri
Michel

In this chapter, we describe the microfluidic chip used throughout this thesis for microbiology applications. The emphasis will be on the technical questions related to the chip: what is measured, how, and what are the limitations when we run an experiment. But we will also try to conceptualize what it means to conduct experiments in a device like this one. We will describe as well the typical growth experiments that will be analyzed in more depth in the following Chapters.

2.1 An original microfluidic chip for parallel measurements of bacterial growth

2.1.1 Design of the chip

The breaking process

The microfluidic chip that is used in this thesis was designed in our lab and is described in-depth in [97]. We are just going to give here its main characteristics for the reader to understand the experiments presented later on.

To be concise, the chip comes down to a big chamber, with two inlets and one outlet, that contains a static array of about 1500 nano-liter sized water-in-oil droplets immobilized in anchors, or traps, in which cells can be encapsulated (Fig. 2.1(A),(B)). Its originality lies in the way by which the droplets are created. Indeed, the droplets are created directly on the anchors thanks to the so-called "breaking" technique [98], which is based on a surface-tension effect. In fact, the anchors are holes that are patterned in the PDMS ceiling of the chamber, and they act as wells of surface energy [99].

This principle was at first designed to keep a droplet trapped in an external oil flow [100]. Let us imagine a big microfluidic chamber, with a small height difference between the glass floor and the PDMS ceiling, and infinite dimensions in the other directions. If a big water-in-oil droplet is introduced in such a chamber, it will be squeezed between the floor and the ceiling and adopt a pancake-like shape. However, the most energetically favorable shape for a droplet is a sphere, which minimizes its surface and thus its surface energy. If now a hole is patterned in the ceiling of the chamber, the the droplet can enter the hole. By doing so, it releases some of its surface energy thanks to the height difference: if the dimensions of the trap are correctly chosen, the surface of the trapped droplet can be smaller than its surface when squeezed in the chamber. Thus it is more energetically favorable for the droplet to remain in the trap, and this trapping force can be quantified.

The same principle is extended in the chip to create droplets in the traps. In fact, let us go back to the big droplet trapped in the anchor that we have described in the last paragraph. If the external oil flow is increased above the trapping force, under certain conditions the droplet will break into two parts. Most of it is flushed away by the oil flow, and a small part remains trapped in the anchor, creating a smaller droplet with a predictable volume [98].

This is then an efficient and robust way of creating droplets directly on the anchors, and the chip used here is the generalization of this simple principle to a an array of 1500 anchors in a big chamber. Hence, an array of 1500 water-in-oil droplets can be created in less than ten minutes.

The method to create the droplets is described in Fig. 2.1(C),(D). First, the chamber is completely filled with fluorinated oil through inlet 1. Then, it is filled with a suspension of bacteria at the desired concentration in bacterial growth medium through inlet 2. This creates a big plug of aqueous solution in the chamber and the anchors, surrounded by an oil lubrication film (thanks to the first step). Finally, an oil flow is applied again through inlet 1. It pushes the aqueous plug and "breaks" it, leaving a droplet in each anchor, encapsulating at the same time the bacteria in these droplets.

The multiple possibilities of this microfluidic device are fully described in [97], we will just give the reader a quick summary here. First of all, this whole "breaking" process can be realized with liquid agarose instead of pure culture medium. After the loading, the agarose can be gelified, and the external oil can be removed by replacing it with an aqueous solution. Hence, the droplets containing bacteria can be perfused with any solution: the microenvironment of the bacteria can be controlled over time. For instance, antibiotics can be added to study the stress response of bacteria. We can also perfuse the chip with a gradient of any compound, with a different concentration for each row of the droplet array. Finally, with the agarose droplets, a particular droplet can be selectively retrieved from the chip: we can precisely melt it with a laser and retrieve its content by pushing it out of the chip with an external oil flow. Its content is then viable and available for any off-chip analysis (sequencing, susceptibility testing..).

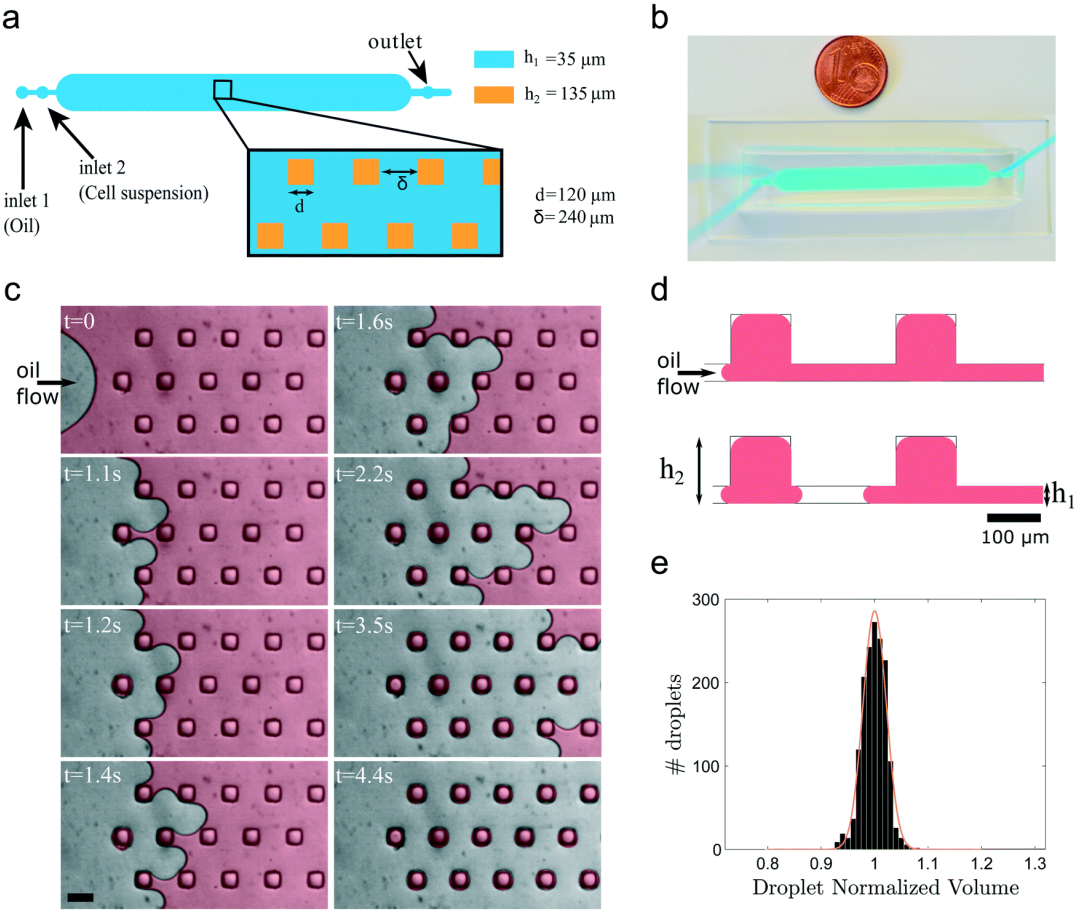


FIGURE 2.1: Taken from [97] Description of the microfluidic device and protocol for droplet production: (a) design of the microdroplet multiwell device. The central chamber has dimensions 0.5 × 4.8 cm and contains a 2D array of 115 × 13 surface-tension anchors. Square anchors have side dimension $d = 120\,\mu$m, spaced by $\delta = 240\,\mu$m. The chamber height is $h_1 = 35\,\mu$m and the anchor height $h_2 = 135\,\mu$m. (b) The device, which fits on a microscope slide, is connected to two inlets and one outlet. (c) Time-lapse of the drop formation process. At t = 0, the cell sample fills the microfluidic chamber entirely, and is being pushed by fluorinated oil (FC40 + 0.5% surfactant) using a hand-pushed syringe. The arrow indicates the oil flow direction. When the interface penetrates between two anchors, it deforms and then breaks, which leaves a well-calibrated droplet in the anchor. The cell sample is colored in red for better visualization. Scale bar: 200 μm. (d) Cross-sectional schematic of the breaking process on anchors. The aqueous sample initially fills large regions and then gets divided into isolated droplets that fill each of the anchors. (e) Experimental histogram of the normalized droplet volumes on one chip. The orange line is the best Gaussian fit to the data, leading a standard deviation $\sigma = 0.02$.

Encapsulation of the bacteria

When bacteria are diluted in suspension, and then encapsulated in small droplets, the number of bacteria per droplet follows a Poisson distribution [101]. This is true as long as there is no interaction between the bacteria.

Theoretically, this Poisson distribution comes from the simple fact that the probability for a given bacterium to be encapsulated in a given droplet of volume v, taken from a bigger volume of medium V, is just the ratio of the volumes v/V. In maths, this is known as a Bernoulli trial: a random experiment with exactly two possible outcomes, "success" and "failure". The probability of getting a success (the bacterium is encapsulated in the droplet) is the ratio of the volumes v/V.

If there are N bacteria passing through our microfluidic chip during loading, then the number of bacteria in a given droplet is the sum of N independent Bernoulli experiments of parameters v/V: it follows a binomial distribution of parameters $(N,v/V)$, and the product of the two parameters is $\lambda = c_0 v$ where c_0 is the concentration of bacteria in the medium. Then, if the bacteria are independent (which we believe to be true as they are diluted), since $N \gg 1$ and $v/V \approx 10^{-3}$, we can apply Le Cam's theorem [102], which states that the number of bacteria per droplet converges in law to a Poisson distribution of parameter λ.

The probability of finding k bacteria in a droplet is then:

$$\mathbb{P}(N_0 = k) = \frac{\lambda^k e^{-k}}{k!}. \tag{2.1}$$

Amselem *et al.* [97] have checked that this was verified in our chip by counting the final number of colonies in each droplet.

One important thing to note about the Poisson distribution is that the probability of having an empty droplet is:

$$\mathbb{P}(N_0 = 0) = e^{-\lambda}. \tag{2.2}$$

As we can see, there is a direct relationship between this probability and the Poisson parameter λ. This is very useful for us, as then the number of empty droplets on the chip can be used as an estimator for finding the Poisson parameter used in a experiment:

$$\hat{\lambda} = -\ln(\hat{p}_0) = -\ln\left(\frac{N_{empty}}{N_{tot}}\right), \tag{2.3}$$

where N_{empty} and N_{tot} are the number of empty droplets and the total number of droplets, respectively, and $\hat{p}_0$ represents the estimated probability of having an empty droplet.

The 95% confidence interval on the estimated value of $\hat{p}_0$ is then [103]:

$$\hat{p}_0 \in \left[\hat{p}_0^+, \hat{p}_0^-\right] \quad \text{where} \quad \hat{p}_0^{\pm} \approx \hat{p}_0 \pm 1.96\sqrt{\frac{\hat{p}_0(1-\hat{p}_0)}{N_{tot}}}, \tag{2.4}$$

which we can translate into a confidence interval on $\hat{\lambda}$:

$$\hat{\lambda}^{\pm} = -\ln\left(\hat{p}_0^{\mp}\right). \tag{2.5}$$

In our case, we have approximately 1000 droplets, and we can plot the ratio of the width of the confidence interval over $\hat{\lambda}$ (see Fig. 2.2). The width of the confidence interval is smaller than 20% of $\hat{\lambda}$ only if we are between $\lambda \approx 0.5$ and $\lambda \approx 4$. This gives us a range of concentration that can be used if we want to be able to estimate λ, as the droplet size is approximately 2 nL: the range of concentration goes from 2.5×10^5 to 2×10^6 cells/mL. Of course, it is not always of interest to be able to evaluate λ, as we will discuss later on.

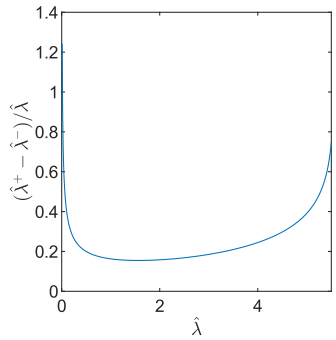


FIGURE 2.2: Relative confidence interval when estimating λ by counting the number of empty droplets

2.1.2 A conceptual approach to microfluidics experiments

Many different experiments can be conducted with this microfluidic platform. However, the interpretation of the results is often delicate because of this Poisson distribution and the fact that the number of cells per droplet is not known exactly but as a statistics. Experiments have to be well designed to avoid mistakes. We will try to explore here the different possibilities.

In the limit of very low λ, the chip allows us to encapsulate almost only single cells. But as λ is very low, the number of empty droplets will be large. For instance, if one wants to statistically achieve that 95% of the droplets contain only one bacteria, she has to choose $\lambda \approx 0.1$. In this case, the expected number of droplets containing bacteria is only ≈ 100. The statistics of such an experiment are then not the same at all as when the whole 1500 droplets are filled with bacteria. This relatively low number of droplets can be sufficient for a lot of applications, such as isolating unknown bacterial species in a sample for further analysis [104]. But it seems to be too low if one wants to test the systematic response of single cells to antibiotics, for instance, where the heterogeneity plays a central role [105] and makes it often preferable to have a lot of data. Studies using other microfluidic technologies, like the well known "mother-machine" [44] seem to be preferable in this case [106].

However, the droplet technology that we have developed has an advantage over technologies resembling the "mother-machine": the progeny of the initial cell is kept in a closed micro-reactor and not thrown away. In a sense,

the lineage of the first cell encapsulated in the droplet can be tracked for a much longer time. However, we are in our droplets in a very peculiar situation: this lineage can be enclosed, but the division tree is not accessible, as the bacteria don't grow in 2D and cannot be individually tracked. The only observable that we have, if we want to take advantage of the relatively high throughput of our droplets, is the number of cells per droplet as a function of time, that is measured indirectly through the fluorescence signal (see next section). Our advantage, once again, is that the whole lineage is kept. The stochastic effects of the division at each generation accumulate in the droplet, even if we will see in the next chapter that the early stages dominate over the later generations. In a 2D system, because of the exponential growth, the bacteria very rapidly invade the whole chamber, and only a few generations (up to 7) can be tracked [107], or part of the descendants has to be thrown away [87]. In our droplets, we can observe, for *E. coli*, we can estimate that more than 12 generations can be observed, simply because the growth is exponential for $\approx$ 250 min, with a mean division time of $\approx$ 20 min.

In our experiments, we often choose the initial concentration of bacteria such that $\lambda \approx 1$. This ensures that only $\approx$ 30% of the droplets are empty, while keeping the number of bacteria per droplet quite low, to still have a trace of the cell-to-cell variability. Indeed, if the initial number of bacteria per droplet is too high, the population-size variability due to the stochasticity of the division process is hidden, as we will see in the next chapter.

As such, experiments in micro-droplets can be seen as an *experimental Monte-Carlo process*: it gives us the ability to repeat and follow in parallel the same stochastic experiment a large number of times. The distribution of the responses in the population can be obtained thanks to the statistics of this system, especially when compared to classical microbiology methods. We know through the Central Limit Theorem that the convergence to the solution in this kind of method goes as $\sqrt{n}$, where n is the number of repetitions of the experiment. Going from a 96 well plate to a chip like ours with $\approx$ 1000 droplets increases then the precision on the final solution by a factor ≈ 3.

There are in fact two ways of applying this Monte-Carlo analogy to our droplets. The first one is to see the droplets as micro-reactors, a little bit like an expanded 96 well plate, with more and smaller wells, and to use them as growth reactors as we have explained above: the stochastic effects accumulate in the droplet, and thanks to the big number of droplets, the distribution of growth responses can be obtained. This is the approach that we will develop in Chapters 3 to 4. The second one is to have a "digital biology" approach [108], and to use the droplets as binary probes, 0 or 1, for instance to test the survival rate of the bacteria to a stress (antibiotic, oxidative..). This kind of approach is already widely used, for instance, for digital PCR [109]. We will see in the last chapter how it can be applied in our case.

2.2 Time-lapse measurements of bacterial growth

We are going to focus now on the experiments of bacterial growth conducted with this microfluidic platform.

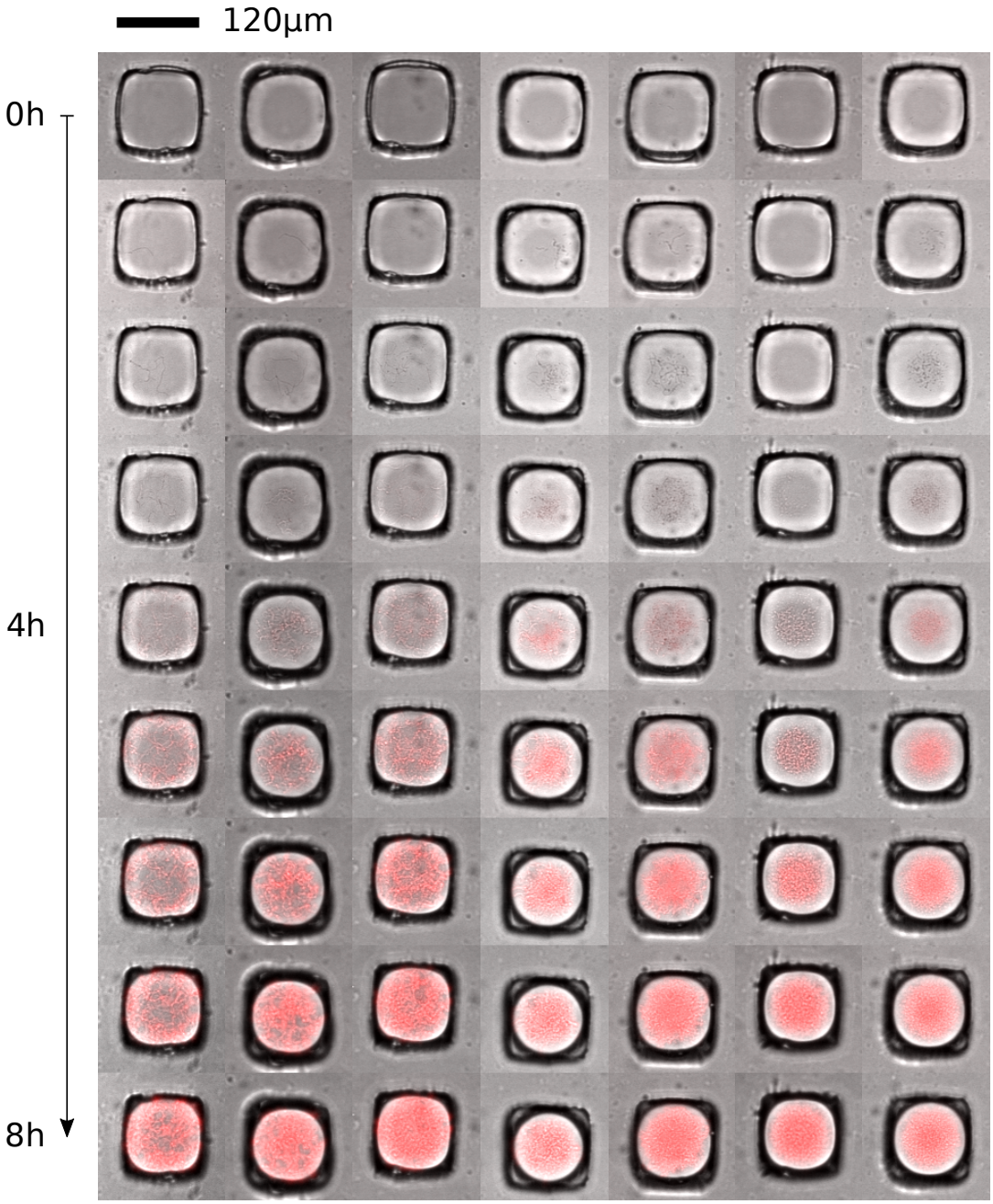


FIGURE 2.3: Example of the growth of *E. coli* in 7 droplets of our microfluidic chip, at 37 °C in LB medium. Fluorescence and bright field images are superposed, with artificial red color for fluorescence.

2.2.1 Principle of the experiment

The principle of the experiment is the following. The chip is loaded with bacteria as described above, with a Poisson parameter λ close to 1. The bacteria come from an overnight culture, and are re-diluted in fresh medium in the morning. They are loaded on the chip in exponential phase ($OD \approx 0.2$). Then the chip is placed in a temperature-controlled motorized microscope at 37°C. The chip is scanned every 5 min, with a 10X magnification. The images are acquired with an EMCCD camera (ANDOR), with an exposure time of 40 ms and a numerical gain of 100. Those values are a compromise between the early stages and the final state. In the early stages, we want to increase the measurement sensitivity as there a few bacteria in the droplets, by increasing the exposure time and/or the gain. However, if we do not want to get completely saturated images at the end, when there are a lot of bacteria in the droplets, we must keep those settings not too high.

By running this experiment, we wish to obtain growth curves of the bacteria in each of the droplets that reflect, as least partially, the cell-to-cell variability of the population, as we will study in the next chapter. We can see on Fig.2.3 that the growth is indeed variable in between the droplets. We can also observe that the droplet are shrinking throughout the experiment, with variability as well. We will come back to this point at the end of this chapter. Similar results can be obtained for *B. subtilis*, and with agarose gel instead of pure liquid medium inside the droplets, as displayed on supplementary movies 1 and 2 (see links in Appendix E).

Note that the whole chip can be scanned, but it takes approximately 12 minutes to do so (for one fluorescent channel plus the bright field images). We were more interested about reducing the time interval between two observations than by scanning the 1500 droplets at each point, which explains why only ≈ 900 droplets were scanned for this experiment.

2.2.2 Measuring the fluorescence

Once the data of our growth experiments have been acquired, we would like to analyze them. For each time point, two images of the chip are taken, one in fluorescence that we will use as a proxy for the number of cells in the drop, and one in bright field that we will use for the detection of the wells. Indeed, to gain time, we do not take one image per trap, but the whole chip is scanned with several traps on one image. We have written a home-built Matlab code to process the bright-field images, detect and track the wells over time, for a full description see [110]. We will focus in this part on the fluorescence measurement.

As we have already said, we want to use the fluorescence measurement as a proxy for the number of cells in the droplets. As each cell is fluorescent, it is intuitive to think that the more cells there are in the droplet, the more intense the fluorescent signal. The fluorescence signal that we measure is then the sum of the contributions of the fluorescence signals of all the cells in the droplet. If we simplify things even more, all cells have the same fluorescent

signal and then we obtain a proportionality relation between the number of cells per droplet and the number of cells in the droplet:

$$Fluo = a_f N, \tag{2.6}$$

where N is the number of cells in the droplet, $Fluo$ is the intensity of the fluorescence signal that we measure, averaged over the droplet, and a_f is the proportionality coefficient between fluorescence and number of cells.

Of course, in real life things are not that simple. First of all, all cells do not have the same fluorescent signal, even in an isogenic population [111]. This heterogeneity in the fluorescence signal of the cell will be discussed later on. Second, we have to deal with the existence of a measurement noise, and we are going to see how in the next paragraph.

Dealing with the background

Indeed, even outside of the droplet, the measured fluorescent signal is not strictly equal to zero, and this background can vary from image to image, but also between time points. If we want to be able to compare one well with another, we have to remove the contribution of this background, which is due to the experimental noise. This is especially important for the beginning of the growth curves, when the number of bacteria is low, and the fluorescent signal of the droplet is quite low as well. The question that then arises is how to cope with that background, and two solutions come to mind: we can divide the fluorescence signal of the droplet by the mean background value, or we can subtract the background value to the signal.

Let us examine an example to see which of the two solutions is better. First of all, we can see (Fig. 2.4(B)) that the background signal changes over time, and that its variations are correlated with variations of the measured signal of the droplet. If we take the difference between the two signals, those variations due to the measurement noise are not completely removed, whereas if we take the ratio, the obtained curve is much smoother (Fig. 2.4(C)). This is even more clear if the same curves are plotted in log scale, which makes sense as the growth of the bacteria is supposed to be exponential, at least while nutrients are abundant in the droplet [112]. We can see a quite noisy curve if we subtract the background (Fig. 2.4(E)), whereas the curve obtained by taking the ratio displays a well defined exponential phase (Fig. 2.4(F)): the noisy variations due to the experimental acquisition are removed from the signal. These observations lead us to think that the ratio is the best way to deal with the experimental noise, which seems to be multiplicative in this case.

However, if we want our signal to be purely proportional to the number of cells in the droplet, then the signal must be translated to zero when there are not bacterial in the droplet, or at least while the number of bacteria is too low to be detected. For now, it is set at 1 as the fluorescence signal of the droplet was divided by the fluorescence signal of the background. The easiest way to translate it back to zero is then to subtract 1 to the obtained signal. We can note that the effective computation that we make in this case is

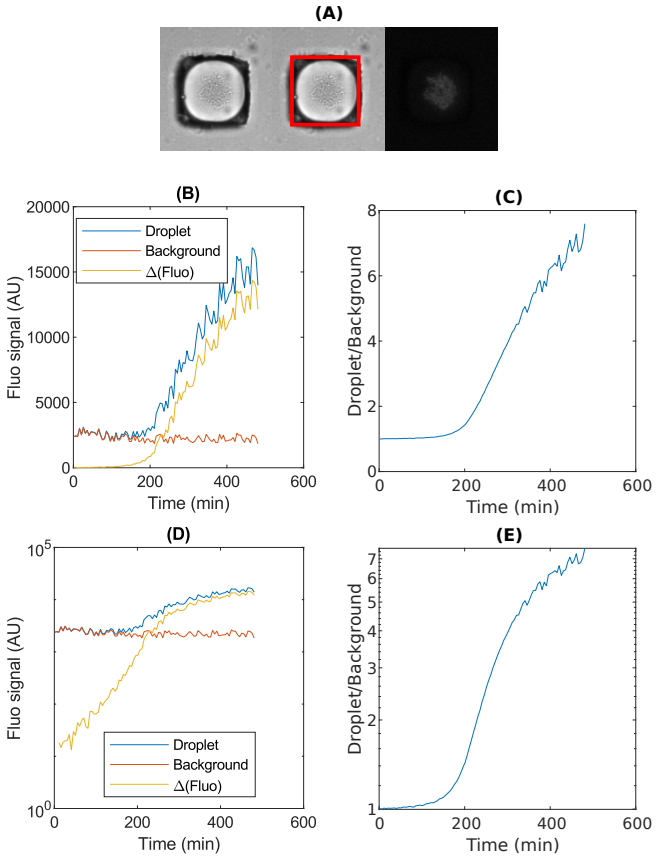


FIGURE 2.4: (A) Example image of one well after cropping of the image: bright field, bright field with detected trap (red), fluorescent image (TRITC) (B) Average of the fluorescent signal inside the red square ("droplet") and outside ("background") (C) ratio of the two signals over time (D),(E) : same, but in log scale. We see that the curves obtained by taking the ratio are much smoother than the one obtained y taking the difference of the background and the droplet fluorescence.

$(Fluo_{drop} - Background)/Background$, which is the fold change of the signal when compared to the background, that is used by other authors to quantify the fluorescence [113]. This is the method that we will apply if we have a single time-point measurement.

If we have a time-lapse of the growth of the bacteria in the droplets, instead of using just 1 as the reference value, we will subtract from the signal of each droplet its value at the beginning of the measurement, before the bacterial growth is detectable.

To summarize, in the rest of this thesis, what we will call the fluorescent signal of a droplet is:

$$Fluo(t) = \frac{< Fluo_{drop}(t) >}{< Background(t) >} - \frac{< Fluo_{drop}(t=0) >}{< Background(t=0) >}. \quad (2.7)$$

Proportionality to the number of cells

We can now check if the fluorescence signal is proportional to the number of bacteria. To check this dependency, we run the following experiment. Bacteria are cultured to exponential phase ($OD \approx 0.2$ in LB liquid medium) from an overnight culture, to be in the same conditions as for the growth experiment on-chip. Then they are put in the freezer for one minute (-18°C) and in the fridge ($+4$°C) for one hour, to stop the division process.

Then, five chips are loaded with successive dilution of this initial culture. As we know the dilution factor for each chip, we can check the proportionality between the mean number of cells per droplet and the measured fluorescence signal (Fig. 2.5).

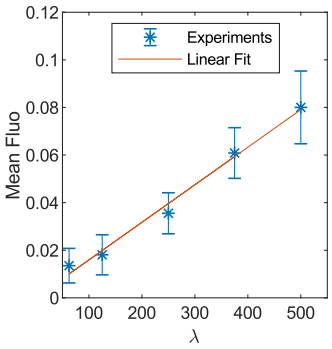


FIGURE 2.5: Calibration curve between the number of cells per droplet and the measured fluorescence signal: mean of the fluorescence signal of the droplets on chips loaded with a series of dilution of the same initial culture stopped in exponential phase, as a function of the estimated mean number of cells per droplet. Error bars: standard deviation of the fluorescent signal of the droplets of each chip.

However, if we want to know the absolute value of the proportionality constant between the fluorescence signal and the number of cells, we have to know the mean number of cells per droplet not only relatively from one chip to another, but as an absolute number. To do this, we load another chip, but with a much higher dilution: by doing this, we can get to the range of λ where λ is estimable by counting the number of empty droplets (as discussed in the previous paragraph). So we incubate this last chip to let the bacteria divide, in order to be able to distinguish easily droplets in which bacteria have grown and empty droplets. We obtain the following estimation for the Poisson parameter in this chip: $\lambda_0 \approx 1.5$.

As we know λ_0 and the dilution factor in between all the chips, we are able to give an absolute value of λ for all the chips, and as we have using eq. 2.7:

$$< Fluo >= a_f \lambda. \tag{2.8}$$

We can get an estimation of the proportionality constant between the measured fluorescence signal and the number of cells in the droplet. We get here:

$$a_f \approx 1.7 \cdot 10^{-4}. \tag{2.9}$$

As such, it is difficult to know whether this value is relevant or not. Of course, other authors have measured such a proportionality ([114], [61]), but the value in itself is very dependent on the acquisition setup, the microfluidic design, and even the bacterial strain and fluorescence construction. Here the settings of the microscope and the camera are the same as for the growth experiments of Fig. 2.7. The camera is an ANDOR EMCCD, with an exposure time of 40 ms an a numerical gain of 100.

There is one thing that we can note though: the values of fluorescence found in the calibration curve above are very low when compared with the values reached by the fluorescence of the droplets when the bacteria are allowed to grow directly on the chip (Fig. 2.7).

To check if this difference in fluorescence signal was due to our protocol for the calibration, we conducted the same calibration experiment, but with bacteria in stationary phase, in order to remove the refrigerating steps. With this culture in stationary phase, we are supposed to have a concentration of bacteria that is similar to the final one reached in the growth experiment on-chip: even of the culture conditions were not exactly the same (microfluidic droplets vs 1 mL tube), we used the same growth medium (LB rich medium) at the same temperature. Besides, the growth the growth of *E. coli* in LB is supposed to be carbon-limited [115]. As the same LB medium is used in the chips or in batch, the concentration of carbon sources is the same, and the bacterial biomass yield should be the same. However, we still obtained a very low fluorescence signal of the droplets in this case when compared with the final fluorescent signal of the droplets in our microfluidic growth experiment (Fig. 2.7). This seems to indicate that the growth, or at least the production of fluorescent proteins, are not the same on chip or in batch. Hence, if the proportionality between fluorescence and number of cells is verified, the absolute value of the coefficient a_f that we have measured here is probably not the relevant one for our on-chip growth experiment.

Exploring the difference between on-chip and batch fluorescence yield

Another very surprising and interesting result is the following. From an overnight culture, we cultured some bacteria to exponential phase and separate the culture into two parts: some of the bacteria were loaded on a chip (chip 1: on-chip growth), which was put in the incubator, and the rest of the culture was left in the 1 mL tube, stirred in the same incubator. The day after, some of these bacteria were loaded directly, without any dilution or replacement of the medium, on a chip (chip 2: batch growth), see Fig. 2.6(A). The fluorescence of the two chips were compared and it was found, not surprisingly considering what we have just said, that the droplets in chip 1 were much more fluorescent than those in chip 2 (Fig. 2.6(B,C)). Then, both chips were re-put in the incubator for one day, and re-scanned 24 hours later. This time, surprisingly, the fluorescence signal of both chips was comparable (Fig. 2.6(B,D)).

This proves once again that there is a big difference between on chip growth and batch growth. This result is very surprising because the growth

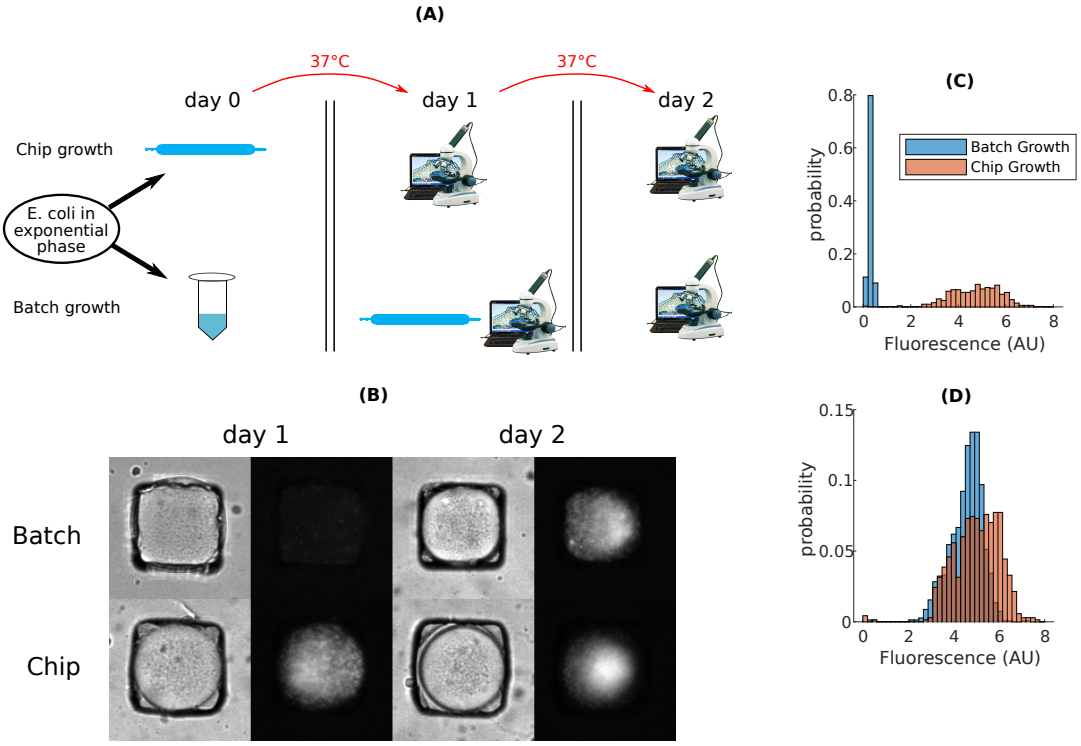


FIGURE 2.6: Parallel experiment of on-chip and batch growth (A) principle of the experiment : at day 0, bacteria in exponential phase are loaded on a chip or let to grow in batch liquid culture (1 mL tube). Everything is put at 37°C. At day 1, the chip is scanned and the batch bacteria are loaded on another chip, which is also scanned. After another 24 hours at 37 °C, both chips are scanned once again (day 2). (B) Fluorescence and bright field images of an example well at day 1 and day 2, for both chips. (C) Fluorescence signal measured at day 1, for both chips (D) Same at day 2. The fluorescence signal of the two chips is similar at day 2, whereas at day 1 the bacteria grown on chip were much more fluorescent than those grown in batch liquid culture.

of *E. coli* in LB is supposed to be carbon-limited [115]. After 24 hours in a 1 mL flask, there should not be any carbon left in the medium for the bacteria to grow when they are put into the droplets. The question that arises is to know if this is due to a difference in biomass yield, or in the production of the fluorescent protein, or even its accumulation in the droplets. To answer the question, the only way would be to be able to measure the number of cells per droplet in a different way than by fluorescence, but we have not yet developed such techniques for our micro-droplets. Many reasons could account for this difference: the oxygen availability, for instance, is known to be a key factor for the growth of *E. coli* [116], and is probably different in the micro-droplets than in batch culture. The PDMS with which the chip is made is very permeable to oxygen [117], which is also highly soluble in the fluorinated oil that surrounds the chip [118]. It has already been observed that with well-oxygenated micro-droplets, the biomass yield could be higher

than in batch liquid culture [119]. The oxygenation could be better in the chip than in a batch culture, but this remains an open question.

To sum things up, we have shown in this section that once the background is correctly removed, the average fluorescent signal of the droplet is proportional to the number of cells contained in the droplet. However, the exact value of the proportionality constant is difficult to measure because of the differences in growth conditions between droplets and classic batch growth.

2.2.3 Analysis of the results

The results of a typical time-lapse measurement of bacterial growth on chip are presented on Fig. 2.7.

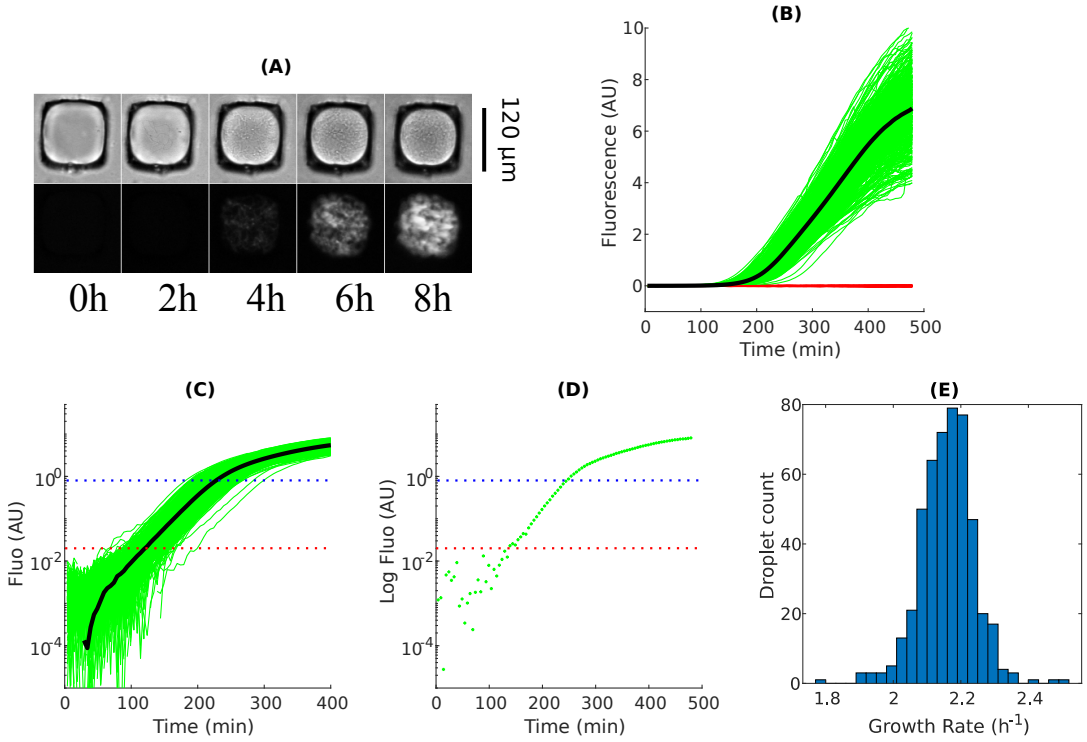


FIGURE 2.7: Results of the growth of *E. coli* on chip (A) Time-lapse in bright-field (top) and fluorescence (bottom) imaging of a droplet that contains bacteria. (B) Fluorescence curves obtained during a growth experiment of *E. coli* at 37°C. Time resolution: 5 min, number of droplets: 915. The green curves are droplets containing bacteria, the red ones are the empty droplets. The blackbold line is the mean of bacteria-containing droplets. (C) Same as (B) but plotted on log-lin scale. The red dotted line represents the limit of detection, the blue one indicates the end of the exponential phase. (D) A typical individual growth curve from the previous set. (E) Distribution of the growth rates in exponential phase measured on each of the droplets.

The first thing that we can say about these growth curves of *E. coli* in droplets is the fact that they display the three classical stages of bacterial

growth (Fig. 2.7(B)), described quite a long time ago by Monod [70]. At first, no growth is apparent, this is the lag phase, that comprises both the microscopic lag phase and the fact that the number of bacteria per droplet is at first to low to be detected [120]. Then the growth is exponential, which can be checked by plotting the curves in log scale (Fig. 2.7(C),(D)): the trajectories are straight lines, which means that the number of cells evolves exponentially with time. Finally, the growth slows down and the cells enter the stationary phase of growth.

An important parameter of the bacterial growth is the growth rate in exponential phase. It is defined as the slope of the log of the number of cells as a function of time during this exponential phase. It is measured routinely in microbiology labs because it quantifies the fitness of a strain [121]. To obtain it from the time-lapse images of bacteria in droplets, we proceed as follows. First, we set once and for all for all the droplets a detection and saturation threshold in log scale (see the red and blue dotted lines in Fig 2.7). Then we fit the trajectory of each droplet in log scale in between these lines with a straight line, and we obtain the growth rate by measuring the slope (Fig 2.7(E)). According to the classical Monod equation [70], as well as more recent works on balanced growth [82, 122], all droplets should display the same growth rate, because its value depends only on the strain and the medium. This is what we see experimentally (the standard deviation of the growth rates that we have measured is only 7%). The mean value that we find is 2.16 h^{-1}, which corresponds to a doubling time of 19 min, close to the value of $\approx$ 20 min found in the literature [115].

Another interesting feature of these growth curves is the variability that they display at each time point. Indeed, if one looks at the distribution of fluorescence signal at a given time, we can see a quite large variability, both in the exponential phase and at the end of the measurement, when the cells enter stationary phase. The variability in the exponential phase will be studied in depth in the next Chapter of this thesis.

The variability at the end is also quite surprising, because the growth of *E. coli* in LB broth is supposed to be limited by the carbon source [115]. Then the only effect that should play a role on the variability of final size is the amount of carbon available for the growth, which is directly proportional to the volume of the droplet. However, the dispersion in droplet volumes was thought to be very low in our system (Amselem et al. [97] measured a standard deviation of 2 to 5%). In the next part, we will try to quantify once again this dispersion in volume and see how we can explore the relation between volume and number of cells in the droplet.

2.2.4 Volume and size of the colony

Measuring the volume of the droplets in-situ

To explore the relation between number of cells and volume in a droplet, we need to be able to measure these two quantities at the same time. There is a trick that will allow us to measure them simultaneously: the LB broth is slightly auto-fluorescent in the green fluorescent channel (excitation at 490

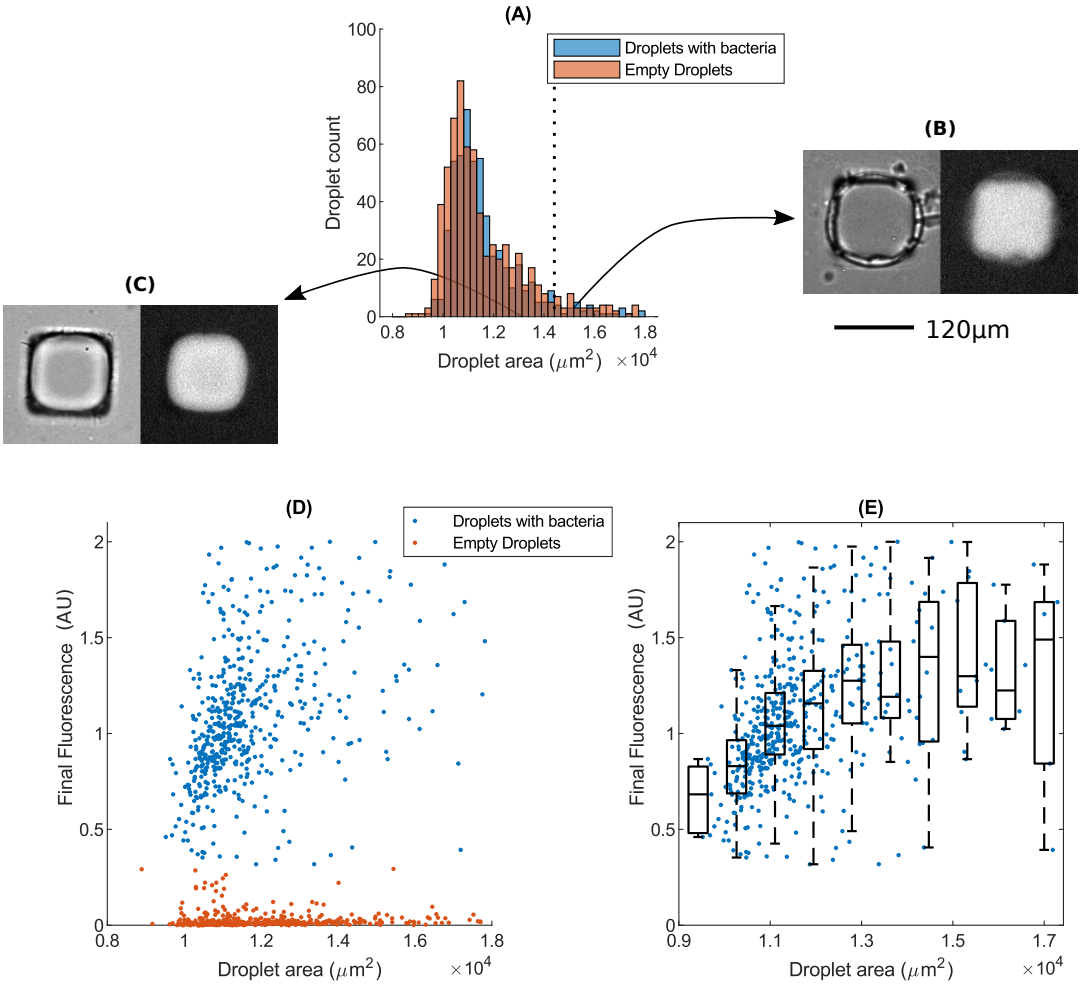


FIGURE 2.8: (A) Histogram of the projected area of the droplets at the beginning of the experiment, for empty droplets and droplets containing bacteria. The black dotted vertical line indicate the projected area of the trap. (B) Example of a droplet that is smaller than the trap, bright field and fluorescence image in FITC. We see the auto-fluorescence of the LB medium. (C) same, but for a droplet that is bigger than the trap. (D) Relation between the final fluorescence intensity of the droplets and the initial projected area of the droplets, for both empty droplets and droplet containing bacteria. (E) Box plot of the same data, but for droplets containing bacteria only. The data is binned into 10 equally spaced bins according to the droplet area. The box indicate the standard deviation, while the central line is the median of each bin. The whiskers indicate the 95% confidence interval. The projected area and the final fluorescence seem to be correlated for small droplets, but for bigger ones the correlation vanishes.

nm, emission at 520 nm). This means that the droplet of LB medium has a fluorescent signal higher than the background noise, because some of its components emit some light in this channel when they are excited by the LED of the microscope, as we can see on Fig. 2.8(B,C). Hence, to get a measure of

the volume of a given droplet, we measure the projected area of the droplet in the green channel, by thresholding the fluorescence image of the droplet in this channel. It does not interfere with the measurement of the fluorescence produced by the bacteria, as this latter one is measured in the red fluorescent channel.

When we do this in one of our chip, we obtain a right-skewed distribution for the area of the droplets, as we can see on Fig. 2.8(A). Most of the droplets are smaller than the projected area of the trap (back dotted line, Fig. 2.8(B)), while some are bigger (Fig. 2.8(C)).

Let us now look at the correlation between this projected area and the final fluorescence value measured in the droplet. We take into account the initial projected area, at the beginning of the experiment, because the droplets shrink during an experiment, as we will discuss at the end of this section. Hence, if we still consider the fact that the growth of *E. coli* in LB broth is limited by the carbon source [115], what matters to account for the variability of bacteria at the end is the initial amount of carbon encapsulated in the droplet. This amount is proportional to the initial volume of the droplet, which is supposed to be related to this initial projected area.

The correlation between the initial area of the droplet and the final fluorescence value at the end of the growth is not really clear if we consider the raw data displayed on Fig. 2.8(D). There is a lot of variability in fluorescence for a given area of the droplet. If we bin the data (Fig. 2.8(E)), we obtain a correlation, especially for the smallest droplets, that are much smaller than the traps. Indeed, we see that the median fluorescence of each box increases with the area of the droplets. For bigger droplets, this correlation seems to vanish. However, what is really striking is that the variability in final fluorescence is huge for a given droplet area. Let us try to understand this by discussing a little bit more the relation between the projected area and the volume of the droplet.

Simulations of the shape of the droplets

To try to better understand this relation, we simulated the shape of the droplet with an open-source software for the modeling of fluid interfaces, called Surface Evolver [123]. We implemented the geometry of our anchors in the software and ran simulations for droplets of different volumes. The results are summarized in Fig. 2.9: we can obtain the shape of the droplets of different volumes in the traps, and the corresponding projected area.

Three regimes can be distinguished in Fig. 2.9(B). First of all, very small droplets do not even touch the vertical walls of the anchor (far-left droplet on Fig. 2.9(A)). In this case, they have a spherical shape (which minimizes their surface energy). We can manually compute the relation between the volume and the projected area in this case. For a sphere of radius R, we have:

$$\begin{aligned} Area &= \pi R^2, \\ Volume &= \frac{4}{3}\pi R^3. \end{aligned} \tag{2.10}$$

Therefore the relation between the projected area and the volume of the droplet is the following:

$$Volume = \frac{4}{3\sqrt{\pi}} Area^{\frac{3}{2}}. \tag{2.11}$$

This theoretical relation for small droplets is indeed verified on the simulations (Fig. 2.9(B), orange dotted line). For those small droplets, we can see that the projected area is indeed a good proxy for the volume of the drop, as the volume increases with the projected area.

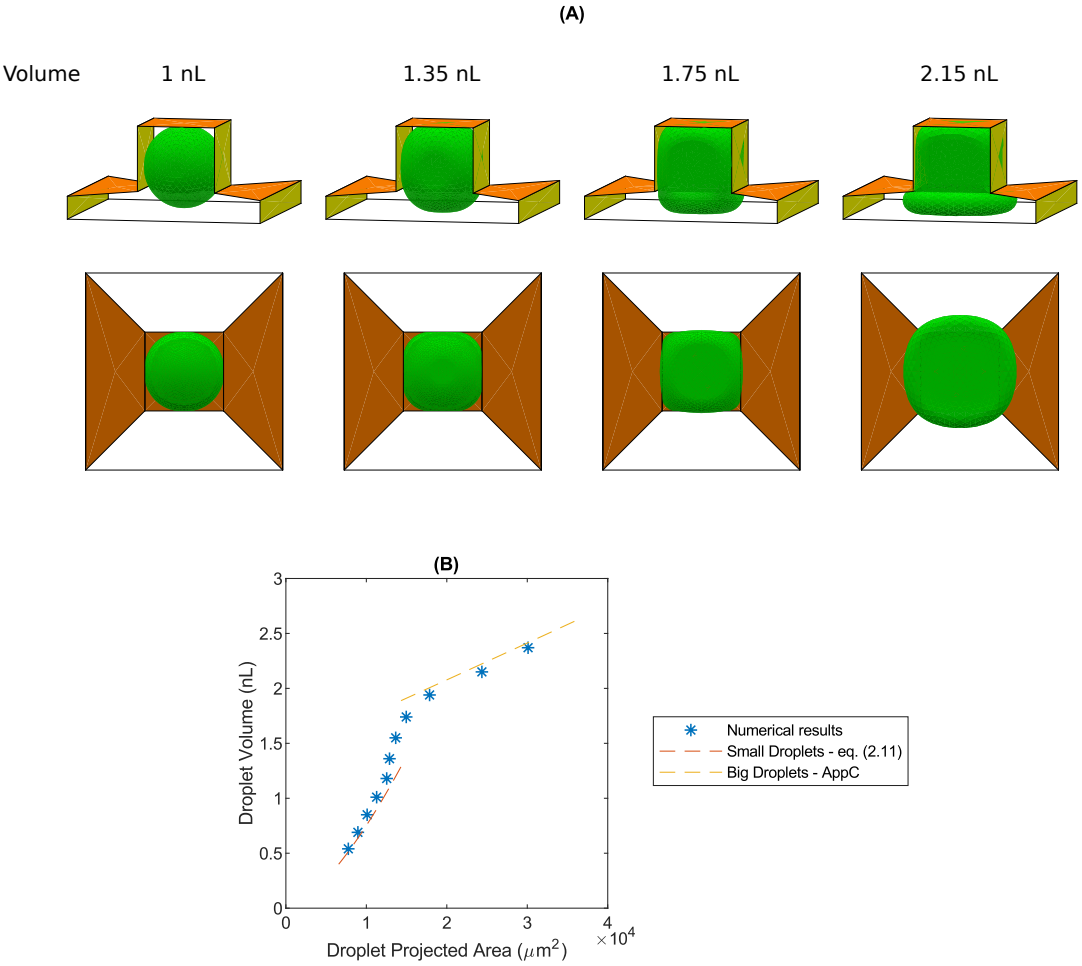


FIGURE 2.9: (A) Simulations of the shape of the droplets for 4 different droplet volumes, from small droplets (left, 1 nL) to bigger ones (right, 2.15 nL), made with Surface Evolver [123], perspective view (top line) and bottom view (bottom line). (B) Simulated volume of the droplet as a function of their maximal projected area in a horizontal plane, with comparison to theoretical volumes in two asymptotic regimes, small droplets (eq. (2.11)), and big droplets (Appendix C).

At the other end of the volume range, for big droplets, the volume also depends nicely on the projected area of the droplet. For these droplets, an approximate relation between these two quantities can be obtained. It would be

a little long to demonstrate it here, but the computation is presented in Appendix C. Rapidly, we can just say that these big droplets almost completely feel the anchor, and the variation in volume and in projected area is mostly due to the pancake-shaped squeezed part that extends in the chamber, as we can see on the far-right droplet in Fig. 2.9(A).

The problematic part happens for droplets that are in between those two extremities. For these droplets, that are close to the size of the trap, the projected area is not a good proxy for the volume. Indeed, we can see on Fig. 2.9(B) that for those droplets, there is a big variation in volume for a small variation in projected area. This is problematic because it means that droplets with the same projected area (measured with fluorescence threshold) can in fact have quite different volumes. Thus, a part of the variability that we observed on Fig. 2.8 could be an artifact of our volume estimation.

To summarize, it is difficult to obtain a trustworthy in-situ measurement of the volume of the droplets. However, it seems that for the range of droplet sizes for which we can have a correct estimation of the volume, the dispersion in final size of the colonies, measured by the fluorescence intensity of the droplet, is related on average to the initial volume of the droplets.

Mathematical modeling of the volume heterogeneity

Even if it is difficult to measure the heterogeneity of the volume of the droplets, we are going to say a few words about how it can affect the distribution of the number of bacteria per cell during loading.

We will model the heterogeneity of the volumes of the droplet by a Gaussian distribution:

$$v \sim v_0 \cdot \mathcal{N}(1, \sigma_v)\,. \tag{2.12}$$

We know that for a given volume v, the number of cells in a droplet follows a Poisson distribution, and the parameter of this Poisson distribution is just proportional to the volume. The Gaussian distribution of the volumes will then yield a Gaussian distribution of Poisson parameters, and the number of cells in the droplets is a mixture of different Poisson laws with normally distributed parameters:

$$\lambda \sim \lambda_0 \cdot \mathcal{N}(1, \sigma_v)\,. \tag{2.13}$$

Hence, we can easily get the mean number of cells per droplet and its variance by using the law of total probability:

$$\begin{aligned} \mathbb{E}(N_0) &= \mathbb{E}_\lambda\left(\mathbb{E}\left(N_0|\lambda\right)\right) \\ &= \mathbb{E}_\lambda\left(\lambda\right) \\ &= \lambda_0. \end{aligned} \tag{2.14}$$

The mean number of cells per droplet is then not affected by the heterogeneity of the volume of the droplet. This remains true even if the volume of the droplet is not Gaussian. Let us now examine the variance:

$$\begin{aligned}\mathrm{Var}(N_0) &= \mathrm{Var}_\lambda\left(\mathbb{E}\left(N_0|\lambda\right)\right) + \mathbb{E}_\lambda\left(\mathrm{Var}\left(N_0|\lambda\right)\right)\\ &= \mathrm{Var}_\lambda\left(\lambda\right) + \mathbb{E}_\lambda\left(\lambda\right)\\ &= \lambda_0^2 \cdot \sigma_v^2 + \lambda_0. \end{aligned} \tag{2.15}$$

So the variance of the number of cells per droplet is increased by an additional term, that grows as λ_0^2. The effect of the heterogeneity of the volume will then be more visible when the mean number of cells per droplet is high. This result would also remain true if the volume distribution was not Gaussian.

We can obtain an approximate value of σ_v thanks to our numerical simulations of the shape of the droplet and our measurements of the droplet size in-situ. Indeed, even if it is difficult to have a trustworthy measurement of the volume for each droplet, our numerical results can still roughly be used to compute a volume distribution of the droplets, using the numerical relation that we have found between the measured projected area and the volume. The measurement may not be very accurate for each droplet, but the standard deviation of the volume distribution can be approximated. We obtain thanks to our fluorescence measurements:

$$\sigma_v \approx 0.25. \tag{2.16}$$

This is a quite high value when compared to the one published in [97], that was below 5%. However, in this reference, the authors only obtained droplets that were bigger than the trap. We do not see this kind of droplets a lot in our experiments, perhaps because the microfluidic molds used for microfabrication were older. Hence the volume distribution is completely different in our case.

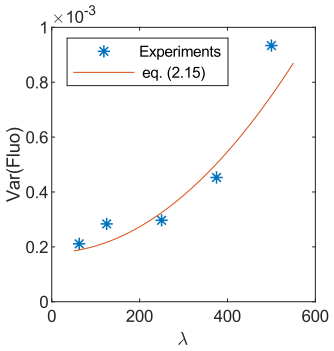


FIGURE 2.10: Variance of the fluorescence signal measured in chips loaded with a series of dilution, as a function of the mean number of cells per droplet, with model of eq. 2.15, that takes into account the dispersion of volume of the droplets.

With this mathematical model, we can take another look at the data of the experiment made to measure the proportionality between the fluorescence and the number of cells (Fig. 2.5), but looking this time at the variance instead of the mean. We can see on Fig. 2.10 that the variance of the fluorescence signal of the droplets does not evolve linearly with the concentration, as it

would in the absence of any volume heterogeneity. Instead, it can be slightly better modeled by eq. 2.15, taking for the volume coefficient of variation the value: $\sigma_v \approx 0.3$. We can note that this value is quite close to the one that we have found in our experiments, that was $\sigma_v \approx 0.25$.

Of course this is only a weak approximation, as we do not have many experimental points here and more would be necessary to really validate the model. However, this seems to indicate that there is a trace of the heterogeneity of the volume of the droplets that cannot be completely neglected. The good thing is that this has an effect only for very high initial number of cells per droplet, and most of the time we will work around $\lambda \approx 1$, so the effect of the heterogeneity of the volume on the initial distribution of the number of cells per droplet is very limited.

Volume shrinkage

In the previous parts, we evoked the initial volume of the droplets, because the volume of the droplets in a chip placed at 37°C, in a water bath, is not constant over time: the droplets are shrinking, as we can see if we plot the projected area of the droplets in the previous experiment (Fig. 2.11(A)). We believe that this is due to the porosity of the PDMS, which is porous to gas and thus to water vapor through the mechanism of pervaporation [124]. We can also see that the droplets on the outskirts of the chip shrink more than those on the center (Fig. 2.11(B)).

The shrinkage of the droplet was limited by placing the chip in water during incubation, but we see here that the efficiency of this technique is quite low. In fact, other people in the group were placing the chip during incubation in a bath of Phosphate-buffered saline (PBS [125]), instead of a water bath, but we can see that this change has a quite low effect: the relative change in projected area is almost the same for the chip placed in a water bath or the one placed in a PBS bath (Fig. 2.11(C,D)). Instead, to really limit the shrinkage of the droplet, the chip has to be pre-soaked in water prior to loading (Fig. 2.11(C,D)); we just put the chip in a water bath at 37°C for three hours prior to loading. By doing this, we saturate the PDMS with water vapor, and the droplet shrinkage is prevented.

2.2.5 Fluorescence Heterogeneity

When we evoked the fluorescence measurement in the droplet, we made the approximation that the contributions of all cells were equal. However, even in balanced growth conditions and for isogenic bacteria, there are fluctuations of the level of expression of proteins (in this case the fluorescent ones) from one cell to another. This is true even if the fluorescent protein is integrated, as in our *E. coli* strain [126], directly in the bacterial chromosome and not in a plasmid [111].

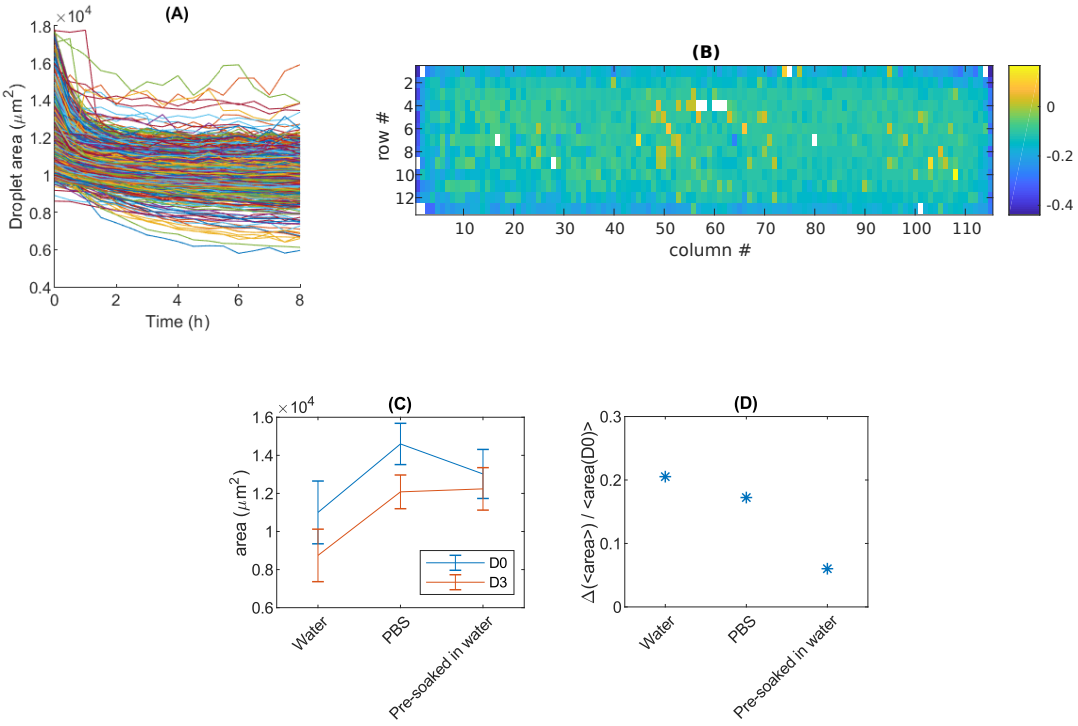


FIGURE 2.11: Shrinkage of the droplets (A) Measured area of the droplets over time, for a chip placed in water at 37°C. (B) Relative diminution of the area of the droplets with respect to the position of the well in the chip. (C) Mean projected area of the droplets in 3 chips loaded simultaneously, then placed at 37°C for 3 days: one in water, one in PBS, and one in water that was pre-soaked in water for 3 hours prior to loading. (D) Relative diminution of the mean area of the droplets in the three cases. We see that the only method to really decrease the relative diminution of the projected area of the droplets, and hence their volume, is to pre-soak them in water prior to loading.

To measure the fluorescence heterogeneity of our strain, we took some cells in exponential phase and observed under the microscope, with a 40X objective (Fig. 2.12(A)). The bright field image were segmented using a machine-learning based segmentation algorithm called Weka [127], that is freely available on Fiji [128]. This algorithm can be trained to recognize rod-like structures like bacteria. Then this segmentation was used for the measurement of the mean fluorescence intensity if the bacteria. It is more robust to do so, as segmenting directly on the fluorescence image could introduce a bias: the fluorescence of the brighter bacteria spreads a little bit on the neighboring pixels, which would be then considered as part of the bacteria, and that impacts then the measured value of the fluorescence.

We apply this method and find a distribution of fluorescence signal of the bacteria that is skewed to the right (Fig. 2.12(B)), as expected from the literature [111]. The coefficient of variation that we found in our case was ≈ 0.18.

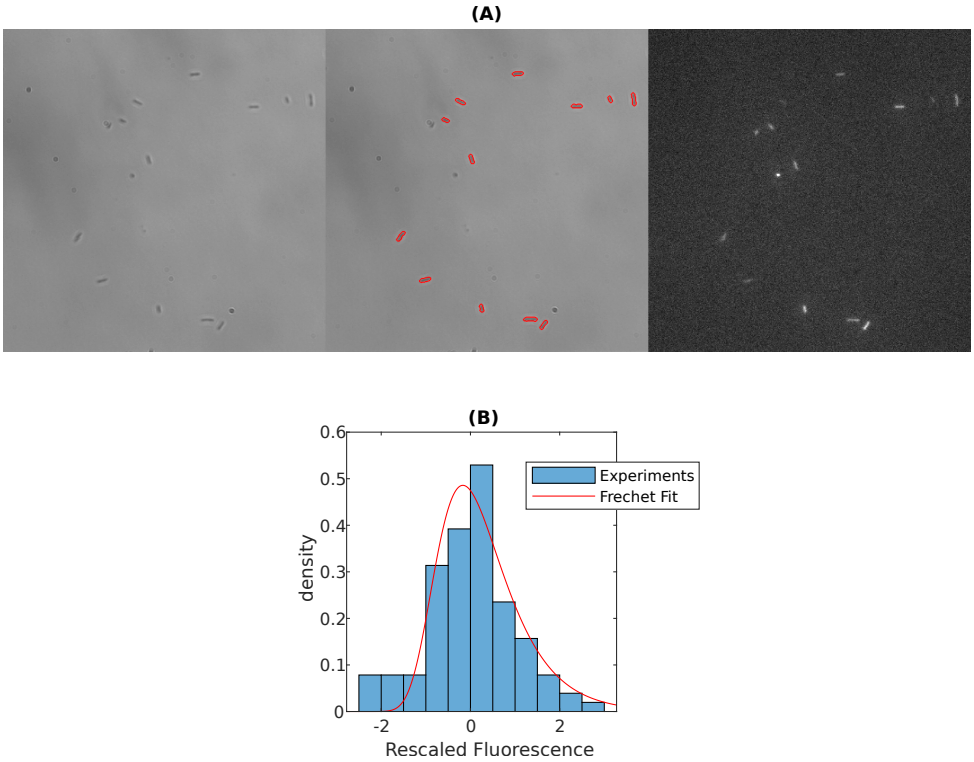


FIGURE 2.12: Heterogeneity of the fluorescence signal of individual cells (A) Bright field image, with example of the segmentation using the Weka [127] algorithm, and fluorescence image. 12 images were analyzed, for a total of 102 bacteria. (B) Distribution of the rescaled fluorescence signal and Frechet fit from [111].

Salman et al. [111] demonstrated that there exists a universal shape of the distribution for protein expression, when rescaled as follows:

$$f(x) = \phi\left(\frac{x-\mu}{\sigma}\right), \tag{2.17}$$

where μ denotes the mean and σ the standard deviation of the distribution of the fluorescence of the cells, and ϕ, which is skewed to the right and could be close to a Gamma or Log-Normal distribution, but is found to be better approximated by a Frechet distribution, of the form:

$$\phi(x;k,m,s) = \frac{1}{ks}\left(\frac{x-m}{s}\right)^{-\frac{1}{k}-1} e^{-((x-m)/s)^{-1/k}}. \tag{2.18}$$

This scaling shape fits indeed our data quite well, as we can see in Fig. 2.12(B). The fluctuations in the fluorescence signal of our cells are then the one commonly expected. But how does it impact the link between the mean fluorescence signal of a droplet and the number of cells in this droplet ?

We consider that the bacteria are all independent. Their individual fluorescence is picked from the above distribution, with mean a_f and standard

deviation σ_f. Let us compute in this case the mean and variance of the fluorescence signal of the droplet, which is just the sum of all the signals from all the bacteria in the droplet:

$$Fluo(t) = \sum_{i=1}^{N} a_f^i, \tag{2.19}$$

where a_f^i is the signal of bacteria i. We have a sum of independent and identically distributed variables, hence we can apply the Lemma of Appendix B.4 to find the mean and the variance:

$$\mathbb{E}\left(Fluo\right) = \mathbb{E}(a_f^i)\mathbb{E}(N), \tag{2.20}$$

$$\mathbb{E}\left(Fluo\right) = a_f\mathbb{E}(N). \tag{2.21}$$

As we could have expected intuitively, the mean fluorescence remains proportional to the mean number of cells even with a fluorescence heterogeneity. The convergence of the measured fluorescence (eq. 2.19) to this mean goes once again as $1/\sqrt{N}$, where N is the number of bacteria in the droplet. For the proportionality coefficient (Fig. 2.5) the number of bacteria per droplet is quite high, so we can consider that the heterogeneity in fluorescence has little effect on the mean. Let us now see what the variance looks like, still using Lemma B.4:

$$\mathrm{Var}\left(Fluo\right) = \mathrm{Var}\left(N\right)\mathbb{E}(a_f^i)^2 + Var(a_f^i)\mathbb{E}(N), \tag{2.22}$$

$$\mathrm{Var}\left(Fluo\right) = a_f^2\mathrm{Var}\left(N\right) + \sigma_f^2\mathbb{E}(N). \tag{2.23}$$

The variance obtained is then the sum of two terms. The first one is just proportional to the variance of the number of cells, and the second one is an additional variance, which is proportional to the fluorescence heterogeneity of the cells. It is proportional to the mean number of cells. In particular, in our calibration experiment, we obtain:

$$\mathrm{Var}\left(Fluo\right) = a_f^2\lambda + \sigma_f^2\lambda. \tag{2.24}$$

We see here that this additional term goes as λ, and not λ^2 as the term due to the heterogeneity of the volume of the droplets (eq. 2.15). It is therefore negligible for the analysis of the variance in our calibration experiment (Fig. 2.10). We will see in the following chapters how this term can affect more accurate measurements of the number of cells in the droplets.

2.2.6 Summary of the Chapter

To summarize, we have presented in this Chapter the microfluidic chip that will be used throughout this thesis for measurements of bacterial growth. It allows us to obtain an array of 1500 micro-reactors in which bacteria can grow, and the growth can be followed in parallel over time. We can thus repeat a growth experiment a huge number of times, which is a kind of experimental Monte-Carlo method. We have quantified and discussed the initial encapsulation of the bacteria in micro-droplets.

The fluorescence of each droplet can then be used as a proxy for the number of cells in the droplet. We have shown that those two quantities are proportional. However, it is difficult to have access to the value of this proportionality constant, because of the differences in growth conditions between growth in batch liquid culture and in nano-liter sized droplets. In this chapter, we have tried to explore those differences experimentally. We have observed, in particular, that a saturated batch culture of *E. coli* could increase its fluorescence signal when encapsulated in our droplets.

Other limitations exist for the repeatable growth of bacteria on this chip: the heterogeneity of the volume of the droplets, for instance, is probably higher in practice than what was previously measured in ref. [97]. We have shown that it is difficult to accurately estimate the volume of the droplets in experimental conditions, with the help of numerical simulations of the shape of the droplets for different droplet volumes. The influence of this volume heterogeneity on the number of bacteria per droplet was modeled and quantified in our experimental conditions.

The fluorescence of the bacteria is also heterogeneous at the single-cell level. We have measured it on our bacterial strain and shown that the fluorescence distribution was in agreement with the literature. We have modeled mathematically the influence of this heterogeneity of fluorescence and shown how it could impact our measurements. This model will be used throughout this thesis and applied to the more advanced data analysis that will be carried out in the following chapters.

Chapter 3

Distribution of the number of cells with time : the Bellman-Harris model

Ma question préférée, qu'est ce jvais faire de toute cette oseille ?
The Booba-Kaaris model

3.1 Introduction

In this chapter, we are going to try to understand the variability observed among the growth curves that we obtained in our growth experiment in microfluidic droplets (see Chapter 2). To recall, for a given time, the droplets displayed a quite big dispersion in fluorescence signal, which is supposed to be proportional the number of bacteria in the droplet. In fact, when we first showed these growth curves to biologists, explaining that these were all obtained starting from a low number of isogenic bacteria in each droplet, they were very surprised by the droplet-to-droplet variability. We will here focus on the droplet to droplet variability in exponential phase and try to see how it can be related to the cell to cell variability. The question is to know whether it is due to the cell to cell variability or is an artifact of our microfluidic experiment.

For this purpose, we will first study the different stochastic models of division at the single cell level, and examine how they affect the population growth. We will show that the outcomes of these models discussed in the literature [129, 130, 45] are similar at the scale of the growth of small populations in droplets, and that a well-known mathematical model can be used to accurately predict the variability of exponential growth at this scale from the cell-to-cell variability: the Bellman-Harris model [76]. We will use this model to quantify the droplet-to-droplet variability, adding to previously known

mathematical results the impact of our experimental constrains. These experimental constrains, that are for instance the Poisson distribution for the initial number of cells per droplet, act as additional sources of stochasticity. Therefore we will show that the variability among droplets can be quantitatively predicted if all those sources of randomness are taken into account.

There is also another objective to this study. Indeed, if the droplet-to-droplet variability carries a trace of the cell-to-cell variability, it could be possible to infer this cell-to-cell variability from the droplets measurements. This is a kind of inverse problem that we have mentioned in the introduction to this thesis. It would be a very interesting way to indirectly measure the cell-to-cell variability as our microfluidic experiments are much easier and cheaper to do than direct single-cell measurements made with more advanced experimental tools, such as the "mother machine" [44] that we have described in the introduction. To tackle this inverse problem, we will try to use both our results obtained from the extended Bellman-Harris model and pseudo-likelihood methods that can be found in the literature [92].

Most of the results of this section were submitted for publication in the Journal of Royal Society Interface, and should be published during the next few weeks.

3.2 Theoretical and numerical results

3.2.1 Comparing the three models of cell division

Bacterial division and growth at the single-cell level have been described by three different models: (i) The “timer” model, in which the cells divide once they’ve reached a given age [129]; (ii) the “sizer” model that supposes that cells divide once they have reached a given size [130]; (iii) the “adder” model in which the cells divide once they have added a certain length to their size at the previous division [45]. However, it is still an active topic of research to know how these different single-cell models determine the growth at the population scale [82]. Here, rather than considering the behavior of each cell in balanced growth, we wish to understand how this single cell behavior affects the growth of a population: if we encapsulate bacteria in a thousand isolated microdroplets according to a Poisson distribution of parameter ≈ 1, and let them divide over a few hours according to the rules of each of the single-cell models, will the distributions of number of bacteria be different for the three models?

In all three models, the volume (or, equivalently, the length, as rod-shaped bacteria like *E. coli* were shown to elongate mostly in one dimension [45]) of individual bacteria grows exponentially with a rate μ. The difference between the three models lies in the criterion for cell division, as described above. The three models are stochastic to account for the individual variability of bacterial cells [131]. There are two sources of noise, both of which are modeled by Gaussian variations: one on the division time, $\xi \sim \mathcal{N}\left(0, \sigma_\xi\right)$, and one on the elongation rate $\mu \sim \mathcal{N}\left(\mu_0, \sigma_\mu\right)$. All three models can be summarized in two equations for the volume of the cell at division v_d and for the

division time τ_d [132, 82]:

$$v_d = 2pv_0 + 2(1-p)v_b, \tag{3.1}$$

$$\tau_d = \frac{1}{\mu}\ln\left(\frac{v_d}{v_b}\right) + \xi, \tag{3.2}$$

where v_b is the volume at birth and v_0 is the constant volume after which division takes place in the adder model. In the equations above the value of the parameter p defines the regulation strategy of a cell: for $p = 0, 0.5, 1$, the cell follows a timer, adder or sizer model respectively. Indeed, if $p = 1$, we get $v_d = 2v_0$, so the volume at division is constant (sizer model). If $p = 0.5$, we get $v_d = v_0 + v_b$ so all cells will add a volume v_0 to their original size v_b (adder model). If $p = 1$, we get $v_d = 2v_b$, and by substituting in eq(3.2), $\tau_d = \ln(2)/\mu + \xi$. The division time of the cells is then fixed at $\ln(2)/\mu$ plus a stochastic noise ξ, which is the timer model.

At the population level, all models lead to an exponentially growing number of bacteria $N(t) \sim e^{\alpha_p t}$, where the growth rate α_p depends on the model. In the case of weak noises ($\mu_0\sigma_\xi \ll 1$), the growth rates of the adder and sizer models are the same, while the growth rate for the timer model is very similar to the one of the other two models, with a difference smaller than 4% [82]. Such small differences are extremely hard, if not impossible, to detect experimentally on measurements at the population scale.

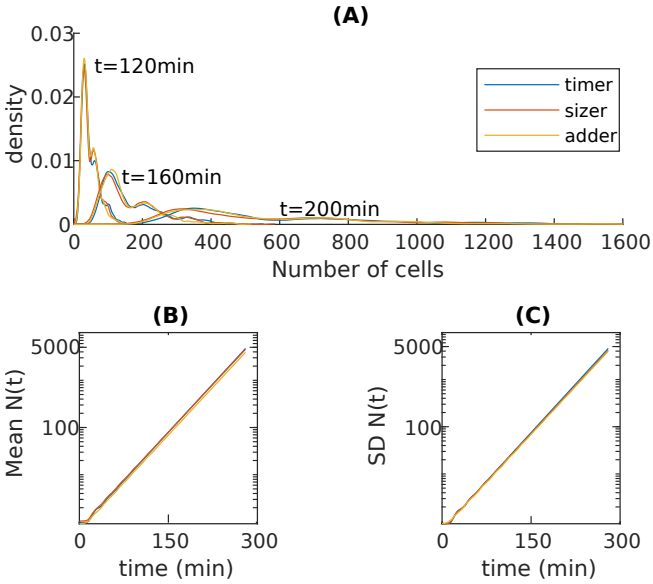


FIGURE 3.1: Comparison between the three models: timer, adder and sizer, with an initial Poisson distribution. $\sigma_t = 0.2\ln(2)/\mu_0, \mu_0 = 0.031$ min^{-1}($\tau_0 \sim 22$ min), $\sigma_l = 0.1\mu_0$, number of realizations= 1000, $\lambda = 0.75$. (A) Distributions of the number of cells at t=120, 160 and 200 min respectively. (B) Mean number of bacteria, and (C) standard deviation of the number of bacteria per simulation, in log scale.

In the general case of arbitrary noise, analytical predictions are difficult to make and one has to rely on numerical simulations to know the growth behavior of the population. Here, we carry out Monte-Carlo simulations mimicking the microfluidic experiment: identical simulations of bacterial growth are run in parallel 1000 times, with the initial number of cells n_0 being chosen according to a Poisson distribution: $P_\lambda(n_0) = \lambda^{n_0} e^{-\lambda}/n_0!$ with a Poisson parameter $\lambda = 0.75$. We choose $\sigma_\xi \sim 0.25\ln(2)/\mu_0$ and $\sigma_\mu = 7\%$, in accordance with our experimental values and those of Ref. [45]. The three models yield very similar distributions of the number of cells as a function of time, see Fig. 3.1.

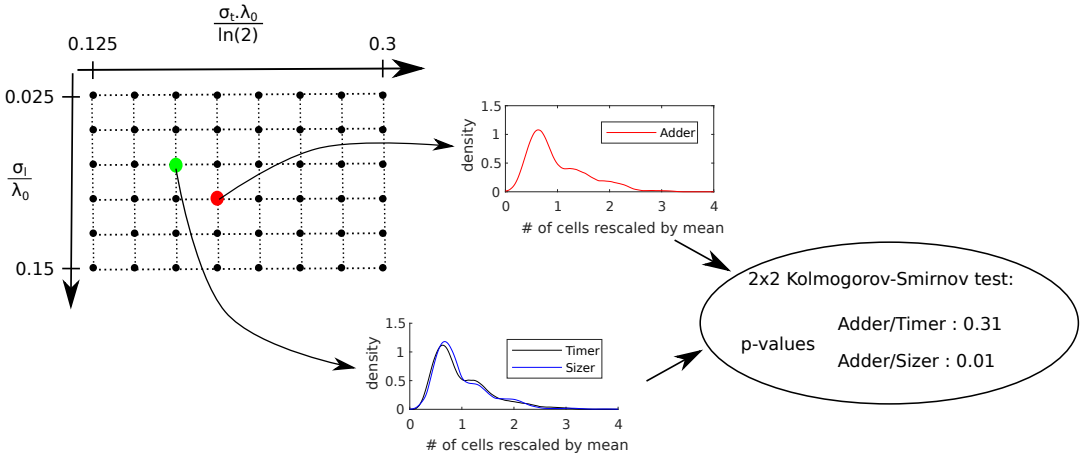


FIGURE 3.2: Principle of the test to compare the three different models at the population scale. First a grid of simulations is built for experimentally possible values of σ_l and σ_t, for all three models. Then each point on the grid is compared to all other points of the grid: the final distributions of number of cells, rescaled by their means, are compared through a two sample Kolmogorov-Smirnov test. As an example, the distribution of the number of cells obtained by an adder model at the position indicated by the red dot is compared to the distributions obtained by the sizer and timer models at the position indicated by the green dot. In this particular example, the results of the adder model at the red position are indistinguishable from that of the timer model at the green position, but distinguishable from that of the sizer ($p < 0.05$). The same method is applied to all points and all models of the grid of simulations. We find that, for each point (σ_l, σ_t) and each model, there is always at least one set of parameters (σ_l, σ_t) of one of the other models leading to similar results with $p < 0.05$.

For further comparison with the microfluidic experiments, we run a complete set of simulations covering the range of experimental values of σ_μ and σ_ξ found by Taheri et al. [45]: $0.1 \leq \sigma_\xi \leq 0.3$, and $0.025 \leq \sigma_\mu \leq 0.15$. The initial number of bacteria is sampled from a Poisson distribution with a Poisson parameter $\lambda = 1$. Bacteria are allowed to grow and divide for ~ 4 h, a time close to the duration of the exponential phase in our experiments. The final distributions of the number of bacteria, rescaled by their mean, are compared

using a two-sampled Kolmogorov-Smirnov test [133] to know whether the outputs of the three models are distinguishable. For each point of the $[\sigma_\mu, \sigma_{\tilde{\xi}}]$ mesh and each of the three models, we find that there always exists a distribution belonging to another model that is statistically indistinguishable at a 5% significance level, as shown in Fig. 3.2. For experimental purposes, the three division models are therefore indistinguishable at the population level and give the same results. Therefore, any model can be used to predict the growth of a population. In the following, we extend the classical timer model of Bellman and Harris [76] to describe a larger set of experimental conditions, which allows us to obtain analytical predictions on the population behavior.

3.2.2 Classical case studied by Bellman and Harris

Bellman-Harris equations

The timer model was studied by Bellman and Harris in their classic paper [76]. Their original assumption is that the division times of all bacteria are independent and picked from a continuous stochastic distribution which is constant over time. Starting from one cell at $t = 0$, the Bellman-Harris (B-H) model derives several properties of the probability distribution of the number of cells as a function of time $N(t)$. For instance it shows that all moments of $N(t)$ grow exponentially with the same growth rate α, such that the $k-$th moment μ_k of the distribution of $N(t)$ is equivalent to $\tilde{n}_k e^{k\alpha t}$ when t goes to infinity, which we will write as $\mu_k \sim \tilde{n}_k e^{k\alpha t}$ in what follows (see AppendixB.1 for precise definition).

In particular, we will note n_1 and n_2 the prefactors of the mean (M_N) and standard deviation (SD_N) of the distribution, respectively:

$$M_N(t) \sim n_1 e^{\alpha t}, \tag{3.3}$$

$$SD_N(t) \sim n_2 e^{\alpha t}. \tag{3.4}$$

It is important for what follows to remark that the coefficient of variation of the number of cells $CV_N(t) = SD_N(t)/M_N(t)$ therefore converges to a constant as t goes to infinity, which is just the ratio n_2/n_1.

The value of the growth rate α is given by the unique solution of the equation:

$$\int_0^\infty e^{-\alpha t} G(t)\, \mathrm{d}t = \frac{1}{2}, \tag{3.5}$$

where G is the density function related to the division time τ_d. We can note that this equation is equivalent to the following:

$$\mathbb{E}(X) = \frac{1}{2} \quad \text{with} \quad X = e^{-\alpha \tau_d}. \tag{3.6}$$

The value of n_1 is given by

$$n_1 = \frac{1}{4\alpha \int_0^\infty t e^{-\alpha t} G(t)\, \mathrm{d}t}, \tag{3.7}$$

and the value of n_2 is:

$$n_2 = n_1 \sqrt{\frac{4\mathbb{E}(X^2) - 1}{1 - 2\mathbb{E}(X^2)}}, \tag{3.8}$$

which implies that the coefficient of the number of cells for long times is then:

$$CV_N(\infty) = \sqrt{\frac{4\mathbb{E}(X^2) - 1}{1 - 2\mathbb{E}(X^2)}}. \tag{3.9}$$

We can remark that the computation of CV_N and α rely solely on the computation of the first two moments of the variable $X = e^{-\alpha\tau_d}$.

Another important property derived by Bellman and Harris is that for every realization of the model $N(t)$, there exists a random variable W such that

$$N(t) \sim W n_1 e^{\alpha t}. \tag{3.10}$$

Several properties of W are derived by Bellman and Harris, but what interests us the most is that every trajectory $N(t)$ should then have the same growth rate α.

Gaussian Division Times

Here we consider that bacteria divide after a time τ_d, which is normally distributed around a mean value τ_0 with standard deviation σ. This distribution allows us to define a microscopic coefficient of variation that describes the variability between individual cells:

$$cv_\mu = \frac{\sigma}{\tau_0}. \tag{3.11}$$

In this case, as τ_d follows a normal law, X follows a rescaled log-normal law and we can easily find the first two moments of X:

$$\mathbb{E}(X) = e^{-\alpha\tau_0 + \frac{\alpha^2\sigma^2}{2}}, \tag{3.12}$$

$$\mathbb{E}(X^2) = e^{-2\alpha\tau_0 + 2\alpha^2\sigma^2}. \tag{3.13}$$

If we insert equation (3.12) into equation (3.6), we get the theoretical value of α:

$$\alpha = \frac{\ln(2)}{\tau_0} \frac{2}{1 + \sqrt{1 - 2\ln(2)cv_\mu^2}}. \tag{3.14}$$

The growth rate is therefore a product of two terms: The first term is the ratio $\ln(2)/\tau_0$, which is the growth rate when the heterogeneity of division times is neglected. The second term is larger than one and increases with cv_μ. It predicts that the cell-to-cell variability of the division times increases the apparent growth rate of the whole population. Although this dependence of

α on cv_μ is verified as shown in Fig. 3.3(A), the net effect on the growth rate is weak, remaining below 5% for any value of cv_μ within a realistic range.

These observations are confirmed by Monte-Carlo simulations, whose results are in excellent agreement with the analytical prediction of eq. (3.14), thus validating the theoretical result, as shown in Fig. 3.3(A). The divergence that is observed between the numerical and theoretical results for $cv_\mu > 0.3$ is due to the increased apparition of negative division times in the theory, which are removed from the simulations.

The complete expressions of n_1 and n_2 can also be obtained using the above equations (3.7,3.8), but we will just give here the asymptotic value of the coefficient of variation :

$$CV_N^2(\infty) = \left(\frac{n_2}{n_1}\right)^2_{BH} = \frac{e^{\alpha^2\sigma^2} - 1}{1 - \frac{e^{\alpha^2\sigma^2}}{2}}. \tag{3.15}$$

This expression depends strongly on cv_μ, as shown in Fig. 3.3(B).Therefore although the cell-to-cell variability has only a small effect on the total growth rate, its signature is clear on the macroscopic coefficient of variation between different realizations. Again, the theoretical prediction is validated by the numerical Monte Carlo simulations, see Fig. 3.3(B).

Note that, if the coefficient of variation of the division time is small ($cv_\mu \ll 1$), we get:

$$\alpha \sim \frac{\ln(2)}{\tau_0}\left(1 + \frac{\ln(2)}{2} cv_\mu^2\right), \tag{3.16}$$

$$\mathbb{E}(X^2) \sim 1 + \frac{\ln(2)^2}{4} cv_\mu^2. \tag{3.17}$$

For the coefficient of variation of the number of cells in exponential phase, we can compute the following approximation:

$$CV_N(\infty) \sim \sqrt{2}\ln(2) cv_\mu \approx cv_\mu, \tag{3.18}$$

because $\sqrt{2}\ln(2) \approx 1$, as observed numerically by Stukalin [78], and proven here theoretically. For small cv_μ, we almost have an equality between the microscopic and macroscopic coefficients of variations, which proves that cell-to-cell variability strongly influences the behavior at the population scale.

Another consequence of the B-H model is that the asymptotic distribution of the number of cells is self-similar in time, when re-scaled by $\exp(\alpha t)$. This means that the distribution of $N(t)/\exp(\alpha t)$ converges to a stationary distribution at late times, which is also recovered by the numerical simulations, as seen on the movie 3 in Appendix E. Similarly to CV_N, the shape of this distribution depends strongly on the single-cell parameters, as plotted on Fig. 3.3(C). Quite intuitively, when cv_μ increases, it also increases the variance of the rescaled distribution.

To say a little bit more about the shape of this distribution, we find that it is close to a log-normal distribution, see Fig. 3.3(D). We get a p-value $p < 0.05$ in a Kolmogorov-Smirnov test for normality [133] for all values of $cv_\mu > 0.1$.

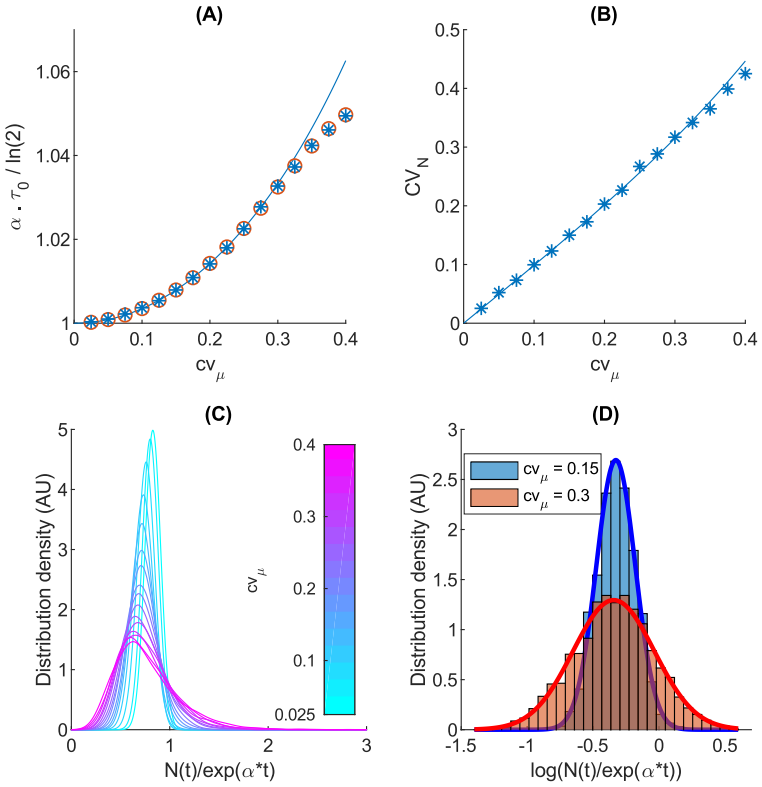


FIGURE 3.3: Theory and simulations for the Bellman-Harris model with a Gaussian distribution for the division times. Simulations were made with $\tau_0 = 21$ min, 2000 simulations going from $t = 0$ to $t = 5$ h. (A) and (B) theory (lines) and simulations (stars) (A) Gain in the exponential growth rate as a function of the cellular-level coefficient of variation cv_μ. Orange circles : growth rate gain computed by numerical solving of eq.(3.6) (B) Macroscopic coefficient of variation CV_N as a function of cv_μ. (C) Asymptotic shape of the distribution of $N(t)/\exp(\alpha t)$. The curves show a kernel fit of the simulated curves. (D) For two values of cv_μ, distribution of the log of $N(t)/\exp(\alpha t)$ and Gaussian fit. We get a p-value $p < 0.05$ in a Kolmogorov-Smirnov test for normality [133] for all values of $cv_\mu > 0.1$.

For smaller values of cv_μ, cells are almost synchronized, and it is more difficult to assess the converging shape of the stationary distribution for t going to infinity, but we believe that the log-normal shape must hold. The shape of the rescaled distribution is then skewed to the right, and as expected for a log-normal distribution, the bigger its variance, the more it is asymmetrical. The shape of the resulting distributions in noisy exponential processes was studied in detail, for a slightly different formal background, in ref. [85].

Hitting Times in a Bellman-Harris model

Another interesting feature of the Bellman-Harris model is the distribution of the hitting times, which is the time required to reach a certain number of cells in the droplet. We have not made analytical predictions for these

hitting times, but we will see how they behave on the simulations and try to relate this behavior to the predictions we have made on the distribution of the number of cells. On the simulations, we find that for long times, or to be more precise, for large values of N, the distribution of the hitting times is constant, but translated because of the mean exponential growth of the population (see Fig. 3.3(A)). In particular, the variance of the time needed to reach 200 or 2000 cells is the same. The difference is just the mean time, which is directly related to the mean exponential growth of the population. The hitting-times distribution seems to be Gaussian (see Fig. 3.3(B)), we get a p-value $p < 0.05$ in a Kolmogorov-Smirnov test for normality [133] for all values of cv_μ. The variance of the distribution of hitting times depends strongly on cv_μ (see Fig. 3.3(C)): the larger cv_μ, the more variable the hitting times.

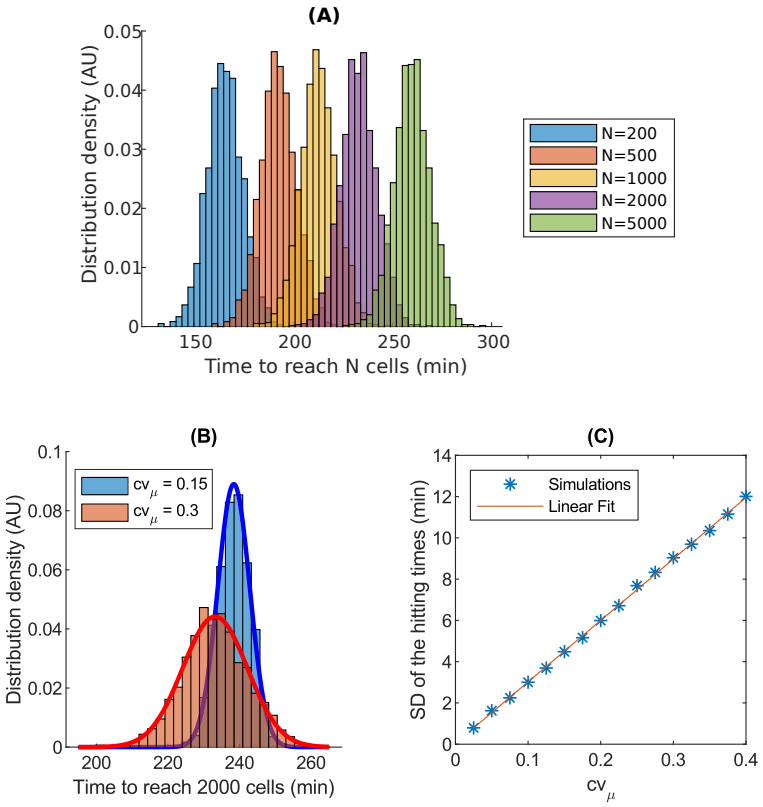


FIGURE 3.4: Hitting Times in a Bellman-Harris simulation (same simulation as Fig. 3.3). (A) Histogram plot of the time needed for each simulation to reach a fixed number of cells. We see that the distribution for the different times is the same, but translated. (B) Distribution of the hitting times to reach 2000 cells for two example values of cv_μ, with Gaussian fit. (C) Evolution of the standard deviation (SD) of the hitting times with cv_μ.

Now that we have seen what the hitting times looked like on the simulations, we can try to explain these results a little bit, not with precise mathematical demonstrations, but appealing to intuition and the results we have for the behavior of the number of cells as a function of time thanks to the

Bellamn-Harris model. One of these results, that is very important, is the one of eq. (3.10): it states that the population in each droplet grows exponentially with the same growth rate. If we look at the trajectories $N(t)$ in log scale, we then have a set of parallel lines, as confirmed numerically. It is then not surprising that the distribution of the hitting times is the same, but translated because of the mean growth: the hitting times are indeed just translated along those parallel lines for each droplet, as we can see in Fig. 3.5(B).

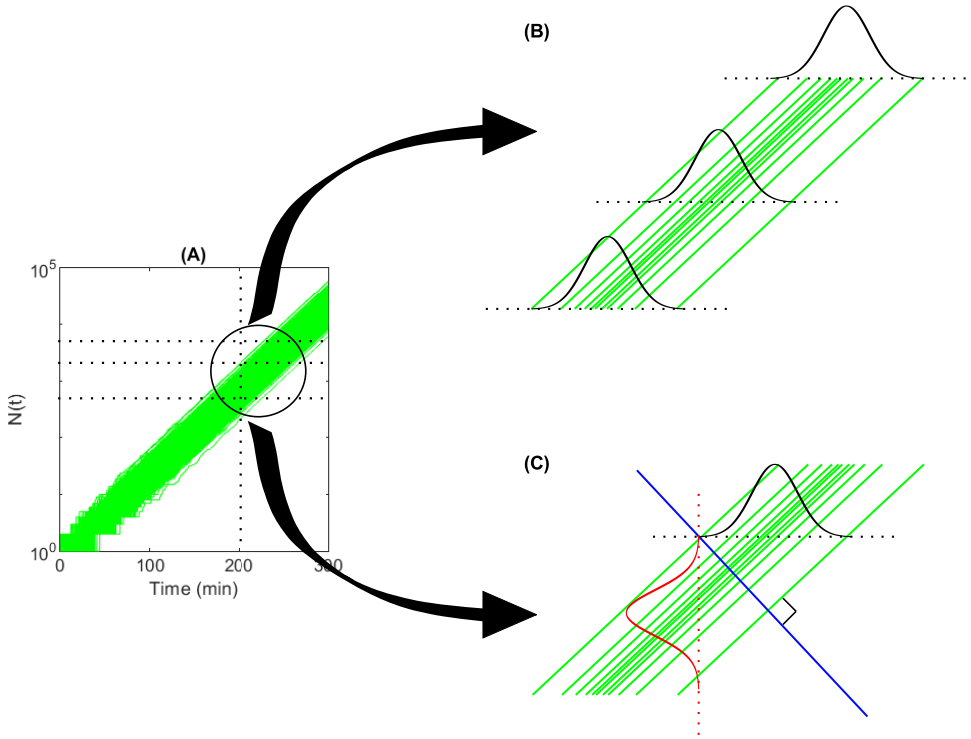


FIGURE 3.5: Geometrical approach to understand the hitting times in a Bellman-Harris model. (A) Individual trajectories in a Bellman-Harris simulation (same as in Fig. 3.3, in log scale). (B) Schematic zoom on some curves, and distribution of the hitting times for three different values of N. As the individual trajectories $N(t)$ are just parallel lines, we see that the distribution of the hitting times is the same, but translated to account for the mean growth. (C) Schematic zoom on some curves, and link between the distribution of $N(t)$ at a fixed time (in red) and distribution of the hitting times (in black). As all the individual trajectories are parallel lines, we see that the two distributions are the same, because of the symmetry with respect to the blue line, which is perpendicular to the trajectories.

Similarly, the same kind of geometrical approach can help us understand why the distribution of the hitting times is Gaussian. Once again, all trajectories, in log scale, are just parallel lines. If we consider the symmetry with respect to a perpendicular line (in blue on Fig 3.5(C)), we see that the distribution of hitting times is just the same as the distribution of $\log(N(t))$, and as we have seen in the previous part that this distribution was Gaussian (as the

distribution of $N(t)$ is log-normal), it is logical to find a Gaussian distribution for the hitting times.

In summary, the hitting times are a kind of dual problem to the distribution of $N(t)$, and both are directly related because of the exponential growth of all the trajectories. In the rest of this chapter, we will then focus on the distribution of $N(t)$, but a lot of the results could be translated to the hitting times. The link between hitting times and individual lag time (i.e. the distribution of the first division for the individual cells) is not straightforward mathematically and was studied in a series of article by Baranyi *et al.*, see [134], [135] and [120].

Taken together, the above results show the effects of the cell-to-cell variability on the population-scale growth. While the impact on the exponent α is always weak, the shape of the asymptotic probability distribution of $N(t)$ is very sensitive to the value of cv_μ. This observation suggests that the variability between cells can be inferred from a large number of population-scale measurements without the need for single-cell resolution. However taking into account other sources of variability in the practical situations will limit such inference, as demonstrated in the following sections.

To be (a little bit more) rigorous

If we want to be rigorous from a mathematical point of view, we have to come back on one point. Indeed, the division times that we consider in this model can only be positive, as it makes no sense for a cell to have a negative division time. When we say that the division times are Gaussian, what we mean in fact is that they are picked from a Gaussian distribution truncated at zero of cumulative probability function $\tilde{G}$, that has to be re-normalized :

$$\tilde{G}(\tau_d) = \frac{G_{\tau_0,\sigma}(\tau_d)}{1 - G_{\tau_0,\sigma}(0)}, \tag{3.19}$$

where $G_{\tau_0,\sigma}$ is the cumulative probability function of the Gaussian law of mean τ_0 and standard deviation σ. But the relative difference, for the positive times, depends only on $G_{\tau_0,\sigma}(0)$, which is very close to 0, especially for small cv_μ, as we can see on Fig. 3.6.

Hence, we cannot rigorously use the fact that the variable $X = \mathbb{E}(e^{-\alpha\tau_d})$ is a log-normal variable in eq.(3.6). But we can still solve numerically this equation and obtain a numerical value for the growth rate α. As we can see on Fig. 3.3(A) (orange circles), the direct numerical value is very close to the B-H theoretical one for small values of cv_μ. For high values of cv_μ, we recover the value of the growth rate found on the simulations, that deviates from the B-H theoretical value.

In fact, one of the hypothesis imposed by Bellman and Harris is not satisfied by the function $\tilde{G}$:

$$\tilde{G}(0^+) = \frac{G_{\tau_0,\sigma}(0)}{1 - G_{\tau_0,\sigma}(0)} \neq 0. \tag{3.20}$$

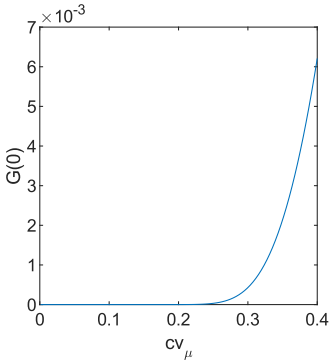


FIGURE 3.6: G(0): probability of finding negative values in a Gaussian distribution of mean $\tau_0 = 21$ min, varying the coefficient of variation cv_μ.

However, we have already seen that $G_{\tau_0,\sigma}(0)$ is very close to zero for reasonable cv_μ (Fig. 3.6), and so $\tilde{G}(0^+)$ as well. The very good match for the interesting observable quantities and the simulations presented on Fig. 3.3 shows anyway that we can still use the Bellman-Harris theory to model what happens in our simulations and experiments.

3.2.3 Poisson distributed initial number of cells

When bacteria are encapsulated in droplets, the initial number of bacteria per droplet follows a Poisson distribution of parameter λ [97]. The probability of having N_0 bacteria in a droplet is given by: $P_\lambda(N_0) = \lambda^{N_0} e^{-\lambda}/N_0!$ [101], where λ is the average number of bacteria per droplet. In a microfluidic setup, the Poisson distribution leads to some droplets being empty, but we solely consider the evolution of droplets containing bacteria. As bacteria are assumed to be independent from one another, this modification of the initial number of bacteria does not change the asymptotic behavior of the distribution: the k-th moment of the distribution of $N(t)$ grows exponentially with the same growth rate $k\alpha$ as previously, but the values of the prefactors in front of the exponential are different from the classical case.

To derive the values of the prefactors, consider a droplet containing initially N_0 bacteria, with $N_0 > 0$. We call $N_k(t)$ the size of the offspring of bacteria k at time t. The total number of bacteria $N_\lambda(t)$ in the droplet at time t is then given by:

$$N_\lambda(t) = \sum_{k=1}^{N_0} N_k(t). \tag{3.21}$$

Since we suppose that bacteria grow independently from one another, the terms of the sum are independent and identically distributed: each $N_k(t)$ follows just a simple Bellman-Harris model, with $N_0 = 1$. We can then use the lemma of Appendix B.4 and we directly get:

$$\mathbb{E}\left(N_\lambda(t)\right) = M_N(t)\mathbb{E}\left(N_0\right), \tag{3.22}$$

$$\mathrm{Var}\left(N_\lambda(t)\right) = SD_N(t)^2 \mathbb{E}\left(N_0\right) + M_N(t)^2 \mathrm{Var}\left(N_0\right), \tag{3.23}$$

where $M_N(t)$ and $SD_N(t)$ are the mean and the standard deviation of the simple Bellman-Harris case, that were defined in the previous section.

To go further, we just have to pay attention to the fact that we want to consider only the droplets that contain bacteria, so N_0 doesn't really follow a Poisson distribution here, but the restriction of a Poisson distribution to $\mathbb{N}^*$. Since the probability of having an empty droplet is $e^{-\lambda}$, we have:

$$\mathbb{P}\left(N_0 = k\right) = \frac{1}{1-e^{-\lambda}} \frac{\lambda^k e^{-\lambda}}{k!}. \tag{3.24}$$

To compute the expected value and the variance of N_0, we can use the probability generating function, that is defined as:

$$G(s) = \mathbb{E}\left(s^{N_0}\right). \tag{3.25}$$

In our case, we get:

$$\begin{aligned} G(s) &= \sum_{k=1}^{\infty} s^k \frac{1}{1-e^{-\lambda}} \frac{\lambda^k e^{-\lambda}}{k!}, \\ &= \frac{e^{-\lambda}}{1-e^{-\lambda}} \sum_{k=1}^{\infty} \frac{(\lambda s)^k}{(k)!}, \\ &= \frac{e^{-\lambda}}{1-e^{-\lambda}} \left(e^{\lambda s} - 1\right). \end{aligned} \tag{3.26}$$

Then the expected value and the variance can be computed from the values in $s = 1$ of the first and second derivative of this function, with the formula:

$$\mathbb{E}\left(N_0\right) = G'(1), \tag{3.27}$$

$$\mathrm{Var}\left(N_0\right) = G''(1) + G'(1) - G'(1)^2. \tag{3.28}$$

Hence, we easily get :

$$\mathbb{E}\left(N_0\right) = \frac{\lambda}{1-e^{-\lambda}}, \tag{3.29}$$

$$\mathrm{Var}\left(N_0\right) = \frac{\lambda}{1-e^{-\lambda}} \frac{1-(\lambda+1)\,e^{-\lambda}}{1-e^{-\lambda}}. \tag{3.30}$$

We can now inject these expressions into equations (3.22) and (3.23), and take the equivalent for t going to infinity. This leads to a mean $M_\lambda(t) \sim n_1(\lambda) e^{\alpha t}$ with:

$$n_1(\lambda) = \frac{\lambda}{1-e^{-\lambda}} n_1, \tag{3.31}$$

while the pre-factor of the standard deviation $SD_\lambda(t) \sim n_2(\lambda)e^{\alpha t}$ is:

$$n_2(\lambda) = \sqrt{\frac{\lambda}{1-e^{-\lambda}}\left({n_2}^2 + \frac{1-(\lambda+1)\,e^{-\lambda}}{1-e^{-\lambda}}{n_1}^2\right)}. \tag{3.32}$$

The expression for $n_1(\lambda)$ is expected because of the independence of the bacterial division times: the mean number of cells at time t is simply the product of the mean number of cells per non-empty droplet and the mean of the classical Bellman-Harris case. These expressions are systematically checked numerically on Fig. 3.7(A).

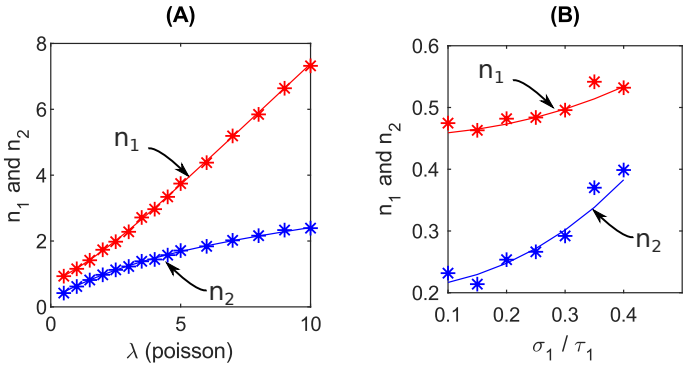


FIGURE 3.7: Comparison between the theoretical predictions (straight lines) of the Bellman-Harris model and Monte-Carlo simulations (stars). The law for the division time of individual bacteria is a Gaussian with mean $\tau_0 = 23$ min and standard deviation $\sigma = 0.25\tau_0$. (A) $n_1(\lambda)$ and $n_2(\lambda)$ as a function of the Poisson parameter λ (2000 simulations for each λ). (B) Effect of an evolving distribution of division times. The first bacterial generation divides with a mean time $\tau_1 = 2\tau_0$, and a standard deviation σ_1. Subsequent division times follow a normal law of mean τ_0 and standard deviation σ_0. n_1 and n_2 depend on σ_1/τ_1. The initial number of cells follows a Poisson distribution with a parameter $\lambda = 0.5$. Result of 2000 independent simulations for each σ_1.

To understand the expression for the standard deviation better, it is useful to divide it by the mean and take the square. Doing so, we obtain the square of the coefficient of variation CV_λ for t going to infinity:

$$CV_\lambda^2(\infty) = \left(\frac{n_2}{n_1}\right)^2_\lambda = \underbrace{\frac{1-e^{-\lambda}}{\lambda}\left(\frac{n_2}{n_1}\right)^2_{BH}}_{(1)} + \underbrace{\frac{1-(\lambda+1)\,e^{-\lambda}}{\lambda}}_{(2)}. \tag{3.33}$$

The variance of the population size is thus the sum of two terms. The first one is the variance due to the stochastic nature of the Bellman-Harris process, corrected to take into account only the wells that contain bacteria. The second term is an additional variance due to the initial Poisson distribution. The contribution of the initial Poisson distribution becomes greater

than the intrinsic variance of the Bellman-Harris process when the Poisson parameter is larger than ≈ 0.5, see black line in Fig. 3.8(A). The dominant effect on the variance of the number of cells is then the variance due to the initial distribution, and not the heterogeneity of the division times. This can be seen clearly in Fig. 3.8(B): the coefficient of variation of the population now depends weakly on cv_μ but cases with different values of λ can clearly be distinguished. Simulations confirm this results, and highlight another surprising effect: the numerically computed values tend to differ more from the theoretical prediction for small values than for large values of λ. Indeed, for small values of λ, fewer droplets contain bacteria, and the stochastic divergence from the prediction is bigger, even with 2000 simulations.

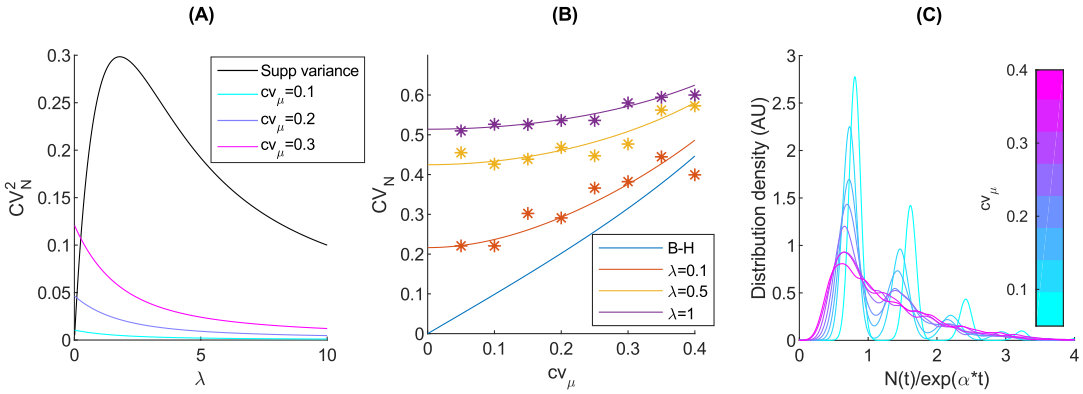


FIGURE 3.8: Theory and simulations for the Bellman-Harris model with a Gaussian distribution for the division times and a Poisson distribution for the initial number of cells. (A) Contributions of the Bellman-Harris process (colored lines, term (1) in eq. (3.33)) and of the initial Poisson distribution (black line, term (2) in eq. (3.33)) to the square of CV_N, as a function of the parameter of the Poisson distribution and for different values of cv_μ. (B) CV_N, asymptotic value coefficient of variation of the number of cells as a function of cv_μ, theory (lines) and simulations (stars). (C) Asymptotic shape of the distribution of $N(t)/\exp(\alpha t)$ (kernel fit, simulations only), for $\lambda = 1$

Note that, as in the classical Bellman-Harris case, the asymptotic distribution of the number of cells remains self-similar (see Supp. Movie 4 in Appendix E). The shape of the self-similar distribution now shows peaks, each peak corresponds to the initial number of bacteria in the droplet, see Fig. 3.8(C). The width of each peak stills depends on the variability of the division times. Indeed, if we consider only the droplets containing initially N_0 bacteria,

$$N_{N_0}(t) = \sum_{k=1}^{N_0} N_k(t). \tag{3.34}$$

All the $N_k(t)$ are independent and identically distributed, thus we have, as N_0 is fixed :

$$\mathbb{E}\left(N_k(t)\right) = \sum_{k=1}^{N_0} \mathbb{E}\left(N_k(t)\right) \sim N_0 n_1 e^{\alpha t}, \tag{3.35}$$

$$Var\left(N_k(t)\right) = \sum_{k=1}^{N_0} Var\left(N_k(t)\right) \sim N_0 n_2^2 e^{2\alpha t}. \tag{3.36}$$

Thus, for the asymptotic coefficient of variation, we get:

$$CV_{N_0}(\infty) = \frac{1}{\sqrt{N_0}} \left(\frac{n_2}{n_2}\right)_{BH}. \tag{3.37}$$

This means that the impact of cell-to-cell variability on the population variance can still be observed in the presence of an initial Poisson distribution of cells, but only if one is able to count the initial number of bacteria in the droplet, and to sort the droplets according to this number. Moreover, because of the term in $\frac{1}{\sqrt{N_0}}$, determining the population variance is only possible for droplets that contain very few bacteria. Note that this dependence in $\frac{1}{\sqrt{N_0}}$ was also found numerically in [78], and is here proven mathematically.

3.2.4 Generation-dependent division time

A final ingredient that must be added to the model is to account for the generation-dependent division time of bacteria. Indeed, when bacteria are set to grow in fresh medium the mean and variance of their division time changes with the generation number [80, 81]. Once again, we consider that bacteria grow independently from each other, so that modifying the first generation does not change the asymptotic exponential behavior of the moments, but the prefactors n_1 and n_2 are modified compared to the classical case.

Let τ_i and σ_i be the mean division time and its standard deviation for the $i-$th generation, where division times are always picked from a Gaussian distribution. To understand how to derive the prefactors, let us study the simpler case where only the first generation time is picked from a different distribution than the steady-state. Let ρ_1 be the density function of the first division time, and ρ_s that of the steady-state division times. The process starts at time $t = 0$ from a single bacterium that divides according to ρ_1. The two daughter bacteria divide according to ρ_s and give birth to an offspring of size N_1 and N_2, as shown in Fig. 3.9.

The total number of bacteria at time t is then $N(t) = N_1(t) + N_2(t)$, where N_1 and N_2 are independent and follow the same law. Therefore, the mean number of bacteria is $\mathbb{E}(N) = \mathbb{E}(N_1) + \mathbb{E}(N_2) = 2\mathbb{E}(N_1)$. Calling t_1 the time at which the initial bacterium has divided, we know that for $t > t_1$, all daughter bacteria divide according to the classical Bellman-Harris process. The asymptotic equivalent of $N_1(t)$ for $t \to \infty$ is then given by:

$$\mathbb{E}(N_1(t)) \sim n_1 e^{\alpha(t-t_1)}. \tag{3.38}$$

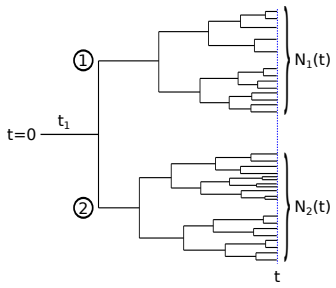


FIGURE 3.9: Schematic view of the division scheme in the case where the first division time is different: a single bacterium divides into 2 bacteria according to a division law with density function ρ_1, with mean τ_1 and standard deviation σ_1. The two daughter bacteria divide according to ρ_s and give birth to an offspring of size N_1 and N_2.

To obtain the expected value $\mathbb{E}\left(N(t)\right)$ of $N(t)$, we multiply by ρ_1 and integrate over all possible values of t_1:

$$\mathbb{E}(N(t)) \sim 2n_1 e^{\alpha t} \int_0^\infty e^{-\alpha t_1} \rho_1(t_1)\, \mathrm{d}t_1. \tag{3.39}$$

Calling $X_1 = e^{-\alpha t_1}$, the mean number of bacteria can be expressed as:

$$\mathbb{E}(N(t)) \sim 2n_1 \mathbb{E}(X_1) e^{\alpha t}. \tag{3.40}$$

Similarly, we obtain:

$$\mathbb{E}(N_1^2) \sim n_1^2 \frac{2\mathbb{E}(X^2)}{1-2\mathbb{E}(X^2)} e^{2\alpha(t-t_1)}, \tag{3.41}$$

which leads to

$$\mathbb{E}(N(t)^2) \sim 2n_1^2 \frac{\mathbb{E}\left(X_1^2\right)}{1-2\mathbb{E}(X^2)} e^{2\alpha t}. \tag{3.42}$$

These expressions are systematically compared to the simulations on Fig. 3.7(B). In the case of Gaussian division times, we know that the random variable X_1 is a log-normal variable, and its moments can be easily computed.

From this last result we obtain the expression of the square of the coefficient of variation in the case where only the first generation is changing, and compare it to the classical Bellman-Harris expression. The coefficient of variation does not depend on τ_1 but solely on σ_1:

$$CV_{\sigma_1}^2(\infty) = \left(\frac{n_2}{n_1}\right)^2_{\sigma_1} = \underbrace{e^{\alpha^2(\sigma_1^2-\sigma^2)}\left(\frac{n_2}{n_1}\right)^2_{BH}}_{(1)} + \underbrace{e^{\alpha^2(\sigma_1^2-\sigma^2)} - 1}_{(2)}. \tag{3.43}$$

This is again the sum of two terms, the first one being directly linked to the intrinsic variance of the Bellman-Harris process, and the second one being an additional variance due to a first generation time that is different

from the following ones. The effect is then similar to the Poisson case: the variance of the distribution of the number of cells is increased when $\sigma_1 > \sigma$, which is generally the case in experiments. When this variance increases, the contribution of the first generation rapidly becomes dominant over the contribution of the Bellman-Harris process (see Fig. 3.10(A)), meaning that the variance of the distribution is mostly due to the early times, and not to the cell-to-cell heterogeneity of division times, see Fig. 3.10(B). Another way to see this result is look at the shape of the self-similar distribution, which is almost independent of the value of the coefficient of variation of the division times, see Fig. 3.10(C).

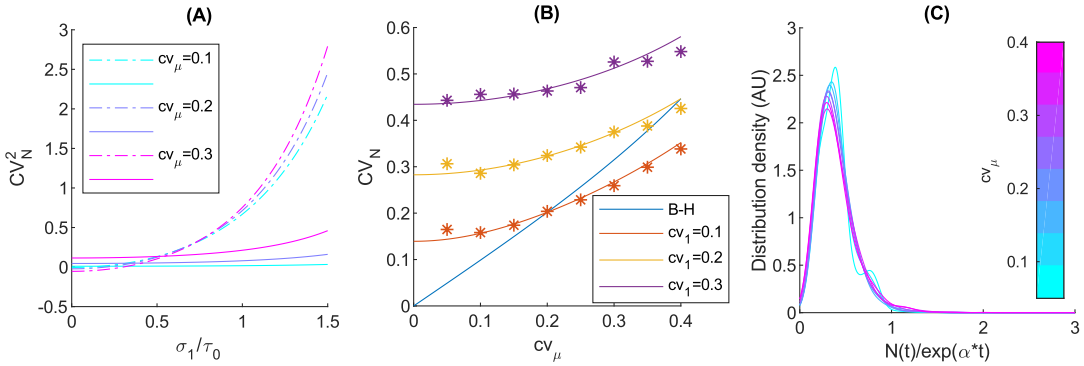


FIGURE 3.10: Bellman-Harris model and simulations for a different first generation with $\tau_1 = 2\tau_0$, varying σ_1. (A) Theoretical plot of the contributions to CV_N^2, when varying σ_1, for different values of cv_μ. Straight lines: contribution of the Bellman-Harris process, term (1) in eq. (3.43). Dotted lines: contribution of the different distribution of division times for the first generation, term (2) in eq. (3.43). (B) CV_N, asymptotic value coefficient of variation of the number of cells as a function of cv_μ: theory (lines) and simulations (stars), for different values of $cv_1 = \frac{\sigma_1}{\tau_1}$. (C) Shape of the distribution of $N(t)/\exp(\alpha t)$ (kernel fit, simulations only), for $\mathrm{cv}_1 = 0.3$.

3.2.5 Putting it all together: Three sources of stochasticity

Now what happens when we mix the two effects, *i.e.* have an initial number of cells per droplet following a Poisson distribution, and the first generation of bacteria having a different generation time? To answer this question, we inject eqs. (3.42) and (3.40) into eq. (3.23). We find that the two effects add up, leading to the following formula for CV_N^2:

$$\begin{aligned} CV^2_{\sigma_1,\lambda}(\infty) = \left(\frac{n_2}{n_1}\right)^2_{\sigma_1,\lambda} =& \frac{1-e^{-\lambda}}{\lambda} e^{\alpha^2\left(\sigma_1^2-\sigma^2\right)} \left(\frac{n_2}{n_1}\right)^2_{BH} \\ &+ \frac{1-e^{-\lambda}}{\lambda}\left(e^{\alpha^2\left(\sigma_1^2-\sigma^2\right)}-1\right) \\ &+ \frac{1-(\lambda+1)\,e^{-\lambda}}{\lambda}. \end{aligned} \tag{3.44}$$

We thus have the sum of three terms. The first one corresponds to the contribution due to the variation of the division times of the cells. The second one is the additional variance due to the first division time, but corrected to take into account only the wells that contain bacteria because of the Poisson distribution. And the third one is the supplemental variance due to this Poisson distribution. In this case, the intrinsic variance of the Bellman-Harris process is now masked by two terms, and as a consequence, the coefficient of variation CV_N is even less dependent on the heterogeneity of the division times cv_μ, see Fig. 3.11(A). The shape of the distribution of $N(t)/\exp(\alpha t)$, which is constant in time, is also almost independent of cv_μ, see Fig. 3.11(B).

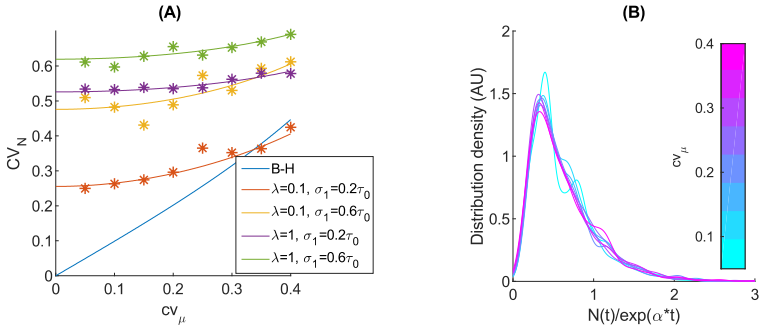


FIGURE 3.11: Theory and simulations for the mixed case where the first generation is different and the cells are initially distributed following a Poisson distribution. (A) CV_N, asymptotic value coefficient of variation of the number of cells as a function of cv_μ, theory (lines) and simulations (stars). (B) Shape of the distribution of $N(t)/\exp(\alpha t)$ (kernel fit, simulations only), for $\lambda = 1$ and $\mathrm{cv}_1 = 0.3$.

The same analysis can be extended to the arbitrary case where the first k generations follow different division laws, and combined with an initial number of bacteria following a Poisson distribution. In the particular case of our experiments, the distribution of division times reaches steady-state at the fourth generation, and we find that at steady-state, the mean $M_{\tau_n,\sigma_n}(t) \sim n_1(\tau_n,\sigma_n)e^{\alpha t}$ and the standard deviation $SD_{\tau_n,\sigma_n}(t) \sim n_2(\tau_n,\sigma_n)e^{\alpha t}$ with:

$$n_1(\tau_n,\sigma_n) = 8n_1\frac{\lambda}{1-e^{\lambda}}\prod_{i=1}^{3} e^{-\alpha\tau_i+\frac{\alpha^2\sigma_i^2}{2}}, \tag{3.45}$$

$$\begin{aligned} CV^2_{\tau_n,\sigma_n}(\infty) &= \left(\frac{n_2}{n_1}\right)^2_{\tau_n,\sigma_n} \\ &= \frac{1-e^{-\lambda}}{\lambda}\left(\frac{1}{8}e^{\alpha^2\sigma_1^2}\left(e^{\alpha^2\sigma_2^2}\left(\frac{e^{\alpha^2\sigma_3^2}}{1-\frac{1}{2}e^{\alpha^2\sigma^2}}+2\right)+4\right)-\frac{\lambda e^{-\lambda}}{1-e^{-\lambda}}\right). \end{aligned} \tag{3.46}$$

Although these formulas are complex, they are the reflection of simple ingredients and can easily and reliably be computed for a given set of parameters. Now let us try to see how they compare with the results of the microfluidic experiments

3.3 Experimental results

In this part, we are going to see how the results of microfluidics time-lapse growth experiments, that were described in the previous chapter, are in agreement with the predictions of the Bellman-Harris model that we have described numerically and theoretically. As explained in the previous chapter, the fluorescence signal of the droplet is used as a proxy for the number of bacterial cells in the droplet, and we will consider first that it is exactly proportional to the number of cells in a droplet (this hypothesis will be questioned in section 3.3.5). We first discuss in depth the results for *E. coli*, and those for *B. subtilis* will be exposed later.

3.3.1 Growth Rate, Mean and Standard Deviation

The first prediction of the Bellman-Harris model that we can check on the experiment is that all droplets containing bacteria should exhibit the same growth rate. As explained in the previous chapter, we set manually and once and for all for all the droplets the detection and saturation limits. Between these two limits the growth rate is obtained for each droplet by fitting the growth curve with an exponential. We obtain a standard deviation of only $\approx 7\%$ among the growth rates of all the droplets, which is very good taking into account the experimental noise and fitting errors. Thus we can say that the first prediction of the B-H model is satisfied by the experiments. This prediction is, of course, also the classical belief in microbiology [112], at least while nutrients are abundant in the droplet.

The second prediction that we can check is that both the mean and the standard deviation should evolve exponentially with the same growth rate. Indeed, if we take the mean and standard deviations of all the fluorescent curves of the droplets that contain bacteria, we obtain a very well defined exponential phase, for approximately 2 hours (see Fig. 3.12).

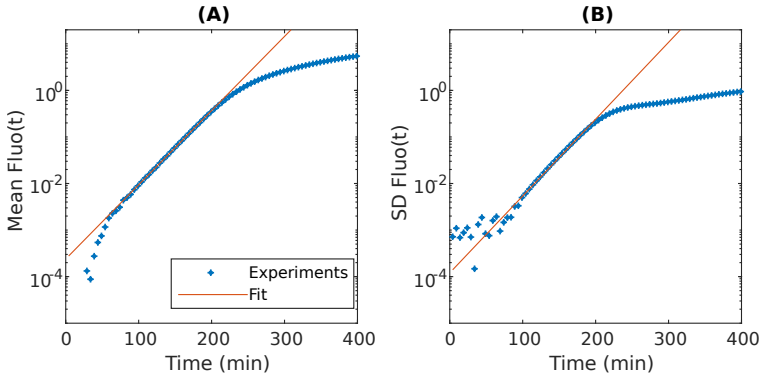


FIGURE 3.12: Mean (A) and Standard Deviation (B) of the fluorescence signal of the droplets in a typical *E. coli* experiment (described in Chapter 2), in log scale, as a function of time, with linear fit (orange).

If we compare the slopes of the two fits, we obtain $\alpha_M \approx 2.19\text{h}^{-1}$ for the mean, and $\alpha_{SD} \approx 2.22\text{h}^{-1}$ for the standard deviation. The relative difference

between these two values is then only of about 1%, we can thus say that the experiments agree with the B-H model on that point. Note that the mean of the growth rates of the droplets is also very close, since its value is $< \alpha_{drop} > \approx 2.14 \pm 0.08 \mathrm{h}^{-1}$.

Hence, we have checked that all droplets grew exponentially with the same growth rate, as well as the mean and standard deviation of the fluorescence signal. To push the comparison further, we need to know if this growth rate that we have found really means something for the bacteria inside the droplets, by having an idea of the distribution of their division times. Thus we could use these division times as inputs for our full Bellman-Harris model. This will allow us to have a comparison for the value of the growth rate, but not only, as the coefficient of variation of the distribution of $N(t)$ could also be computed and compared with the experiments.

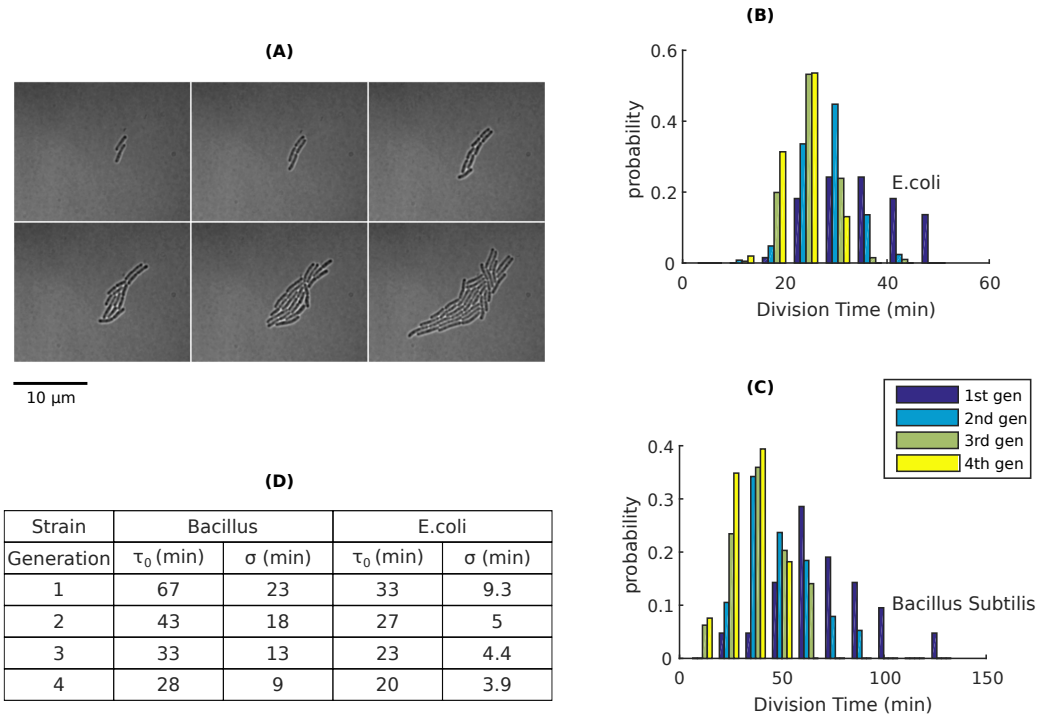
3.3.2 Individual division times

To find the individual division times and their distributions, we confine bacteria between a glass slide and a thin homemade agarose pad (mixed with LB2X at 50/50), and then observe them under the microscope every 3 minutes for 3 hours, with a 90X magnification, kept at 37°C (see Fig. 3.13(A) for an example). The division times were determined with a segmentation algorithm [136]. The pre-culture conditions were the same as on the chip, in order to obtain the more relevant division times to use as inputs for our Bellman-Harris model. However, since the bacteria are here growing in agarose under mechanical constrain and not in liquid, we know that the comparison with the microfluidic experiment is not perfect, but this will give us an idea of the mean division time and its variability.

We find that the distribution of division times can be well approximated by a Gaussian with mean τ and standard deviation σ, see Fig. 3.13(B,C). Moreover, the parameters of the Gaussian distribution evolve with the generation number, a sign of the adaptation of bacteria to their new environment [137]. For *E. coli*, the average generation time decreases from 33 to 21 minutes between the first and the fourth generation, while the coefficient of variation $cv_\mu = \sigma/\tau$ decreases from 0.28 to 0.19 over the same time. This is in agreement with results previously reported in the literature on *E. coli* [81] and *Salmonella enterica* [80]. In full exponential phase, Taheri-Araghi *et al.* [45] reported coefficients of variations ranging from 0.14 to 0.22 for *E. coli*.

Some references in the literature, such as refs. [78, 87], suggest to take gamma distributions instead of Gaussian distributions for the cell division times. However, we have not measured enough division events to statistically distinguish between the two distributions, and switching from one to another does not qualitatively affect our results, see Fig. 3.14.

We can use then these distributions of division times as inputs for our Bellman-Harris model. We will consider that after the fourth generation, the cells are in exponential phase and the distribution of their division times is constant. We could not really check this hypothesis because after four or five division, the field of view of the camera was filled with bacteria, and those



(D)

Strain	Bacillus		E.coli	
Generation	τ_0 (min)	σ (min)	τ_0 (min)	σ (min)
1	67	23	33	9.3
2	43	18	27	5
3	33	13	23	4.4
4	28	9	20	3.9

FIGURE 3.13: (A) Time-lapse image of *E. coli* cells dividing under a 90X objective. The images correspond to t=0, 21, 42, 63, 84, and 105 min respectively. (B) Density of the division times obtained from the time-lapse images, for *E. coli*, and for the first four generations. (C) Density of the division times obtained from the time-lapse images, for *B. subtilis* and for the first four generations. (D) Fitted values of the mean division times and their standard deviations, for both strains.

begun to grow in 3D, out of focus... Note though that even between the third and fourth generation, the difference is already quite low (Fig. 3.13(B)). This simulation scheme as already been used in the literature ([80],[81]).

3.3.3 Full comparison

The distribution of division times from the single-cell experiments can now be used as inputs to obtain theoretical and numerical predictions. The Poisson parameter is obtained directly from the chip experiment by counting the number of positive droplets, as explained in Chapter 2.

Growth rate

The first thing that we can compare is the value of the growth rate. If we take the theoretical formula (3.14), and inject the values of tau_0 and cv_μ that we have found for the individual cells, we obtain $\alpha_{\text{B-H}} = 2.01\text{h}^{-1}$. This is very close to the values discussed above (to recall, $\langle \alpha_{drop} \rangle \approx 2.14 \pm 0.08\text{h}^{-1}$ for the mean of all the droplets). This bolsters us in our hope that the division times that we have found in our agar pads are relevant for the comparison with

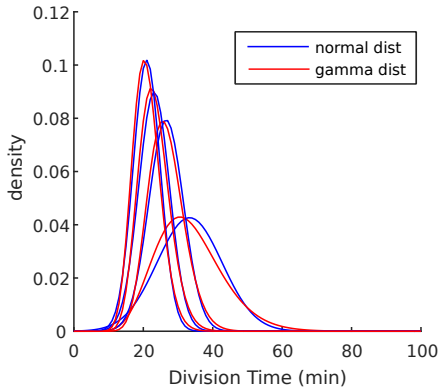


FIGURE 3.14: Comparison of fits for the experimental distribution of division times shown in Fig. 3.13, for the first 4 generations. Blue: fit to a Gaussian distribution. Red: fit to a Gamma distribution.

the on-chip growth. Note though that we find a slightly bigger growth rate when measuring in the droplets, this could indicate that the cells had not completely reached their full exponential phase at their fourth division.

This good match in terms of growth rate is clear if we plot on the same graph the experimental, numerical and theoretical mean and standard deviation (Fig. 3.15(A),(B)). However, the experimental data is in fluorescence and not in number of cells. To gather the curves on the same plot, we had to manually adjust vertically the experimental curves, which gives us in fact a measure of the proportionality between the fluorescence of the droplets and the number of cells. The same coefficient was used for the mean and the standard deviation. Here, we find that $a_f \approx 7 \cdot 10^{-4}$, where $Fluo(t) = a_f N(t)$.

We also recover an exponential growth for higher moments of the distribution, at least up to the fifth central moment, as shown on Fig. 3.16. The fluorescence coefficient used here is the one found on the mean. We can see that the growth rate found on those higher moments is also correct, even if it starts to deviate for the fifth moment. Higher (>5) moments don't present this exponential growth, as they are more sensitive to the saturation limit.

Coefficient of variation

However the value of this growth rate cannot distinguish between the simple Bellman-Harris model of Section 3.2.2 and the complete model of Section 3.2.5. This distinction can be made by measuring the coefficient of variation CV_N of the number of bacteria, which is predicted to be constant in the exponential phase. Such a behavior is indeed observed in the simulations (see Fig. 3.15(C)). The experimental coefficient of variation is also approximately constant, but only for a short time of 1 hour see Fig. 3.15(C), even though the exponential growth of the mean and standard deviation makes it difficult to obtain a real plateau: a slight difference in the observed growth rates of theses two quantities leads to an exponentially increasing difference

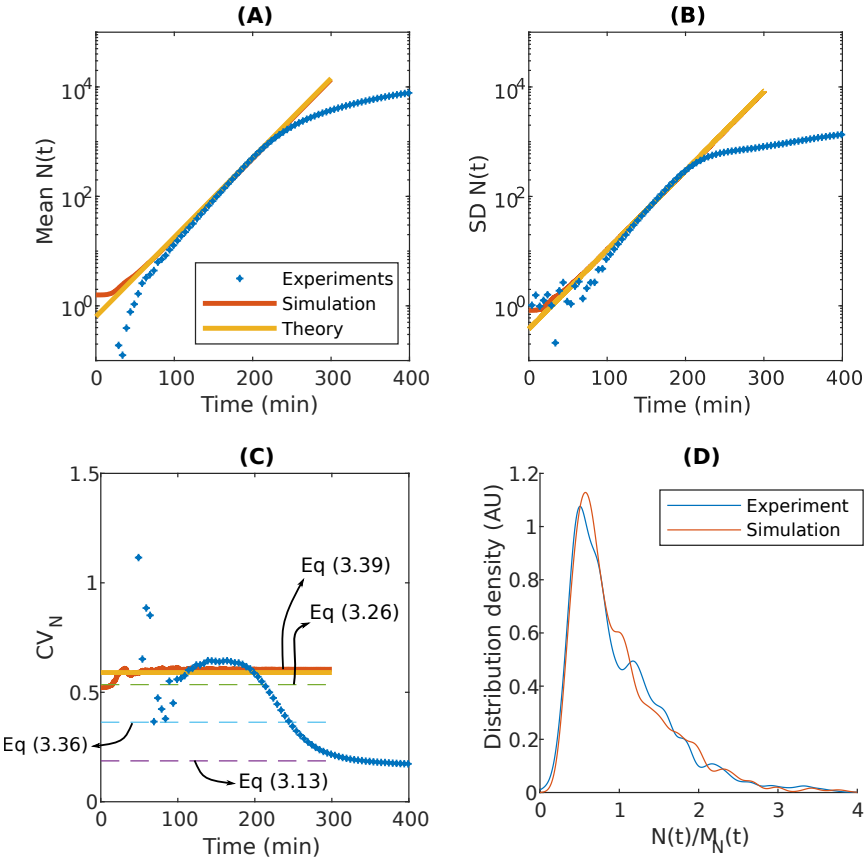


FIGURE 3.15: Comparison between the model (yellow), numerical (orange) and experimental results (blue)(same experiment as in Chapter 2). (A) Mean number of cells as a function of time $M_N(t)$, (B) Standard deviation of the number of cells as a function of time $SD_N(t)$, (C) Coefficient of variation of the number of cells as a function of time CV_N. Dashed lines correspond to the asymptotic value of CV_N when the sole source of variability is the cell division time (classical B-H theory, purple), in the presence of an initial Poisson distribution of cells (cyan), and in the presence of both a Poisson distribution of cells and an evolving distribution of division times (green). (D) Shape of the distribution of $N(t)/\exp(\alpha t)$, numerical and experimental, kernel fit.

between the values of the mean and standard deviation. Yet, the experimental value of the plateau during this period is in good agreement with the expected theoretical value, see Fig. 3.15(C). Note that all sources of variability need to be taken into account to obtain a good agreement between the experimental and theoretical CV_N. Considering solely stochastic cell division, as in the classical Bellman-Harris model, significantly underestimates the experimental CV_N (dashed purple line). As more sources of variability are added to the model, the coefficient of variation increases. The effect of both the initial Poisson distribution (cyan dashed line), and of the generation-dependent division times (green dashed line) on the value of the coefficient of variation are shown in Fig. 3.15(C).

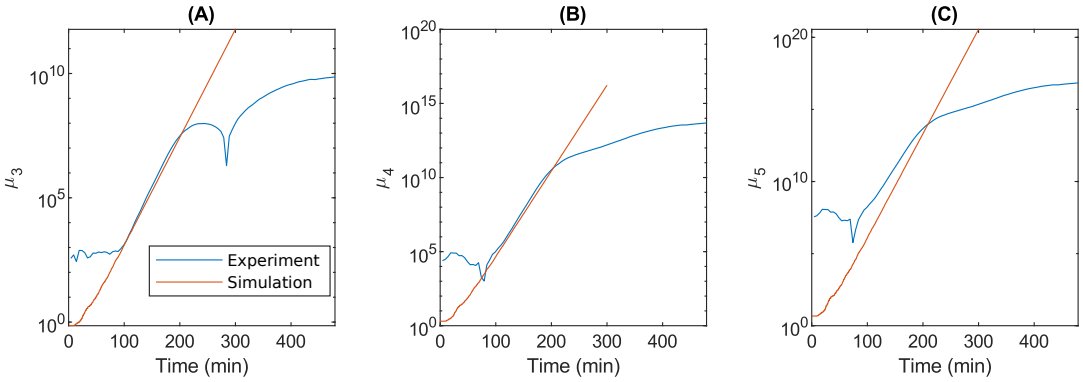


FIGURE 3.16: Comparison between experiments and simulations for the evolution of the 3rd (A), 4th (B) and 5th (C) central moments of the distribution of the number of cells. Data shown are for *E. coli*.

Note that the proportionality constant between the fluorescent signal and the number of bacteria cancels out when computing the ratio $CV_N = n_2/n_1$. As such raw experimental data can be directly compared with the simulations and the theory with no fitting parameter. The agreement that is observed between experimental data and theory shows that all sources of variability have to be taken into account to understand the variability in behaviors between droplets: the initial Poisson distribution, the first generations times and the cell-to-cell variability through a Bellman-Harris process. When one of these sources of variability is not taken into account, the error between experimental results and the theoretical ones is larger, see dashed lines in Fig. 3.15(C).

Self-similarity

Finally, we verify that the shape of the experimental distribution of fluorescence is self-similar in time, when rescaled by its mean value (see Supplementary Movie 5 in Appendix E). The self-similar behavior is limited to the exponential phase, before saturation of bacterial growth. Furthermore the shape of this distribution is in very good agreement with the numerical prediction of the full model, as shown in Fig. 3.15(D). The distribution is here rescaled by the mean and not by $\exp(\alpha t)$ to get rid of the coefficient of proportionality between the fluorescence and the number of cells. Thus, the distribution of the number of cells in the microfluidic droplet is indeed the one expected from the microscopic cell-to-cell variability and the additional sources of randomness.

Results for *B. subtilis*

We have presented above only the on-chip comparison for *E. coli*, but we obtain similar results for *B. subtilis*, as we can see on Fig. 3.17. We do obtain an exponential growth phase for the mean and standard deviation, with the expected growth rate (3.17(A),(B)). The results are a bit less satisfying for

the coefficient of variation: it deviates more form the expected theoretical value (3.17(C)) than what we had for *E. coli* (3.17(A),(B)). Part of the error is due to the fact that we have used Gaussian division times instead of Gamma-distributed ones. As we can see on Fig. 3.13(C), the division times of *B. subtilis* are more skewed to the right than those of *E. coli*, especially for the first generation, so if we use Gamma-distributed division times in our Bellman-Harris model, it yields slightly better results (3.17(D)). The theoretical computation of the Bellman Harris model with a Gamma law are very easy to make, they are even presented in the original paper[76] (see Appendix B). However, part of the error remains even with that modification. Perhaps *B. subtilis*, which is a bacteria normally living in the soil [138], has a more different growth in liquid, and thus the individual division times obtained with the agar pad experiment are less comparable to those happening on-chip, contrary to what happens with *E. coli*, where they seem to be a good input for our Bellman-Harris model.

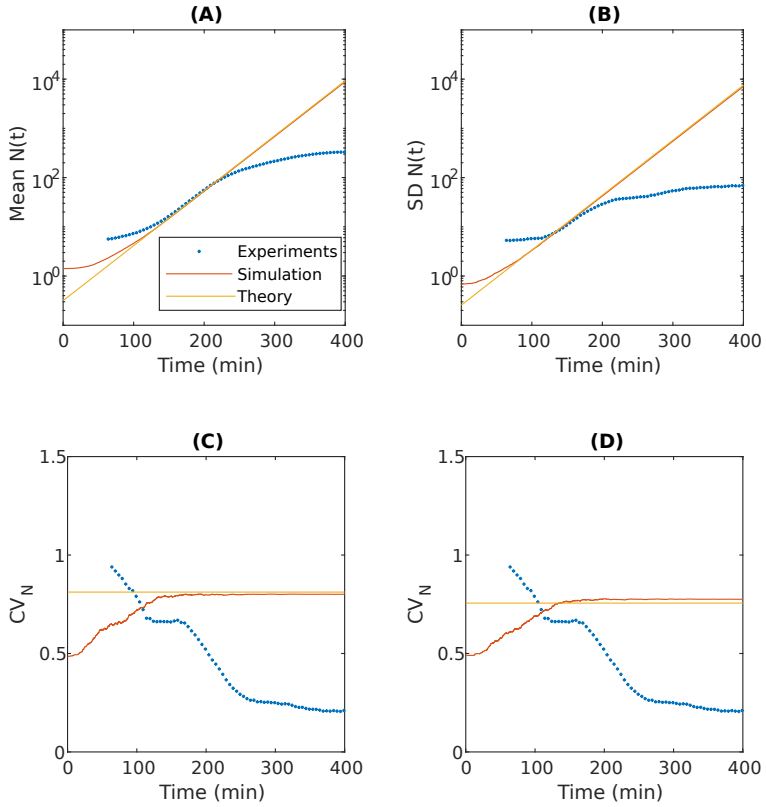


FIGURE 3.17: Comparison between the model (yellow), numerical (orange) and experimental results (blue)(same experiment as in Chapter 2). (A) Mean number of cells as a function of time $M_N(t)$, (B) Standard deviation of the number of cells as a function of time $SD_N(t)$, (C) Coefficient of variation of the number of cells as a function of time CV_N (D) Same, but using Gamma division times instead of Gaussian.

3.3.4 Towards inference ?

However, if we obtain a good match between our microfluidic experiments and the theory and simulations of the Bellman-Harris model, those necessitated to make single cell measurements of the division times. Those experiments are quite painful to realize when compared with microfluidic experiments. Another great possibility of the Bellman-Harris model is the following: it can theoretically be used the other way round, in the sense that the observations made on the population can be used to infer data on the single cells. Indeed, with Gaussian division times, there are only two unknowns that characterize the bacteria in the exponential phase : the mean division time τ_0 and the coefficient of variation cv_μ. And we also have two observables that characterize the growth of the population in the droplets, which are the growth rate α and the coefficient of variation CV_N, and we know how they depend on the microscopic parameters. Thus the equations (3.14) and (3.44) can theoretically be inverted, and the microscopic values characterizing the growth of the bacteria be found from the distribution of the population in the droplets.

Unfortunately, it does not work that well in practice. As we have discussed in the theoretical part, the variability that we observe among the droplets depends much less on the cell-to-cell variability than on the other sources of variability such as the initial Poisson distribution or the different first generation time. This is indeed very clear on Fig. 3.8(B), 3.10(B), and 3.11(A), where we see that the dependence in cv_μ is quite low for CV_N. This is also true for the self-similar shape of the distribution (Fig. 3.10(C) and Fig. 3.11(B)), which depends weakly on cv_μ. Hence, even if the equations can be theoretically inverted, the result will be of no use in practice: because of this loose dependence, an error of a few percent in terms of CV_N will give a huge error in terms of cv_μ.

Of course, this is only true for cv_μ, and an approximation of the mean division time of the bacteria can be found, using the growth rate, which depends mostly on τ_0 (Fig. 3.3(A)), but this can be made routinely with classical OD measurements. In our case, using just the classical formula $\alpha = \log(2)/\tau_0$, we can estimate $\tau_0 \approx 19.5$min, which is very close to the value of 21 min found in the single cell measurements.

Getting rid of these external influences on CV_N can be quite challenging. To decrease the influence of the Poisson distribution, we would have to lower the Poisson parameter λ (see Fig. 3.8(A)). But as a consequence, as our droplet number is limited, the deviation from the model would be bigger, simply because of the stochasticity, as we discussed previously using the simulations (Fig. 3.8(B)), and it is not sure that reliable measurements of CV_N could be made. To decrease the influence of the first generations, the bacteria should be collected in exponential phase and not diluted before being loaded on to the chip, and the loading made at 37°C. This is doable, but even by doing this, there would still be an issue: the cells would not be synchronized when loaded on to the chip, and even this is sufficient to lower the dependence of CV_N on cv_μ, as we can see on Fig. 3.18. Solutions to synchronize cells exist in the literature, but as we want to keep the cells in exponential phase, the

only way would be the so-called "baby-machine" [139]. However, it would be difficult to get it to a sufficient efficiency and to couple it with a microfluidic device.

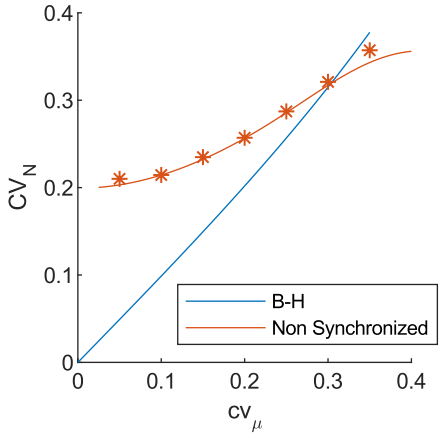


FIGURE 3.18: Effect of the initial non-synchronization of the cells on the relation between the cell to cell variability and the heterogeneity among droplets: CV_N as a function of cv_μ, simulations (stars) and theory (line) for a fully non-synchronized Bellman-Harris model (orange) and a simple Bellman-Harris model (blue).

The theoretical computation for the non-synchronized case can be made quite easily. If the cells are not synchronized, then the first division time is the product of two terms: the Gaussian law that we normally use in our Bellman-Harris model is multiplied by an independent uniform variable between 0 and 1, that represents the moment in the cell cycle at which the cell was encapsulated in the droplet. To compute what happens in this case, we can use the same trick as for the different first generation, because after the first generation, every cell behaves as in a simple Bellman-Harris model. We condition the variable N by this first division time to get:

$$\mathbb{E}\left(N(t)|T_1 = t_1\right) \sim 2n_1 e^{\alpha(t-t1)}, \tag{3.47}$$

which implies, by the law of total probability,

$$\mathbb{E}\left(N(t)\right) \sim 2n_1 e^{\alpha t}\mathbb{E}\left(e^{-\alpha T_1}\right). \tag{3.48}$$

Once again, it all comes down to computing the expected value of $X_1 = e^{-\alpha T_1}$, where T_1 is random variable for the first division time. But we know that T_1 is the product of an independent uniform variable U by a Gaussian law of distribution $\mathcal{N}(\tau_0, \sigma)$. Hence, T_1 knowing U is just a normal law of distribution $\mathcal{N}(u\tau_0, u\sigma)$. Therefore, the variable X, conditioned by U is once again a rescaled log-normal law, and we have:

$$\mathbb{E}\left(e^{-\alpha T_1}|U = u\right) = e^{-\alpha u\tau_0 + \frac{\alpha^2 u^2 \sigma^2}{2}}. \tag{3.49}$$

Then, we use the law of total probability to get:

$$\mathbb{E}\left(e^{-\alpha T_1}\right) = \int_0^1 e^{-\alpha u \tau_0 + \frac{\alpha^2 u^2 \sigma^2}{2}} \mathrm{d}u. \tag{3.50}$$

This last integral can be computed numerically. Exactly the same analysis can be conducted with N^2 instead of N, and thus we obtain the theoretical value of CV_N that is plotted on Fig. 3.18.

Another try: log-likelihood method

Considering what we have described above, it seems difficult to use the direct equation of the Bellman-Harris model and to use them for inference. However, this is not the only possibility for inferring the parameters for individual cell division, namely τ_0 and cv_μ, using the measured distribution of $N(t)$. A classical tool for statistical inference are maximum likelihood estimators (MLE) [140].

Those methods are based on the following principle: we consider a model that depends on a set of parameters θ, and we have a set of observed empirical data x. The idea is to estimate the correct parameters θ by selecting the value of θ that maximizes the probability of observing the data set x in the model. To do this, a likelihood function is computed and maximized over the range of possible parameters θ. The value of the likelihood function is the probability of observing x in the model, given a set of parameters θ. For practical reasons, the log of the likelihood function is often considered.

Such estimators have already been built for binary processes resembling the Bellman-Harris process considered in this Chapter (see for instance refs. [91, 141, 142]), but they only apply to models where there is an explicit expressions of the distribution of the number of cells as a function of time. We do not have access to such an explicit expression in our Bellman-Harris process, and another difficulty is that we only have a few available observation points, that are not the same for all the droplets, because of the experimental conditions: for each droplet, we have to be in between the detection and saturation limits, see Chaper 2.

However, Hyrien [92] developed a method that matches our experimental constrains. It is called a pseudo-likelihood method, as the estimation is based on a numerical approximation of the likelihood function, to bypass the unavailability of an explicit expression of the distribution of the number of cells. This method also takes into account the fact that the observation points are not the same for every droplet. We will not enter too much into details here, but the method is based on a numerical estimation of the mean and variance-covariance cell counts by Monte-Carlo simulations. Therefore, a numerical likelihood function can be computed, and maximized also numerically.

Therefore, we tried to apply this estimation method in our case. But before trying with the experimental data, we tried on a simulation. The principle is the following. On the one had, we consider a test simulation of the growth of *E. coli*, that mimics what can be observed on our chip, thanks to

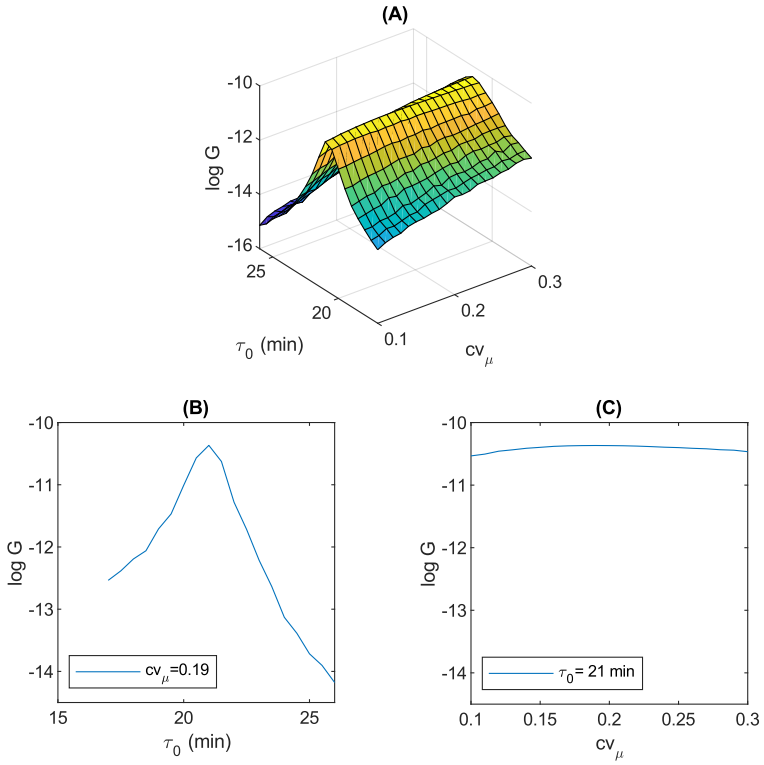


FIGURE 3.19: Simulated results of the pseudo log-likelihood method described in [92]. A full simulation of the growth of *E. coli* (same as on Fig. 3.11) was used as a test for inference of the growth parameters τ_0 and cv_μ, using the pseudo-likelihood method, and a grid of Monte-Carlo simulations. For each point the (τ_0, cv_μ) grid, 2000 simulations were computed. (A) General shape of the likelihood estimation function G, as a function of the parameters τ_0 and cv_μ. If we consider the maximum of G, we find the correct values of τ_0 and cv_μ that were used in the test simulation, respectively $\tau_0 = 21$ min and $cv_\mu = 0.19$. (B) Variation of $\log(G)$ as a function of τ_0 when cv_μ is fixed at 0.19. (C) Variation of $\log(G)$ as a function of cv_μ when τ_0 is fixed at 21 min. We see that the function G is much more peaked when τ_0 is varied than when cv_μ is varied.

our full Bellman-Harris model, already used for Fig. 3.11. Our goal is to estimate the values of τ_0 and cv_μ that were used for the simulation. On the other hand, we run simulations of this full Bellman-Harris model for a grid of the parameters (τ_0, cv_μ) that covers the possible values of these parameters. This grid is used for the computation of the pseudo-likelihood function, and we look at the value of the parameters that maximize this function given the test simulation. For this, we set ourselves a number of observation points that mimic those that we have experimentally.

The results are presented on Fig. 3.19. The positive result is that we are able to estimate the correct value of the parameters: the maximum of the pseudo-likelihood function is found for $\tau_0 = 21$ min and $cv_\mu = 0.19$, which are the values that were indeed used for the test simulation. However, if

we look more closely at the shape of the likelihood function that we get (Fig. 3.19(A)), we note that the estimation is very selective on the value of τ_0 (Fig. 3.19(B)), but not on cv_μ. Indeed, for varying cv_μ, the shape of the likelihood function is almost flat (Fig. 3.19(C)), whereas the maximum is very peaked for τ_0. In a sense, this is not surprising, as we have already explained that the shape of the distribution of $N(t)$ was weakly dependent on cv_μ.

As a consequence, it is almost impossible that this method will work for the experimental estimation of cv_μ, because a slight noise on the data, for instance will probably have huge consequences on the estimation of cv_μ. Indeed, we could never get this method to work with our experimental data, the estimation that we got on cv_μ was always far from the one measured on single cells on an agar pad. Another difficulty lies in the fact that we only have an indirect estimation of the number of cells through the measurement of fluorescence, and the link in between those two quantities can be difficult to quantify, as we have discussed in Chapter 2.

As a whole, these results on the inference of cv_μ through the inversion of the Bellman-Harris equation or pseudo-likelihood methods suggest that because of our experimental constrains, it might be difficult to infer cv_μ from the experiment using only the distribution of the number of cells at the different observation points. In the next Chapter, we will try an inference method rather based on following the individual trajectories of each droplet with time.

3.3.5 Experimental noise

If we take another look at the comparison between the experimental and theoretical coefficient of variation, we see that there is a little error in between them. The experimental one is indeed constant, but slightly bigger than the expected value. This error could come from a bad estimation of the individual division times of the bacteria: as we have already said, they were estimated in different culture conditions than the on-chip ones. However, they could also come from the measurement noise, particularly when measuring the fluorescence of the droplets. This is what we are going to explore theoretically and numerically in this part.

Additive noise

The first kind of noise that we can think of is a simple additive noise, which can be written as follows:

$$Fluo(t) = a_f N(t) + \eta(t). \tag{3.51}$$

Where η models the noise and can be at first approximation taken as a Gaussian: $\eta(t) \sim \mathcal{N}\left(0, \sigma_\eta^2\right)$. We will make two hypotheses on η. First, we consider that η is independent of N. Second, we consider that, as a random noise, it is also not correlated in time, which means that for our observations $\eta(t_i)$ and $\eta(ti+1)$ are independent.

In this case, we simply have, by independence of N and η:

$$\begin{aligned}\mathbb{E}\left(Fluo(t)\right) &= a_f \mathbb{E}(N(t)) + \mathbb{E}(\eta(t)) \\ &= a_f \mathbb{E}(N(t)),\end{aligned} \tag{3.52}$$

$$\begin{aligned}\mathrm{Var}\left(Fluo(t)\right) &= a_f^2 \mathrm{Var}(N(t)) + \mathrm{Var}(\eta(t)) \\ &= a_f^2 \mathrm{Var}(N(t)) + \sigma_\eta^2.\end{aligned} \tag{3.53}$$

Therefore, if we compute the coefficient of variation, we get:

$$CV_{fluo}^2 = CV_N^2 + \frac{\sigma_\eta^2}{a_f^2 \mathbb{E}(N(t))^2}. \tag{3.54}$$

This is the sum of two terms. The first one is just the coefficient of variation for the number of cells, and the second one is the corrective term due to the noise. As we can see, it will rapidly vanish, because the numerator is constant but the number of cells evolves exponentially. This is confirmed by the simulations (Fig. 3.20(A)). There is then little chance that it can account for a constant increase of the experimental coefficient of variation in the exponential phase.

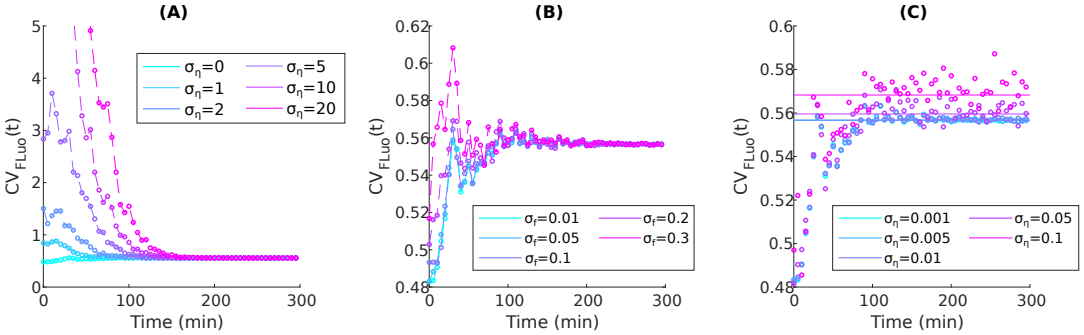


FIGURE 3.20: Effect of different kind of noise on the measured coefficient of variation of the fluorescence. The simulation used is the full simulation of *E. coli* growth presented on Fig. 3.15, with division times taken from the single cell experiment, then the noise is added with $a_f = 1$, and varying its variance (A) Additive noise (B) Heterogeneity of the fluorescence (C) Multiplicative noise, with theory (straight lines) and simulations (circles).

Heterogeneity of the fluorescence

Another source of noise in the experiment can come from the heterogeneity of the fluorescence signal, which is not homogeneous among the cells but varies from one bacteria to another, even if the cells are isogenic [111]. In this case, the relation between the fluorescence and the number of cells can be

written as:

$$Fluo(t) = \sum_{i=1}^{N(t)} a_f^i(t), \tag{3.55}$$

where $a_f^i(t)$ represents the fluorescence signal of cell i at time t. To simplify the computations, we will consider that the fluorescence signals of the cells are independent, and that they are also independent in time. These are of course simplifying hypotheses, as the fluorescence signal of two sister cells is probably related for instance. The fluorescence is also probably correlated for one cell from time t_i to time t_{i+1}, but as we average on a lot of cells, these effects are negligible compared to the heterogeneity in itself. We will note σ_f^2 the variance of a_f^i and a_f their mean.

Then we have, using Appendix B.4:

$$\mathbb{E}\left(Fluo(t)\right) = a_f \mathbb{E}(N(t)), \tag{3.56}$$

$$\mathrm{Var}\left(Fluo(t)\right) = \mathbb{E}(N(t))\sigma_f^2 + a_f^2 \mathrm{Var}(N(t)). \tag{3.57}$$

This yields for the coefficient of variation:

$$CV_{fluo}^2 = CV_N^2 + \frac{cv_f^2}{\mathbb{E}(N(t))}. \tag{3.58}$$

We can see that we have the same effect as for the additive noise : the correction term due to the noise is vanishing at long times, even if it vanishes a little bit less rapidly here (in $e^{-\alpha t}$ instead of $e^{-2\alpha t}$). This is confirmed by the simulations, see Fig. 3.20(B). Once again, the heterogeneity of the fluorescence among cells is unlikely to account for the observed increase of the coefficient of variation.

Multiplicative noise

The experimental noise could be not only additive, but also multiplicative, meaning that we could have :

$$Fluo(t) = a_f \eta(t) N(t). \tag{3.59}$$

With $\eta(t)$ being a random variable, independent of $N(t)$, with mean 1 and standard deviation σ_η. This multiplicative modeling of the experimental noise may sound stranger than the more common additive one, but it could make sense if we think that the background in the fluorescent images had to be divided and not subtracted to get the correct signal. We will also consider that the noise is not correlated in time.

In this case, we simply have, by independence of N and η:

$$\mathbb{E}\left(Fluo(t)\right) = a_f \mathbb{E}(N(t)), \tag{3.60}$$

$$\mathrm{Var}\left(Fluo(t)\right) = a_f^2 \mathrm{Var}(N(t))\sigma_\eta^2 + a_f^2 \mathrm{Var}(N(t)) + a_f^2 \mathbb{E}(N(t))^2 \sigma_\eta^2. \tag{3.61}$$

and for the coefficient of variation:

$$CV^2_{fluo} = CV^2_N + CV^2_N\sigma^2_\eta + \sigma^2_\eta. \tag{3.62}$$

In this case, we can see that the measured coefficient of variation is indeed increased when compared with the coefficient of variation of the population. It could then be the explanation of the observed increase. But as we can see on Fig. 3.20(C), this increase is very low and has another effect that we don't see at all in the experiments : it increases the dispersion of the obtained CV around the theoretical value a lot. There are then few chances that this is the correct explanation for what we see experimentally.

Heterogeneity of the droplet initial volume

Another source of experimental variability that we have not taken into account at all for now is the little heterogeneity that exists in the droplet initial volume, as described in the previous chapter.

To recall, we will consider that the initial volume of the droplets follows a Gaussian distribution, which implies that the number of cells in each droplet follows a Poisson distribution, but with a Gaussian distribution of Poisson parameters λ.

$$\lambda \sim \mathcal{N}(\lambda_0, \sigma_\lambda). \tag{3.63}$$

To see how that impacts the droplet to droplet variability in exponential phase, we will use the lemma of Appendix B.4, just as we did for the case of the initial Poisson distribution, to recall we got the following equations (eq. (3.22),(3.23)) :

$$\mathbb{E}(N(t)) = M_N(t)\mathbb{E}(N_0), \tag{3.64}$$

$$\mathrm{Var}(N(t)) = SD_N(t)^2\mathbb{E}(N_0) + M_N(t)^2\mathrm{Var}(N_0). \tag{3.65}$$

The only difference with this simple Poisson case is that we don't have an exact Poisson law for N_0 any more but instead a mixture of Poisson law of normally distributed parameters. By using the law of total probability, we get:

$$\mathbb{E}(N_0) = \mathbb{E}_\lambda(\mathbb{E}(N_0|\lambda)), \tag{3.66}$$

and we know that N_0 knowing λ is the restriction of a Poisson law to the positive numbers (as we consider the droplet that contain bacteria), so:

$$\mathbb{E}(N_0) = \mathbb{E}_\lambda\left(\frac{\lambda}{1-e^{-\lambda}}\right). \tag{3.67}$$

The expression of this expected value is not that simple, as it depends on the density of the Gaussian distribution of λ, we will just state the integral form here, and then it can be computed numerically for a given set of parameters $(\lambda_0, \sigma_\lambda)$:

$$\mathbb{E}_\lambda\left(\frac{\lambda}{1-e^{-\lambda}}\right) = \int_{-\infty}^{\infty} \frac{\lambda}{1-e^{-\lambda}} \frac{e^{-\frac{(\lambda-\lambda_0)^2}{2\sigma^2_\lambda}}}{\sigma_\lambda\sqrt{2\pi}} \mathrm{d}\lambda. \tag{3.68}$$

Similarly, the expression of the variance of N_0 can be obtained and computed numerically:

$$\begin{aligned}\mathrm{Var}\left(N_0\right) &= \mathbb{E}_\lambda\left(\mathrm{Var}\left(N_0|\lambda\right)\right) + \mathrm{Var}_\lambda\left(\mathbb{E}\left(N_0|\lambda\right)\right) \\ &= \mathbb{E}_\lambda\left(\frac{\lambda}{1-e^{-\lambda}}\frac{1-(\lambda+1)\,e^{-\lambda}}{1-e^{-\lambda}}\right) + \mathrm{Var}_\lambda\left(\frac{\lambda}{1-e^{-\lambda}}\right).\end{aligned} \tag{3.69}$$

Now that we have the expressions of the expected value and variance of N_0, we can go back to the expression of the variance and expected value of $N(t)$ to find the square of the coefficient of variation, and we get :

$$CV^2_{\lambda_0,\sigma_\lambda} = \frac{1}{\mathbb{E}\left(N_0\right)}CV^2_{BH} + \frac{\mathrm{Var}\left(N_0\right)}{\mathbb{E}\left(N_0\right)^2}. \tag{3.70}$$

We can study how this coefficient of variation varies with the variability of the initial volume of the droplets, and the result is that the initial variability of the volume increases the coefficient of variation when compared with a simple Poisson case ($\sigma_\lambda = 0$), but the relative increase is very low (see Fig. 3.21). Even with quite high variations of the initial volume ($\sigma_\lambda = 0.2\lambda_0$), we still get a relative increase of only 1%. The heterogeneity of the volume of the droplet thus have a very low impact on the coefficient of variation.

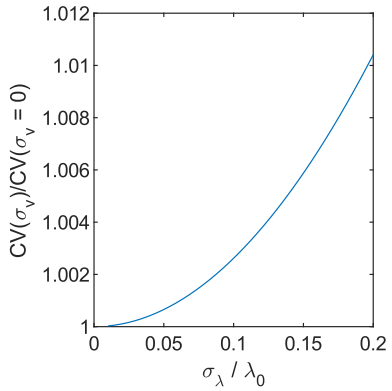


FIGURE 3.21: Theoretical effect of the heterogeneity of the volume of the droplets on the coefficient of variation of the number of cells in exponential phase, relatively to the case where the Poisson parameter is fixed.

3.4 Summary and conclusion

In summary, we showed first that the three models of bacterial growth (adder, sizer, timer) were indistinguishable at the scale of the small populations that grow in our droplets (from a few individuals to a few thousands). This is not surprising since single-cell measurements were required to discriminate between the three models [45]. Realizing this however allows us to develop a theoretical description based on the Bellman-Harris model, which has the advantage of providing several analytical predictions. The mean growth and the variance in population size among droplets in which bacteria divide from a single individual with a defined cell-to-cell variability can be predicted. These predictions are extended here to account for additional sources of variability in the droplet experiments, including a variable number of cells per droplet initially and a division time that varies with the generation number. The analytics are validated with Monte-Carlo simulations that show excellent agreement for the different parameters. This is, to our knowledge, a quite new modeling approach, that takes into account both experimental constrains and advanced stochastic results to quantify the distribution of the number of bacteria exponentially growing in microfluidic droplets. Indeed, if the analytic predictions of the Bellman-Harris model have already been extensively studied numerically [78, 79, 80, 81], the comparison with experimental data, in particular for the variance of the distribution of population size, was lacking in the literature.

The four predictions of the expanded Bellman-Harris model where then verified in our microfluidics experiments: (i) The exponential growth in each droplet with a constant exponential growth rate; (ii) the exponential growth of the mean, standard deviation, and higher moments of the cell number distribution, all with the same exponent; (iii) a constant value of the coefficient of variation CV_N of the population size that shows a strong dependence on the model details; (iv) a self-similar shape of the distribution of number of bacteria in all droplets. Although the value of the growth rate (3.14) is unmodified compared with the classic Bellman-Harris model, recovering the shape of the self-similar distribution and therefore the coefficient of variation requires the full model presented in eqs. (3.45) and (3.46).

Our analysis shows that the effect of cell-to-cell variability on the distribution of population sizes is dominated by other sources of randomness in the experiments. This is quite intuitive, as these other sources of stochasticity come into play at the beginning of the growth: the differences get amplified by the exponential growth afterwards. This can be improved in the future by reducing the uncertainty due to the variable division times (e.g. by synchronising the cells initially) or from the Poisson distribution (e.g. by working at smaller values of λ).

Because the effect of the cell-to-cell variability on the growth of populations is dominated by these external sources or randomness, it seems difficult to infer the cell-to-cell variability from the population measurements. Both a direct inference method using our results and a pseudo-likelihood method

found in the literature [92] proved to be difficult to apply to our experimental system. In the following Chapter, we will try to develop a new inference method to tackle this inverse problem.

More interestingly, however, the current results provide a baseline for the expected behavior when the bacterial phenotype is normally distributed. This baseline can then be compared with measurements of growth under stress conditions or by looking at interactions between different strains in order to detect departures from the normal behavior. This will enable measurements of the heterogeneity of cellular response to external factors, without the need to observe the cells individually.

Chapter 4

Following individual trajectories: the residuals

Des figures s'enchaînent mais des souvenirs m'enchaînent, gros
J'envoie des manchettes, wesh, l'armée des enfers rôde
VALD, Résidus

4.1 Introduction

In this chapter, we will develop a new method to infer the cell-to-cell variability of division times from the measurement of the growth of small bacterial populations in droplets. As we have already discussed in the introduction of this thesis, this is interesting for several reasons. The growth variability of bacteria is related to many other phenomena, such as the stochasticity in gene expression [75] for instance, and this in turns creates phenotypic variability among genetically identical cells. This variability plays a key role in many clinically important cases, such as biofilm formation [16], expression of virulence factors during infections by pathogens [34]...

We have seen in the previous chapter that is was almost impossible to infer this cell-to-cell variability of division times in exponential growth from the study of the distribution of the number of cells per droplet at given observation times, because the shape of this distribution is mostly due to external sources of variability, occurring at the early stages of the growth in microfluidic droplets, and not to the cell-to-cell biological variability.

A few methods exist in the literature to address this inference problem. However, most of them are not suited for our experimental data. In particular, the methods based on a continuous observation [89, 90] of the lineage tree can not be used here, as we only observe the droplets at discrete observation points. Moreover the whole tree is not observable in our case, we just have access to the number of bacteria in each droplet. A method based on pseudo-likelihood estimation [92], adapted to this observation pattern, was

tested in the previous chapter, but could not be applied in our case, once again because of the impact of the early stages of the growth.

We have then developed a new method, based on following the trajectories of individual droplets in time, and analyzing the deviation from pure exponential growth. Droplets are followed throughout the exponential phase of growth, and this allows us to get rid of the deleterious influence of the randomness at early times. In this chapter, we expose how this theoretical and numerical method works, sketch its mathematical proof, and demonstrate its efficiency on numerical simulations. We will then study its applicability in our experimental case.

4.2 Evolution of the residuals with time

4.2.1 Idea and definition

The idea developed here is to follow the individual trajectories of the droplets with time. By trajectory, we mean the number of cells as a function of time $N(t)$. Indeed, these trajectories are not purely exponential, as would be the case in a simple deterministic growth model. The deviation from the pure exponential growth is due to the stochasticity of the division process and thus is linked to the variability of the individual division times. To quantify this deviation from the pure exponential growth, we are going to measure certain quantities that we will call the residuals of the trajectories. Our hope is to be able to infer the individual variability of the division times from this measurement of the residuals. As this measure relies only on the stochastic evolution of the trajectory, we can intuitively think that it will not be impacted by the events happening at the initial stages of the division process that caused the inference to fail in the previous chapter, such as the initial Poisson distribution or the different division times for the first generations.

The idea that we will implement to measure the residuals is the following: for a given trajectory $N(t)$, we can estimate the exponential growth rate α by fitting the log of the curve with a straight line. Then, to measure the residual at some time t, we will measure the difference between the value of the trajectory at time $t + \Delta t$ and the expected value if the growth was purely exponential, which is $N(t)e^{\alpha \Delta t}$, as we can see on Fig. 4.1. The value of the residual at time t is then:

$$\boxed{Res(t) = N(t + \Delta t) - N(t)\exp(\alpha \Delta t)\,.} \tag{4.1}$$

In the experiments, we have a set of observation points for each trajectory: $\{t_i\}_{i \in 1..n}$, that are equally spaced in time, with $\Delta t = t_{i+1} - t_i$, and we will define the residual at each observation point t_i by:

$$\boxed{Res_i = Res(t_i) = N(t_{i+1}) - N(t_i)\exp(\alpha \Delta t)\,.} \tag{4.2}$$

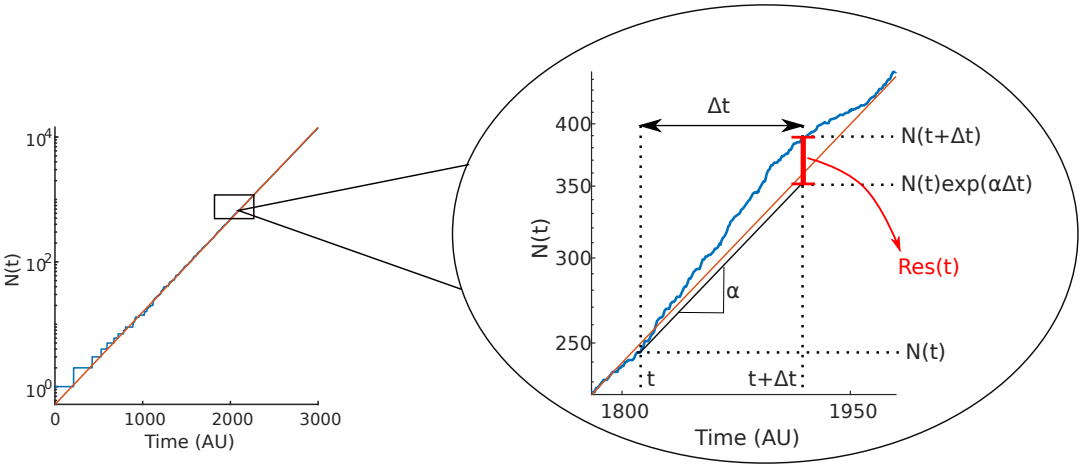


FIGURE 4.1: [Left] Simulated example of a trajectory for one droplet: number of cells $N(t)$ as a function of time(blue), and exponential fit (orange), in log scale. [Right] Zoom and definition of the residual of the trajectory at time t, for a given time step Δt. We first fit the growth rate α on the whole trajectory: it is just the slope of the trajectory in log scale. Then, if the growth was purely exponential, with rate α, from $N(t)$, the value of the trajectory in $t + \Delta t$ would be $N(t) \exp(\alpha \Delta t)$. The residual is the difference between this pure exponential growth and the actual value of the trajectory in $t + \Delta t$, which is $N(t + \Delta t)$.

4.2.2 Mathematical Properties

The residuals of the trajectories have two main properties, that we are going to prove in this section for the simpler case of a Yule process (see definition below). The proof of the more complicated Bellman-Harris case will be sketched, and the properties only checked on simulations. The two properties are the following:

1. The residuals are normally distributed, and centered around 0.
2. The variance increases exponentially with time, with rate α, and the prefactor depends on the variability of the individual division times.

These two properties can be mathematically summarized as follows:

$$Res_i \sim \mathcal{N}\left(0, f\left(cv_\mu\right) \exp\left(\alpha t_i\right)\right). \tag{4.3}$$

Note that the form of the function f is unknown for the moment, and will be discussed in the next section.

Exponential division times: Markov case

Let us first prove those properties in the simpler case of a Yule process [143]. A Yule process is a particular case of the more general Bellman-Harris process that we have described in the previous chapter. In this case, the division times of the bacteria are not Gaussian. Instead, the division times are picked

from an exponential distribution of density $g(t) = a\exp(-at)$, where a is a positive number. This case is simpler because the Yule process verifies the Markov property, which states that the process is memory-less: the probability for a bacteria to divide at time t does not depend on its age. This hypothesis is not very biologically relevant, as we have seen in the previous chapter, however it simplifies a lot the mathematical proof.

To study the residuals $Res(t) = N(t+\Delta t) - N(t)\exp(\alpha\Delta t)$, we are going to re-write the number of bacteria at time $t+\Delta t$ as the sum of the number of descendants of the $N(t)$ bacteria that were present at time t:

$$N(t+\Delta t) = \sum_{i=1}^{N(t)} X_{i,t,\Delta t}, \tag{4.4}$$

where $X_{i,t,\Delta t}$ is the number of descendants of cell i at time $t+\Delta t$, with $i \in [1..N(t)]$. Thanks to the independence of all cells, the random variables $X_{i,t,\Delta t}$ are all independent (branching property). And because we are in the Markov case, they are all identically distributed, and independent of t: $X_{i,t,\Delta t} = X_{i,\Delta t}$. Thus we have:

$$Res(t) = \sum_{i=1}^{N(t)} \left(X_{i,\Delta t} - e^{\alpha\Delta t} \right), \tag{4.5}$$

which can be re-written as:

$$\frac{Res(t)}{\sqrt{N(t)}} = \frac{1}{\sqrt{N(t)}} \sum_{i=1}^{N(t)} \left(X_{i,\Delta t} - e^{\alpha\Delta t} \right). \tag{4.6}$$

Where $X_{i,\Delta t} - e^{\alpha\Delta t}$ are independent, and identically distributed. We can then apply the Central Limit Theorem. This well-known theorem of stochastic calculus states that, if (Y_i) is a sequence of independent and identically distributed random variables, then with $\mathbb{E}(Y_i) = \mu$ and $\mathrm{Var}(Y_i) = \sigma^2$, then the random variable defined as:

$$S_n = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} (Y_i - \mu), \tag{4.7}$$

converges in distribution as n goes to infinity to a Gaussian variable of variance σ^2. The definition of the convergence in distribution can be found in Appendix B.3. We will note this as:

$$S_n \xrightarrow[n\to\infty]{d} \mathcal{N}\left(0, \sigma^2\right). \tag{4.8}$$

We can then come back to our case, and apply the Central Limit Theorem for $Y_i = X_{i,\Delta t} - e^{\alpha\Delta t}$. By definition of the growth rate α, we have $\mathbb{E}(Y_i) = 0$. We get:

$$\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \left(X_{i,\Delta t} - e^{\alpha\Delta t} \right) \xrightarrow[n\to\infty]{d} \mathcal{N}\left(0, \mathrm{Var}\left(X_{\Delta t} - e^{\alpha\Delta t} \right)\right). \tag{4.9}$$

We are almost done with the demonstration. The only remaining issue is that for the computation of the residual at time t (eq. (4.6)), the number of terms in the sum is itself a random variable $N(t)$, and not any natural integer n as in the statement of the Central Limit Theorem. However, as $N(t)$ goes to infinity when t goes to infinity, and $N(t)$ is independent of $X_{i,\Delta t}$, we can show that the result of the Central Limit Theorem is conserved, see lemma B.4 in Appendix. We can then conclude that:

$$\frac{Res(t)}{\sqrt{N(t)}} \xrightarrow[t\to\infty]{d} \mathcal{N}\left(0, \mathrm{Var}\left(X_{\Delta t} - e^{\alpha \Delta t}\right)\right). \tag{4.10}$$

We can see here that the residuals are Gaussian, not exactly for a fixed time, but if we fix the number of cells N at which they are observed. We will see on the simulations, for the non-Markov case, that even for a fixed time t, they are close to being normally distributed.

In this Markovian case, the distribution of $N(t)$ is completely known [143]: it is a geometrical law of parameter $\exp(-\alpha t)$. We can then have an explicit expression of the right-side variance of the equation below:

$$\mathrm{Var}\left(X_{\Delta t} - e^{\alpha \Delta t}\right) = \frac{1 - e^{-\alpha \Delta t}}{e^{-2\alpha \Delta t}}. \tag{4.11}$$

And we see that it is linked to the variance of the individual division times, which is in this case $1/\alpha^2$.

Gaussian division times

For the non Markov case, the proof is much more difficult, because the age of the cells at time t influences the number of daughter cells that they produce at time $t + \Delta t$. We are just going to sketch here how the proof could be done, and check on the simulations that these properties are conserved in the full Bellmann-Harris model.

To prove these properties in the non-Markov case, we would have to work with the age structure of the population.

Let $(A_i)_{i\in[1,N(t)]}$ be the ages of the $N(t)$ cells at time t. It can be shown that the age distribution is stationary at long times [144]. The ages are also almost independent, as two sister cells of course have correlated ages, but the probability to pick up two sister cells (or even closely related cells) decreases very fast because of the exponential increase of the population size. Thus, we also almost recover the independence and identical distribution, and:

$$Res(t) = \sqrt{N(t)} \frac{1}{\sqrt{N(t)}} \sum_{i=1}^{N(t)} \left[X_{i,t,\Delta t} - h_i e^{\alpha \Delta t}\right]. \tag{4.12}$$

Here h_i depends on the age of the parent cell i. By conditioning on the ages, we still obtain a sum of independent and centered variables on the right side of the equation, and we can still apply a form of the Central Limit Theorem. The key is that the variance of the right-side terms is a function of the ages,

which are distributed following a stationary distribution. The problem is much more complicated as we have to condition on the ages, which are continuous. We would then have to work in infinite dimension to recover a kind of Markov property applying to the ages, but this is beyond our mathematical capabilities. In the end, we still obtain the same shape and time-evolution of the residuals.

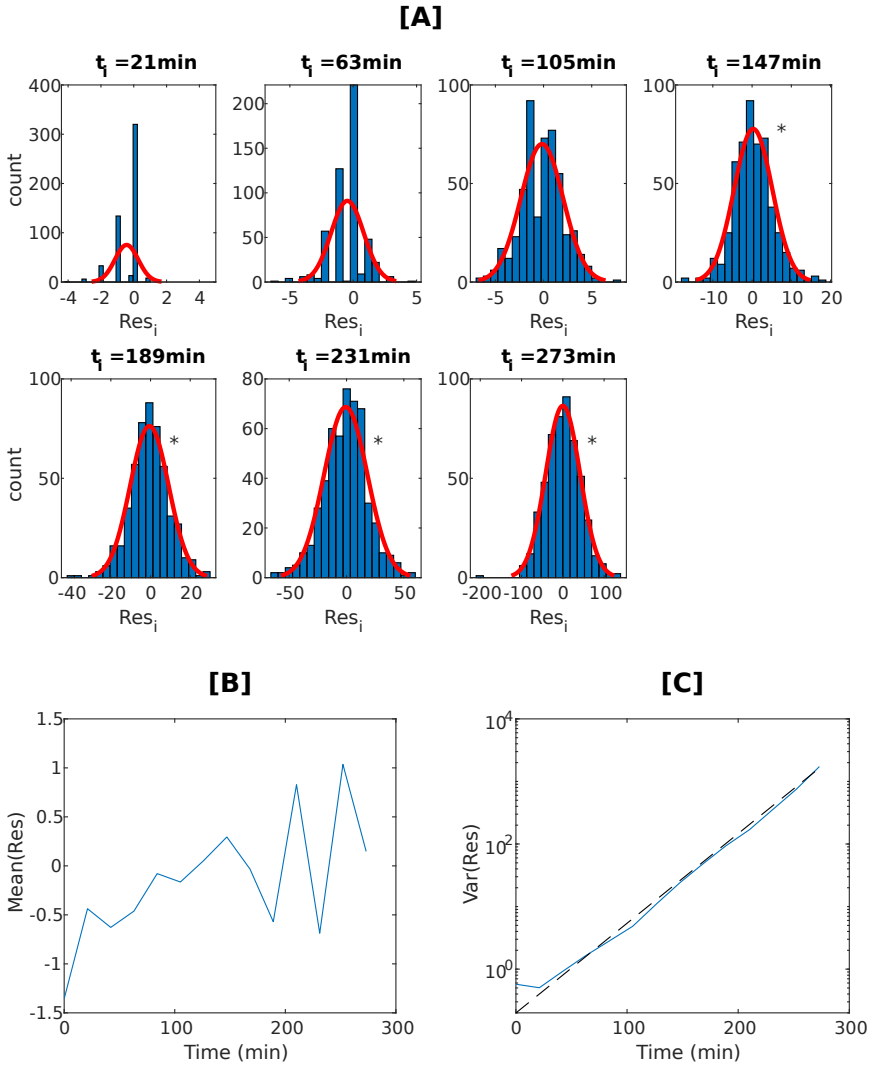


FIGURE 4.2: Simulation: [A] Distribution of the residuals for different times t_i, with Gaussian fit in red. The stars indicate a p-value $p < 0.05$ in the Kolmogorov-Smirnov test for normality. [B] Mean of the residuals as a function of time. [C] Variance of the residuals as a function of time, in log scale. The black dashed line has a slope α. The simulation is the one already described in the previous chapter, for the full *E. coli* experimental parameters and $\lambda = 0.75$ (1000 traps).

We can check that these properties are verified on the Bellman-Harris simulations, directly on a full simulation of the growth of *E. coli*, with an initial Poisson distribution ($\lambda = 0.75$) and the division times changing with the generation, as measured in the previous chapter. The results are presented

on Fig. 4.2. As we can see, the residuals are rapidly (after $\approx$ 100 min) very close to being normally distributed when t is fixed (this is confirmed, for $t \geq 147$ min, by a Kolmogorov-smirnov test [133]), and their mean remains close to 0. Their variance, on the other hand, evolves exponentially with a rate equal to the growth rate of the population α.

4.2.3 Choosing Δt

Synchronicity

Now that we know that the residuals are normally distributed, and that their variance increases exponentially with rate α, we can begin to study the prefactor function f in front of the exponential, which encompasses the dependence of the residuals on the individual division times of the bacteria. We will do so by computing the residuals of simulated growth experiments, that were already described in the previous chapter (Bellman-Harris model with truncated Gaussian laws for the division times of the bacteria). We have to keep in mind that what we want is to be able to infer the variability of the individual division times from the measurement of the residuals. Thus we would like this function f to be very dependent on the variability of the division times of the bacteria cv_μ. Moreover, it would be make sense for f to be an increasing function of cv_μ, as the more individual variability there is, the more widely spread the residuals should be.

However, these two requirements are not that easy to fulfill. Indeed, the choice of the parameter Δt is critical for f to have the desired properties. To understand this, let us take a look at the example presented on Fig. 4.3.

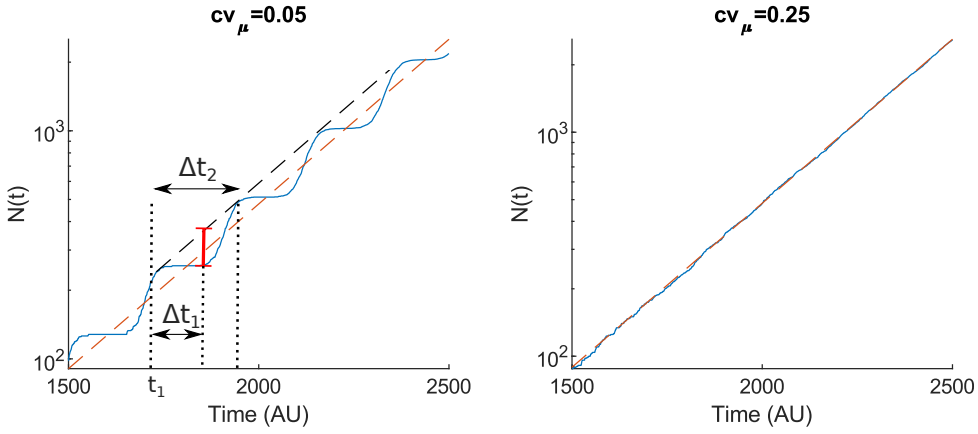


FIGURE 4.3: Simulated trajectories $N(t)$ (plain blue), for two values of cv_μ, with exponential fit (dashed orange): [Left] $cv_\mu = 0.05$, [Right] $cv_\mu = 0.25$. We see that if $\Delta t = \Delta t_1$, the value of the residuals can be quite big for the smaller value of cv_μ, whereas if $\Delta t = \Delta t_2$, the residuals will be much smaller.

Two simulated trajectories are presented here, with cv_μ that is the only difference in between them. Comparing these two graphs shows that the trajectory with the smallest cv_μ deviates in fact more from the exponential

fit then the one with the bigger individual variability. This is due to a synchronicity effect: as the variability of division times is very low, all the cells divide more or less at the same time. A bad choice of Δt can then lead to a very big value for the residual, as shown in red on the graph. To avoid this, we can see that the only possibility is to take a time step Δt that is equal to a multiple of the period of synchronicity of the cells, as for instance Δt_2 on the graph. This period is of course the mean division time of the cells τ_0, which can be approximated by $\tau_0 \approx \log(2)/\alpha$, see Chapter 3.

To check these intuitive guesses, we can compare the variance of the residuals on different Bellman-Harris simulations, with the same τ_0 but varying cv_μ, for two values of the time step Δt. If we take a time step that is equal to τ_0, the variance of the residuals are straight lines with a slope close to $\alpha \approx \log(2)/\tau_0$, and the bigger cv_μ, the bigger the variance of the residuals (Fig. 4.4[B]). At the opposite, if we take a different time step, for instance here $\Delta t = \tau_0/2$, we get completely different results (Fig. 4.4[A]). For small cv_μ, because of the synchronicity problems, the slope of the variance of the residuals is not equal to α any more. For larger cv_μ, the variance of the residuals is not very sensitive to cv_μ, which is not good for the inference problem. Moreover, the variance of the residuals is bigger for small cv_μ, which is not intuitive at first sight, and is also not good for the backwards inference.

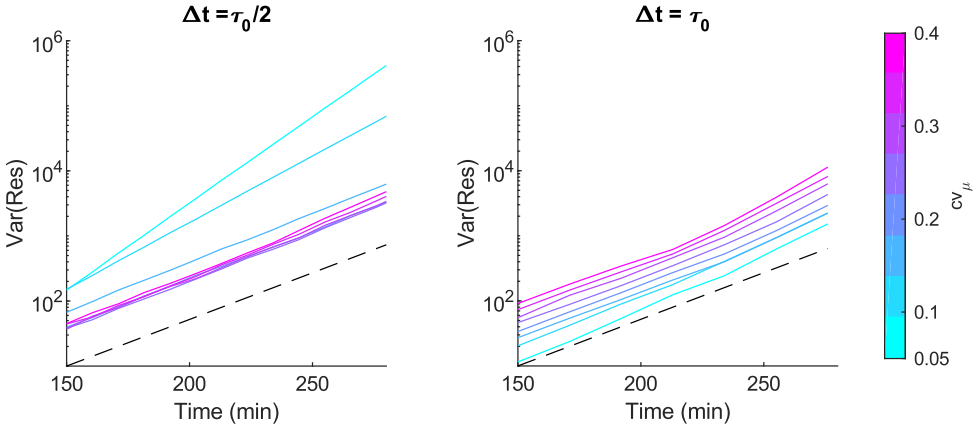


FIGURE 4.4: Variance of the residuals (log scale) for two values of the time step Δt, $\Delta t = \tau_0/2$ [Left] and $\Delta t = \tau_0$ [Right], and varying $cv_\mu \in [0.05, 0.4]$. The parameters for the simulation are $\tau_0 = 21$ min, 2000 droplets in each simulation, $N_0 = 1$. The black dashed line indicates the slope $\alpha = \log(2)/\tau_0$.

As the variance of the residuals evolves with time with an exponential rate α, the ratio of this variance to the mean number of cells $\langle N \rangle$ is a constant, and this ratio is directly linked to the variability of the division times of the cells. If we choose a good Δt, as explained above, this ratio evolves almost linearly with cv_μ, see Fig. 4.5[A], which makes it a very good candidate as the quantity to use for the inference of cv_μ.

Another question that arises from this choice is how close to τ_0 do we have to choose Δt if we want the inference to work ? The answer can be read on Fig. 4.5[A], where Δt was varied around the value τ_0. For small

values of $cv_\mu < 0.15$, the inference does not work any more as soon as Δt is different from τ_0. This is due once again to the synchronicity of the cells for those small variability values. However, those values are quite unlikely to happen in the experiments, see Chapter 3, Fig. 3.13. For the values that we expect ($cv_\mu > 0.15$), the inference is still possible even if we approximate τ_0 by $\pm 5\%$, which is a precision that we can achieve in the experiments. This method of inference seems therefore to be promising.

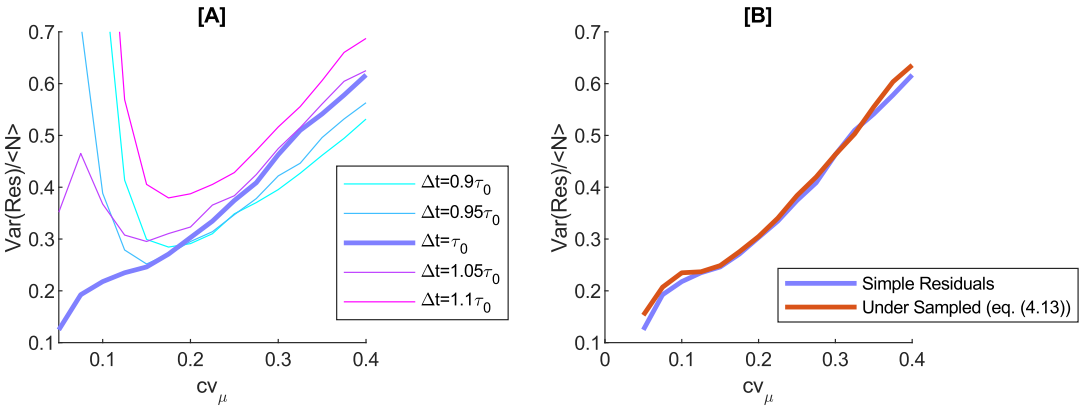


FIGURE 4.5: Mean of the ratio of the Variance of the residuals to the mean number of cells. [A] Varying the time step for the residuals Δt. [B] Comparison between the simple evaluation of the residuals (eq. (4.2), with $\Delta t = \tau_0$) and the under-sampled evaluation (eq. (4.13), with $\Delta t = 0.25\tau_0, k = 3$)

Changing the Sampling Rate

However, in the experiments the minimum interval between two observations is set by the scanning speed of the microscope. For a typical experiment, we have $\Delta t = 5$ min, and the mean division time of *E. coli* is around 21 min. We could of course increase the time interval between two observations to this value, but then we would have very few observations in the exponential phase for each droplet, as we are limited by the detection threshold at the beginning and the saturation at the end.

The simplest solution to still get a nice inference curve is the following: we keep the time interval between two points at its lowest experimental value in order to get as many observations as we can, but we slightly change the computation of the residuals. Instead of computing the residual between t_i and t_{i+1}, we are going to compute the residual between t_i and t_{i+k}, such that $k\Delta t \approx \tau_0$. For our example with *E. coli* for instance, we will set $k = 4$, and thus the effective time interval in the residuals will be 20 min, which is close to the estimated value of τ_0. The exact formula for the residual is then:

$$Res_i = N(t_{i+k}) - N(t_i)\exp(\alpha k \Delta t)\,. \tag{4.13}$$

We can see on Fig. 4.5 that this does not change the shape of the inference curve, and thus this is the method that we are going to use for the rest of this section.

At this point, we have proven that the residuals are Gaussian and that their variance increases exponentially with the same growth rate as the population. We have also seen on the simulations that if we want this variance to depend nicely on the microscopic variability of division times, we had to choose a time step close to τ_0, or to change the sampling rate of the residuals such that $k\Delta t \approx \tau_0$. We can now go back to the experiments.

4.3 Experimental residuals

4.3.1 A slope problem

Let us now examine the residuals in the experiments. If the fluorescence signal that we measure is proportional to the number of cells in the droplet, it is easy to see that the residuals of this fluorescent signal are also proportional to the residuals in terms of number of cells. They should then display a Gaussian and centered distribution as well, and their variance should evolve exponentially with the same growth rate as the mean of the fluorescence.

To compute these residuals, we have to be sure though that we are in the exponential phase in the droplet. We will apply for this the same method as in Chapter 2. The detection and saturation threshold are set manually and once and for all for all the droplets. Then, for each of them, only the points in between those limits are taken into account for the computation of the residuals. For each observation point t_i, the distribution of the residuals encompasses only the droplets that are in exponential phase. As we can see on Fig. 4.6, this distribution is close to a Gaussian, as expected. It deviates a little bit from the distribution towards the end of the considered time points, probably because the effect of the saturation can already be observed. It can also explain why the center of the distribution changes from zero to negative values at the end: the growth rate is slightly over-estimated, and the residuals become negative.

If we now look at the variance of those experimental residuals, it seems indeed to evolve exponentially with time (see Fig 4.7, top line). However, if we re-normalize this variance by the mean of the fluorescence, we don't obtain a constant value as expected, but a time decrease(Fig 4.7, bottom line): the exponential rate of growth of the variance of the experimental residuals is lower than the one expected with the theory and simulations. This is reproducible for several monitoring conditions and the two bacterial strains that we have studied (see the three different plot in Fig. 4.7). There is then no chance of inferring the variability of division times with those experimental data as such, and we will try to understand why the variance growth rate is low in the next section.

4.3.2 Residuals and noise

The first source of error that comes to mind when we try to account for a deviation of the experimental residuals from the theory is the experimental noise. The detailed computations for different kinds of measurement noise can be

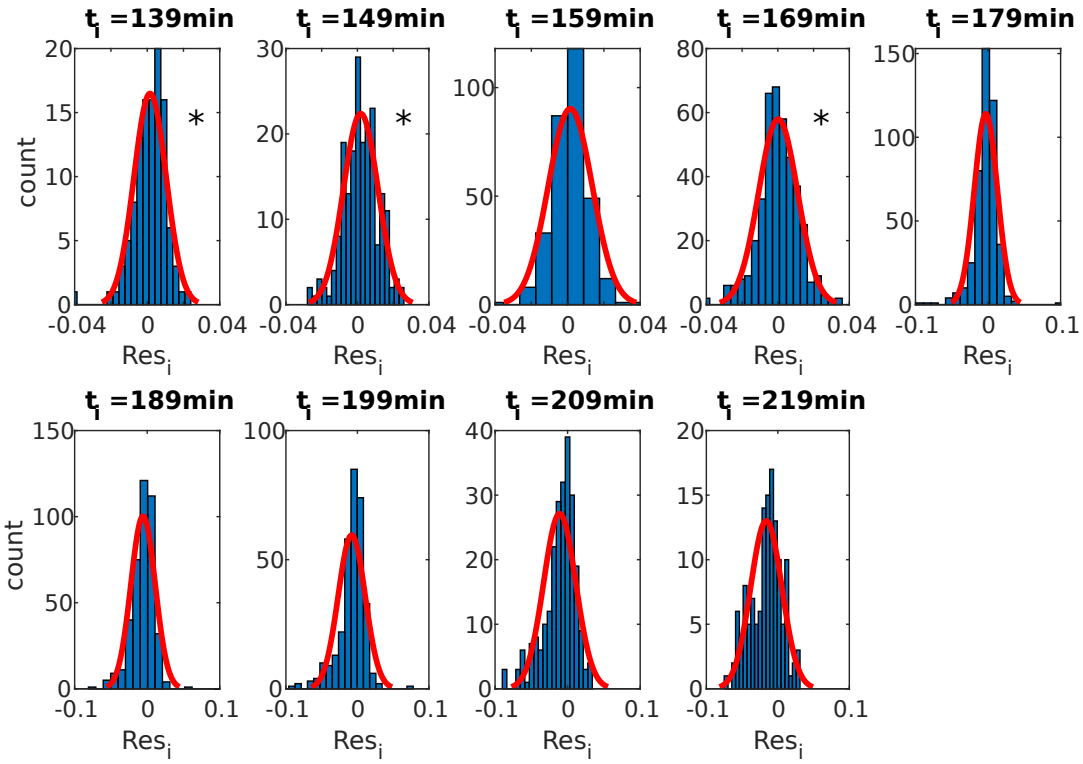


FIGURE 4.6: Histogram of the residuals, in a typical experiment (same as Chapter 2, *E. coli* in droplets of LB medium), at different time points, with a time step such that $k\Delta t \approx \tau_0$ ($k = 4$), with Gaussian fit (red). The stars indicate a p-value $p < 0.05$ in the Kolmogorov-Smirnov test for normality.

found in Appendix D. To summarize, the existence of an experimental noise does affect the residuals. However, it does not account for the slope decrease that we have mentioned. Indeed, if we take into account the heterogeneity of the fluorescence, we find that the variance of the residuals is increased by a constant factor, but its rate of growth is unchanged. If we consider an additive noise, the exponential rate of growth of the variance of the residuals is indeed lowered for a specific time-window that could match the one of the experiments. However, the fact that the decrease in the exponential rate of growth of the variance is reproducible for different bacterial strains and settings of the camera makes us doubt that an additive noise could be responsible. Moreover, the measurement noise was found to be multiplicative in Chapter 2. Finally, a multiplicative noise increases the exponential rate of growth of the variance instead of decreasing it.

We then have to try to find another reason to account for the deviation of the variance of the experimental residuals from the theory mentioned in the previous paragraph.

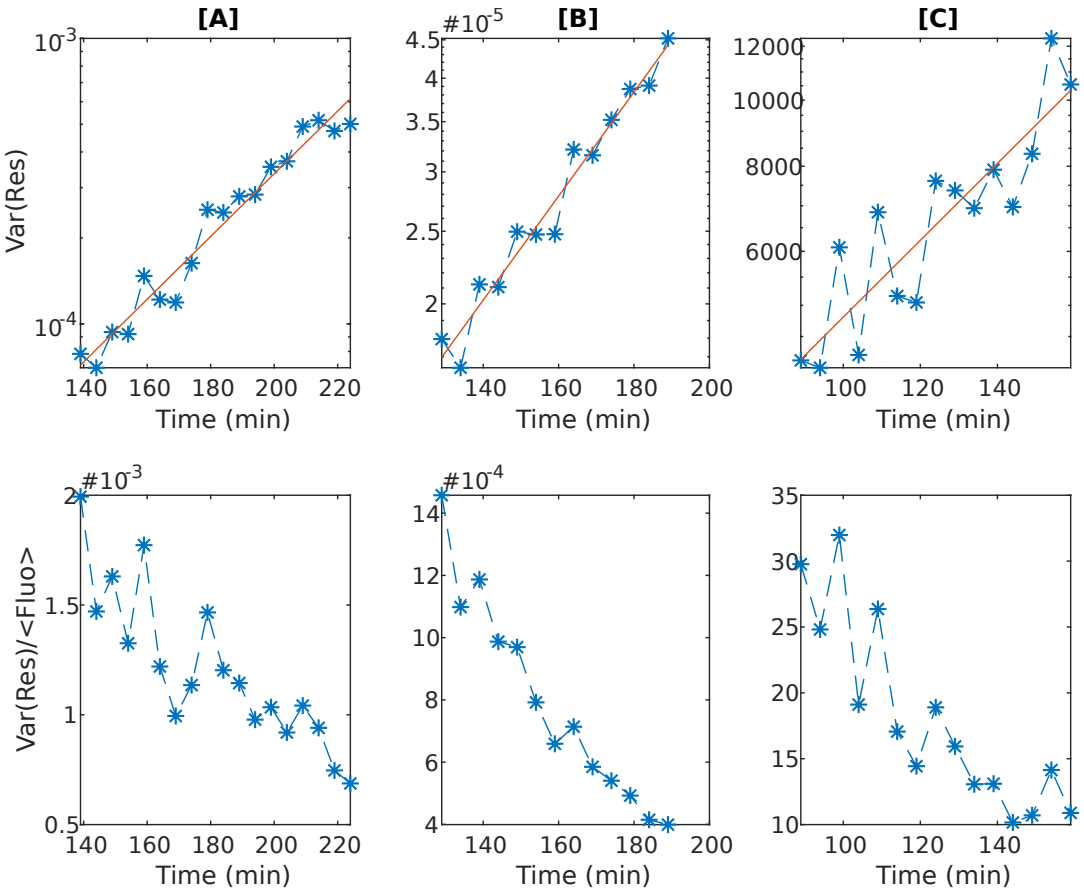


FIGURE 4.7: Variance of the residuals for three independent growth experiments on chip. Experimental conditions: [A] *E. coli* in LB (same experiment as Chapter 2), [B] same strain, same medium, but different settings of the camera [C] *B. subtilis* in LB (same experiment as in 3). Top line: Variance of the experimental residuals (log scale,blue) as function of time, with linear fit in orange. Bottom line: same, but divided by the mean of the fluorescence signal.

4.3.3 Experimental sampling

If the noise or the heterogeneity of the fluorescence are not enough to explain the behavior of the experimental residuals, there is something else that differs from the simulations in the experiments: the way we sample the data. Indeed, in the experiments, at each observation time t_i, we consider only the droplets that are in exponential phase. The droplets are considered to be in exponential phase if and only if $Fluo(t_i)$ is between the limits of detection and saturation, which are set once and for all for every droplet. We can do the same thing in the simulations, setting for instance $N = 100$ and $N = 2500$ as the lower and upper limits for the computation of the residuals (Fig. 4.8[A]). The residuals seem to behave very closely to what happens in the experiments: they are still close to being Gaussian (Fig. 4.8[B]), and their variance increases exponentially with time (Fig. 4.8[C]), but its rate of growth is lower than the growth rate of the bacteria, which causes the ratio of the variance of

the residuals to the mean number of cells to be time-decreasing (Fig. 4.8[D]).

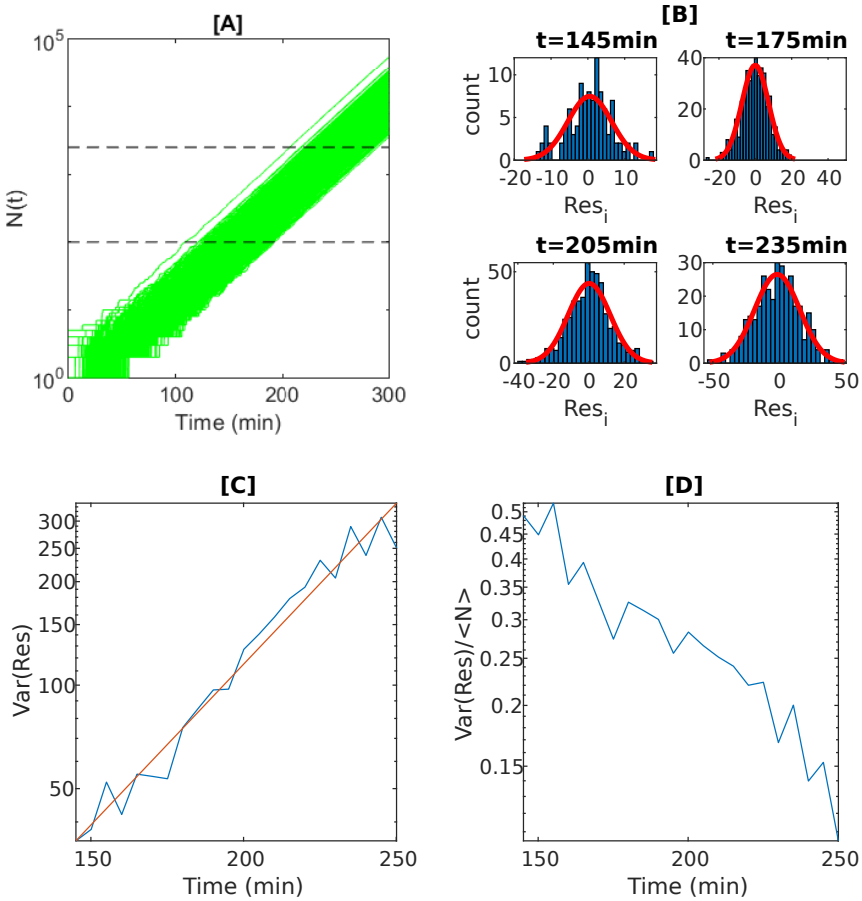


FIGURE 4.8: Residuals in a simulation of *E. coli* (same as Fig. 4.2) when we set fake limits for the detection and saturation, to mimic the experimental situation. [A] Growth curves as a function of time, in log scale, with the lower and upper limit in black. [B] Histograms of the residuals, with Gaussian fits, for several time points. [C] Variance of the residuals (log scale,blue) as function of time, with linear fit in orange. [D] Variance of the residuals divided by the mean number of cells as a function of time, in log scale. This ratio decreases with time instead of being constant.

Thus just by sampling the simulated data like we do in the experiments, we can see that this sampling causes the variance of the residuals to deviate from the theory. However, we cannot completely remove this sampling in the experimental part, because we have to study only the droplets in the exponential phase. We are going to see in the next section a method to keep the sampling but recover the correct exponential growth of the variance.

4.4 Binning the residuals by the number of cells

4.4.1 Definition, numerical results and inference method

Idea and definition

The sampling of the exponential data causes the variance of the residuals to be underestimated if we look at the time-distribution of the residuals. This is indeed intuitive, because the sampling truncates the distribution for a fixed t. Let us take the example of Fig. 4.8[A], and consider only the growth curves in between the two black lines. If we look at the distribution of the residuals for $t \approx 150$ min, only some of the trajectories will be taken into account, and this can reduce the rate of growth of the variance when compared with the theory.

The idea in this part is then to stop working on time-distribution of the residuals, but rather to bin the residuals by the value of N. Then, as the selection of the exponential phase is based on the value of N as well, the distribution of the residuals for a fixed N will be unaffected by this selection. To go back to the example, if you consider horizontal slices of the curves on Fig. 4.8[A] between the two black lines, all the curves are taken into account for all the slices.

The principle is then the following: if we set N_{det} and N_{sat} as the detection and saturation limits, respectively, we are going to choose a finite sequence (N_j) such that:

$$\cup_j \left[N_j, N_{j+1}\right[= [N_{det}, N_{sat}[. \tag{4.14}$$

Then, for every observation time t_i, and every trajectory $N^k(t)$, there exists a unique j such that:

$$N^k(t_i) \in \left[N_j, N_{j+1}\right[. \tag{4.15}$$

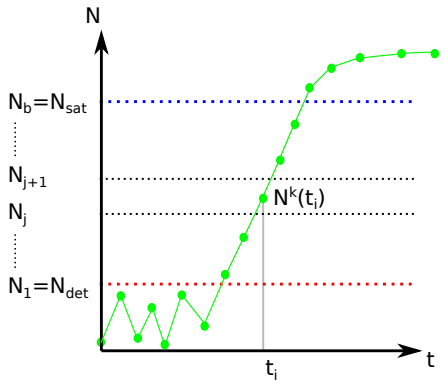


FIGURE 4.9: Schematic definition of the principle for binning the residuals. The limits for the detection N_{det} and the saturation N_{sat} are set once and for all for all the droplets. In between these two limits, we set bins $\left[N_j, N_{j+1}\right[$ such that every observed point of every trajectory can be indexed in one of those bins.

We can define, for every trajectory $N^k(t)$, the instantaneous growth rate:

$$\alpha_j^k = \frac{1}{\Delta t}\log\left(\frac{N^k\left(t_{i+1}\right)}{N^k\left(t_i\right)}\right). \tag{4.16}$$

We are mainly going to study the mean of those instantaneous growth rates for all the bins, that we will note α_j:

$$\alpha_j = \langle \alpha_j^k \rangle_k. \tag{4.17}$$

According to Bellman and Harris [76], we should have $\alpha_j^k \sim \alpha$ in exponential phase.

The other quantity that we are going to study is the residuals, as defined by eq. (4.2), and we will index them by j as well:

$$Res_j^k = N^k(t_{i+1}) - N^k(t_i)\exp\left(\alpha\Delta t\right). \tag{4.18}$$

All the properties that we have shown for the residuals in time still hold for the residuals binned by the number of cells. In fact, the mathematical proofs are even more straightforward, as eq. (4.10) gives us directly that for a fixed N, the residuals are Gaussian and their variance increases as N, with a prefactor depending on the variability of division times.

As a consequence, if we look at the mean:

$$Res_j = \langle Res_j^k \rangle_k, \tag{4.19}$$

we should have $Res_j \approx 0$ in exponential phase.

We will note $\mathrm{Var}\left(Res\right)_j$ the variance of the residuals in the $j-th$ bin:

$$\mathrm{Var}\left(Res\right)_j = \mathrm{Var}\left(Res_j^k\right). \tag{4.20}$$

We have, for j such that we are in exponential growth,

$$\mathrm{Var}\left(Res\right)_j \sim N_j f(cv_\mu). \tag{4.21}$$

where f is an unknown function of the microscopic coefficient of variation cv_μ. In particular, we see that if we rescale the variance of each bin by computing the ratio $\mathrm{Var}\left(Res\right)_j / N_j$, we get a quantity that is constant over the bins and depends only on the microscopic coefficient of variation. This is what we will use for inference.

Numerical results

To match the exponential growth of the number of bacteria, we will choose logarithmically spaced bins. We can go back to the numerical example of the last section to check the properties that we have just listed.

The results are plotted on Fig. 4.10. First of all, the instantaneous growth rate Fig. 4.10[A] does converge to a constant value, but it is slightly lower than the one expected by the theory of Bellman and Harris. This is due to the numerical error in the simulation, and the error is below 2%. As we can see on Fig. 4.10[B], the binned residuals are Gaussian, and their mean is close to zero (Fig. 4.10[C]). If we look at the variance as a function of the N-bin value (Fig. 4.10[D]), we do obtain a straight line with the desired slope, which is 1.

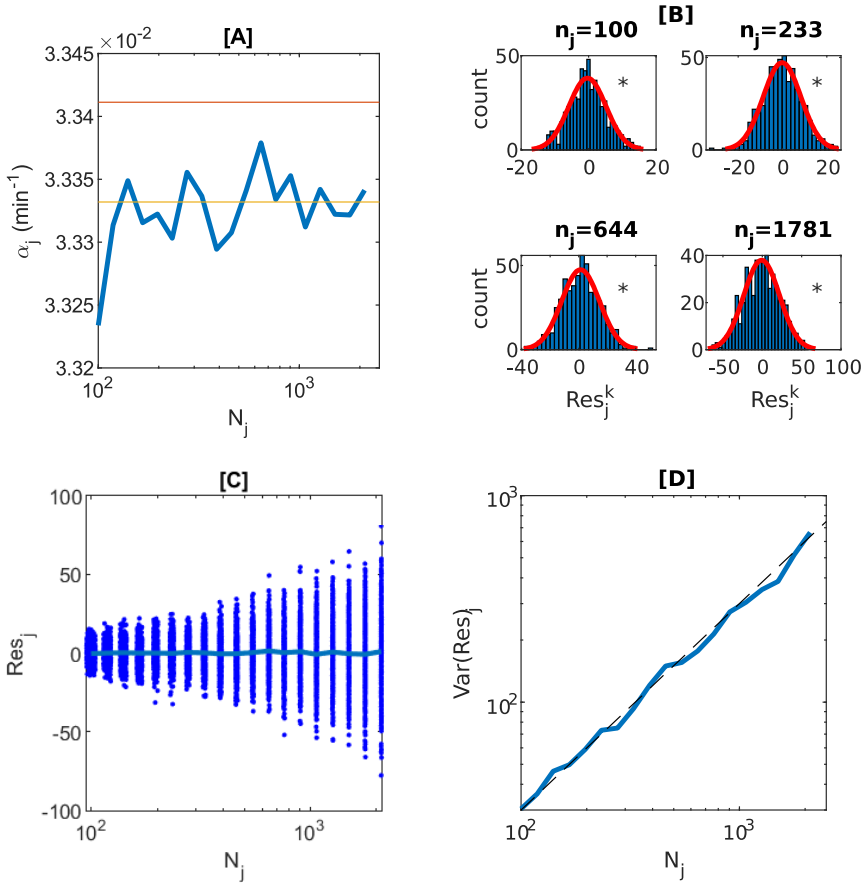


FIGURE 4.10: Residuals in a simulation of *E. coli* (same as Fig. 4.2), binned by N. We have here 20 bins, logarithmically spaced between $N_{det} = 100$ and $N_{sat} = 2500$. [A] Mean Instantaneous Growth Rate, with the Belmman-Harris theoretical value (orange, see chapter 3) and the growth rate of the simulated mean. [B] Histograms of the residuals for 4 bins, with Gaussian fit (red). The stars indicate a p-value $p < 0.05$ in the Kolmogorov-Smirnov test for normality. [C] Residuals in each bin and their mean as a function of N_j (log scale for N_j). [D] Variance of the residuals (log scale,blue) as function of N_j, with a black line indicating the slope 1 (log-log scale).

Hence, we have recovered the theoretical rate of growth of the variance, even with a sampling of the trajectories between a detection and saturation limit. We can check now that this variance is still suitable for inference of the variability of the individual division times: we want the variance to depend strongly on the variability on the division times, as on Fig. 4.3. The results

are plotted on Fig. 4.11. On the left-side graph, we can see that the variance mostly behaves as predicted, except for very low cv_μ. This is due once again to the synchronicity of the cells in those cases, and we are not going to focus on them as they are not likely to happen in real experiments with bacteria (see Fig. 3.13). If we focus on $cv_\mu \geq 0.15$, then the variance of the residuals of each bin behaves as predicted ($\mathrm{Var}\,(Res)_j \propto N_j$). We can rescale the variance by dividing it it by the bin value N_j, and we obtain a ratio that does not depend on j any more. The value of this rescaled variance strongly depends on the microscopic coefficient of variation cv_μ, which gives us a nice curve to infer cv_μ from the computed variance of the residuals (Fig. 4.11[B]). This inference curve, contrary to what we obtained in the previous Chapter, is almost insensitive to the initial conditions (Poisson distribution, different first generation). It is then a very good candidate for the inference of cv_μ in our experimental conditions.

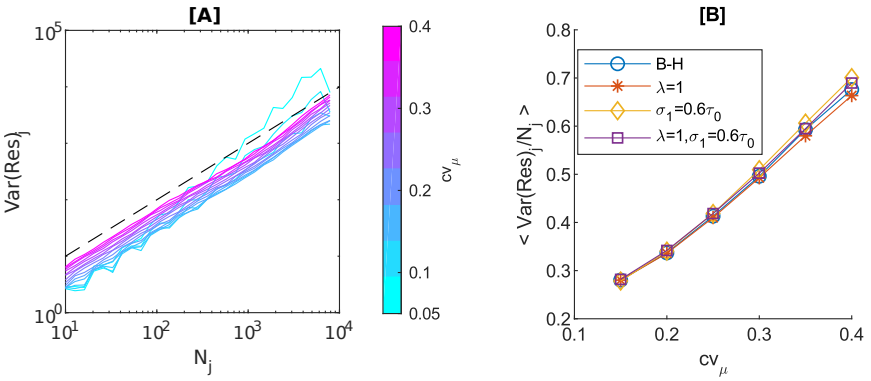


FIGURE 4.11: Variance of the residuals binned by N, when varying cv_μ: same simulations as on Fig. 4.3, with 30 bins logarithmically spaced between $N_{det} = 10$ and $N_{sat} = 10000$. [A] Variance of each bin as a function of the bin value (log-log scale). The black dotted line indicates the line $y = x$. [B] Mean value of $\mathrm{Var}\,(Res)_j / N_j$ as a function of the microscopic coefficient of variation cv_μ, for $cv_\mu \in [0.15, 0.4]$, for the Bellman-Harris classical case (blue circles), and adding a Poisson distribution (orange stars), a different first generation-time (yellow diamonds) or both (purple squares)

Summary of the inference method

If we sum up what we have obtained until now, we can describe an inference method for the microscopic variability of the division times based on the residuals, for given observation points t_i spaced by Δt:

1. Define for all curves a detection and saturation limit N_{det} and N_{sat}.
2. Fit all curves in the exponential phase to get the growth rate α, and thus an approximate value of $\tau_0 \approx \frac{\log(2)}{\alpha}$.

3. Use this value of τ_0 to determine how to evaluate the residuals, choosing k such that $k\Delta t \approx \tau_0$.
4. Compute the residuals at each time point.
5. Choose a finite sequence (N_j) that covers $[N_{det}, N_{sat}[$ and bin the residuals of all the trajectories.
6. Compute the variance of the binned residuals, divide it by the value of the bin to rescale it and take the mean of this ratio across the bins.
7. Read on the inference curve Fig. 4.11[B] the value of cv_μ .

For this last point, we can use a linear approximation of the relation between the ratio and the microscopic cv_μ. If we fit the curves of Fig. 4.11[B] with a linear curve, we obtain the equation:

$$y = 1.6x + 0.02, \tag{4.22}$$

with an excellent $r^2 = 0.99$. We can use then this relation for the inference.

An example

We can very easily apply this, for instance, to the full simulation of the *E. coli* growth used on the previous figure 4.2.

We set $N_{det} = 100$ and $N_{sat} = 2500$. Then we fit the trajectories between these two limits to obtain $\alpha \approx 0.033\text{min}^{-1}$ (yellow line on Fig. 4.2[B]), which gives us $\tau_0 \approx 20.8$ min.

To mimic the experiments, we set $\Delta t = 5$ min. Then to have $k\Delta t \approx \tau_0$, we choose $k = 4$.

We compute the residuals, bin them in 20 N-bins, and if we compute the mean ratio of the variance to N_j (rescaled variance), we obtain

$$\langle \frac{\text{Var}\,(Res)_j}{N_j} \rangle_j \approx 0.295 \tag{4.23}$$

Which gives us, using Fig. 4.11[B]:

$$cv_\mu \approx 0.17. \tag{4.24}$$

We used in the simulations was $cv_\mu = 3.9/21 \approx 0.185$, so our inference method works very well, with an error of $\approx 7\%$.

We can also try to estimate the confidence interval for the estimation that we just made. The only parameter that we used for the estimation was finally the ratio of the variance of the residuals in each bin divided by the bin value N_j, which we call the rescaled variance of the residuals. The estimation was made using the mean value of this parameter, which is theoretically constant. Hence a confidence interval can be found by looking at how this rescaled

variance varies across the bins. Let us compute the standard deviation of this ratio over the bins, we get:

$$\Delta_j \left(\frac{\text{Var}\,(Res)_j}{N_j} \right) \approx 0.019, \tag{4.25}$$

where Δ is the standard deviation. This yields a precision on the microscopic coefficient of variation of, using the approximated linear inference relation:

$$\frac{\Delta cv_\mu}{cv_\mu} \approx 2\%. \tag{4.26}$$

The confidence interval of the inference is then very narrow.

Convergence to the solution

A question that arises when studying this example is the speed of convergence to the final result: how many simulated trajectories do we need to get the correct estimation of cv_μ ? To answer this question, we vary the number of parallel growth trajectories that we take into account for the inference. By analogy with the experimental case, in which each trajectory corresponds to a droplet in a trap of the microfluidic array, we will use the word "trap" for a trajectory. We will then vary the number of traps ($Ntraps$) taken into account for the inference, from 10 traps to the full simulation ($Ntraps = 1000$).

The more traps are taken into account, the more the ratio of the variance of the residuals to the number of cells is close to a constant, see Fig. 4.12[A].

There are then two quantities to study to see how accurate our inference is. The first one is the error made when we estimate cv_μ with fewer traps: how much does the estimated value deviate from the correct value of cv_μ? The second one is the confidence interval of this estimation: what is the range in which this estimated cv_μ lies ?

First, let us study the relative error made on the estimated cv_μ, which compares the cv_μ estimated with $Ntraps$ to the cv_μ estimated with the full simulation ($Ntraps = 1000$):

$$\frac{\epsilon(cv_\mu)}{cv_\mu}(Ntraps) = \frac{cv_\mu(Ntraps) - cv_\mu(1000)}{cv_\mu(1000)}. \tag{4.27}$$

The results are presented on Fig. 4.12[B], and it seems that with only 100 traps, we could get a correct estimation of cv_μ, since the relative error drops below 10% for $Ntraps > 100$, which is already a good accuracy.

However, the relative error is not the only parameter to take into account: the precision of the inference also depends on the confidence interval that we have for this estimated cv_μ, as we vary the number of traps taken into account for the inference. As we have seen on Fig. 4.12[A], if we don't have a lot of traps, the ratio of the variance of the residuals to the bin value is not really a constant, and the confidence interval of the estimation is quite wide.

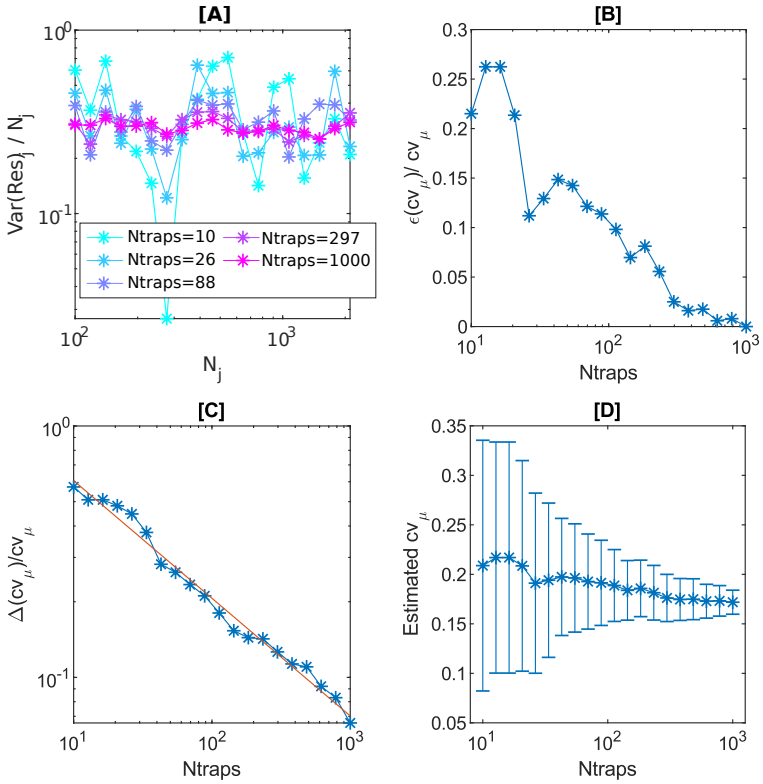


FIGURE 4.12: Influence of the number of simulations ($Ntraps$) on the inference of cv_μ. [A] Variance of the residuals divided by N_j for 5 different number of traps. [B] Relative error made when estimating cv_μ as a function of the number of traps taken into account for the inference (x-log scale). [C] Relative confidence interval on the estimated cv_μ as a function of the number of traps taken into account for the inference (log-log scale), with linear fit. The slope of the fit is ≈ -0.5. [D] Estimated cv_μ as a function of the number of traps taken into account for the inference, with error bars corresponding to the confidence interval of plot [C] (x-log scale)

We will quantify this confidence interval as the standard deviation of this ratio over the bins:

$$\frac{\Delta cv_\mu}{cv_\mu} = \Delta\left(\frac{\mathrm{Var}\,(Res)_j}{N_j}\right)\frac{1}{\langle\frac{\mathrm{Var}(Res)_j}{N_j}\rangle_j}. \tag{4.28}$$

The symbol Δ is here used to note the standard deviation. If we accept the linear inference relation (4.22), the confidence interval that we have on cv_μ is indeed directly related to the confidence interval that we have on the ratio of the variance of the residuals to N_j, which is how much this ratio remains constant across the bins.

This confidence interval seems to decrease as $1/\sqrt{Ntraps}$(see Fig. 4.12[C]), which is a quite slow decrease with $Ntraps$, and this forces us to increase the number of traps way above 100 to have a trustworthy estimation.

This appears even more clearly when we gather these last two results

on the same graph (Fig. 4.12[D]). We can see that we rapidly get a correct estimation of the value of cv_μ, but if we want this estimation to be reliable, we really have to get to high values of $Ntraps$, at least above 500. This is why we keep the number of traps as high as $\approx$ 1000 in the experiments, knowing that this confidence interval will probably be widened because of the experimental noise.

Now let us try to see if we can apply this method to the experiments.

4.4.2 Back to the experiments

The method of inference has been designed to be directly applicable for our microfluidic experiments, using the fluorescence instead of the number of cells $N(t)$. We can in the same way define a finite sequence $\left(Fluo_j\right)_j$ such that:

$$\cup_j \left[Fluo_j, Fluo_{j+1}\right[= \left[Fluo_{det}, Fluo_{sat}\right[. \tag{4.29}$$

Then for every fluorescence trajectory $Fluo^k(t)$ and for every observation time t_i, there exist a unique j such that:

$$Fluo^k(t_i) \in \left[Fluo_j, Fluo_{j+1}\right[. \tag{4.30}$$

Instantaneous Growth Rate

One of the main differences between our experiments and the simulations is the existence of a real detection threshold at the beginning of the trajectory, and a saturation limit and the end. This is taken into account by our inference method, which allows us to set these limits and infer on the exponential phase. But even in the exponential phase, the experimental trajectories are not pure exponential, or only for a very short time. This can be seen by plotting the instantaneous growth rates of all the trajectories as a function of the bin value (see Fig. 4.13[A]). It remains constant only for a very small number of bins, that are in the middle of the exponential phase. As a consequence, the α that we use in the formula of the residuals is not well-estimated, especially at the end of the exponential phase, and the residuals are not centered around zero, but rather negative (see Fig. 4.13[B]).

Corrected residuals

To solve this problem, we will slightly change the way that we compute the residuals. Instead of using α, we will use the mean of the instantaneous growth rates of the $j-th$ bin:

$$Res_j^k = Fluo^k(t_{i+1}) - Fluo^k(t_i)\exp(\alpha_j \Delta t). \tag{4.31}$$

The experimental residuals computed with this correction are centered around zero (Fig. 4.14[A]). They are almost Gaussian (Fig. 4.14[B]). Their variance plotted as a function of $Fluo_j$ increases with a slope 1 (Fig. 4.14[C]), as expected theoretically, but only for a small number of bins. This is confirmed

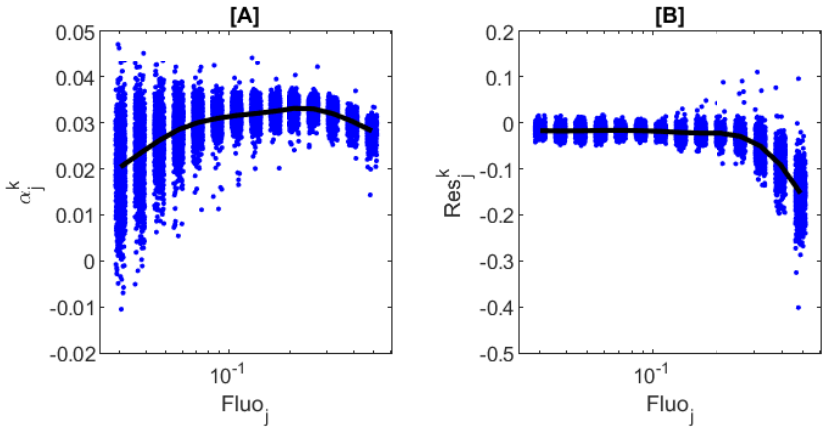


FIGURE 4.13: [A] Instantaneous growth rate in an *E. coli* growth experiment in droplets (same as in Chapter 3), binned by fluorescence value, with mean of each bin (dark bold line). [B] Residuals in the same experiment, binned by fluorescence value, with mean of each bin (dark bold line)

if we plot directly the ratio of the variance over $Fluo_j$, we obtain an approximately constant ratio in the middle of the exponential phase (Fig. 4.14[D]). Indeed for the 6 bins in between $Fluo_f = -2.43$ and $Fluo_j = -1.36$, we get:

$$\Delta\left(\frac{\text{Var}\left(Res^{Fluo}\right)}{Fluo}\right) \approx 0.053\langle\frac{\text{Var}\left(Res^{Fluo}\right)}{Fluo}\rangle_j. \tag{4.32}$$

Which allows us to say that this ratio is approximately constant for the considered fluorescence values. Note that this accuracy is comparable to the one we had in the simulations (eq. (4.25)), where we had an accuracy of $\approx 6\%$.

What about the inference ?

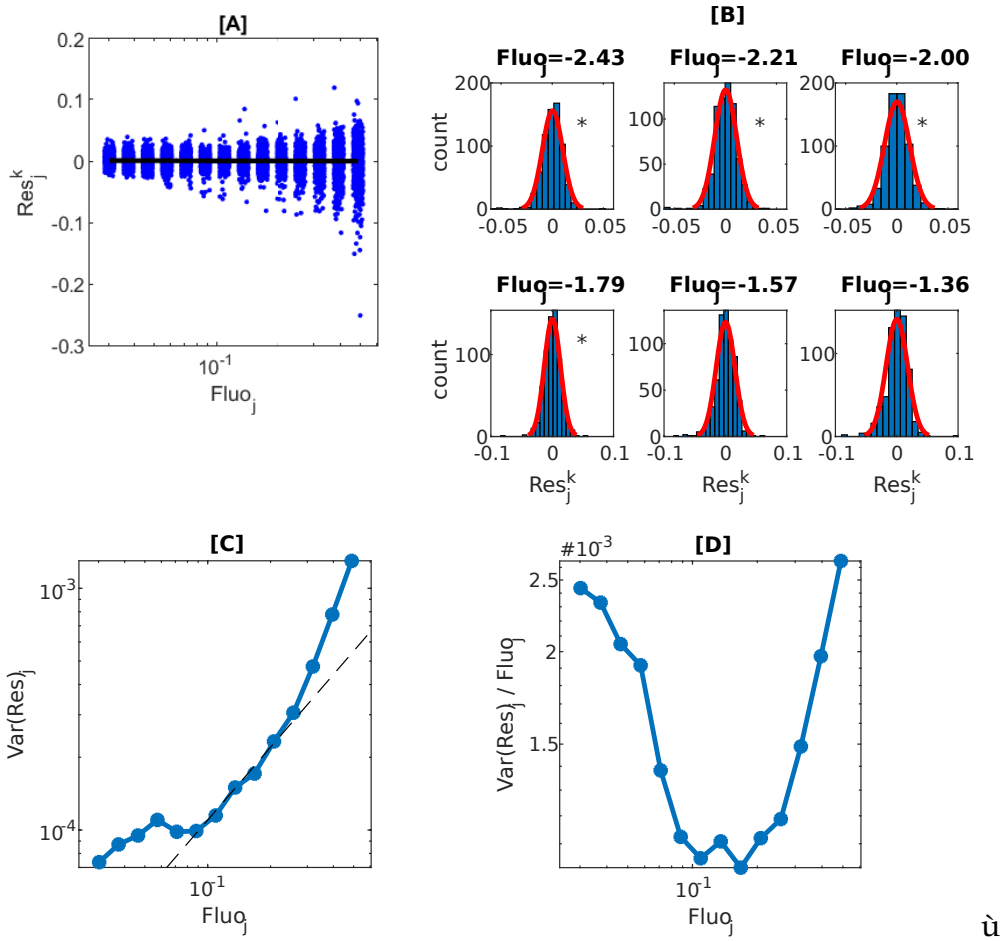
At this point, we have been able to compute experimental residuals that behave in agreement with the classical branching theory. So we can in principle use them to infer the microscopic variability of division times. But a problem remains: the residuals that we have computed are computed on the fluorescence value and not directly on the number of cells in the droplets.

If we consider that the fluorescence is exactly proportional to the number of cells, we have:

$$Fluo(t) = a_f N(t). \tag{4.33}$$

It is easy to see with eq. (4.2) that the residuals computed on the fluorescence are proportional to the residuals computed on the number of cells:

$$Res_{Fluo}(t) = a_f Res_N(t). \tag{4.34}$$



ù

FIGURE 4.14: [A] Residuals obtained by eq. (4.31) in a typical *E. coli* experiment (same as previously), binned by the fluorescence value, with mean value for each bin (dark bold line). [B] Histograms of the residuals for 6 bins, with Gaussian fit (red). The stars indicate a p-value $p < 0.05$ in the Kolmogorov-Smirnov test for normality. [C] Variance of the residuals in each bin. The dark line indicates a slope 1. [D] Variance of the residuals in each bin divided by the fluorescence value of the bin.

Then if we compute the ratio of variance to the fluorescence, we get the same proportionality:

$$\frac{\mathrm{Var}\left(Res^{Fluo}\right)}{Fluo} = a_f \frac{\mathrm{Var}\left(Res_N\right)}{N}. \tag{4.35}$$

Therefore, if we want to use the residuals that we have computed to read the value of cv_μ on the curve of Fig. 4.11[B], we have to know precisely the value of a_f. Unfortunately, as we have seen in Chapter 2, it is very difficult to accurately estimate the value of a_f.

If we use for instance the value that we have inferred using the Bellman-Harris model, we have $a_f \approx 7 \cdot 10^{-4}$. If we compute the mean ratio of the

variance of the residuals to the fluorescence in the experiments, we get:

$$\frac{\mathrm{Var}\left(Res^{Fluo}\right)}{Fluo} \approx 1.1 \cdot 10^{-3}. \tag{4.36}$$

Thus, we get the following estimation for the rescaled variance of the residuals:

$$\frac{\mathrm{Var}\left(Res_N\right)}{N} \approx 1.57 \tag{4.37}$$

This value obtained from the experiment is completely out of range if we want to use the inference relation of Fig. 4.11[B]. In fact, we have a factor 10 here between the rescaled variance of the residuals in the experiments and the estimated variance in the simulations.

Of course, the difference is not only due to the fluorescence coefficient a_f. The heterogeneity of the fluorescence signal of the cells could also play a role, as we have seen that it artificially increased the variance of the residuals measured in fluorescence. Indeed, according to eq. (D.26) of Appendix C, it would lead us to decrease the rescaled variance by the quantity:

$$\frac{\mathrm{Var}\left(Res_N\right)}{N} = \frac{1}{a_f}\frac{\mathrm{Var}\left(Res^{Fluo}\right)}{Fluo} - \frac{\sigma_f^2}{a_f^2}e^{k\alpha\Delta t}\left(e^{k\alpha\Delta t}+1\right). \tag{4.38}$$

If we want to estimate this, we can use the estimation of the heterogeneity of the fluorescence made in Chapter 2, which yielded $\frac{\sigma_f}{a_f} \approx 0.2$. As here we have chosen k such that $k\Delta t \approx \tau_0$, which implies $e^{k\alpha\Delta t} \approx 2$, we get:

$$\frac{\sigma_f^2}{a_f^2}e^{k\alpha\Delta t}\left(e^{k\alpha\Delta t}+1\right) \approx 0.24 \tag{4.39}$$

As we have previously estimated the first term of eq. (4.38) to be ≈ 1.57, we can see that the contribution of the heterogeneity of the fluorescence is not negligible.

But the main problem remains to accurately measure the coefficient of proportionality between the measured fluorescence of a droplet and the number of cells in this droplet. It prevents us from using our inference method based on following the individual trajectories of the droplets to get an absolute measurement of the microscopic coefficient of variation of the individual division times cv_μ. However, this method could still be used to measure the relative difference in between two growth conditions: for instance, how does the use of an antibiotic at sub-lethal concentrations changes cv_μ ? The problem though remains that one has to be sure that a_f is not affected by the change of growth condition, which makes the antibiotic example evoked here a little unrealistic, but other biological questions could maybe take advantage of this relative measurement.

The theoretical and numerical framework could also be used in other situations, in which the relation between the observable quantity (fluorescence, optical density [145]...) and the number of cells is better controlled. It is not

specific to our microfluidic system, although it requires to have a lot of observations of parallel bacterial growth.

4.5 Summary and Conclusion

We have developed in this chapter a novel theoretical and numerical framework for the inference of the cell-to-cell variability of division times from the growth curves of populations. This inference method is based on following individual growth trajectories: the stochasticity of the individual division times makes them deviate from a pure exponential growth, and this deviation can be quantified through quantities that we have called the residuals. The variance of these residuals can be linked back to the individual variability of division times, and thus used for inference. This inference method allows us to free ourselves from the experimental constrains that caused inference methods based on measurements of the distribution of number of cells in the droplets to fail in the previous chapter, in particular the initial stochasticity in the number of cells per droplet and the generation dependent division times in the early stages of the growth.

The analysis carried out in this chapter showed also that one has to be quite careful if she wants to apply this theoretical framework to experimental data. The way that experimental data are sampled, in particular, plays a great role. But this can be solved by correctly binning the experimental data, and we have proven our inference method to be successful to recover the cell-to-cell variability in a simulation of bacterial growth that mimicked our experimental constrains. We also quantified how the experimental noise could impact the measurement of the residuals, and showed that our inference model was quite robust, at least for an experimental noise without any time correlation.

We retrieve the theoretically predicted behavior of the residuals in the experiments. In particular, the experimental residuals are normally distributed, and their variance increases exponentially with the fluorescence signal of the droplets, with the same growth rate as the bacteria. This is a very nice agreement of experimental data with advanced statistical predictions from the theory of branching processes.

Unfortunately, the actual state of our knowledge on our experimental system does not allow us to completely apply this inference method. In our case, it is indeed quite complicated to accurately know the relation between our fluorescence measurements and the number of cells contained in the droplets, as we have demonstrated in Chapter 2. This inference scheme requires a precise quantification of this relation, and hence could not be successfully applied to our experiments.

However, the analysis developed here could maybe be used in a different manner in the future experiments of the lab. For instance, we have done recently preliminary experiments of yeast growth in the microfluidic chip. Yeast cells are bigger than bacteria, and can be directly counted in

the droplets, without any fluorescence measurements. The theoretical framework developed here could thus be employed to infer the cell to cell variability of yeast cells, without any direct or continuous observation of the lineage tree, as required by existing methods [89, 90], even if work still needs to be done on the image analysis side. Division in yeast is also known to be asymmetrical, which could complicate a little bit the model.

This theoretical framework could also be of great use for other microfluidic systems, in which the measurement of bacterial growth is made with other means than fluorescence. For instance, the growth of the colony in microfluidic droplets can be measured by monitoring changes in optical density, and some authors have already proven such systems to be useful for the study of bacterial growth [146, 145]. In particular, the growth of *Actinobacteria* was proven to be suitable for this kind of measurement [146]. Our theoretical inference method could hence be used for a biotechnologically important system, since the majority of modern antibiotics are derived from these microorganisms [146]. This brings us to our next chapter, which focuses on how our microfluidic system could be used to study the effects of antibiotics on bacteria.

Chapter 5

Antibiotics and SoS response

... — ...
E. coli on the Titanic

5.1 Introduction

Since penicillin was discovered in the 1940s and antibiotics became clinically used, bacterial resistance has spread in a spectacular way [147]. This has become a critical issue for some infections, such as tuberculosis, because of the emergence of multi-drug resistant pathogens that are much more difficult to eradicate [148]. This is, in a way, the dark side of Darwinian evolution: as more bacteria are exposed to antibiotics, more resistance pathways are evolved, and given the rate of dissemination of resistance mechanisms [149], and the poor rate at which new antibiotics are being discovered [150], we are unlikely to win this evolutionary arms race against bacterial resistance to antibiotics [151].

There is then a need for a more reasoned use of antibiotics. This involves in particular new experimental tools to test the susceptibility of a bacterial strain to an antibiotic more rapidly: patients could be treated quicker, and more efficiently [152]. Classical methods of antibiotic susceptibility testing often require a few days of incubation [153] before yielding a result. New PCR methods have been developed to reduce the diagnosis time to 1 to 4 hours [154], but they only provide a genetic profile of the strain, and not its direct phenotypic susceptibility to a given antibiotic.

Microfluidics has emerged as a natural tool for this kind of application. Indeed, thanks to the very low volumes that are used, and the high degree of confinement [113], the time needed to determine the susceptibility of a strain to an antibiotic can be drastically reduced compared to classical methods. For instance, in the case of Urinary Track Infections (UTIs), that are of critical clinical importance [155], but often overlooked as they mostly concern the oppressed half of humanity, namely women, it has been shown that

the antibiotic testing time could be reduced to 30 minutes using single-cell microfluidic technology [46]. More generally, a lot of microfluidic devices to test microbial susceptibility have been developed in the past few years, ranging from single-cell technologies to millifluidic systems [156], while some authors put their efforts in developing very elegant and easy to use devices. A device to test the susceptibility of a strain in only five pipetting steps has hence been designed [157]. In the first part of this chapter, we are going to see how our microfluidic chip can be used in this context.

However, antibiotic susceptibility testing is not the only application of microfluidics for the study of interactions between antibiotics and bacteria. They have proven to be useful, for instance, for the characterization of persister cells [158, 11], that are able to survive an antibiotic treatment thanks to their dormant state [159], and not thanks to a hereditary resistance gene. Single-cell approaches are often needed to characterize this kind of phenotypic survival, as they are often due to heterogeneity in the population, both for persister cells [105] or non-dormant tolerant cells [32], that are able to survive an antibiotic treatment while being still metabolically active. In the context, evoked above, of the global antibiotic crisis, a better understanding of these persistence states is critical, as they often pave the way to the emergence of resistant mutants [160]. One bacterial response that is known to induce persistence to antibiotics, at least in some cases [161], is the SoS response. In the second part of this chapter, we will try to see how our chip can be useful for the study of this specific pathway.

5.2 Using the chip to test the response of bacteria to antibiotics

5.2.1 Multi-Chip MIC testing and digital measurements

The first measurement that is routinely made in microbiology labs to characterize the susceptibility of a bacterial strain to an antibiotic is the Minimal Inhibitory Concentration (MIC) [162]. This is the minimal concentration of the antibiotic for which no growth of the strain is observed. In classical microbiology labs, it is measured by plating a bacterial culture on a series of agar plates infused with different concentrations of antibiotic. The plates are then set to incubate to see on which ones bacteria can grow [162]. The same can also be made with liquid cultures instead of agar plates, this is called the broth dilution method [163]: liquid tubes containing a series of dilution of the antibiotic are inoculated with the same bacterial culture. These classical techniques are still commonly used today in hospitals [164]. More advanced techniques can also be used, such as the Etest® commercialized by BioMerieux [165]. It is based on a strip containing a step-like gradient of antibiotics that is put on an agar plate inoculated with bacteria. The MIC is read as the point of the strip where bacteria cannot grow on the agar plate, see Fig. 5.1. For a review of these techniques, see Ref. [153]. In this section, we are going to see how this measurement can be made with our chip, and

how this technique would compare with the classical techniques described above.

We do not mean to conduct here a systematic measurement of the MIC for a particular strain and a particular antibiotic, but rather to expose in a conceptual way the advantages and drawbacks of a measurement of the MIC with our chip. For this purpose, we will use the results of two experiments made with two antibiotics during this PhD, both with *E. coli*. The first antibiotic is Gentamicin (Gm), which belongs to the Aminoglycoside class. It is a bactericidal antibiotic that irreversibly binds the 30S subunit of the bacterial ribosome, preventing protein synthesis. The second antibiotic is Ciprofloxacine (Cip), belonging to the Fluoroquinolone class. Its mechanism of action is to inhibit the bacterial DNA gyrase, hence preventing the replication of the DNA. It is a crucial antibiotic for clinical use, and is on the World Health Organization's list of essential medicines [166]. We will also use it further down as an inducer of the SoS response.

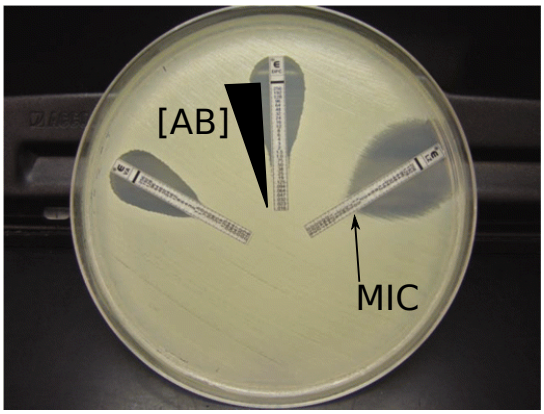


FIGURE 5.1: Taken and adapted from [153]. A Staphylococcus aureus isolate tested by the Etest gradient diffusion method with vancomycin (VA), daptomycin (DM), and linezolid (LZ) on Mueller-Hinton agar. The minimum inhibitory concentration of each agent is determined by the intersection of the organism growth with the strip as measured using the scale inscribed on the strip.

Estimating the MIC

How can we measure the MIC on our chip ? The simplest way is to load several chips in parallel with the same concentration of bacteria, but different concentrations of antibiotics. The bacteria all come from the same antibiotic-free pre-culture. The chips are then incubated at 37°C overnight to see in which chips bacteria have grown. This is still done by fluorescence measurement. A first example of the results that we can get with this method is found in Fig. 5.2(A), featuring *E. coli* and Gentamicin. We can thus read the MIC on this graph, which is the minimal concentration for which no droplet exhibits bacterial growth. For this strain, the MIC is between 5 and 10 μg/mL.

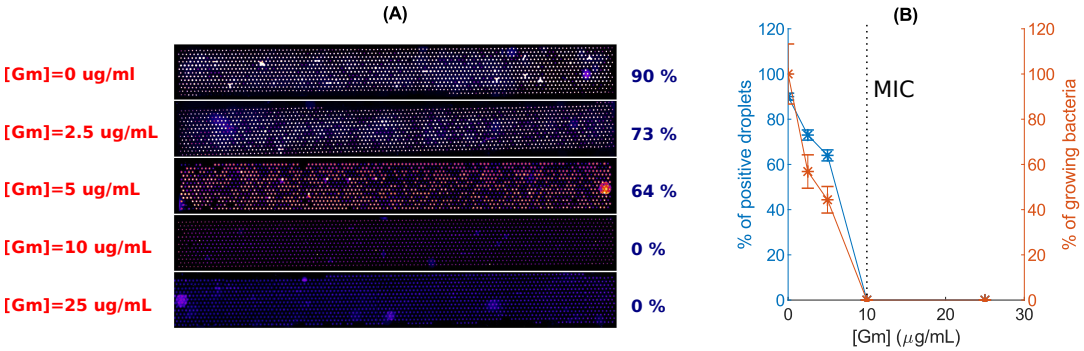


FIGURE 5.2: Multi-Chip Antibiogram. 5 chips are loaded with the same bacterial suspension of *E. coli*, but with different concentrations of Gentamicin (Gm) in LB medium (A) Fluorescence images obtained with a slide scanner (SensoSpot®-Fluorescence, Sensovation AG, Germany), and percentage of positive droplets in each chip, after a 24h hour incubation at 37°C (B) Measured percentage of positive droplets (blue), and computed percentage of growing bacteria (orange), with eq. (5.5)), as a function of the concentration of Gentamicin

How does this measurement compare with the classical measurements of the MIC ? The main difference between the measurement done with our chips and the classical methods of determining the MIC is the inoculum size, which is the number of bacteria on which the antibiotic is tested at the beginning of the experiment. It is known that the inoculum size influences the results of an antibiotic susceptibility test: this is called the inoculum effect [167]. Intuitively, it comes down to the fact that the more bacteria there are in a culture when the antibiotic is applied, the more there is a chance that one of them will grow despite the antibiotic. Indeed, the more bacteria there are, the more there is a chance that one of them will have a particular phenotype or a particular genotype that will allow it to grow despite the antibiotic stress. The cell-to-cell heterogeneity discussed in the introduction plays a key role here.

Because of this inoculum effect, standards have been set to measure the MIC. For the plating method, it is recommended to inoculate the agar plates with 10^4 colony forming units (CFU) [162]. For the broth dilution method, each tube has to be inoculate with 10^5 CFU/mL [162]. For the Etest, the inoculum size is even higher, as all the bacteria growing at a lower concentration of antibiotic on the agar plate can try to make the bacterial front grow in the higher concentration.

In our chip, the inoculum size is lower, at least on the experiment presented on Fig. 5.2. On the one hand, we have $\approx$ 1500 droplets on our chip. On the other hand, we can estimate the mean number of bacteria per droplet from the chip in which no antibiotic was used, using the method that we have described in Chapter 2. We get for the mean number of bacteria per droplet: $\lambda \approx 2$. This means that our inoculum size is here $\approx 3 \cdot 10^3$, which is less than what is commonly used to determine the MIC. Because of the inoculum effect, we expect that the MIC found in our chip is lower than what can be

found with the classical susceptibility methods.

To illustrate this, we can examine the results of an experiment where both measurements are made in parallel. The drug used here is Ciprofloxacine. We allow some *E. coli* to grow to exponential phase ($OD_{600} \approx 0.2$) in an antibiotic free medium. Then we inoculate them at 3/1000 in a series of six 1 mL tubes containing different concentrations of another drug called Ciprofloxacin. These tubes are directly used to load 6 chips. Then the tubes and the chips are put at 37°C for incubation. The bacteria are then initially at the same concentration in the tubes and in the chips. After 24 hours of incubation, the chips are scanned to know the number of positive droplets, and we also measure the optical density of the liquid cultures in the tubes. The results are presented on Fig. 5.3.

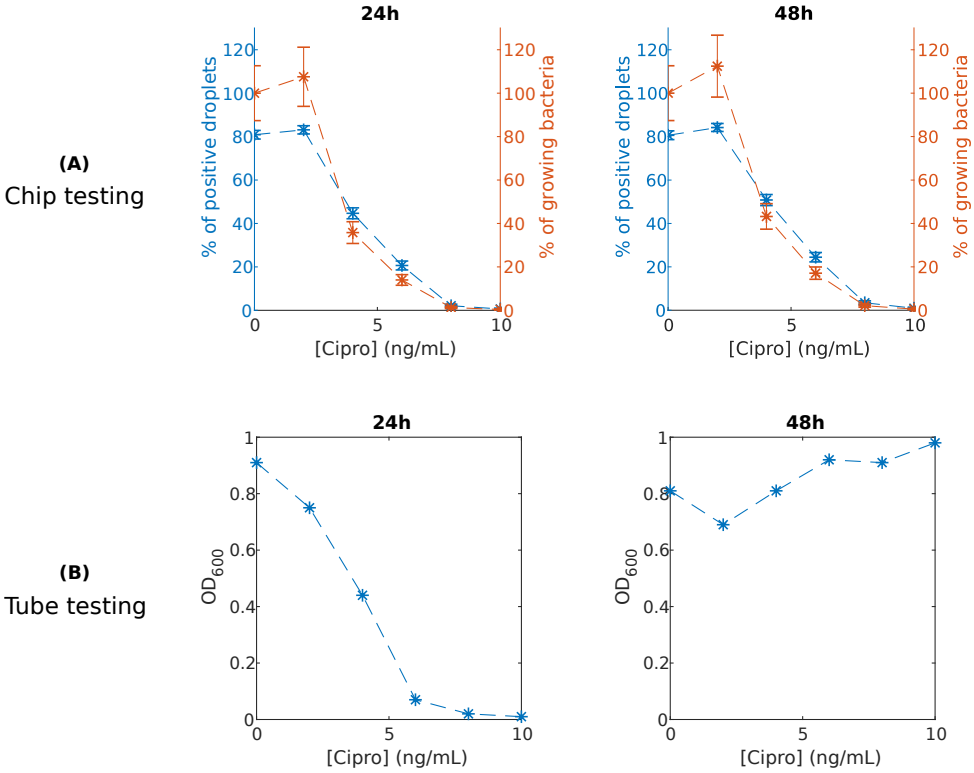


FIGURE 5.3: Comparison between a multi-chip MIC test and a MIC test in 1 mL tubes (broth dilution method). (A) Upper row : (Blue) Percentage of positive droplets in a series of chips loaded with the same bacterial suspension of *E. coli*, and different concentrations of Ciprofloxacin, after 24h (left) or 48h (right) of incubation at 37°C. (Orange) percentage of growing bacteria, computed from the percentage of positive droplets with eq. (5.5). (B) Lower row : optical density of the tubes used originally to load the chips, after 24h (left) or 48h (right) of incubation at 37°C

The results are slightly more complicated than what we expected. If we look at the results after 24 hours, we see that the chips with a Ciprofloxacine concentration lower than 8 ng/mL have some positive droplets, in which bacteria have grown. The same thing is found in the tube, with all tubes

below 8 ng/mL displaying a non-negligible optical density which indicates that bacteria have grown in this tube. Both for the chips and the tube, the MIC would be 8 ng/mL. However, if we look at the results after 48 hours, we see that while the chip results are unchanged, bacteria have finally grown in all tubes. The concentration needed to really inhibit the growth in tube is then superior to 10 ng/mL. Interesting conclusions can then be drawn from this experiment:

- First, the growth dynamics are impacted by the antibiotic stress. Bacteria grow more slowly when this stress is increased, this is why they take longer to be detected in tubes at concentrations approaching the MIC. The static measure of the MIC is therefore not a sufficient measurement to characterize the effect of an antibiotic. This idea will be developed in the next section.

- We see the impact of the inoculum effect when we measure the MIC in our chip, with a higher value of the MIC found in tube. The MIC measured in our chip is under-estimated.

- However, the measurement of the MIC in our chips yields a final result after 24 hours, contrary to the broth dilution method that requires 48 hours. We see here an effect that we have already evoked a few times in this thesis. Thanks to the low volumes used with microfluidic technologies, the final state is reached quicker, and measurements are made faster.

Let us come back to the inoculum effect. The MIC found for the same strain and Ciprofloxacine, using an Etest is 16 ng/mL, which confirms that the value found with our chip is underestimated. If we want to determine the MIC using our chips with a inoculum size that is comparable to the ones recommended for the classical methods of MIC testing, we can just increase the initial concentration of bacteria loaded on our chips. Indeed, we just have to tune our dilution of the pre-culture used to load the chips so that the mean number of bacteria per droplet is around 10. By doing this, we will have around 10^4 bacteria loaded on each chip. This is an inoculum size comparable to the one recommended for testing the MIC with the plating method [162]. However, everything comes at a cost, and we will see in the next section that this increase in inoculum size can be prejudicial if we want to retrieve more information on the bacteria from those multi-chip experiments.

Digital measurement of the growth probability under antibiotic stress

Indeed, with these multi-chip experiments, we can not only estimate the MIC but also the probability for a bacteria to grow at a given concentration of antibiotic. Indeed, for each concentration, we don't only have a global binary response (growth or no growth on the chip): we can also count the number of droplets in which bacteria have grown. This is a typical digital approach:

the droplets are used as binary probes for the response to antibiotics, to infer an individual cell response.

We can note that the percentage of growing bacteria is directly accessible as well if the plating method is used to determine the MIC. In this case, the number of colony forming bacteria can just be counted on each plate, and compared to the number of bacteria growing when plating the same culture on an antibiotic-free agar plate. An example of this kind of measurement can be found on Fig. 5.4, which was taken from ref. [168]. On this figure, the percentage of growing bacteria as a function of the antibiotic concentration is found for two *E. coli* strains, a wild-type strain and a resistant mutant. We can see that the profile of this percentage of viable cells as a function of the antibiotic concentration is not the same at all for the two strains. For the wild-type strain, there is a sharp decrease near the MIC, whereas for the resistant strain, the decrease in percentage of viable cells is smoother. This observation was used as a basis in ref. [168] to demonstrate that the growth phenotype of the resistant mutant under antibiotic stress was bistable: there exist a range of concentration of the antibiotic for which these mutants can stochastically grow or not. We will not go deeper into those concepts here, but it shows that the percentage of viable cells is a valuable information.

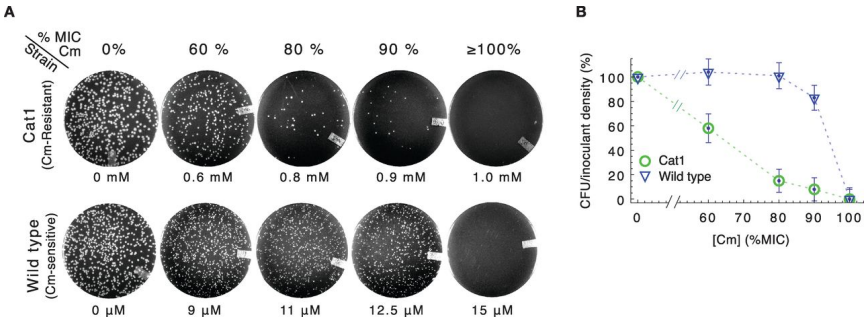


FIGURE 5.4: Taken from [168] Heterogeneous response of Chloramphenicol (Cm) resistant cells.*E. coli* cells were diluted from log phase batch cultures lacking Cm, and were spread onto LB agar at densities of several hundred cells per plate before overnight incubation at 37°C. (A) Typical plate images of Cm-resistant Cat1 (top row) and Cm-sensitive wild type (bottom row) cells, with Cm concentration indicated below each plate and also given above as approximate fraction of the empirically determined MIC plate for each strain. (B) Percentage of viable cells grown on Cm-LB plates; CAT-expressing cells (Cat1, green) and wild type cells (EQ4, blue). Error bars estimate SD of CFU, assuming Poisson-distributed colony appearance

At the opposite, if the MIC is measured with the broth dilution method or with an Etest, this percentage of viable cells for each concentration cannot be obtained. Indeed, one only obtains a binary information on the growth at each concentration: growth in a tube or no observable growth; growth at a given concentration on the agar plate or no observable growth. This is therefore an advantage of the plating method or of our chip-based method. But

let us go back to our multi-chips measurements and see how the probability for a cell to grow can be inferred.

Let us call μ the probability for a cell to grow into a colony despite the antibiotic stress, that we will consider to be the same for all cells, and p^+ the probability for a droplet to be positive, i.e. to contain bacteria at the end of the experiment. We will consider here that if a cell grows, all the daughter cells grow as well. We want to get the relation between p^+ and μ. Let us first use the law of total probability, conditioning p^+ by the number of cells encapsulated in the droplet N_0:

$$p^+ = \sum_{k=0}^{\infty} p(N_0 = k) p^+_{N_0=k}, \tag{5.1}$$

where $p^+_{N_0=k}$ is the probability of having a positive droplet if the number of bacteria in the droplet is exactly k. We know that the number of cells per droplet when the chip is loaded follows a Poisson distribution, hence we have:

$$p^+ = \sum_{k=0}^{\infty} e^{-\lambda} \frac{\lambda^k}{k!} p^+_{N_0=k}, \tag{5.2}$$

If we know that $N_0 = k$, we have, since a droplet is positive if and only if at least one cell that has grown despite the antibiotic treatment:

$$\begin{aligned} p^+_{N_0=k} &= 1 - p^-_{N_0=k} \\ &= 1 - (1-\mu)^k. \end{aligned} \tag{5.3}$$

Therefore, by inserting this expression into the previous sum, we get:

$$\begin{aligned} p^+ &= \sum_{k=0}^{\infty} e^{-\lambda} \frac{\lambda^k}{k!} \left(1 - (1-\mu)^k\right) \\ &= e^{-\lambda} \left(\sum_{k=0}^{\infty} \frac{\lambda^k}{k!} - \sum_{k=0}^{\infty} \frac{\left((1-\mu)\lambda\right)^k}{k!} \right) \\ &= e^{-\lambda} \left(e^{\lambda} - e^{\lambda(1-\mu)} \right) \\ &= 1 - e^{-\lambda\mu}. \end{aligned} \tag{5.4}$$

Thus, by counting the number of positive droplet N^+, we can estimate $p^+ \approx N^+/N_{tot}$. We can also estimate λ by counting the number of positive droplets when there are no antibiotics, as we have explained in Chapter 2. Hence we can estimate to the probability for bacteria to grow despite the antibiotic stress:

$$\mu = -\frac{1}{\lambda} \ln\left(1 - p^+\right). \tag{5.5}$$

We can therefore apply this method to our previous experiments presented on Fig. 5.2 and 5.3. For each chip, we can count the number of positive droplets and hence estimate μ with eq. (5.5). For clarity, this probability is plotted as a percentage on the figures 5.2 and 5.3 (orange axes). As we can

see, in our experiment, the percentage of positive droplet is superior to the percentage of surviving bacteria. This is simply due to the fact that there is more than one bacteria encapsulated per droplet when the chip is loaded.

What is the error that we make when making this estimation ? There are in fact two sources of error. The first one is the error made when estimating λ by counting the number of positive droplets on the chip without antibiotic, that has been computed in Chapter 2. The second one is the error made by estimating p^+ by counting the number of positive droplets, which is in fact of the same nature: it is the error made when estimating a probability with a digital measurement. We can also estimate it by using the Central Limit Theorem, and obtain a confidence interval on p^+. We can combine the two effects to obtain the confidence intervals plotted on Fig. 5.2(B). We can see that the error made with the agar method (Fig. 5.4(B)) is similar.

However, this confidence interval depends on the value of μ and λ. As we have explained in Chapter 2, if λ is too big, all the droplets are filled with bacteria, and the mean number of cells per droplet cannot be estimated any more. The same thing happens for μ. If μ is very close to 1, then all the droplets contain growing bacteria, and p^+ can not be used to estimate μ. The two effects are in fact mixed in this case, as the value of λ also influences the estimation of μ. Indeed, if λ is increased, droplets are more likely to be positive, as there are more bacteria per droplets, the probability that one of them will grow is higher. The estimation of μ will then be less precise if λ increases.

This sounds a bit technical as it stands, so let us illustrate this with a theoretical example inspired by the results presented on Fig. 5.4. The question is the following: can we retrieve the two profiles of the percentage of growing bacteria using our digital measurements? To answer, we impose two profiles for the percentage of growing cells when the antibiotic concentration is varied: one very sharp profile, with the percentage of growing bacteria dropping suddenly at a critical concentration, and a linear profile, with the percentage of growing bacteria linearly decreasing with the concentration. This mimics the different viability percentages of, respectively, wild-type and resistant cells of Fig. 5.4. Then, for these two profiles, we compute the resulting number of positive droplets on the chip according to eq. (5.4). We repeat the theoretical computation for two values of the Poisson parameter λ: $\lambda = 2$, which is close to what we had on Fig. 5.2, and $\lambda = 10$, which is closer to the standards for MIC determination.

If the percentage of surviving cells decreases sharply close to the MIC, as for the wild type cells in Fig. 5.4(B), the two values of λ are almost equivalent. However, at 60% of the MIC, for instance, all droplets are still positive if $\lambda = 10$: the determination of the percentage of growing bacteria will then be less accurate with this higher value of λ. This is even clearer if we consider a linear decrease of the percentage of growing cells with the concentration of antibiotics, which mimics the case of resistant cells (Cat1) on Fig. 5.4(B). The decrease in percentage of surviving cells is not detected at all on the percentage of positive droplets for the case $\lambda = 10$, until the concentration of antibiotic reaches 75% of the MIC. Indeed, there are so many bacteria per droplet

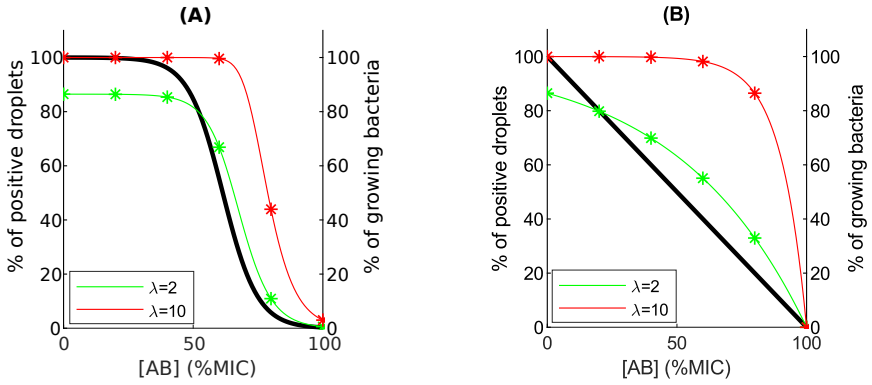


FIGURE 5.5: Theoretical percentage of positive droplets for two profiles of the percentage of growing bacteria when varying the antibiotic concentration (black bold line): (A) sigmoidal shape with rapid decrease of the percentage of growing cells, like on Fig. 5.4(B), for the wild type cells; (B) linear decrease of the percentage of growing cells, like on Fig. 5.4(B), for the Cat1 cells. The percentage of positive droplets is then computed with eq. (5.4) for two values of the Poisson parameter λ, which corresponds to the mean number of cells per droplet. The results are plotted in green for $\lambda = 2$, and in red for $\lambda = 10$. Some particular points of the obtained theoretical curves are highlighted with stars. These are the points that would be obtained if an experiment was run with 6 chips in parallel, as in Fig. 5.3.

that there is always one that statistically grows. It is therefore impossible in this case to use the digital measurement of positive droplets for inference of the percentage of growing bacteria, whereas if $\lambda = 2$, this is doable.

There is then a trade-off between the measurement of the MIC and the measurement of the percentage of bacteria growing under an antibiotic stress. To measure the MIC according to the classical standards, it is better to increase λ, whereas to measure the survival rate of antibiotics, it is preferable to decrease it.

At this point, we have demonstrated that the MIC of a strain can be determined with our chip. However, the MIC only quantifies the ability of a strain to grow under a constant antibiotic stress, but as we have seen on Fig. 5.3, the dynamics also plays a great role. In the next section, we will see how we can use our chips to study the efficiency of an antibiotic treatment over time.

5.2.2 Exposure time to antibiotics

Indeed, the MIC does not provide any indication on the dynamics of the action of the antibiotics [169]. It just quantifies the intuitive fact that the more antibiotic there is in a culture, the more it is difficult for bacteria to grow. It is also intuitive to think that the longer bacteria are exposed to an antibiotic, the more bacteria are killed. This is traditionally quantified by time-kill curves, which are the percentage of surviving bacteria as a function of the exposure time to the antibiotic. These measurements that contain information

about the dynamics of the response have proven to be useful for designing efficient treatments against pathogens [170], and have also been studied theoretically [171]. Contrary to the measurement of the MIC, they are useful for the study of bacterial persistence, which we have already defined. Balaban [11] used for instance time-kill curves to show that bacterial persistence was due to a phenotypic switch. As we can see on Fig. 5.6, the time-kill curves of wild-type bacteria are well described by a double exponential kinetics. Indeed, the number of viable cells in a culture under antibiotic stress decays exponentially with time, but at some point there is a change in the rate of decay. If we fit the exponential decay in number of viable cells with an exponential function $\exp(-t/\tau_0)$, the majority of the cells is killed with a characteristic time τ_0 of 25 min, whereas after the slope shift the characteristic time is 6 hours. This indicates that there are some cells that are much more difficult to kill with an antibiotic than the vast majority of the cells. The same behavior is observed for *hipA7* mutants, that are known to exhibit a higher proportion of persisters, but the change in killing rate is attained faster, and the proportion of persisters is indeed higher. We will now compare these findings with data obtained on our chips.

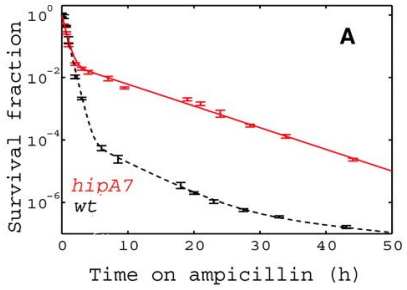


FIGURE 5.6: Taken from [11] Killing curves of wild-type and hipA7 mutant cells. Black symbols and dashed curve indicate wild type (wt); red symbols and solid curve, hipA7 mutant. HipA7 mutants are known to display a higher proportion of persister cells.

In microbiology labs, these time-kill measurements rely once again on plating [172]: bacteria are exposed to a drug in a liquid culture, and at the desired time points, the culture is sampled, diluted, and plated on agar plates to count the number of CFU and thus the number of viable cells in the culture. This method is very demanding in terms of space (a lot of agar plates are needed) and time (colonies take usually a few days to be countable). A simpler and faster method with chips would be then very valuable.

To measure time-kill curves with our chip, we take advantage of a feature that we have not described yet in this PhD, which is the possibility to control the environment of the bacteria over time by replacing the oil surrounding the droplets with an aqueous medium. In fact, the loading process can be performed with a mixture of ultra-low gelling agarose and liquid medium instead of pure liquid medium. Then the chip is placed in a fridge for 30min

to allow the agarose to gelify. As a consequence, the oil phase can be removed and replaced by any aqueous phase: the droplets and their content will remain stuck in the traps because of the gelification.

Therefore, to measure time-kill curve with our chips, we can load several chips in parallel with the same bacterial culture and the same concentration of antibiotic. The chips are then all put at 37°C for incubation with the antibiotic. Each chip corresponds to a time point in the time kill curve: at each time point, we take one chip, remove the surrounding oil and replace it with antibiotic-free medium. The droplets remain in the traps because they were made with agarose gel. Then, once enough fresh medium has been flowed through the chip to remove the antibiotic, we reintroduce oil to isolate each droplet from one another. In fact, the aqueous phase is pushed by the oil and droplets are generated once again through a "breaking" process, as described in Chapter 2. Throughout this whole process, the bacterial content of the droplets does not change, because bacteria are embedded in the agarose gel. The chips are put back in the incubator, and after 24 hours, the number of positive droplets in each chip can be measured (see Fig. 5.7(A)). Each positive droplet corresponds to at least one bacteria that has survived the antibiotic treatment.

Using eq. (5.5), we can once again obtain an estimation of the time-kill curves (Fig. 5.7(B)). The variation of the fraction of viable cells is close to the expected exponential decay with time, with a characteristic time $\tau_0 \approx 90$ min (Fig. 5.7(C)), even if more time points would be necessary to confirm it. However, we don't see a decrease of the rate of decay, which would indicate the presence of persister cells, as we have explained previously. This once again would need confirmation with more time points. The absence of persister cells in the chip is confirmed by the fact that there was no positive droplet on the chip incubated for 12 hours with Gentamycin. We can explain that by the fact that the frequency of persister cells in wild-type *E. coli* is known to range between 10^{-4} and 10^{-6} [159]. We are therefore not likely to see them on our chip if we don't increase the mean number of cells per droplet up to $\lambda \approx 10$.

However, with this relatively high value of λ, all the droplets of the antibiotic-free chip would be positive. As explained in Chapter 2, the estimation of λ is then impossible. This challenge can be easily overcome, as already discussed in Fig. 2.5: another chip can be loaded in parallel with a higher dilution of the same pre-culture. We can estimate λ in this chip, and as the dilution ratio is known, λ can also be retrieved for the other chips.

Besides, for the early times of the time-kill curves, many bacteria can survive. If a high value of λ is used, we will once again run into the same estimation problem. To solve this issue, the concentration of bacteria used for loading can be adjusted for the different time points of the time-kill curves, which correspond in our case to different chips. On the one hand, for the early times, a low number of bacteria per droplet is preferable, to not have only positive droplets. On the other hand, for the late times, few bacteria can survive. It is then preferable to increase the number of bacteria per droplet, in order to obtain positive droplets. As long as the relative dilution between the different chips is known, the whole time-kill curve can be reconstructed.

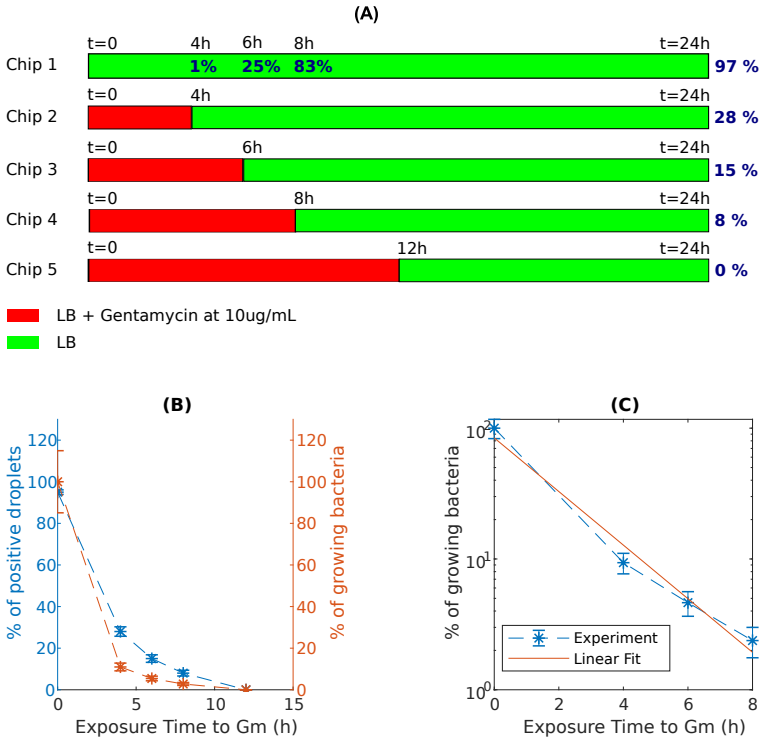


FIGURE 5.7: Multiple chip time-kill curve of *E. coli* when the exposure time to Gentamycin is varying. (A) Principle of the experiment: all the chips were loaded with bacteria coming from the same pre-culture of *E. coli*. Chip 1 is the control chip, for which no Gentamycin was used. The blue number indicate the number of positive droplets (B) Results for the percentage of positive droplets (blue) and computed percentage of growing bacteria for each exposure time with eq. (5.5) (orange). (C) In log scale, result for the percentage of growing bacteria and linear fit.

Overall, these results show that the time-kill curves of an antibiotic can be obtained with our chips. The concentration of bacteria used to load the chip has to be adjusted depending on the regime that we want to observe.

5.2.3 Growth under antibiotic stress

Until now, we have only studied the effect of antibiotics as a matter of life or death of the bacteria: the study of the MIC or the time-kill curve both quantify the ability for the bacteria to survive (or be killed) by the antibiotic, when the concentration or the exposure time is varied. But another approach matters at least as much: if some bacteria are not killed, what is the nature of the growth of the bacteria under an antibiotic stress ? This plays a crucial role, for instance, in the emergence of antibiotic resistance: even at concentrations of drug far below the MIC, *de novo* resisting mutants can be selected [173].

Furthermore, it has been known for a long time that growth and antibiotic susceptibility are correlated: a slow bacterial growth is often correlated with a decreased sensitivity to antibiotics [174]. The most extreme case is the one

of bacterial persister cells, that are able to survive an antibiotic treatment by being in a dormant state [159].

However, although these effects have been known for a long time, there is still a lack of mathematical models relating the molecular mechanism of action of antibiotics to the growth rate of the bacteria under antibiotic stress. These models could be of key importance for designing optimal antibiotic treatments, in particular for drug combinations, that can be used to fight the emergence of resistance [175, 176].

One of the first predictive models for the growth rate under antibiotic stress was proposed only a few years ago by Greulich *et al.* [74] for ribosome targeting antibiotics. These antibiotics can be classified into two classes, those binding reversibly to the ribosomes and those binding irreversibly. From this detailed study of the mechanism of action, and using the growth law relating the ribosomal content of the cell to its growth rate [177], the growth rate of the bacteria could be predicted under antibiotic pressure. It was shown that the two classes of antibiotics had different effects on the growth rate, with increased susceptibility with the growth rate for reversibly binding antibiotic and the inverse effect for irreversibly binding antibiotic.

Nevertheless, this model only considered a fixed dose of antibiotic, although the dynamics can be of crucial importance to better design antibiotics treatments [178]. The same authors recently extended their model to take into account a time-dependent antibiotic concentration [179], and showed that the two classes of antibiotics that have been described also displayed different responses to time-varying antibiotic concentrations, see Fig. 5.8. They studied several time profiles of antibiotic exposure, and different concentrations of the drug. Antibiotics that have a low binding affinity with the ribosomes were found to cause a sharp decrease in growth rate when a step concentration is suddenly applied. At the opposite, for the high affinity antibiotics, the decrease in growth rate was found to be slower. For a time-finite antibiotic pulse, the dynamics are also very different (Fig. 5.8). However, this second work was only numerical and theoretical, and there is still a need, to our knowledge, for experimental confirmation: a device is needed in which the growth rate can be monitored while the micro-environment of the bacteria, in particular the antibiotic stress, is controlled. A few devices have been developed [180], but we believe that our chip could be of great interest here.

Indeed, we have already mentioned the fact that we can control over time the concentration of antibiotic thanks to a phase change. The different time profiles of exposure to antibiotic could thus be tested, as the chip can be placed under a motorized and temperature-controlled microscope, with the concentration of antibiotic controlled by flow rate controllers, and scanned for growth rate measurement at the same time. That imposes though the use of a strain producing a fluorescent protein for growth rate measurement as in Chapter 2.

There is another feature of the chip that we have not mentioned yet in this manuscript that could be of great use here: a gradient of antibiotics can be created on the chip, each row of the two-dimensional array of anchors being perfused by a different concentration of antibiotic. This only requires a slight

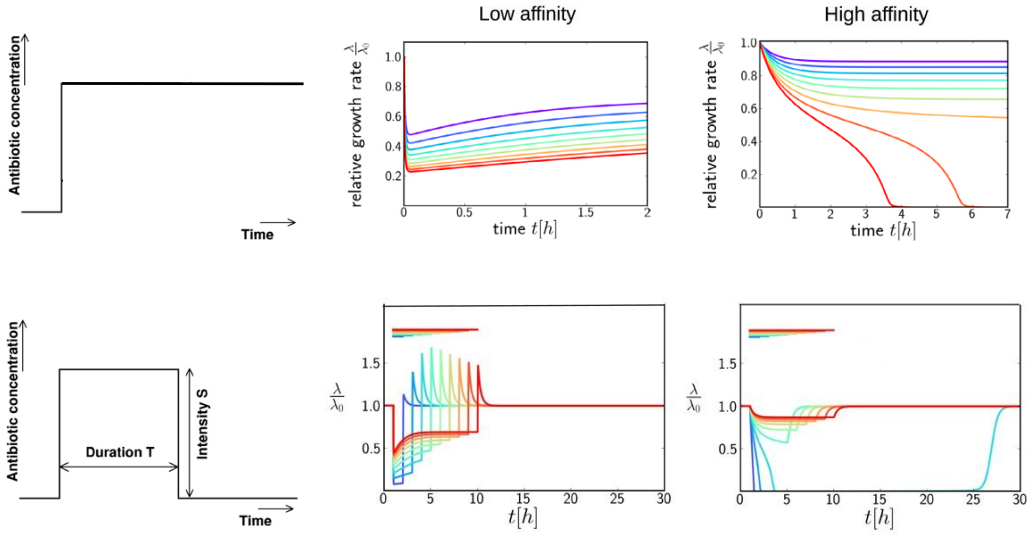


FIGURE 5.8: Taken and adapted from [179] Numerical trajectories showing the dynamics of growth inhibition after a step increase (upper row), or a transient step increase (lower row), for antibiotics that have a low affinity with ribosomes (left) or a high affinity (right). What is plotted is the relative growth rate $\lambda(t)/\lambda_0$ as a function of time after the step increase in antibiotic concentration. For the step increase, the final antibiotic concentration a_{ex} final is indicated by the line colour, ranging from $0.4 \cdot \mathrm{IC}_{50}$ (purple) to $1.3 \cdot \mathrm{IC}_{50}$(red), in steps of $0.1 \cdot \mathrm{IC}_{50}$. The IC_{50} is the concentration of antibiotic needed to halve the growth rate of the strain in steady-state, see [74]. For the transient step increase (lower row), the antibiotic concentration is switched suddenly to a value S at the start of the dose, and is switched back to zero after a time T. The total dose $S \times T$ is fixed at $4 \cdot \mathrm{IC}_{50}$

modification of the design of the chip, as described on Fig. 5.9. This concept has proven to be useful, for instance, for the determination of MIC directly on-chip [97].

Hence, with this modified chip for gradient production, different concentrations could be tested on the same chip, which is a great advantage compared to existing systems such as the morbidostat [180]. Different concentrations for a step profile of antibiotic concentration could be tested, to be compared with the theoretical results of Fig. 5.8. However, one has to be a little bit careful with the experimental protocol if she wants to compare the results with the theoretical predictions: the bacteria have to be in exponential phase before the antibiotic is applied. As the chip has to be put in a fridge prior to phase change for gelification of the droplet, the bacteria may take longer than 3 or four generations, as found in Chapter 3, to be in full exponential phase. But by waiting a little bit more, the effect of several different doses of antibiotic could be tested and compared with the theoretical predictions of [179]. The whole chip can be scanned in approximately 12 min, so the

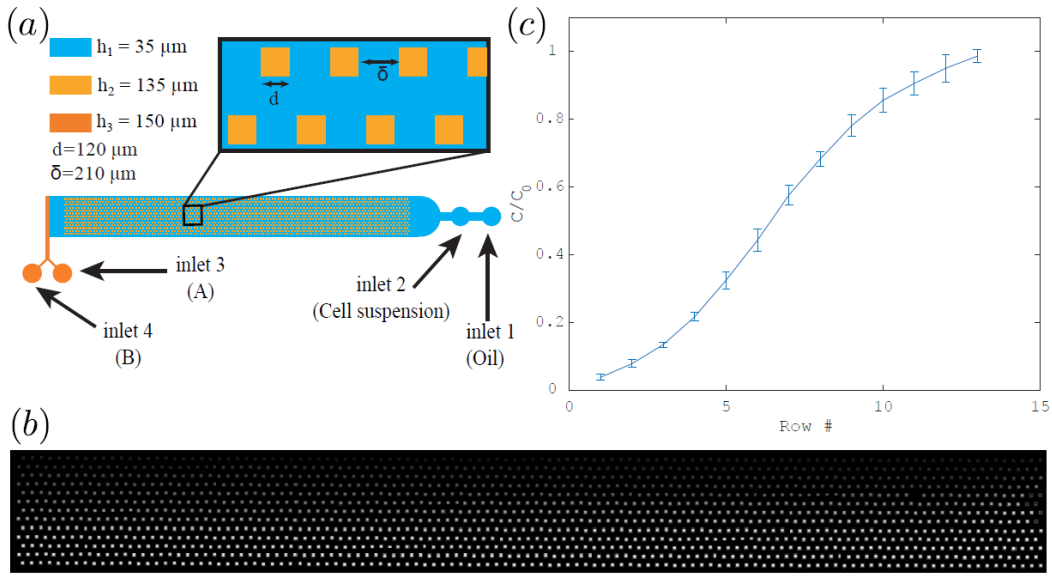


FIGURE 5.9: Taken from [110] a) General chip geometry for the production of a continuous concentration gradient within the chip. Inlets 1 and 2 are used for breaking the droplets on the anchors. The modification relies on the addition of inlets 3 and 4, which are used to produce the antibiotic concentration gradient thanks to the big height difference between the inlets and the chamber, as described in [181]. We used different heights for the main chamber: both 35μm and 15μm can be used. b) large scan of a concentration gradient produced on chip with pure water and a solution of fluorescein. The fluorescein concentration increases gradually from top to bottom rows of the array. c) Calibration curve for the 15 μm chamber presented in b) with an inlet flow rate of 3 μL/min. The error bars are calculated for the standard deviation of each concentration along each row.

main dynamic features presented in this paper should be observable experimentally, even if some of them will probably be missed, such as the peaks in the case of a transient step-like dose. If one wants to have a better temporal resolution, fewer rows can be scanned. Instead of 13 different concentrations, 8 can be scanned every 5 minutes, for instance.

This dynamical growth response of bacteria to ribosome targeting antibiotics is then one case where the experimental possibilities of the chip used throughout this PhD could be of great utility. Time was missing during this present PhD to conduct the experiments, but it is a suggestion for further experimental work in the group.

In the first part of this Chapter, we have shown how our chip could be used to carry out routine measurements made in microbiology labs, such as the MIC and the time-kill curves of antibiotics. Through a careful digital analysis, the use of a chip can lead to a gain of space and time compared to classical plating techniques. We have also tried to suggest a way to use

the full possibilities of the chip in order to address very recent experimental challenges for the understanding of the dynamics of the growth inhibition of bacteria by antibiotics. In the next part, we will expose some preliminary results showing how our chip can be used to study a challenging bacterial stress response called the SoS response.

5.3 Using microfluidics to study the SoS Response of bacteria

5.3.1 A little introduction

The SoS response of bacteria has now been known for more than 40 years [182], but is still challenging for researchers nowadays. It is a broad and transient response that is activated when the DNA of the bacterial cell is damaged. A short overview of the SoS response will be exposed here, for an excellent review read [183].

To go more into details, the SoS response is triggered when there is an abnormally high amount of single-strand DNA (ssDNA) in the cell. Those single-strand DNA are created following DNA damage, from double-strand DNA breaks, but can also be due to stalled replication [184]. The induction pathway is then the following [185]: this triggers the binding of RecA to these single-strand filaments, which in turns catalyzes the cleavage of the LexA repressor. This causes the derepression of the SoS genes, which are about 40 in *E. coli* for instance [186].

The induction of the SoS response can be due to multiple factors: UV radiation [187], reactive oxygen species [188], or even high pressure [189]. What interests us more is the fact that antibiotics can also induce the SoS reponse. This is not surprising for antibiotics that directly target the replication of DNA, such as Fluoroquinolones [190]. But there are also more complex pathways that can come into play. For instance, β-lactams, that target the cell wall, have been shown to be able to induce the SoS response of *E. coli* [191]. The activation of the SoS response can also be due to doses of antibiotics that are far below the MIC, and don't even affect the growth of the bacteria [192].

What are then the effects of this SoS reponse? The two main effects are growth inhibition [193] and DNA repair. Indeed, the LexA regulon includes a lot of genes involved in DNA repair, such as the recombination and repair genes recA, recN, but also genes that codes for DNA polymerases that are known to be error-prone [5, 194]. Therefore, it has been shown that the SoS response increases the rate of mutagenesis [6], paving the way for the emergence of resistant mutants to antibiotics [195].

However, this is not the only way by which the SoS response increases the resilience of bacteria to antibiotics treatments. For instance, it can also induce the formation of dormant persister cells [161]. Some resistance genes to Fluoroquinolones have also been proven to be under the direct control of LexA [196]. The SoS reponse is also involved in stress-induced biofilm

formation [197], that are known to affect the efficiency of antibiotic treatment [198]. Finally, the SoS response has been shown to promote the horizontal gene transfer of resistance genes [199]. Molecules involved in the SoS response have then naturally become a potential target for new treatment strategies [200].

A better understanding of the SoS response is then crucial in the context of the global antibiotic resistance crisis. A lot is already known, but a lot of the studies referenced here are based on population-scale approaches. A single-cell approach is globally lacking, although it has been shown to be useful, for instance, for the discovery of the dynamics of the SoS response [201], which encourages for the use of stochastic models to describe this response [202]. Other bacterial stress responses, and the ability of bacteria to survive to a stress, have been shown to depend a lot on the cell to cell variability [106]. It is intuitive to believe that the SoS response will behave similarly, and thus single-cell approaches are necessary.

Our microfluidic chip, in this perspective, is a useful tool because it enables us to isolate individual cells in droplets. This enables us not only to characterize the SoS response of individual cells, but also to relate it to the phenotypic ability to survive an antibiotic treatment. Indeed, as the descendants of the mother cell stay in the droplet, the ability for a single cell to re-grow after an antibiotic stress inducing the SoS response can be quantified and linked to the intensity of the SoS response. We will present in the next part the experiments that we conducted in this direction in collaboration with the Mazel group of Institut Pasteur. They have been working on the SoS response in *E. coli* and *Vibrio Cholerae* for a long time, and provided strains expressing fluorescence when the SoS response is activated. In the first part, we will show preliminary results obtained with an *E. coli* strain that they kindly let us use in 2016. These preliminary findings led to the obtention of a grant to pursue the collaboration between the two research groups. Another strain was constructed, with more fluorescent reporters, and the results of more recent experiments will be discussed in the second part.

5.3.2 Preliminary results

The experiments that presented here were conducted using a strain engineered from *E. coli* MG1655, with a GfP fluorescent reporter for SoS response (GfP under the control of recN, see complete description in Ref. [6]). The objective of these experiments was to demonstrate the utility of our microfluidic chip to study the SoS response of individual cells, particularly, as we have already mentioned, the ability to survive an antibiotic treatment in relation with the intensity of the SoS response of the cell. We will use Ciprofloxacin (Cip) as the inducer of SoS response. It is known to induce the SoS response [161], even at sub-MIC concentrations [203]. The experiments were conducted in LB medium.

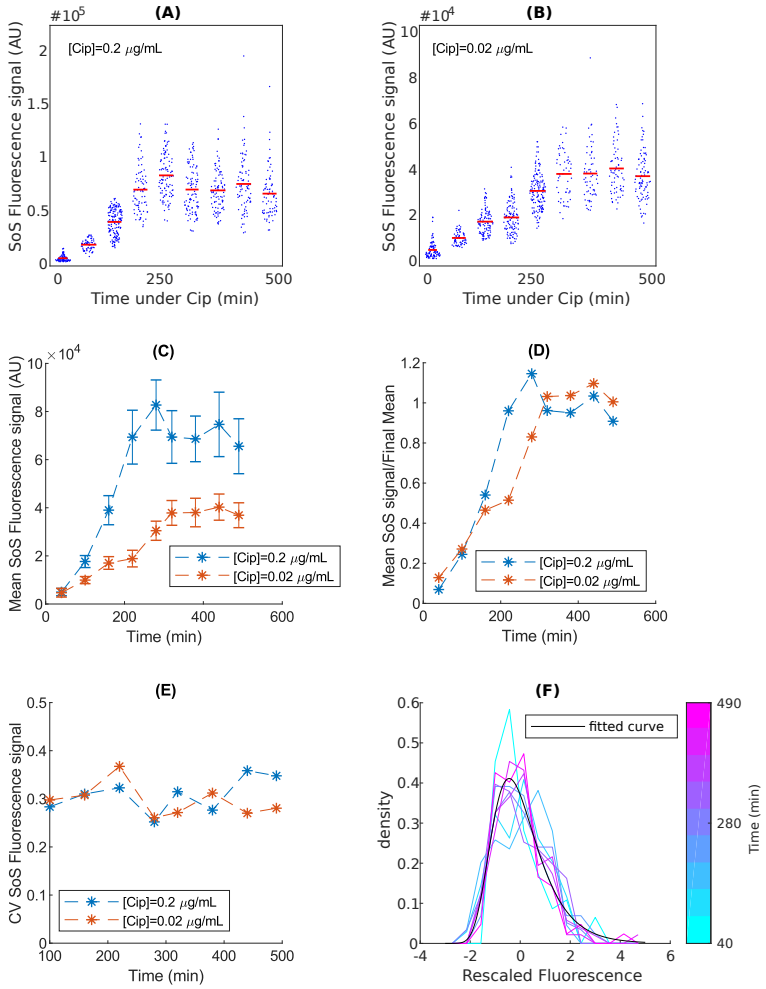


FIGURE 5.10: Fluorescence measurement of the SoS response of single cells. Two liquid cultures are submitted to an antibiotic stress (Ciprofloxacine) at two different concentrations, and sampled over time to measure the SoS response of single cells between a glass slide and a cover slip (A),(B) Fluorescence signal of the cells sampled at different times, with mean in red, for (A) [Cip] = 0.2 μg/mL (MIC), and (B) [Cip] = 0.02 μg/mL. (C) Comparison of the mean fluorescence measured over time at the two concentrations of Cip. (D) Mean fluorescence rescaled by the value of the final plateau. Error bars are $\pm$ standard deviation (SD) (E) Comparison of the coefficient of variation of the fluorescence, which is the ratio of the standard deviation to the mean, for the two concentrations. (F) Rescaled fluorescence distribution: (Fluo-Mean)/SD, as in [111], and best Frechet Fit, for the [Cip] = 0.2 μg/mL case.

SoS response of single cells

Before studying how the ability of cells to recover from an antibiotic stress is related to their individual SoS response, we need first to study this SoS

response at the single-cell level. For this purpose, we use classical measurements between a glass slide and a cover slip. We culture some *E. coli* to exponential phase ($OD_{600} \approx 0.2$), and then add Ciprofloxacin at two different concentrations, 0.2 μg/mL, which is the MIC, and 0.02 μg/mL. The cultures were incubated at 37°C. Then every hour, the cultures were sampled and the SoS response of individual cells was measured by thresholding the image in the green channel and measuring the mean fluorescence of each cell. The results are presented on Fig. 5.10.

We see that for both concentrations of the antibiotic, the SoS response is indeed induced, as the mean fluorescence signal of the cells increase over time (Fig. 5.10(A),(B)). This confirms in particular that it can be induced below the MIC. The dynamics of the induction of the SoS response for the two concentrations is similar in the sense that the mean fluorescence signal increases before reaching a final plateau (Fig. 5.10(C)). Of course, the value of this plateau is higher when the concentration of antibiotics is higher. However, if we re-scale the mean fluorescence by the value of this plateau, we notice that the dynamics are not exactly the same: the final value is reached more rapidly when the antibiotics concentration is higher (Fig. 5.10(D)). We can then deduce that the SoS response is induced more rapidly and more intensely when the antibiotic stress increases.

Until now, we have only used the properties of the mean fluorescence signal. But these measurements were made to start to study the heterogeneity of the SoS response. Let us then examine what the distribution of fluorescence signal looks like over time. A very interesting result is that although the variability of the measured fluorescence signal increases over time, it is constant when rescaled by its mean (Fig. 5.10(E)). These measurements indicate that the relative dispersity of the individual SoS response of the cells does not increase when the SoS response is induced. The shape of the distribution of the SoS responses of individual cells is rightly skewed, as expected for any protein, see [111] and Chapter 2 (Fig. 2.12). When we rescale it as in Chapter 2, by subtracting the mean and dividing by the standard deviation, we do obtain a collapse of the rescaled distributions at the different times (Fig. 5.10(F)), and the shape of these rescaled distribution is very well approached by a Frechet distribution [111]. The distribution of expression of the recN protein involved in SoS response that we measure here seems then to be similar to the one expected for other proteins.

Re-growth after SoS induction

Now that we have studied the distribution of the SoS response among single cells, we can try to see how it is correlated with the ability of a cell to survive the antibiotic stress and re-start to divide once the antibiotic is removed. For this, we use our microfluidic chip to take advantage of the ability to encapsulate individual bacteria in droplets.

The protocol is the following. We culture some bacteria to exponential phase in bulk and like previously, we add Ciprofloxacine at 0.2 μg/mL. After 3.5 hours of growth under antibiotic stress, which is the time needed for

the SoS induction to reach its final state, see Fig. 5.10(C), we remove the antibiotic by centrifugation and re-suspension in fresh LB medium, and load the bacteria in a chip with agarose. Then, we manually detect wells containing one single bacteria and take an image with a 30X magnification. About 30 wells could be imaged. We take one initial image and one image after 6 hours of incubation at 37°C.

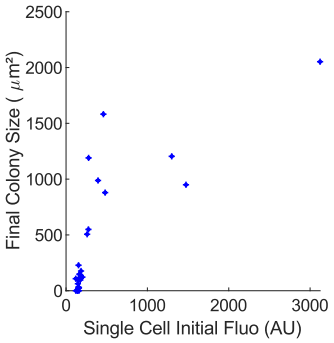


FIGURE 5.11: Final size of the colony grown from a single cell, after 6 hours of incubation at 37°C, as a function of the initial fluorescence intensity of this single cell, which is linked to its SoS response. Prior to loading, the cells were exposed to Cip at 0.2 μg/mL for 3 and a half hours to induce the SoS response. Then the antibiotic was removed and the cells were allowed to re-start growing in pure LB media.

The results are presented on Fig. 5.11. It is of course difficult to draw definitive conclusions because of the small amount of experimental points, but it seems from this experiment that the more intense the SoS response of the single-cell initially encapsulated in the droplet, the more chance it has to re-start growing and form a large colony. It would mean that a more intense SoS response actually means a more efficient repair of the DNA, and not that the DNA was more damaged by the antibiotic. Of course, controls would be needed, but those initial experiments gave us the will to push this approach further.

Limitations of these preliminary experiments

However, these experiments have a few limitations that we have to take into account if we want to get more definitive results. First of all, they were made in LB medium, which is not a very well controlled medium. In order to get more repeatable results, we decided to switch to MOPS rich defined medium [204].

Secondly, we only have one fluorescent reporter on this strain. It would be very useful to be able to measure growth separately from the SoS response, particularly if we want to obtain the dynamics of the recovery after an antibiotic treatment. For instance, if we want to know if the lag time or the growth rate of the colony growing from a cell that has survived the antibiotic treatment is correlated with its SoS response, we would need a second fluorescent reporter to quantify the growth. This would also allow us to get a

better measurement of the SoS response of single cells, as the thresholding of the fluorescence images could be made on this second fluorescent reporter. Indeed, in those preliminary experiments, as the detection of the cells and the measure of this SoS response was made with the same channel, bacteria that had a very low level of expression of the SoS response were not detected.

5.3.3 Pursuing experiments

A new strain

A new strain was built in the Mazel group, still engineered from *E. coli* MG1655. This time, two fluorescent reporters were integrated, directly on the bacterial chromosome. For the measurement of the SoS response, a PsfiA-mGFP integration plasmid was integrated at HK022 site of the E.coli MG1655 genome, following the protocol described in [205]. SfiA (or SulA) is the SOS-induced division inhibitor that depolymerises the FtsZ ring and thus triggers filamentous growth when cells are exposed to DNA damage [206]. It has already been used for single-cell studies of the SoS response [207], as it is known to be induced 100-fold during the SoS reponse [208].

The second fluorescent reporter (mCherry) was placed at the Lac site of the genome, and is inducible by addition of IPTG to the medium. It will allow us to measure the size and the growth of the colonies.

This strain, for the following experiments, is cultured in MOPS rich defined medium, supplemented with 0.4% glucose and IPTG at 1mM.

The MIC of this train in MOPS rich was measured to be at 16 ng/mL, with the broth dilution method (series of dilution in 1 mL liquid culture [163]).

Individual SoS response

We can then run the same experiment that we did with the previous strain, in paragraph 5.3.2, in order to measure the SoS response of single cells and its distribution. This time, the cells are detected by thresholding in the red channel .The measure of the SoS response is still done in the green channel, using the segmentation of the images made in the red channel. This segmentation is also used to measure the size of each cell, by taking the major axis length of each segmented region, as shown in Fig. 5.12. We still use the concentrations corresponding to the MIC and MIC/10, but this time the MIC is around 16 ng/mL. The results are presented on Fig. 5.13.

Let us compare these results with Fig. 5.10. For the lethal concentration, we see a increase in the SoS response of the cells (Fig. 5.13(A)). However, with this new strain, for an antibiotic concentration that is one tenth of the MIC, the increase in the SoS response is very low (Fig. 5.13(B)), at least for the mean response (Fig. 5.13(C)). But at his sub-MIC concentration, we notice the presence of a few cells with a very high SoS response in the population, that we did not see that clearly with the previous strain.

The relative dispersion is once again almost constant over time, for both concentrations (Fig. 5.13(D)). It is higher than what we had with the previous strain and cell detection method. This was expected, as we are now detecting

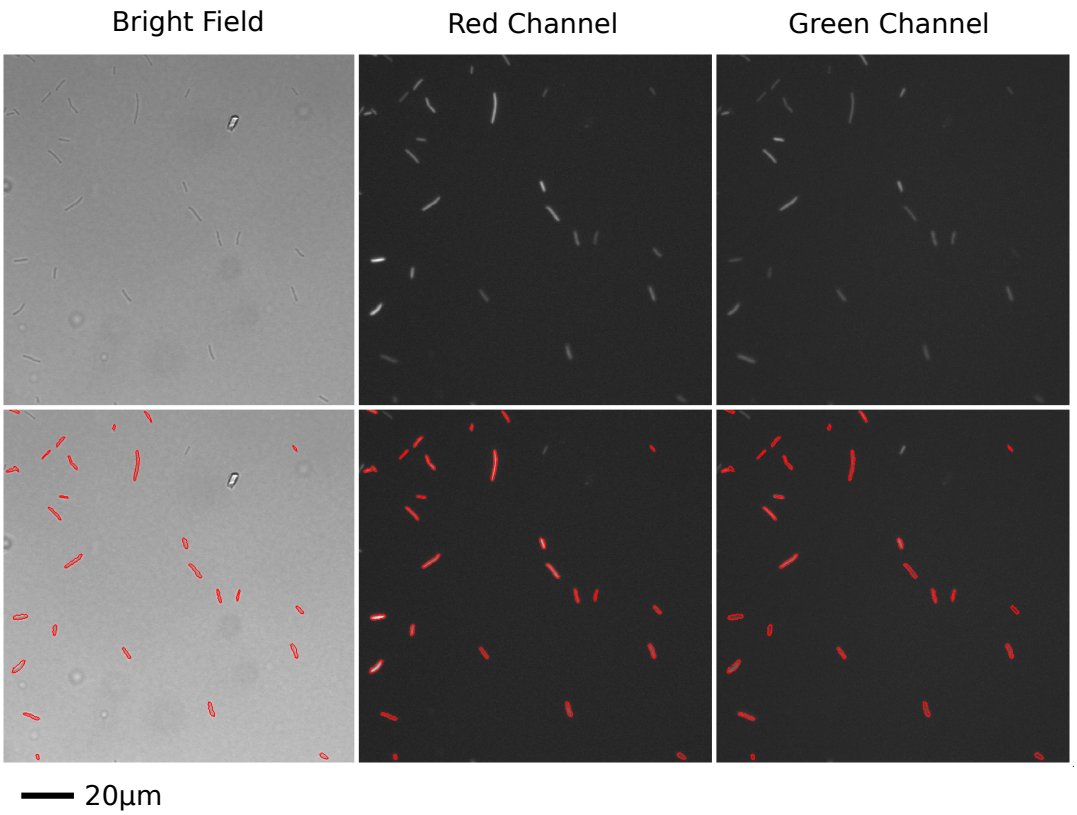


FIGURE 5.12: Example images of the measurement of the SoS responses of single cells, taken here after 1 hour exposure to Cip at 16ng/mL, at 37 °C. Top row: raw images. Bottom row: segmented images. The segmentation is done by thresholding the images in the red channel, with the bacteria expressing a mCherry fluorescent reporter induced by the addition of IPTG to the culture medium.

cells using another fluorescent channel. Cells with a very low SoS induction can now be detected, and the variability is globally increased.

Finally, we see that the distribution of the SoS response among the cells, once rescaled by subtracting the mean and dividing by the standard deviation, collapse once again onto a constant distribution that is rightly skewed and corresponds quite well to a Frechet fit, as described in ref [111].

With this experiment, we also have access to the distribution of size of the cells in the stressed population. For the lethal concentration, we see that the size increases with time (Fig. 5.14(A)). This was expected, as the SoS response is known to inhibit cellular division and causes the appearance of long filamentous cells [193]. However, if we compare the dynamics of the mean size increase (Fig. 5.14(C)) with the mean increase of the SoS fluorescent reporter Fig. 5.13(C)), we notice that the size increase is slower. There is therefore a delay between the induction of the SoS response and its effect on the mean size of the cells.

For the sub-MIC concentration tested here, the increase in size is almost null (Fig. 5.14(B)), which makes sense as the SoS response is not really induced in this case. But just as for the SoS response distribution, we notice

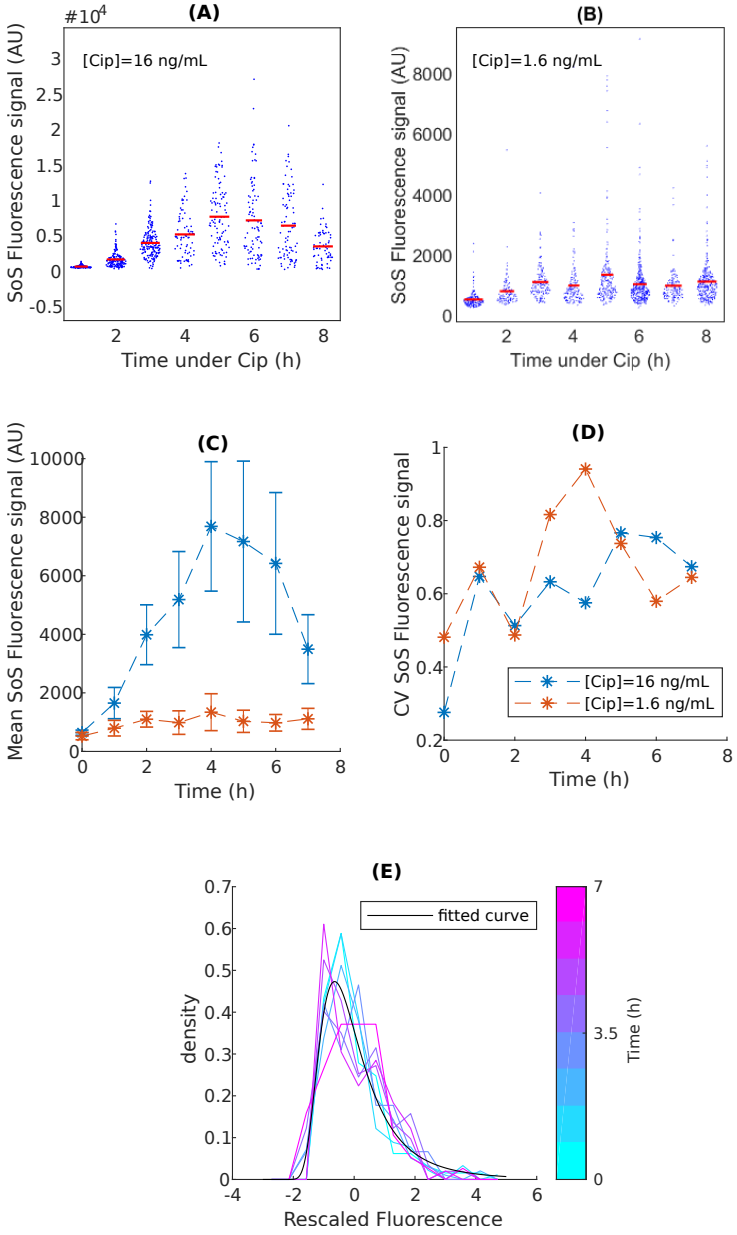


FIGURE 5.13: Fluorescence measurement of the SoS response of single cells, with the new strain. Two liquid cultures are submitted to an antibiotic stress (Ciprofloxacine) at two different concentrations, and sampled over time to measure the SoS response of single cells between a glass slide and a cover slip (A),(B) Fluorescence of the cells sampled at different times, with mean in red, for (A) [Cip] = 16 ng/mL (MIC), and (B) [Cip] = 1.6 ng/mL. (C) Comparison of the mean fluorescence measured over time at the two concentrations of Cip. (D) Comparison of the coefficient of variation of the fluorescence, which is the ratio of the standard deviation to the mean, for the two concentrations. (E) Rescaled fluorescence distribution: (Fluo-Mean)/SD, as in [111], and best Frechet Fit, for the [Cip] = 16 ng/mL (MIC) case.

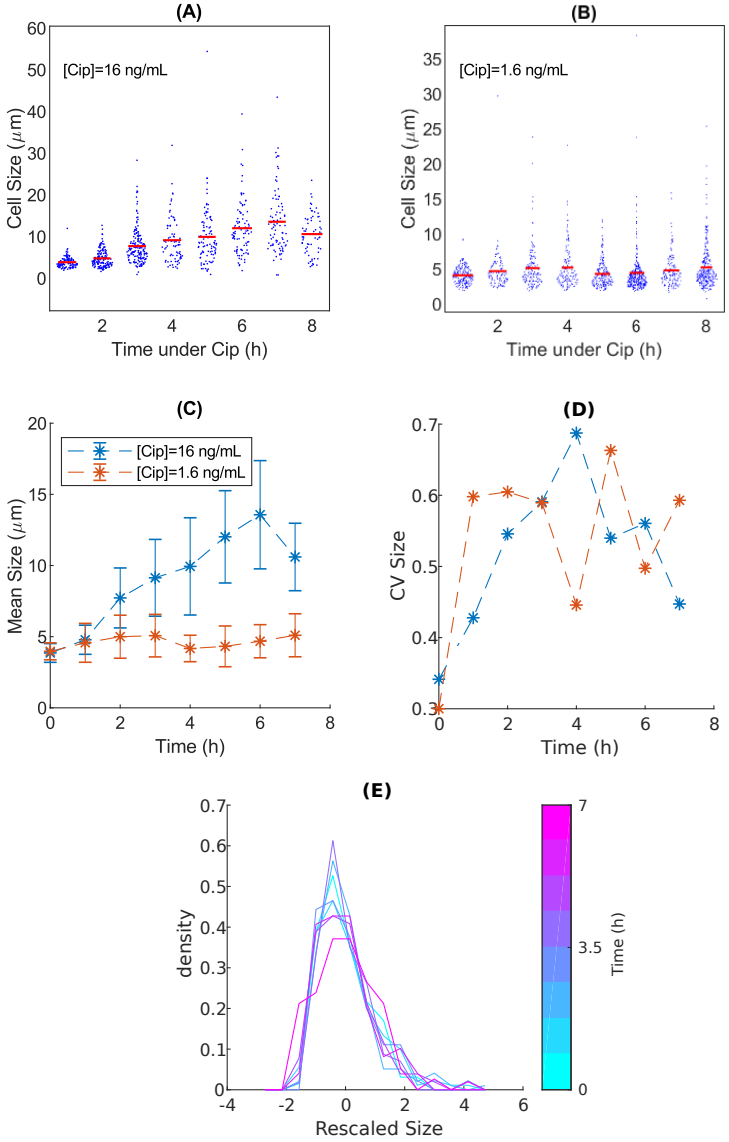


FIGURE 5.14: Size of the cells under antibiotic stress. Two liquid cultures are submitted to an antibiotic stress (Ciprofloxacine) at two different concentrations, and sampled over time. (A),(B) Size of the cells sampled at different times, with mean in red, for (A) [Cip] = 16 ng/mL (MIC), and (B) [Cip] = 1.6 ng/mL. (C) Comparison of the mean size measured over time at the two concentrations of Cip. (D) Comparison of the coefficient of variation of the size of the cells. (E) Rescaled size distribution: (Size-Mean)/SD, for the [Cip] = 16 ng/mL (MIC) case.

that there are some cells that are much longer than the others : the right hand queue of the distribution is quite large.

The relative variability of the sizes is once against almost constant with time (Fig. 5.14(D)), and if we rescaled the distribution of sizes as we have already explained, we get once again a collapse of the distributions at different times (Fig. 5.14(E)). The distribution of sizes is rightly skewed, just like the

distribution of the SoS response.

There is then a lot of similarities between the SoS response distribution and the size distribution. The question that arises is to know whether these quantities are correlated at the single-cell level. The answer, singularly, seems to be negative: the correlation coefficient between the size and the intensity of the fluorescence of the cells in green is quite low (Fig. 5.15(A),(C)). We see here an interesting case of correlation at the scale of the population but not at the single-cell level, at least not in our measurements. Other authors have noticed this increase in mean cell size when the SoS response is induced without any clear correlation at the single-cell level [207].

Even if the correlation is not true for all cells, we have noticed, both for the size and the SoS response, that the right-side queue of the distribution was quite long, especially for the sub-MIC case. We can then wonder if those cells are the same: are the very long cells also the very bright cells ? To answer this question, we need to select "long" or "bright" cells. We do so by selecting the cells that have a size (respectively a fluorescence) that is higher than the sum of the mean and the standard deviation of the sizes (respectively fluorescence). It seems then that only $\approx$ 20% of the "very long" cells are also "very bright" (Fig. 5.15(B)). The answer is then once again negative: the very long cells are not the ones where the SoS response is induced the most. Looking at the percentage of "very long" cells among the "very bright" cells yields similar results.

However, we have to be aware that something is lacking in those measurements to be able to draw definitive conclusions, which is the ability to follow single cells over time. With this experiments, we only have samples of the culture at every time, but the dynamics of the single-cell SoS induction and size increases are completely missed. For instance, the very bright cells at the beginning could become the very long cells a few hours later because of the effect of sulA. The dynamics of the induction of the SoS response are known to be complex and very precise at the single-cell level [201]. This could explain the apparently counter-intuitive results that we get for the correlation between the size and SoS response of the cells.

Cell Recovery after antibiotic treatment

Let us come back now on the question of the recovery after treatment. We would like to see if the preliminary results that we got (see Fig. 5.11) are reproducible. We would also like to see if the dynamics of the recovery, such as the lag time or the growth rate of the colony emerging from a cell that has survived are correlated with the initial SoS response of the cell.

Unfortunately, with the new strain, those measurements turned out to be much more difficult. Indeed, the individual fluorescence of the cells was on average much lower than with the previous strain, which makes them impossible to detect with a 30X magnification in the wells. To overcome this challenge, we had to switch to a bigger magnification, to enhance the fluorescence signal. The problem is that it meant switching to a 60X oil objective, that has a much lower working distance. Therefore, to use this objective, the

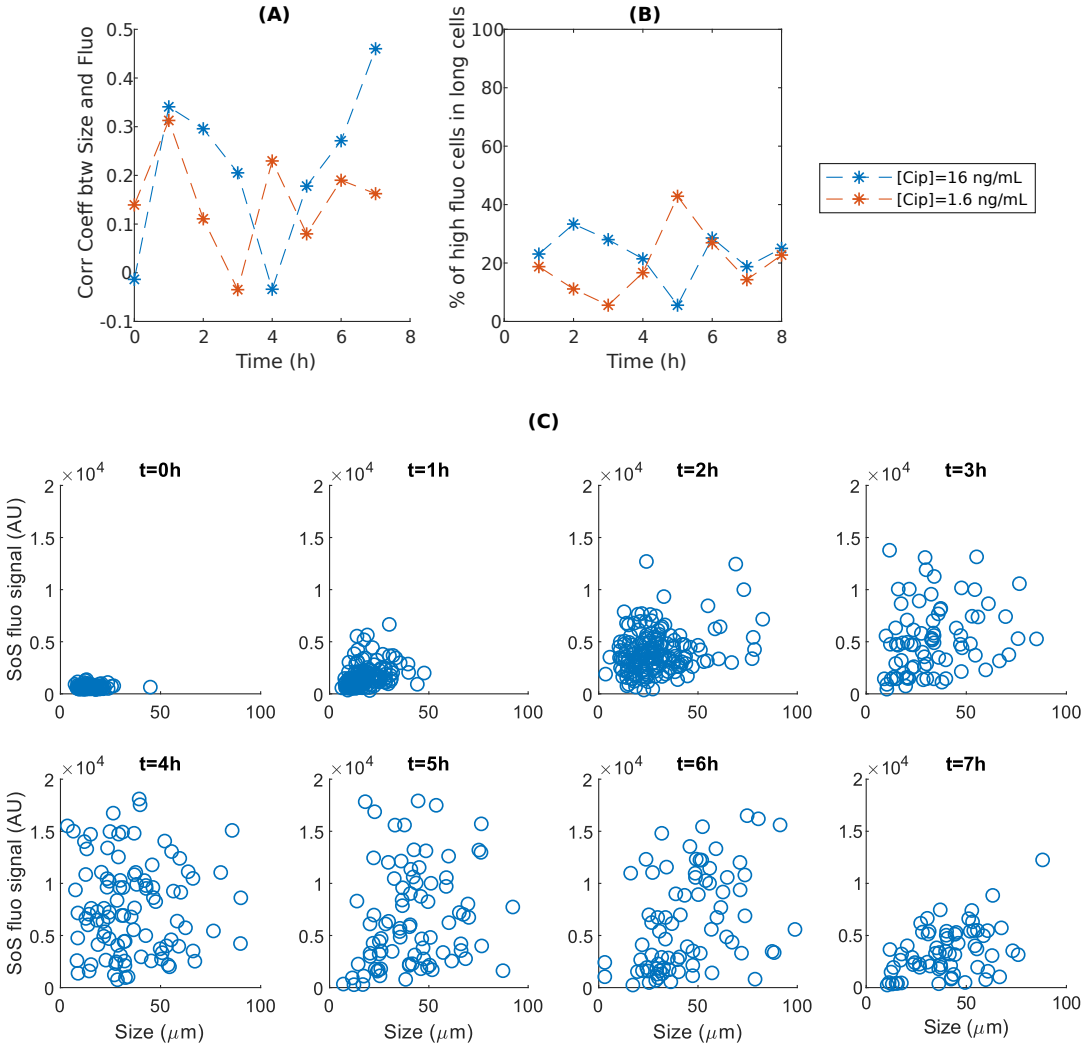


FIGURE 5.15: Correlation between size and fluorescence of the single cells. (A) Correlation coefficient between the sizes and the SoS fluorescence intensity of the single-cells over time. (B) Percentage of "very bright" cells among the "very long" cells. Cells are supposed to be very bright when their fluorescence is superior to the sum of mean fluorescence plus the standard deviation, and the same criterion is applied for the size. (C) Scatter plot of the individual cell SoS fluorescence signal as a function of their size, for the experiment with [Cip]=16 ng/mL, for the different observation points.

PDMS part of the chip has to be attached onto a 100 μm cover slip, because the 1 mm glass slides that we were usually using are too thick.

Moreover, if we want to record images of the growth of the colony at 37°C, we have to put the chip in a water bath to prevent the evaporation of the droplets. This was usually made by placing the chip in a petri dish filled with water, but with a 60X objective the thickness of the petri dish makes this method unusable. We solved this issue by laser cutting a rectangular hole in the bottom of the petri dish. Then, the PDMS chip, attached on the cover slip, was bonded to the bottom of the petri dish, from below, by UV-activated glue

that was deposed on the coverslip, see Fig. 5.16. Hence, the petri dish could be filled with water, and the distance between the objective and the chip was reduced enough to use the 60X oil objective.

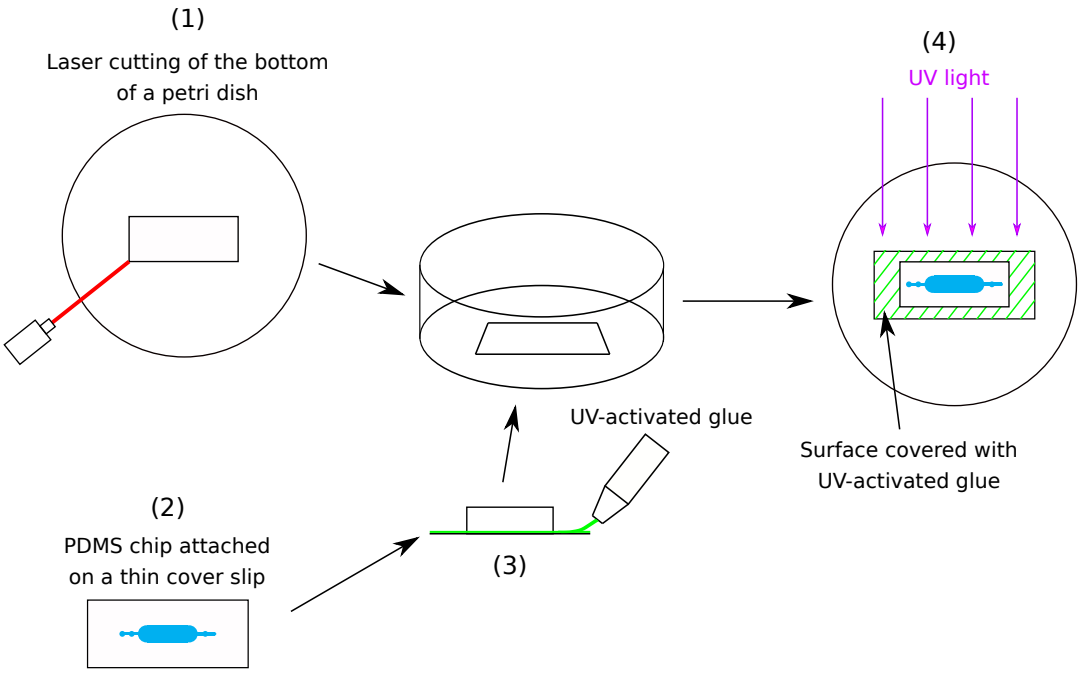


FIGURE 5.16: Protocol to make the device for the recovery experiment with a 60X objective. (1) A hole is cut with a laser in the bottom of a petri dish. (2) A chip is attached onto a thin cover slip, which is then covered with UV-activated glue on the sides of the PDMS and (3) put from below in the hole of the petri dish (4) UV light is shined on the whole system to stick the cover slip to the bottom of the petri dish, making the whole system impermeable when the petri dish is filled with water. As the coverslip was sticked below the petri dish, it is accessible for direct wetting of the optical oil necessary to grab images with a 60X objective.

Thanks to this experimental improvement, we could take images of the growth in the chip with a 60X objective. We repeated then the experiment described in section 5.3.2: we cultured some *E. coli* to exponential phase, and then exposed them to Cipro at 5 ng/mL for three hours, before washing the antibiotic and recording the growth. Not all of the bacteria were growing into colonies after the treatment (see Fig. 5.17(A)). In this experiment, 19 of the 34 monitored droplets exhibited bacterial growth. The growth after the treatment was exponential (Fig. 5.17(B)).

However, it turns out to be difficult to find any correlations between the growth parameters and the SoS response of the mother cell after the antibiotic treatment. Neither the growth rate of the colony growing from this mother cell (Fig. 5.18(A)), nor its lag time (Fig. 5.18(B)) seem to be correlated with the initial green fluorescent signal of the mother cell, used as a proxy for the SoS response. Even for the final size of the colony, that we have previously found to be correlated in section 5.3.2, no correlation is found with this new

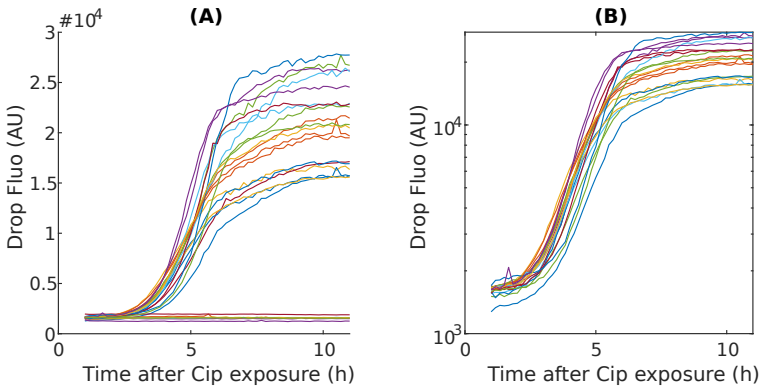


FIGURE 5.17: Growth of the colonies after antibiotic treatment, recorded with a 60X objective. The growth is measured using the red fluorescence channel. (A) Fluorescence signal of the drops over time. (B) Same, but in log scale for drops with growing cells only

bacterial strain (Fig. 5.18(C)). The correlations found previously could not be confirmed.

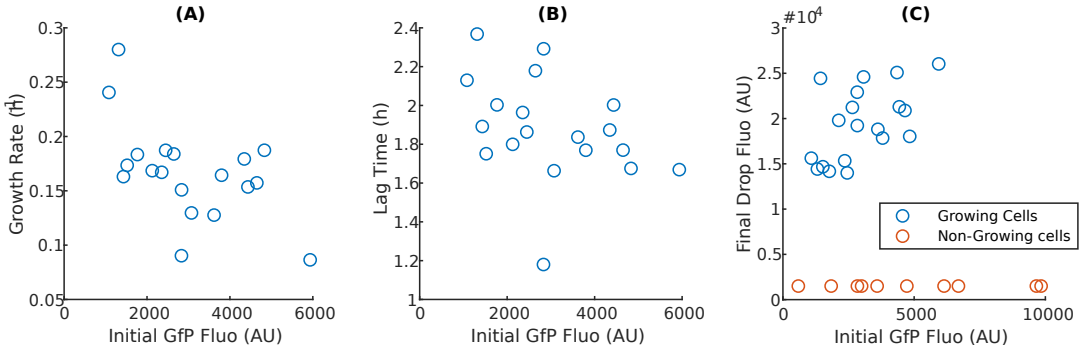


FIGURE 5.18: Correlation between the growth parameters and the initial GfP fluorescence of the cell after antibiotic treatment, which is related to the intensity of the SoS response of the cell. (A) Growth rate as a function of the initial GfP fluorescence. (B) Lag Time as a function of the initial GfP fluorescence. (C) Final fluorescence signal of the drop (in the red channel) as a function of the initial GfP fluorescence of the cell.

To see if some correlations were appearing, we tried to increase the concentration of antibiotic. However, the number of surviving and growing bacteria after the antibiotic treatment was then very low (for 10ng/mL, only 2 bacteria out of 30 were surviving). The number of bacteria that we can follow in this experiment is very limited because we have to manually detect them, and this is a very tedious task because their level of individual fluorescence is low. Moreover, as the focal depth is low with the 60X objective, it is frequent not to detect of a second bacteria is in the same droplet, but in another focal plane, and then the growth parameters of the two colonies interfere and the data have to be thrown away. For the same reason, bacteria can grow out of the thin focal plane even if they grow in agarose, and it is hard to make

sure that the whole colony is taken into account in the final measurement, which could also be a reason for the lack of correlation. Overall, this experiment may have to be re-thought to yield better results. For instance, bacteria could be loaded on the chip together with the antibiotic. The fluorescence signal of single bacteria could thus be monitored over time in the agarose droplets during the antibiotic treatment. The antibiotic would afterwards be directly removed in the chip through a phase change, and the monitoring of the bacterial recovery would follow. By doing this, we could see if the individual increase in fluorescent signal has to be taken into account instead of the raw value at the end of the antibiotic treatment. A correlation between this increase in fluorescent signal and the cell survival and recovery to the antibiotic treatment may be found. The traps could also be made slightly smaller to reduce the microscopy-related problems.

5.4 Summary and Conclusion

In this last Chapter, we have shown how our chip can be used to test the response of bacteria to an antibiotic stress. First of all, we have discussed how the minimal inhibitory concentration (MIC) of an antibiotic can be found by loading several chips in parallel with the same bacterial inoculum but different concentrations of antibiotics. This methods represents a gain of space and time when compared to classical MIC determination methods, such as plating or the broth dilution method. Indeed, the chip measurement yields a final result faster than those classical methods thanks to the low volumes that are used. However, we have to increase the concentration of bacteria with respect to what was used in the previous chapters in order to respect the standardized inoculum size for MIC testing [162].

These experiments where several chips are loaded in parallel with different concentrations of antibiotics can also be used to infer the chance for a bacteria to grow despite the antibiotic stress. This can be made using basic stochastic calculus and a digital analysis, which means counting the number of droplets where bacteria have grown despite the antibiotics. This digital measurement is accurate only for a finite range of bacterial concentration and for a finite range of probability to survive the antibiotic stress. This limitation have been illustrated with an example inspired by the literature [168].

Thanks to the ability offered by our microfluidic technology to control the environment of the bacteria over time,the time-kill curves of an antibiotic with our chip can also be obtained. Once again a digital analysis is used to infer the fraction of viable cells as a function of the exposure time to the antibiotics from the chips measurements. Another great experimental possibility offered by our chip is to perfuse the droplet array with a concentration gradient of any compound, each row of the droplet array being submitted to a different concentration. We have also suggested how this possibility could be used in the future in the lab to confirm some theoretical and numerical results of the literature [179] about the dynamical effects of antibiotics on the growth rate of bacteria.

Finally, some preliminary experiments to study the SoS response of bacteria in our chip have been conducted, in collaboration with the Mazel lab in Institut Pasteur. We have shown that our chip could be used to study the recovery of individual bacteria to an antibiotic treatment and correlate it with the individual SoS response of the bacteria. Experimental improvements were made, both on the microbiology side in the Mazel group and on the microfluidics and microscopy side in our group. The preliminary results have yet to be confirmed, but this project has led to obtain a common grant to continue the collaboration between the two labs on this subject.

Chapter 6

Conclusion

Mais oui mais ouiiii, l'école est finie !
Sheila

It is now the time to conclude this thesis. We will try first to give an overview of what has been done in this PhD. Then we will propose some ideas of future research in the group, and finally expose personal perspectives on some problems that have been discussed throughout this work.

Let us begin with the advances that have been completed during this PhD. My work in the lab begun a few months after the microfluidic platform used throughout this whole project had begun to be used to cultivate bacteria. The design was already in its final state and some data had already been obtained for the parallel growth curves of bacteria in the droplets [97]. On the pure technological microfluidic side, the work had in some way already been done before I begun my PhD, and no major improvements were still needed. My personal contribution was to precisely quantify and discuss some aspects of this technology that are important to conduct microbiology experiments in the chip. These aspects were a little bit left aside in the previous works of the group as the focus was to develop the microfluidic chip, which is not simple, and to establish a proof-of-concept of its efficiency. In particular, we have seen in Chapter 2 that the relation between the measured droplet fluorescence and the number of cells per droplet was deciphered, leading to new questions on how the growth of bacteria in droplets could be compared with batch growth in liquid flasks. The image analysis aspect was also one of the constant investigation themes in this work, notably the background noise question that is evoked in Chapter 2. The question of the volume heterogeneity of the droplets was also investigated both numerically and theoretically. Its consequence on the distribution of the number of bacteria per droplet was hence modeled. Last but not least, the question of the importance of the cell-to-cell fluorescence heterogeneity of the bacteria for the droplet measurement was also treated quantitatively and related to single-cell experiments. As a whole, even if they do not represent ground breaking scientific advances, it

was important for the group to treat these technical questions, in order to be able to serenely push further the microbiology experiments in droplets.

Nevertheless, the core of this thesis was the modeling of the variability of the growth curves obtained in the microfluidic droplets. As we have already mentioned, some growth curves had already been obtained in the group before, but their measurement was systematized and improved in terms of reproducibility and time-resolution. The variability in terms of population size between the droplets was a mystery, and it was resolved during this PhD. We have shown that using a pre-existing mathematical model encompassing the cell-to-cell variability of division times, and adding to this stochastic model other sources of randomness due to our experimental system, the variability between the growth curves could be quantitatively explained, at least during the exponential phase of the growth. Moreover, not only the variability, but the shape of the distribution of number of cells per droplet could be predicted. Our extended stochastic growth model could be used to model the growth of bacteria in micro-droplets, which is itself a booming field [55]. As a phenomenological model, it can also be used as a basis for any situation in which the growth of a quantity is exponential and stochastic. There are countless examples, from the spread of epidemics to the growth of cancerous tissues [85]. We have also shown numerically that the models of single-cell division discussed in the literature yield similar results at the scale of relatively small populations of bacteria microfluidic in droplets, which could interest both the microfluidcs and microbiology communities working with these models. It is indeed a question of intense current research to try to bridge the gap between stochastic individual division and population dynamics [82, 86].

The initial objective of this study, however, was to go backwards in this scheme: if the droplet-to-droplet variability is related to the cell-to-cell variability, can we quantify the cell-to-cell variability from the droplet measurements ? It turns out to be difficult, because the variability among droplets in population size is dominated by the other sources of stochasticity that we have evoked. The equations of our stochastic model relating the cell-to-cell variability to the population variability in exponential phase are then not usable for inference. We have also tried some advanced mathematical methods existing in the literature, based on pseudo-likelihood method [92], that matched our experimental observation scheme, but they failed for the same reason.

To overcome these difficulties, and try by other means to retrieve the cell-to-cell variability of exponential times from the growth measurement in droplets, we developed a *de novo* inference scheme based on following each individual droplet in time. In fact, the growth trajectory of each droplet deviates from a pure exponential growth because of the stochasticity of division times at the single-cell level. We quantified this deviation by computing quantities that we called the residuals of the trajectory, and showed that the variance of these residuals, if properly measured, can be used for inferring the variability of division times at the single cell scale. This inference model

proved to be efficient despite our experimental constrains such as the initial Poisson distribution for the number of cells per droplet: we validated it on simulations of bacterial growth mimicking these constrains. Unfortunately, we could not get it to work on our experiments because of the uncertainty that we have on the relation between fluorescence and number of cells. Notwithstanding, as we have already said, this new inference method could still be used by other groups that have different measurement methods for the growth of bacteria, or even for populations of cells for which a direct measurement of the number of cells could be available, such as yeast.

Last but not least, some experiments with antibiotics were conducted during the course of this PhD. It was the beginning of microbiology in the lab, so there was a lot to learn, particularly on the methodology side. A lot of measurements were therefore done without systematic protocols, particularly at the beginning of the thesis.

Having our microfluidic chip is in a sense both a advantage and a drawback. An advantage, because it allows us to collect a lot of data in a brand new way, so previously unknown data about bacteria and antibiotic can be produced. However, as it is a new way to look at those interactions, it is sometimes hard to compare those results with the literature. Nevertheless, we have proved that using our chip could result in a gain of space and time for testing the minimal inhibitory concentration of a strain, for instance. We have also shown how to compare our digital measurements with direct plating measurements of the number of viable cells, for instance, showing that if the correct range of dilution was chosen, using the chip could result once again in a gain of space and time.

The chip was also used to study the SoS response of bacteria, which is a broad bacterial stress response induced when the DNA of the cell is damaged. A fruitful collaboration between our group and Didier Mazel's group, which has been studying these questions at Institut Pasteur for a long time, was started. The chip could be used to encapsulate single cells, monitor their stress response and see how this influences the colony that can grow, or not, from this single cell once the antibiotic stress was removed. While some preliminary results seemed to indicate that there indeed was a correlation between the SoS response of the mother cell, measured at the end of the antibiotic treatment, and the size of the colony growing from this cell, more systematic experiments done at the end of the PhD could not confirm this correlation. But along the way, new strains have been engineered on the microbiology side and new experimental techniques to observe them on the microfluidic side, and this project is on-going in the lab.

This brings us to the future projects that could be made in continuation of this PhD. We have already mentioned in Chapter 5 how the chip could be used to confirm numerical results [179] about the dynamical growth of response of bacteria submitted to a sub-lethal antibiotic stress. The chip could also be used to study other stress responses than the SoS response, based on the same method (encapsulation of individual cells and correlation of its stress response to the growth parameters of the colony growing from this stress response). New strains would need to be engineered, but for instance

the general stress response of the bacteria, mediated by *RpoS*, could hence be studied [209].

In fact, other interesting works have been conducted during this PhD, but did not make it to the final contents of the thesis because they are still at the stage of preliminary results and would need continuation. In particular, a collaboration with the lab of Guy-Franck Richard of Institut Pasteur has been started in order to study yeast cells with the microfluidic chip. The aim is to characterize the dynamics of DNA double strand breaks repair in yeast cells. The cells are engineered so that they become fluorescent when they have repaired their DNA, as shown in an example time-lapse image in Fig. 6.1. The characteristics of the DNA repair already have been studied at a population level [210], and the objective was to study it at the cell level in the chip. The growth in itself in microfluidic droplets turned out to be much more variable than what was expected, with some cells that randomly stop growing. New questions could thus emerge from these preliminary results.

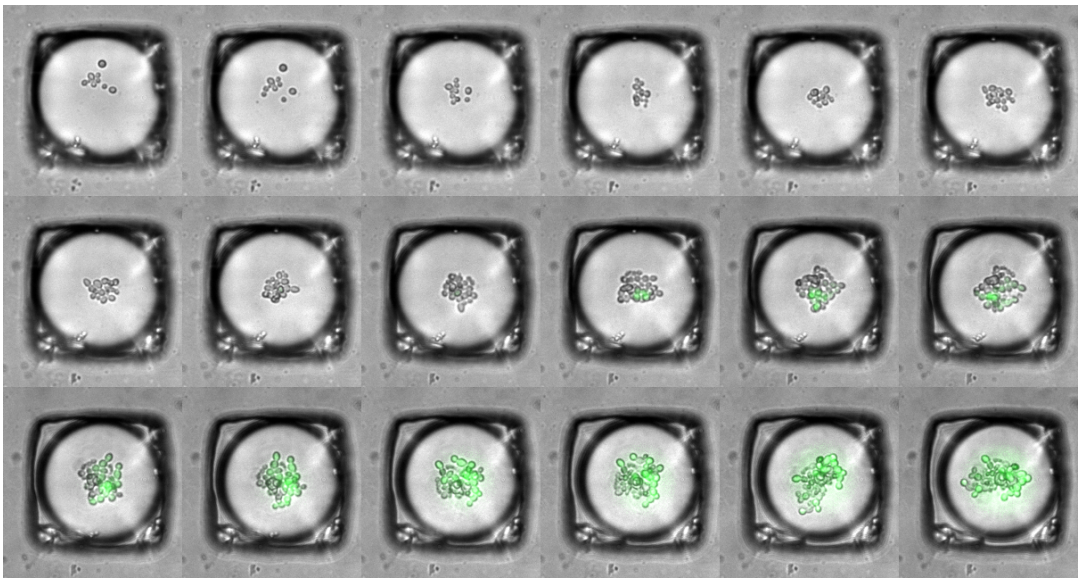


FIGURE 6.1: *Saccharomyces cerevisiae* yeast cells dividing and repairing their DNA in a microfluidic droplet. One image every 2 hours for 34 hours.

This yeast cell example also illustrates how the skills acquired during this PhD by studying bacteria in micro-droplets were easily transferable to other biological subjects. The image analysis techniques and the statistical analysis of the data developed initially for the growth of bacteria turned out to be useful, for instance, for the analysis of the transfection of mammalian cells cells with liposomes [211]. They were also useful to characterize the link between spatial organization and regulation of important proteins in 3D human cells aggregates [212] (submitted to Science Advances).

Let us finish by sharing some personal thoughts about some subjects treated in this PhD. First, about the use of microfluidics. The approach developed in this PhD was to use microfluidics as a tool to increase the statistics of classical microbiology experiments. Hence, data about the variability of

the very classical exponential growth of bacteria in a well defined medium could be obtained, and inference methods developed. However, I think that microfluidics can be used in a different and maybe more fruitful way: it can be used to mimic more complicated micro-environments than well-stirred lab-defined culture medium. This kind of work is for instance developed in the lab of Robert Austin. They have shown that wild-type *E. coli* could evolve resistance to Ciprofloxacin in a very short time (a few hours), provided that they are cultivated in a microfluidic system called the "Death Galaxy" [214]. This system consists of poorly connected micro-chambers. On both sides of the micro-chamber array, culture medium flows, with pure culture medium on one side an antibiotics on the other. This medium can diffuse slowly in the array through nanoslits that are 100nm deep, see Fig. 6.2. It mimics how small ecological niches, in nature, can evolve resistance much more rapidly

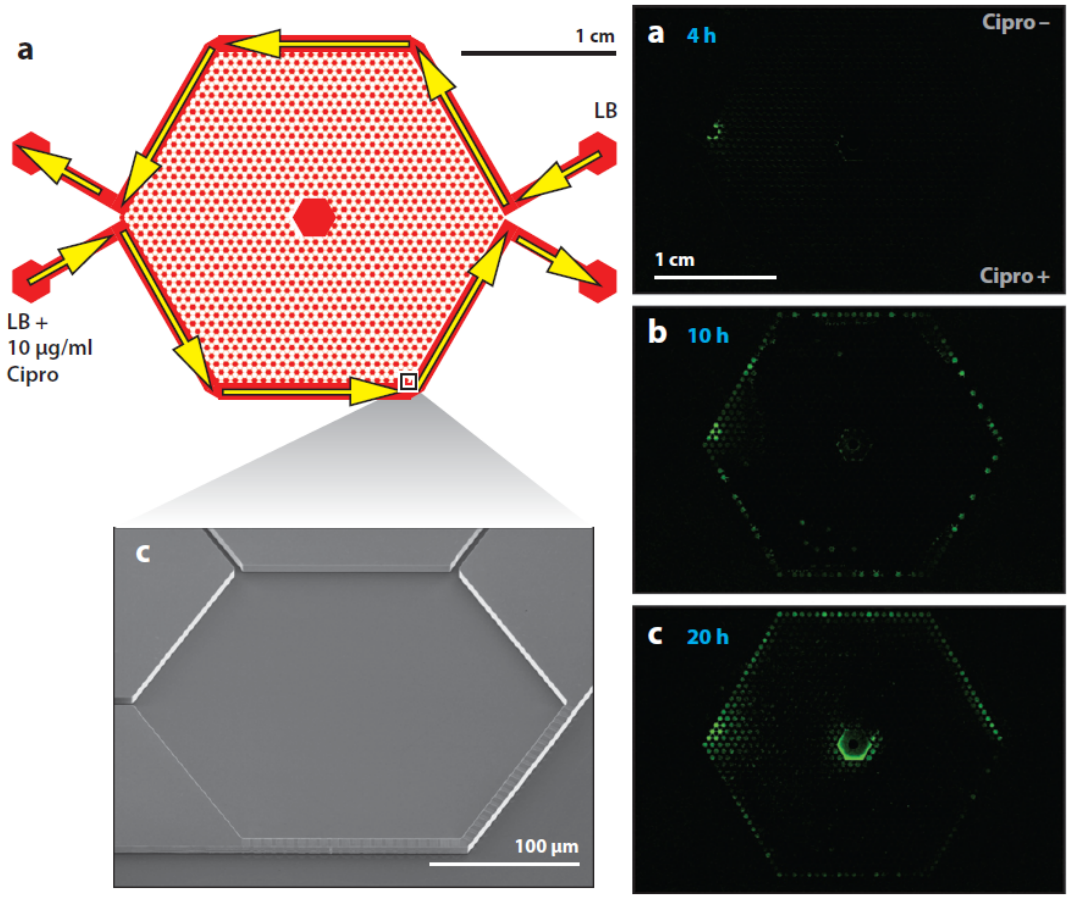


FIGURE 6.2: Taken and adapted from [213]. The "death galaxy" microfluidic system developed in Austin's lab. Left Column: design of the system. Small chambers (10 microns height), in red, are connected by small channels that have the same height but are only 10 microns wide. Bottom: Scanning electron microscope image of the array. Right Column: apparition of the resistance to Ciprofloxacine after only 4 hours in the device, and subsequent spreading in the system (the green color indicates the presence of bacteria thanks to a fluorescent reporter).

then a globally mixed population. It is due to the fact that the fitness landscape is explored very differently. Very interesting concepts about evolution are therefore illustrated by this experiment [213]. Microfluidics are here used to test concepts rather than increasing statistics.

This is, to my opinion, a direction that would be interesting for the use of microfluidics for fundamental research. But microfluidics are not only used for this purpose. These technologies are also more and more transferred and becoming available for clinical use. With a market that is thought to be valued at more than 25 billion dollars in 2025 [215], it is a booming sector. Unfortunately, it seems that the goal of the vast majority of people working on getting microfluidic technologies out of the research labs is to make their share in this big pie by patenting and commercializing these technologies at a high cost. Multiple examples can be found, such as the now multi-million valued microfluidic start-up Fluidigm. I personally think that this capitalistic approach is a bit of a shame. Other approaches could be much more valuable for everyone but financial investors. Open-source, community-driven approaches exist though, such as Metafluidic [216].

Another subject that we have evoked in this PhD was the antibiotic crisis: there is a very fast and global spread of antibiotic resistance [147]. we need to reduce our misuse of antibiotics, because the more we use them, the more new resistant strains emerge. That's why new susceptibility tests, based for instance on microfluidic technologies such as ours, could be of great help. But if we take another look, the biggest misuse of antibiotics does not come from human therapeutic use. More than 50% of the antibiotics used in the US are used in agriculture [217], and not even to treat animals but mostly as growth promoting factors. This is very dangerous, as these antibiotics spread in the environment, and resistant mutants can emerge, that can be harmful to humans as well [218]. To reduce these risks, as classical country farming cannot provide an exponentially growing human population with enough animal products, we have no choice but to reduce our dependence to animal products in our diet. This would, moreover, also have a positive impact on climate change, as we recall that agriculture alone is responsible for $\approx$ 15% of the global emissions of greenhouse gases [219], while switching to a less animal-dependent diet could reduce these emissions by as much as 80% [220].

Appendix A

Materials and Methods

A.1 Bacterial strains and media

Chapters 2, 3 and 4

Two strains were used in these chapters, one *Escherichia coli* (referred to as *E. coli*) and one *Bacillus subtilis*.

The strain *E. coli* AV04 is a gift from Antoine Vigouroux at Institut Pasteur (Paris, France) and constitutively expresses an mCherry fluorescent reporter [126].

Bacillus subtilis strain GM2938, a kind gift from Dominique Le Coq (Micalis Institute, INRA, Université Paris-Saclay), has an mCherry fluorescent reporter integrated by double crossing in amyE, and a resistance gene for spectinomycin [amyE::PhyperSpank-mCherry "Rudner", lacI/SpecR]. The fluorescence is induced by the addition of IPTG (SIGMA-ALDRICH, St. Louis, Missouri). All growth media used for this strain contain IPTG at a concentration of 1 mM.

Bacteria are grown in rich LB medium (SIGMA-ALDRICH L3022), supplemented with IPTG at a concentration of 1 mM in the case of *B. subtilis*. For all experiments, bacteria are grown overnight in fresh liquid medium from an LB-agar plate, then re-diluted at 1/200 in fresh medium until they reach an $OD_{600} \sim 0.2$ (Amersham, Little Chalfont, UK). Bacteria are then used for the on-chip or single cell experiments. For *B. subtilis*, we also added spectinomycin (SIGMA) at 0.1 mg/ml in the overnight culture to avoid contaminations. All culturing steps were conducted at 37°C.

Chapter 5

Two other *E. coli* strains were used in this chapter.

The first one is a nice gift from Zeynep Baharoglu at Institut Pasteur (Paris, France). It was engineered from *E. coli* MG1655, with a GfP fluorescent reporter for SoS response (GfP under the control of recN, see complete description in Ref. [6]). It was used for all experiments with Gentamicin and the preliminary experiments of SoS response. Those were conducted in LB medium.

The second one was still engineered from *E. coli* MG1655 in the group of Didier Mazel, but by Julia Bos. This time, two fluorescent reporters were integrated, directly on the bacterial chromosome. For the measurement of the SoS response, a PsfiA-mGFP integration plasmid was integrated at HK022

site of the E.coli MG1655 genome, following the protocol described in [205]. SfiA (or SulA) is the SOS-induced division inhibitor that depolymerises the FtsZ ring and thus triggers filamentous growth when cells are exposed to DNA damage [206]. It has already been used for single-cell studies of the SoS response [207], as it is known to be induced 100-fold during the SoS reponse [208]. The second fluorescent reporter (mCherry) was placed at the Lac site of the genome, and is inducible by addition of IPTG to the medium. It will allow us to measure the size and the growth of the colonies.

This second strain was used for the later experiments about the SoS response and the comparison between tube and chip MIC testing with Ciprofloxacine. These experiments were made using MOPS rich defined medium, supplemented with 0.4% glucose [204].

A.2 Microscopy

For all experiments, the chips were placed on the motorized stage of an epi-fluorescence Nikon-Ti microscope, with a temperature control at 37°C. Fluorescence images of the droplets are taken at a 10X magnification with an EMCCD Camera (Andor Technologies, Belfast, UK). About 800 droplets can be scanned in 5 minutes. The resolution is 1.6μm/pixel.

For binary counting of the droplets in the MIC testing experiments of Chapter 5, instead of using an epi-fluorescence motorized microscope, we can use a simpler and faster slide scanner (SensoSpot®-Fluorescence, Sensovation AG, Germany). We can obtain a low-resolution image of the chip in a few minutes, at a resolution of 6.5μm/pixel, and up to 4 chips can be scanned at the same time.

Appendix B

Mathematical complements

B.1 Equivalence

Throughout this thesis, the notation $\sim$ is used in its mathematical sense, "is equivalent to", meaning that :

$$f(t) \sim g(t) \Leftrightarrow \frac{f(t)}{g(t)} \xrightarrow{t\to\infty} 1. \tag{B.1}$$

B.2 Moments of a distribution

The moments of a random variable X are defined as follows:

$$\mu_k = \mathbb{E}\left[(X-\mu)^k\right]. \tag{B.2}$$

B.3 Convergence in distribution : definition and lemma

Let $(X_n)_{n\in\mathbb{N}}$ be a sequence of random variables. $(X_n)_{n\in\mathbb{N}}$ convergences in distribution to a random variable X if and only if for every function f that is continuous and bounded, $\mathbb{E}(f(X_n)) \xrightarrow[n\to\infty]{} \mathbb{E}(f(X))$. We will note this:

$$X_n \xrightarrow[n\to\infty]{d} X. \tag{B.3}$$

We are also going to state here a lemma that is useful for the demonstration of the Gaussian distribution of the residues.

If

- $(X_n)_n$ is a sequence of random variables such that $X_n \xrightarrow[n\to\infty]{d} X$
- N_t is a random variable with integer values, such that $N_t \xrightarrow[t\to\infty]{d} Y$
- N_t,X_n are independent

Then

$$X_{N_t} \xrightarrow[t\to\infty]{d} X_Y. \tag{B.4}$$

The proof is very easy to do, it is enough to understand that as X_n and N_t are independent

$$\mathbb{E}(f(X_{Nt})) = \mathbb{E}(f(X_{N_t})|N_t) \quad \text{for every bounded and continous function f}$$

Then we can apply to the right side the convergence in distribution of (X_n), and we obtain naturally the convergence in distribution of $(X_{N_t})_t$.

B.4 Sum of independent and identically distributed variables

Let's consider a discrete random variable N, and $(X_i)_{i\in(1..\mathbb{N})}$, a collection of i.i.d random variables, that are also independent of N. We want to study the random variable Y defined as:

$$Y = \sum_{i=1}^{N} X_i. \tag{B.5}$$

We would like to compute the variance and expected value of Y. Let's start with the expected value, and use the law of total expectation to write:

$$\mathbb{E}(Y) = \mathbb{E}\left(\mathbb{E}(Y|N)\right) = \sum_n \mathbb{E}(Y|N=n)\mathbb{P}(n), \tag{B.6}$$

$$\text{with} \quad \mathbb{E}(Y|N=n) = \mathbb{E}\left(\sum_{i=1}^{n} X_i|N=n\right), \tag{B.7}$$

and as the X_i are identically distributed, and thanks to the conditioning, we get simply:

$$\mathbb{E}(Y|N=n) = n\mathbb{E}(X). \tag{B.8}$$

Therefore we get the value of the expected value:

$$\mathbb{E}(Y) = \sum n\mathbb{P}(n)\mathbb{E}(X) = \mathbb{E}(X)\mathbb{E}(N). \tag{B.9}$$

Now let's examine the expected value of Y^2:

$$\mathbb{E}\left(Y^2\right) = \mathbb{E}\left(\sum_{i=1}^{N}\sum_{j=1}^{N} X_i X_j\right), \tag{B.10}$$

and by linearity:

$$\mathbb{E}\left(Y^2\right) = \mathbb{E}\left(\sum_{i=1}^{N} X_i^2\right) + \mathbb{E}\left(\sum_{i,j,i\neq j}^{N} X_i X_j\right). \tag{B.11}$$

The first term of the sum can be simplified by using once again the law of total expectation:

$$\begin{aligned}\mathbb{E}\left(\sum_{i=1}^{N} X_i^2\right) &= \sum_n \mathbb{E}\left(\sum_{i=1}^{n} X_i^2 | N=n\right)\mathbb{P}(n)\\ &= \sum_n \mathbb{E}\left(X^2\right) n\mathbb{P}(n)\\ &= \mathbb{E}\left(X^2\right)\mathbb{E}(N). \end{aligned} \tag{B.12}$$

The second term is slightly more complicated, but we sill apply the sale rule, keeping in mind that X_i and X_j are independent as $i \neq j$:

$$\begin{aligned}\mathbb{E}\left(\sum_{i,j,i\neq j}^{N} X_i X_j\right) &= \sum_n \mathbb{E}\left(\sum_{i,j,i\neq j}^{n} X_i X_j | N=n\right)\mathbb{P}(n)\\ &= \sum_n \left(\sum_{i,j,i\neq j}^{n} \mathbb{E}(X)^2\right)\mathbb{P}(n)\\ &= \mathbb{E}(X)^2 \sum_n \left(n(n-1)\right)\mathbb{P}(n)\\ &= \mathbb{E}(X)^2\left(\mathbb{E}\left(N^2\right) - \mathbb{E}(N)\right). \end{aligned} \tag{B.13}$$

We can gather the two terms to obtain:

$$\begin{aligned}\mathrm{Var}(Y) &= \mathbb{E}\left(Y^2\right) - \mathbb{E}(Y)^2 = \mathbb{E}\left(Y^2\right) - \mathbb{E}(X)^2\mathbb{E}(N)^2\\ &= \mathbb{E}\left(X^2\right)\mathbb{E}(N) + \mathbb{E}(X)^2\left(\mathbb{E}\left(N^2\right) - \mathbb{E}(N)\right) - \mathbb{E}(X)^2\mathbb{N}^2\\ &= \mathbb{E}(N)\left(\mathbb{E}\left(X^2\right) - \mathbb{E}(X)^2\right) + \mathbb{E}(X)^2\left(\mathbb{E}\left(N^2\right) - \mathbb{N}^2\right), \end{aligned} \tag{B.14}$$

which gives us the final result:

$$\mathrm{Var}(Y) = \mathbb{E}(N)\mathrm{Var}(X) + \mathbb{E}(X)^2\mathrm{Var}(N). \tag{B.15}$$

B.5 Gamma division times in a Bellman-Harris model

We will present here what the Bellman-Harris model yields when we use division times τ that follow a Gamma distribution. The Gamma distribution has the following density function that relies on two parameters a and b, such that:

$$\Gamma_{a,b}(t) = \frac{1}{b^a\Gamma(a)} t^{a-1} e^{\frac{-t}{b}}, \tag{B.16}$$

where $\Gamma(a)$ is the value of the Gamma function in a:

$$\Gamma(a) = \int_0^\infty t^{a-1} e^{-t} \mathrm{d}t. \tag{B.17}$$

Foe the computation of the results of the Bellman-Harris model, we know that a lot depends on the random variable $X = e^{-\alpha\tau}$. And in this case, it is very easy to compute the expected value of X, because:

$$\begin{aligned}
\mathbb{E}(X) &= \int_0^\infty e^{-\alpha t}\Gamma_{a,b}(t)\mathrm{d}t \\
&= \int_0^\infty e^{-\alpha t}\frac{1}{b^a\Gamma(a)}t^{a-1}e^{\frac{-t}{b}}\mathrm{d}t \\
&= \frac{(1/b+\alpha)^a}{b^a}\int_0^\infty \frac{1}{(1/b+\alpha)^a\,\Gamma(a)}t^{a-1}e^{-t(1/b+\alpha)}\mathrm{d}t \\
&= \frac{(1/b+\alpha)^a}{b^a}\int_0^\infty \Gamma_{a,b/(1+b\alpha)}(t)\mathrm{d}t \\
&= \frac{(1/b+\alpha)^a}{b^a}.
\end{aligned} \tag{B.18}$$

Therefore we have, for the value of the growth rate α, that is solutione of $\mathbb{E}(X) = 1/2$:

$$\alpha = \frac{1}{b}\left(2^{\frac{1}{a}} - 1\right). \tag{B.19}$$

The same trick as the one used in the above computation (exhibiting expected values of other Gamma distribution) can be also used to compute $\mathbb{E}(X^2)$ and the value of n_1, and hence we get all the desired integral computations of the Bellman-Harris model.

We will just show the results here, we have

$$n_1 = \frac{1}{2a}\frac{2^{\frac{1}{a}}}{2^{\frac{1}{a}} - 1}, \tag{B.20}$$

$$\mathbb{E}(X^2) = \frac{1}{(2\alpha b + 1)^a}. \tag{B.21}$$

This last equation can be injected into eq. (3.9) to get the coefficient of variation for the number of cells in the exponential phase.

The same trick works also very well for the more complicated cases of different division times, the same kind of integral computation appears.

Appendix C

Approximated volume of a big droplet

In this Appendix, we are going to compute an approximation for the volume of droplets that are bigger than the trap. Numerical simulations of the shape of this kind of droplet, along with the parameters of the problem, can be found on Fig. C.1.

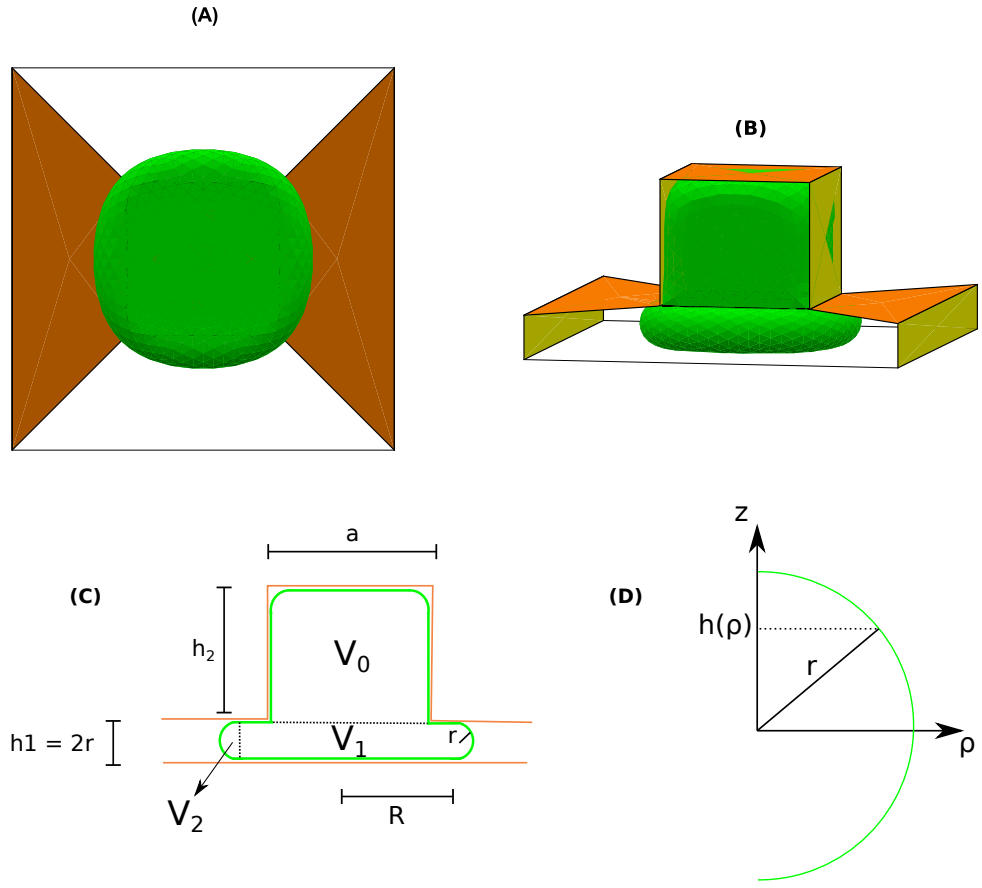


FIGURE C.1: (A) and (B) numerical simulation of a droplet bigger than the trap (volume 2.3 nL): (A) bottom view and (B) side view. (C) Geometrical parameters of the problem. (D) Zoom on the semi-spherical crown and definition of the cylindrical coordinates.

For this purpose, we are going to make some approximations. We will first consider that the volume variation of this kind of droplet is mostly due to the squeezed part of the droplet that extends in the chamber. The volume of the part of the droplet that is in the trap will be considered to be constant. With our parameters, that means that V_0 is a constant. Moreover, we will approximate the value of V_0 by the volume of the trap, which has a cuboid shape, we will therefore use:

$$V_0 \approx a^2 \cdot h_2. \tag{C.1}$$

The other part of the droplet is the pancake-like part in the chamber of the chip. Its real shape is complex, as we can see on the simulations C.1. We will approximate it as a circularly symmetric shape composed of two parts. First, a cylindrical center of volume V_1. Its radius is R and its height is the height of the chamber h_1. Then, a crown that surrounds this cylinder, and has a semi-circular section of radius $r = h_1/2$. This crown has volume V_2.

The volume of the cylindrical central part can be easily computed:

$$V_1 = \pi R^2 h_1. \tag{C.2}$$

The last part of the droplet is the semi-spherical crown around the cylinder. The volume of this part is a little bit more difficult to compute. We will place ourselves in cylindrical coordinates, and we have:

$$V_2 = \int_0^{2\pi} \int_R^{r+R} \int_{-h(\rho)}^{h(\rho)} \rho \mathrm{d}\rho \mathrm{d}z \mathrm{d}\theta, \tag{C.3}$$

where ρ is the radial distance, θ the azimuth angle and z the height, see Fig. C.1(D). Since the crown shape is circularly symmetric, but also symmetric with respect to the plane z=0, we can get rid of the integrate over θ, and consider only the upper part of the crown:

$$V_2 = 4\pi \int_R^{r+R} \int_0^{h(\rho)} \rho \mathrm{d}z \mathrm{d}\rho. \tag{C.4}$$

As we consider that the shape of the crown is semi-spherical, we have the following relation between the height of the crown and the radial coordinate ρ:

$$h(\rho) = \sqrt{r^2 - (\rho - R)^2}. \tag{C.5}$$

We can then integrate over z to get:

$$V_2 = 4\pi \int_R^{r+R} \rho \left[z\right]_0^{\sqrt{r^2-(\rho-R)^2}} \mathrm{d}\rho. \tag{C.6}$$

We substitute the coordinate ρ by $s = \rho - R$ to get:

$$V_2 = 4\pi \int_0^r (s + R) \sqrt{r^2 - s^2} \mathrm{d}s, \tag{C.7}$$

$$V_2 = 4\pi \left(\int_0^r s\sqrt{r^2 - s^2}\mathrm{d}s + R \int_0^r \sqrt{r^2 - s^2}\mathrm{d}s \right). \tag{C.8}$$

To compute the second term of the sum, we have to set $u = s/r$:

$$V_2 = 4\pi \left(\left[-\frac{1}{3}(r^2 - s^2)^{\frac{3}{2}} \right]_0^r + Rr^2 \int_0^1 \sqrt{1 - u^2}\mathrm{d}u \right), \tag{C.9}$$

and with $u = sin(t)$:

$$V_2 = 4\pi \left(\frac{r^3}{3} + Rr^2 \int_0^{\pi/2} \cos^2(t)\mathrm{d}t \right). \tag{C.10}$$

To compute $\int_0^{\pi/2} \cos^2(t)\mathrm{d}t$, we use the well-known formula:

$$\cos^2(t) = \frac{1 + \cos(2t)}{2}, \tag{C.11}$$

which yields the final result:

$$V_2 = 4\pi \left(\frac{r^3}{3} + Rr^2 \frac{\pi}{4} \right). \tag{C.12}$$

Therefore the total volume of the droplet is the following:

$$V = \frac{4\pi}{3} r^3 + \pi^2 R r^2 + 2\pi R^2 r + a^2 \cdot h_2, \tag{C.13}$$

and the projected area measured from below the droplet is the projected area of the squeezed part, which is:

$$A = \pi \left(R + r \right)^2. \tag{C.14}$$

R can be varied and those two equations combined to get the dotted yellow line of Fig. 2.9(B).

Appendix D

Residuals and Noise

D.1 Additive noise

The first kind of noise that we can think of is a simple additive noise, which can be written as follows:

$$Fluo(t) = a_f N(t) + \eta(t). \tag{D.1}$$

Where η models the noise and can be at first approximation taken as a Gaussian: $\eta(t) \sim \mathcal{N}\left(0, \sigma_\eta^2\right)$. We will make two hypotheses on η. First, we consider that η is independent of N. Second, we consider that, as a random noise, it is also not correlated in time, which means that for our observations $\eta(t_i)$ and $\eta(t_{i+1})$ are independent.

In this case, the residuals can be simply computed as follows:

$$Res_i^{Fluo} = Fluo(t_{i+1}) - Fluo(t_i). \tag{D.2}$$

$$Res_i^{Fluo} = a_f N(t_{i+1}) + \eta(t_{i+1}) - a_f N(t_i) e^{\alpha \Delta t} - \eta(t) e^{\alpha \Delta t}. \tag{D.3}$$

And, because of the independence hypothesis that we mentioned above, if we take the variance of the last equation, we get:

$$\mathrm{Var}\left(Res_i^{Fluo}\right) = a_f^2 \mathrm{Var}\left(Res_i\right) + \sigma_\eta^2 \left(1 + e^{2\alpha \Delta t}\right). \tag{D.4}$$

The additive noise thus adds a constant part to the expression of the residuals. This constant part is directly proportional to the variance of the noise. As the residuals increase exponentially with time, at long times the additional variance due to the noise will be completely negligible, but for short times, it can overcome the contribution of the residuals, see Fig. D.1. There are thus two asymptotic regimes: at short times, the variance of the residuals is constant and proportional to the variance of the noise, and at long times, it is exponentially growing with rate α. At the transition between those two regimes, the residuals do not evolve exponentially, but if, unfortunately, this is the only region that we can see in the experiments, this could explain the decrease of the slope that we see in the experiments.

Hence, if we choose manually the value of a_f, the experimental results match these numerical ones quite closely (Fig. D.1[A],[B], black plain line),

for an estimated $\sigma_\eta \approx 2$. However, the fact that this difference of slope happens in a similar way for different bacterial strains and different settings of the microscope, which are parameters that would change the importance of the additive noise, leads us to think that their might be another explanation.

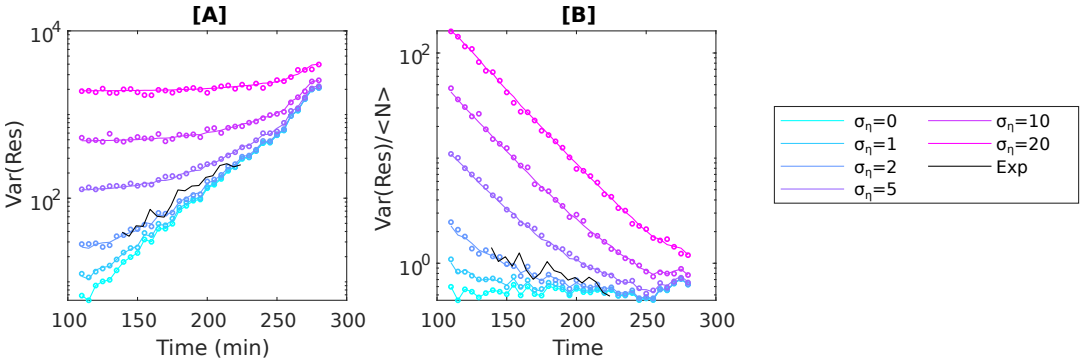


FIGURE D.1: Colors: Influence of the additive noise on the residuals: for a full *E. coli* simulation (same as Fig. 4.2), when varying the variance of the noise σ_η. Plain line: theory (eq. D.4), circles: simulations. [A] Variance of the residuals in log scale [B] Variance of the residuals divided by the mean number of cells. In black, the results of a typical experiment (same as in chapter 2, 3), with an arbitrary chosen value of a_f.

D.2 Multiplicative noise

The experimental noise could be not only additive, but also multiplicative, meaning that we could have:

$$Fluo(t) = a_f \eta(t) N(t). \tag{D.5}$$

With $\eta(t)$ being a random variable, independent of $N(t)$, with mean 1 and standard deviation σ_η. This multiplicative modelization of the experimental noise may sound stranger than the more common additive one, but it could make sense if we think that the background in the fluorescent images had to be divided and not subtracted to be correctly removed, see Chapter 2. We will also consider that the noise is not correlated in time.

If we measure the residuals in this case, we would get:

$$Res_i^{Fluo} = a_f \eta(t_{i+1}) N(t_{i+1}) - a_f \eta(t_i) N(t_i) e^{\alpha \Delta t}. \tag{D.6}$$

We are going to use the two following lemmas: if X and Y are independent random variables, then the expected value of the product is:

$$\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y). \tag{D.7}$$

And the variance of the product is:

$$\mathrm{Var}(XY) = \mathrm{Var}(X)\mathrm{Var}(Y) + \mathrm{Var}(X)\mathbb{E}(Y)^2 + \mathrm{Var}(Y)\mathbb{E}(X)^2. \tag{D.8}$$

With lemma D.7, it is easy to see that the residuals keep the same mean if we multiply them by an independent noise. Let us examine now their variance:

$$\begin{aligned}\mathrm{Var}\left(Res_i^{Fluo}\right) = a_f^2 \Big[&\mathrm{Var}(\eta(t_{i+1})N(t_{i+1})) + e^{2\alpha\Delta t}\mathrm{Var}(\eta(t_i)N(t_i)) \\ &-2e^{\alpha\Delta t}\mathrm{cov}\left(\eta(t_i)N(t_i), \eta(t_{i+1})N(t_{i+1})\right)\Big]. \end{aligned} \tag{D.9}$$

For the first two terms, it is easy to see with lemma D.8 that we get, since $\mathbb{E}(\eta(t)) = 1$:

$$\mathrm{Var}(\eta(t_i)N(t_i)) = \sigma_\eta^2 \mathrm{Var}\left(N(t_i)\right) + \sigma_\eta^2 \mathbb{E}(N(t_i))^2 + \mathrm{Var}\left(N(t_i)\right). \tag{D.10}$$

For the covariance term, we get:

$$\begin{aligned}\mathrm{cov}\left(\eta(t_i)N(t_i), \eta(t_{i+1})N(t_{i+1})\right) = &\mathbb{E}\left(\eta(t_i)N(t_i)\eta(t_{i+1})N(t_{i+1})\right) \\ &- \mathbb{E}\left(\eta(t_i)N(t_i)\right)\mathbb{E}\left(\eta(t_{i+1})N(t_{i+1})\right),\end{aligned} \tag{D.11}$$

and because of the independence:

$$\begin{aligned}\mathrm{cov}\left(\eta(t_i)N(t_i), \eta(t_{i+1})N(t_{i+1})\right) &= \mathbb{E}(\eta(t_i))\mathbb{E}(\eta(t_{i+1}))\mathbb{E}\left(N(t_i)N(t_{i+1})\right) \\ &- \mathbb{E}(\eta(t_i))\mathbb{E}\left(N(t_i)\right)\mathbb{E}(\eta(t_{i+1}))\mathbb{E}\left(N(t_{i+1})\right) \\ &= \mathrm{cov}\left(N(t_i), N(t_{i+1})\right).\end{aligned} \tag{D.12}$$

We have to use here the expression of the residuals in terms of the covariance:

$$\mathrm{Var}\left(Res_i\right) = \mathrm{Var}(N(t_{i+1})) + \mathrm{Var}(N(t_i))e^{2\alpha\Delta t} - 2\mathrm{cov}\left(N(t_{i+1}), N(t_i)\right). \tag{D.13}$$

We gather the three terms of the sum to get:

$$\begin{aligned}\mathrm{Var}\left(Res_i^{Fluo}\right) &= a_f^2 \Big[\mathrm{Var}\left(Res_i\right) + \sigma_\eta^2 \Big(e^{2\alpha\Delta t}\mathbb{E}(N(t_i))^2 + \mathbb{E}(N(t_{i+1}))^2 \\ &\quad + e^{2\alpha\Delta t}\mathrm{Var}\left(N(t_i)\right) + \mathrm{Var}\left(N(t_{i+1})\right)\Big)\Big] \\ &= a_f^2 \Big[\mathrm{Var}\left(Res_i\right) + \sigma_\eta^2 \Big(e^{2\alpha\Delta t}\mathbb{E}(N(t_i)^2) + \mathbb{E}(N(t_{i+1})^2)\Big)\Big].\end{aligned} \tag{D.14}$$

Here we will have to use equivalents and not real equalities, but for long times, we know thanks to Bellman and Harris (see Chapter 3) that:

$$\mathbb{E}(N(t_{i+1})^2) \sim \tilde{n}_2 e^{2\alpha t_{i+1}} \sim \tilde{n}_2 e^{2\alpha(t_i+\Delta t)} \sim \mathbb{E}(N(t_i)^2)e^{2\alpha\Delta t}. \tag{D.15}$$

Therefore, we obtain:

$$\mathrm{Var}\left(Res_i^{Fluo}\right) \sim a_f^2 \Big(\mathrm{Var}\left(Res_i\right) + 2\sigma_\eta^2 e^{2\alpha\Delta t}\mathbb{E}(N(t_i)^2)\Big). \tag{D.16}$$

We thus obtain a corrective term that is not constant, contrary to the additive noise, but rather proportional to the expected value of the square of $N(t)$, which evolves in $e^{2\alpha t}$. We can check the influence of this noise terms numerically (Fig. D.2)

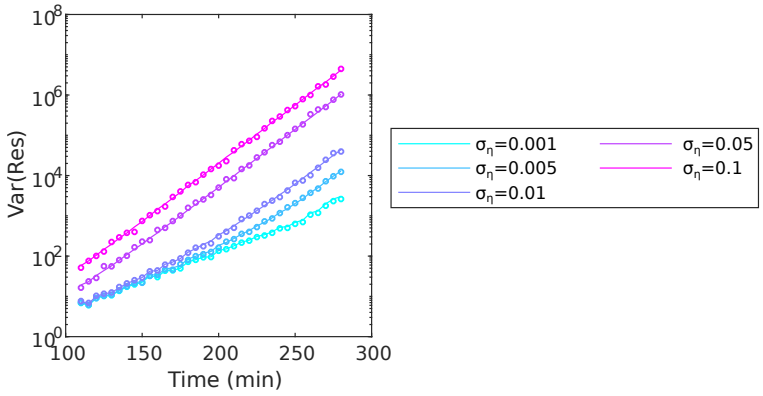


FIGURE D.2: Influence of the multiplicative noise on the residuals: for a full *E. coli* simulation (same as Fig. 4.2), when varying the SD of the noise σ_η. Plain line: theory (eq. D.16), circles: simulations. Variance of the residuals in log scale

The multiplicative noise will then increase the rate of growth of the variance of the residuals (that is proportional to $e^{\alpha t}$) rather then decreasing it. Hence it cannot explain what we see in the experiments.

D.3 Heterogeneity of the fluorescence

Another source of noise in the experiment can come from the heterogeneity of the fluorescence signal, which is not homogeneous among the cells but varies from one bacteria to another, even if the cells are isogenic [111]. In this case, the relation between the fluorescence and the number of cells can be written as:

$$Fluo(t) = \sum_{i=1}^{N(t)} a_f^i(t). \tag{D.17}$$

Where $a_f^i(t)$ represents the fluorescence signal of cell i at time t. To simplify the computations, we will consider that the fluorescence signals of the cells are independent, and that they are also independent in time. These are of course simplifying hypotheses, as the fluorescence signal is probably correlated for one cell from time t_i to time t_{i+1} for instance, but as we average on a lot of cells, these effects are negligible compared to the heterogeneity in itself.

We can then compute the residuals as follows:

$$Res_i^{Fluo} = Fluo(t_{i+1}) - Fluo(t_i)e^{\alpha \Delta t}. \tag{D.18}$$

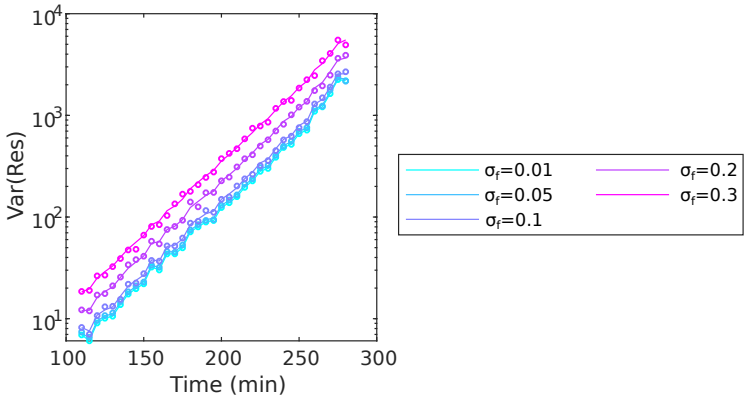


FIGURE D.3: Influence of the fluorescence heterogeneity on the residuals: for a full *E. coli* simulation (same as Fig. 4.2), when varying the SD of the fluo σ_f. Plain line: theory (eq. D.26), circles: simulations. Variance of the residuals in log scale

If we compute the variance of the residuals, we get

$$\begin{aligned}\mathrm{Var}\left(Res_i^{Fluo}\right) &= \mathrm{Var}\left(Fluo(t_{i+1})\right) + \mathrm{Var}\left(Fluo(t_i)\right) e^{2\alpha\Delta t} \\ &\quad - 2\mathrm{cov}\left(Fluo(t_{i+1}), Fluo(t_i)\right). \end{aligned} \tag{D.19}$$

For the first two terms, we can just use the formula of the sum of i.i.d variables (Appendix B.4) to get:

$$\mathrm{Var}\left(Fluo(t_i)\right) = \mathrm{Var}\left(\sum_{i=1}^{N(t_i)} a_f^i(t_i)\right) \tag{D.20}$$

$$= a_f^2 \mathrm{Var}(N(t)) + \sigma_f^2 \mathbb{E}(N(t)). \tag{D.21}$$

For the covariance term, we are going to use the law of total probability. First we can note that, because the $a_f^i(t)$ are independent, if we condition by $N(t), N(t+1)$, we have:

$$\begin{aligned}\mathbb{E}\left(Fluo(t_{i+1})Fluo(t_i)|N(t_{i+1}), N(t_i)\right) &= \mathbb{E}\left(Fluo(t_{i+1})|N(t_{i+1}), N(t_i)\right) \\ &\quad \times \mathbb{E}\left(Fluo(t_i)|N(t_{i+1}), N(t_i)\right) \\ &= a_f^2 N(t_i) N(t_{i+1}).\end{aligned} \tag{D.22}$$

Then we see that with the law of total probability:

$$\mathbb{E}\left(Fluo(t_{i+1})Fluo(t_i)\right) = \mathbb{E}\left(\mathbb{E}\left(Fluo(t_{i+1})Fluo(t_i)|N(t_{i+1}), N(t_i)\right)\right) \tag{D.23}$$

$$= a_f^2 \mathbb{E}\left(N(t)N(t+1)\right). \tag{D.24}$$

Therefore we have simply:

$$\mathrm{cov}\left(Fluo(t_{i+1}), Fluo(t_i)\right) = a_f^2 \mathrm{cov}\left(N(t_{i+1}, N(t_i)\right). \tag{D.25}$$

And we can inject this expression in the formula of the variance, using eq. (D.13):

$$\mathrm{Var}\left(Res_i^{Fluo}\right) = a_f^2\left[\mathrm{Var}\left(Res_i\right) + \frac{\sigma_f^2}{a_f^2}\mathbb{E}(N(t))e^{\alpha\Delta t}\left(e^{\alpha\Delta t}+1\right)\right]. \tag{D.26}$$

Therefore, the heterogeneity of the fluorescence does not change the rate of growth of the residuals, since the second term of the sum is proportional to $\mathbb{E}(N(t))$, but rather adds a constant value (in log scale). This is verified on the simulations (Fig. D.3). The heterogeneity of the fluorescence is also not a good candidate to account for what we wee in the experiments.

Appendix E

URLs for Supplementary Movies

The movies presented in the text of this PhD can be found at the following links:

Supplementary Movie 1 https://figshare.com/s/44346599fd0cf67b50db
Time-lapse movie of the growth of *B. subtilis* in microfluidic droplets (LB medium). One image every 10 minutes, 8 hours of growth are displayed. Bright field and fluorescence images are superposed with fake red colors for fluorescence images.

Supplementary Movie 2 https://figshare.com/s/26a198938d6ea166e387
Time-lapse movie of the growth of *E. coli* in microfluidic droplets (LB medium with agarose gel). One image every 30 minutes, 12 hours of growth are displayed. Bright field and fluorescence images are superposed with fake green color for fluorescence images.

Supplementary Movie 3 https://figshare.com/s/0c874766dac94012ce2f
Numerical shape (kernel fit) of the distribution of the number of cells rescaled by $\exp(\alpha t)$, for a time varying from 0 to 300 min, with a 5 min time step. Bellamn-Harris classical case with Gaussian division times: same simulation as on Fig. 3.3, with a coefficient of variation for the division times of 0.2.

Supplementary Movie 4 https://figshare.com/s/e7f3062444e272cfda41
Numerical shape (kernel fit) of the distribution of the number of cells rescaled by $\exp(\alpha t)$, for a time varying from 0 to 300 min, with a 5 min time step. Bellamn-Harris classical case with Gaussian division times: same simulation as on Fig. 3.8, with a coefficient of variation for the division times of 0.2.

Supplementary Movie 5 https://figshare.com/s/aa0c94a1f04c9067e42f
Description: Shape (kernel fit) of the distribution of the fluorescence of the droplets in the experiment described in Fig 2.7 rescaled by the mean fluorescence signal, for a time varying from 74 to 274 min, with a 5 min time step.

Bibliography

[1] H. RUBIN. "The significance of biological heterogeneity". In: *Cancer and Metastasis Review* 9.1 (1990), pp. 1–20.

[2] M. SHIBATA and M. M. SHEN. "The roots of cancer: stem cells and the basis for tumor heterogeneity". In: *BioEssays: News and Reviews in Molecular, Cellular and Developmental Biology* 35.3 (2013), pp. 253–260.

[3] K. M. DAVIS and R. R. ISBERG. "Defining heterogeneity within bacterial populations via single cell approaches". In: *BioEssays* 38.8 (2016), pp. 782–790.

[4] A. R. WALLACE. *Darwinism : an exposition of the theory of natural selection, with some of its applications*. London : Macmillan, 1889.

[5] M. PATEL, Q. JIANG, R. WOODGATE, M. M. COX, and M. F. GOODMAN. "A New Model for SOS-induced Mutagenesis: How RecA Protein Activates DNA Polymerase V". In: *Critical reviews in biochemistry and molecular biology* 45.3 (2010), pp. 171–184.

[6] Z. BAHAROGLU and D. MAZEL. "Vibrio cholerae Triggers SOS and Mutagenesis in Response to a Wide Range of Antibiotics: a Route towards Multiresistance". In: *Antimicrobial Agents and Chemotherapy* 55.5 (2011), pp. 2438–2441.

[7] M. W. VAN DER WOUDE. "Phase variation: how to create and coordinate population diversity". In: *Current Opinion in Microbiology*. Cell regulation 14.2 (2011), pp. 205–211.

[8] E. ROTMAN and H. S. SEIFERT. "The genetics of Neisseria species". In: *Annual Review of Genetics* 48 (2014), pp. 405–431.

[9] M. ACKERMANN. "A functional perspective on phenotypic heterogeneity in microorganisms". In: *Nature Reviews Microbiology* 13.8 (2015), pp. 497–508.

[10] N. E. FREED, O. K. SILANDER, B. STECHER, A. BÖHM, W.-D. HARDT, and M. ACKERMANN. "A Simple Screen to Identify Promoters Conferring High Levels of Phenotypic Noise". In: *PLOS Genetics* 4.12 (2008), pp. 1–6.

[11] N. Q. BALABAN, J. MERRIN, R. CHAIT, L. KOWALIK, and S. LEIBLER. "Bacterial Persistence as a Phenotypic Switch". In: *Science* 305.5690 (2004), pp. 1622–1625.

[12] J. W. BIGGER. "Treatment of Staphylococcal Infections with Penicillin by Intermittent Sterilisation". In: *The Lancet* 244.6320 (1944), pp. 497–500.

[13] L. HALL-STOODLEY, J. W. COSTERTON, and P. STOODLEY. "Bacterial biofilms: from the Natural environment to infectious diseases". In: *Nature Reviews Microbiology* 2.2 (2004), pp. 95–108.

[14] P. GILBERT, J. DAS, and I. FOLEY. "Biofilm susceptibility to antimicrobials". In: *Advances in Dental Research* 11.1 (1997), pp. 160–167.

[15] A. FOREMAN, G. HOLTAPPELS, A. J. PSALTIS, J. JERVIS-BARDY, J. FIELD, P.-J. WORMALD, and C. BACHERT. "Adaptive immune responses in Staphylococcus aureus biofilm-associated chronic rhinosinusitis". In: *Allergy* 66.11 (2011), pp. 1449–1456.

[16] V. BERK, J. C. N. FONG, G. T. DEMPSEY, O. N. DEVELIOGLU, X. ZHUANG, J. LIPHARDT, F. H. YILDIZ, and S. CHU. "Molecular architecture and assembly principles of Vibrio cholerae biofilms". In: *Science* 337.6091 (2012), pp. 236–239.

[17] P. GUPTASARMA. "Does replication-induced transcription regulate synthesis of the myriad low copy number proteins of Escherichia coli?" In: *BioEssays: News and Reviews in Molecular, Cellular and Developmental Biology* 17.11 (1995), pp. 987–997.

[18] D. T. GILLESPIE. "Exact stochastic simulation of coupled chemical reactions". In: *The Journal of Physical Chemistry* 81.25 (1977), pp. 2340–2361.

[19] M. B. ELOWITZ, A. J. LEVINE, E. D. SIGGIA, and P. S. SWAIN. "Stochastic Gene Expression in a Single Cell". In: *Science* 297.5584 (2002), pp. 1183–1186.

[20] T. M. NORMAN, N. D. LORD, J. PAULSSON, and R. LOSICK. "Stochastic Switching of Cell Fate in Microbes". In: *Annual Review of Microbiology* 69.1 (2015), pp. 381–403.

[21] P. J. CHOI, L. CAI, K. FRIEDA, and X. S. XIE. "A stochastic single-molecule event triggers phenotype switching of a bacterial cell". In: *Science* 322.5900 (2008), pp. 442–446.

[22] M. PTASHNE. *A Genetic Switch: Gene Control and Phage Lambda*. Cell Press & Blackwell Scientific Publications, 1986.

[23] A. OCHAB-MARCINEK and M. TABAKA. "Bimodal gene expression in noncooperative regulatory systems". In: *Proceedings of the National Academy of Sciences* 107.51 (2010), pp. 22096–22101.

[24] J. L. CHERRY and F. R. ADLER. "How to make a biological switch". In: *Journal of Theoretical Biology* 203.2 (2000), pp. 117–133.

[25] T. S. GARDNER, C. R. CANTOR, and J. J. COLLINS. "Construction of a genetic toggle switch in Escherichia coli". In: *Nature* 403.6767 (2000), pp. 339–342.

[26] A. T. NIELSEN, N. A. DOLGANOV, T. RASMUSSEN, G. OTTO, M. C. MILLER, S. A. FELT, S. TORREILLES, and G. K. SCHOOLNIK. "A bistable switch and anatomical site control Vibrio cholerae virulence gene expression in the intestine". In: *PLoS pathogens* 6.9 (2010), pp. 1–23.

[27] L. VAN MELDEREN and M. SAAVEDRA DE BAST. "Bacterial Toxin–Antitoxin Systems: More Than Selfish Entities?" In: *PLoS Genetics* 5.3 (2009), pp. 1–6.

[28] I. KEREN, D. SHAH, A. SPOERING, N. KALDALU, and K. LEWIS. "Specialized Persister Cells and the Mechanism of Multidrug Tolerance in Escherichia coli". In: *Journal of Bacteriology* 186.24 (2004), pp. 8172–8180.

[29] L. GELENS, L. HILL, A. VANDERVELDE, J. DANCKAERT, and R. LORIS. "A General Model for Toxin-Antitoxin Module Dynamics Can Explain Persister Cell Formation in E. coli". In: *PLoS Computational Biology* 9.8 (2013), pp. 1–17.

[30] L. BUTS, J. LAH, M.-H. DAO-THI, L. WYNS, and R. LORIS. "Toxin-antitoxin modules as bacterial metabolic stress managers". In: *Trends in Biochemical Sciences* 30.12 (2005), pp. 672–679.

[31] J. W. LITTLE and C. B. MICHALOWSKI. "Stability and instability in the lysogenic state of phage lambda". In: *Journal of Bacteriology* 192.22 (2010), pp. 6064–6076.

[32] T. BERGMILLER, A. M. C. ANDERSSON, K. TOMASEK, E. BALLEZA, D. J. KIVIET, R. HAUSCHILD, G. TKAČIK, and C. C. GUET. "Biased partitioning of the multidrug efflux pump AcrAB-TolC underlies long-lived phenotypic heterogeneity". In: *Science* 356.6335 (2017), pp. 311–315.

[33] E. SCALLAN, R. M. HOEKSTRA, F. J. ANGULO, R. V. TAUXE, M.-A. WIDDOWSON, S. L. ROY, J. L. JONES, and P. M. GRIFFIN. "Foodborne illness acquired in the United States–major pathogens". In: *Emerging Infectious Diseases* 17.1 (2011), pp. 7–15.

[34] M. DIARD, V. GARCIA, L. MAIER, M. N. P. REMUS-EMSERMANN, R. R. REGOES, M. ACKERMANN, and W.-D. HARDT. "Stabilization of cooperative virulence by the expression of an avirulent phenotype". In: *Nature* 494.7437 (2013), pp. 353–356.

[35] M. ARNOLDINI, I. A. VIZCARRA, R. PEÑA-MILLER, N. STOCKER, M. DIARD, V. VOGEL, R. E. BEARDMORE, W.-D. HARDT, and M. ACKERMANN. "Bistable Expression of Virulence Genes in Salmonella Leads to the Formation of an Antibiotic-Tolerant Subpopulation". In: *PLOS Biology* 12.8 (2014), pp. 1–8.

[36] E. KUSSELL and S. LEIBLER. "Phenotypic Diversity, Population Growth, and Information in Fluctuating Environments". In: *Science* 309.5743 (2005), pp. 2075–2078.

[37] O. X. CORDERO and M. F. POLZ. "Explaining microbial genomic diversity in light of evolutionary ecology". In: *Nature Reviews. Microbiology* 12.4 (2014), pp. 263–273.

[38] LURIA S.E. and DELBRUCK M. "Mutations of bacteria from virus sensitivity to virus resistance". In: *Genetics* 28.491 (1943), pp. 491–511.

[39] D. A. KESSLER, R. H. AUSTIN, and H. LEVINE. "Resistance to Chemotherapy: Patient Variability and Cellular Heterogeneity". In: *Cancer Research* 74.17 (2014), pp. 4663–4670.

[40] T. M. SQUIRES and S. R. QUAKE. "Microfluidics: Fluid physics at the nanoliter scale". In: *Reviews of Modern Physics* 77.3 (2005), pp. 977–1026.

[41] G. M. WHITESIDES. "The origins and the future of microfluidics". In: *Nature* 442 (2006), pp. 368–373.

[42] L. POTVIN-TROTTIER, S. LURO, and J. PAULSSON. "Microfluidics and single-cell microscopy to study stochastic processes in bacteria". In: *Current Opinion in Microbiology* 43 (2018), pp. 186–192.

[43] S. COOKSON, N. OSTROFF, W. L. PANG, D. VOLFSON, and J. HASTY. "Monitoring dynamics of single-cell gene expression over multiple cell cycles". In: *Molecular Systems Biology* 1 (2005), pp. 1–6.

[44] P. WANG, L. ROBERT, J. PELLETIER, W. L. DANG, F. TADDEI, A. WRIGHT, and S. JUN. "Robust Growth of Escherichia coli". In: *Current Biology* 20.12 (2010), pp. 1099–1103.

[45] S. TAHERI-ARAGHI, S. BRADDE, J. SAULS, N. HILL, P. LEVIN, J. PAULSSON, M. VERGASSOLA, and S. JUN. "Cell-Size Control and Homeostasis in Bacteria". In: *Current Biology* 25.3 (2015), pp. 385–391.

[46] O. BALTEKIN, A. BOUCHARIN, E. TANO, D. I. ANDERSSON, and J. ELF. "Antibiotic susceptibility testing in less than 30 min using direct single-cell imaging". In: *Proceedings of the National Academy of Sciences* 114.34 (2017), pp. 9170–9175.

[47] L. POTVIN-TROTTIER, N. D. LORD, G. VINNICOMBE, and J. PAULSSON. "Synchronous long-term oscillations in a synthetic gene circuit". In: *Nature* 538.7626 (2016), pp. 514–517.

[48] B. OKUMUS, C. J. BAKER, J. C. ARIAS-CASTRO, G. C. LAI, E. LEONCINI, S. BAKSHI, S. LURO, D. LANDGRAF, and J. PAULSSON. "Microfluidics-Assisted Cell Screening (MACS): an automated platform for single-cell microscopy of suspension cultures". In: *Nature Protocols* 13.1 (2018), pp. 170–194.

[49] B. OKUMUS, D. LANDGRAF, G. C. LAI, S. BAKSHI, J. C. ARIAS-CASTRO, S. YILDIZ, D. HUH, R. FERNANDEZ-LOPEZ, C. N. PETERSON, E. TOPRAK, M. EL KAROUI, and J. PAULSSON. "Mechanical slowing-down of cytoplasmic diffusion allows in vivo counting of proteins in individual cells". In: *Nature Communications* 7.11641 (2016).

[50] W. WANG, M.-J. ZHANG, and L.-Y. CHU. "Microfluidic approach for encapsulation via double emulsions". In: *Current Opinion in Pharmacology* 18 (2014), pp. 35–41.

[51] J. CLAUSELL-TORMOS, D. LIEBER, J.-C. BARET, A. EL-HARRAK, O. J. MILLER, L. FRENZ, J. BLOUWOLFF, K. J. HUMPHRY, S. KÖSTER, H. DUAN, C. HOLTZE, D. A. WEITZ, A. D. GRIFFITHS, and C. A. MERTEN. "Droplet-based microfluidic platforms for the encapsulation and screening of Mammalian cells and multicellular organisms". In: *Chemistry & Biology* 15.5 (2008), pp. 427–437.

[52] P. GRUNER, B. RIECHERS, L. A. CHACÒN ORELLANA, Q. BROSSEAU, F. MAES, T. BENEYTON, D. PEKIN, and J.-C. BARET. "Stabilisers for water-in-fluorinated-oil dispersions: Key properties for microfluidic applications". In: *Current Opinion in Colloid & Interface Science* 20.3 (2015), pp. 183–191.

[53] J. LEDERBERG. "A simple method for isolating individual microbes". In: *Journal of Bacteriology* 68.2 (1954), pp. 258–259.

[54] LWOFF, ANDRE and GUTTMAN, ANTOINETTE. "Recherches sur un baccilus megatherium lysogene". In: *Annales de l'institut Pasteur* 78 (1950), 711:739.

[55] T. S. KAMINSKI, O. SCHELER, and P. GARSTECKI. "Droplet microfluidics for microbiology: techniques, applications and challenges". In: *Lab on a Chip* 16.12 (2016), pp. 2168–2187.

[56] E. BROUZES, M. MEDKOVA, N. SAVENELLI, D. MARRAN, M. TWARDOWSKI, J. B. HUTCHISON, J. M. ROTHBERG, D. R. LINK, N. PERRIMON, and M. L. SAMUELS. "Droplet microfluidic technology for single-cell high-throughput screening". In: *Proceedings of the National Academy of Sciences* 106.34 (2009), pp. 14195–14200.

[57] J. LIM, O. CAEN, J. VRIGNON, M. KONRAD, V. TALY, and J.-C. BARET. "Parallelized ultra-high throughput microfluidic emulsifier for multiplex kinetic assays". In: *Biomicrofluidics* 9.3 (2015), pp. 1–11.

[58] M. KIM, M. PAN, Y. GAI, S. PANG, C. HAN, C. YANG, and S. K. Y. TANG. "Optofluidic ultrahigh-throughput detection of fluorescent drops". In: *Lab Chip* 15.6 (2015), pp. 1417–1423.

[59] H.-D. XI, H. ZHENG, W. GUO, A. M. GAÑÁN-CALVO, Y. AI, C.-W. TSAO, J. ZHOU, W. LI, Y. HUANG, N.-T. NGUYEN, and S. H. TAN. "Active droplet sorting in microfluidics: a review". In: *Lab on a Chip* 17.5 (2017), pp. 751–771.

[60] D. COTTINET, F. CONDAMINE, N. BREMOND, A. D. GRIFFITHS, P. B. RAINEY, J. A. G. M. DE VISSER, J. BAUDRY, and J. BIBETTE. "Lineage Tracking for Probing Heritable Phenotypes at Single-Cell Resolution". In: *PLOS ONE* 11.4 (2016), e0152395.

[61] S. P. DAMODARAN, S. EBERHARD, L. BOITARD, J. G. RODRIGUEZ, Y. WANG, N. BREMOND, J. BAUDRY, J. BIBETTE, and F.-A. WOLLMAN. "A Millifluidic Study of Cell-to-Cell Heterogeneity in Growth-Rate and Cell-Division Capability in Populations of Isogenic Cells of Chlamydomonas reinhardtii". In: *PLOS ONE* 10.3 (2015), pp. 1–28.

[62] Y.-J. EUN, A. UTADA, M. F. COPELAND, S. TAKEUCHI, and D. B. WEIBEL. "Encapsulating bacteria in agarose microparticles using microfluidics for high-throughput cell analysis and isolation". In: *ACS chemical biology* 6.3 (2011), pp. 260–266.

[63] W. POSTEK, P. GARGULINSKI, O. SCHELER, T. S. KAMINSKI, and P. GARSTECKI. "Microfluidic screening of antibiotic susceptibility at a single-cell level shows the inoculum effect of cefotaxime on E. coli". In: *Lab on a Chip* 18.23 (2018), pp. 3668–3677.

[64] S. JAKIELA, T. S. KAMINSKI, O. CYBULSKI, D. B. WEIBEL, and P. GARSTECKI. "Bacterial growth and adaptation in microdroplet chemostats". In: *Angewandte Chemie (International ed. in English)* 52.34 (2013), pp. 1–10.

[65] A. GRODRIAN, J. METZE, T. HENKEL, K. MARTIN, M. ROTH, and J. M. KÖHLER. "Segmented flow generation by chip reactors for highly parallelized cell cultivation". In: *Biosensors & Bioelectronics* 19.11 (2004), pp. 1421–1428.

[66] E. J. STEWART. "Growing Unculturable Bacteria". In: *Journal of Bacteriology* 194.16 (2012), pp. 4151–4160.

[67] S. Y. LIM, B.-J. KIM, M.-K. LEE, and K. KIM. "Development of a real-time PCR-based method for rapid differential identification of Mycobacterium species". In: *Letters in Applied Microbiology* 46.1 (2008), pp. 101–106.

[68] A. M. SIDORE, F. LAN, S. W. LIM, and A. R. ABATE. "Enhanced sequencing coverage with digital droplet multiple displacement amplification". In: *Nucleic Acids Research* 44.7 (2016), pp. 1–9.

[69] F. C. NEIDHARDT. "Bacterial Growth: Constant Obsession with dN/dt". In: *Journal of Bacteriology* 181.24 (1999), pp. 7405–7408.

[70] J. MONOD. "The Growth of Bacterial Cultures". In: *Annual Review of Microbiology* 3 (1949), pp. 371–394.

[71] M. SCHAECHTER, O. MAALOE, and N. KJELDGAARD. "Dependency on Medium and Temperature of Cell Size and Chemical Composition during Balanced Growth of Salmonella typhimurium". In: *Journal of General Microbiology* 19 (1958), pp. 592–606.

[72] M. SCOTT, C. W. GUNDERSON, E. M. MATEESCU, Z. ZHANG, and T. HWA. "Interdependence of Cell Growth and Gene Expression: Origins and Consequences". In: *Science* 330.6007 (2010), pp. 1099–1102.

[73] M. SCOTT and T. HWA. "Bacterial growth laws and their applications". In: *Current Opinion in Biotechnology* 22.4 (2011), pp. 559–565.

[74] P. GREULICH, M. SCOTT, M. R. EVANS, and R. J. ALLEN. "Growth-dependent bacterial susceptibility to ribosome-targeting antibiotics". In: *Molecular Systems Biology* 11.3 (2015), pp. 796–796.

[75] D. J. KIVIET, P. NGHE, N. WALKER, S. BOULINEAU, V. SUNDERLIKOVA, and S. J. TANS. "Stochasticity of metabolism and growth at the single-cell level". In: *Nature* 514.7522 (2014), pp. 376–379.

[76] R. BELLMAN and T. HARRIS. "On Age-Dependent Binary Branching Processes". In: *The Annals of Mathematics* 55.2 (1952), pp. 280–295.

[77] E. O. POWELL. "Growth Rate and Generation Time of Bacteria, with Special Reference to Continuous Culture". In: *Journal of General Microbiology* 15.3 (1956), pp. 492–511.

[78] E. B. STUKALIN, I. AIFUWA, J. S. KIM, D. WIRTZ, and S. X. SUN. "Age-dependent stochastic models for understanding population fluctuations in continuously cultured cells". In: *Journal of The Royal Society Interface* 10.85 (2013), pp. 1–11.

[79] A. A. ALONSO, I. MOLINA, and C. THEODOROPOULOS. "Modeling Bacterial Population Growth from Stochastic Single-Cell Dynamics". In: *Applied and Environmental Microbiology* 80.17 (2014), pp. 5241–5253.

[80] K. P. KOUTSOUMANIS and A. LIANOU. "Stochasticity in Colonial Growth Dynamics of Individual Bacterial Cells". In: *Applied and Environmental Microbiology* 79.7 (2013), pp. 2294–2301.

[81] C. PIN and J. BARANYI. "Kinetics of Single Cells: Observation and Modeling of a Stochastic Process". In: *Applied and Environmental Microbiology* 72.3 (2006), pp. 2163–2169.

[82] J. LIN and A. AMIR. "The Effects of Stochasticity at the Single-Cell Level and Cell Size Control on the Population Growth". In: *Cell Systems* 5.4 (2017), pp. 358–367.

[83] S. IYER-BISWAS, G. E. CROOKS, N. F. SCHERER, and A. R. DINNER. "Universality in stochastic exponential growth". In: *Physical Review Letters* 113.2 (2014), pp. 1–5.

[84] S. IYER-BISWAS, C. S. WRIGHT, J. T. HENRY, K. LO, S. BUROV, Y. LIN, G. E. CROOKS, S. CROSSON, A. R. DINNER, and N. F. SCHERER. "Scaling laws governing stochastic growth and division of single bacterial cells". In: *Proceedings of the National Academy of Sciences* 111.45 (2014), pp. 15912–15917.

[85] D. PIRJOL, F. JAFARPOUR, and S. IYER-BISWAS. "Phenomenology of stochastic exponential growth". In: *Physical Review E* 95.6 (2017), pp. 1–12.

[86] F. JAFARPOUR, C. S. WRIGHT, H. GUDJONSON, J. RIEBLING, E. DAWSON, K. LO, A. FIEBIG, S. CROSSON, A. R. DINNER, and S. IYER-BISWAS. "Bridging the Timescales of Single-Cell and Population Dynamics". In: *Physical Review X* 8.2 (2018), pp. 1–14.

[87] M. HASHIMOTO, T. NOZOE, H. NAKAOKA, R. OKURA, S. AKIYOSHI, K. KANEKO, E. KUSSELL, and Y. WAKAMOTO. "Noise-driven growth rate gain in clonal cellular populations". In: *Proceedings of the National Academy of Sciences* 113.12 (2016), pp. 3251–3256.

[88] D. G. HOEL and K. S. CRUMP. "Estimating the Generation-Time Distribution of an Age-Dependent Branching Process". In: *Biometrics* 30.1 (1974), pp. 125–135.

[89] K. B. ATHREYA and N. KEIDING. "Estimation Theory for Continuous-Time Branching Processes". In: *Sankhyā: The Indian Journal of Statistics, Series A (1961-2002)* 39.2 (1977), pp. 101–123.

[90] M. HOFFMANN and A. OLIVIER. "Nonparametric estimation of the division rate of an age dependent branching process". In: *Stochastic Processes and their Applications* 126.5 (2016), pp. 1433–1471.

[91] J. NEDELMAN, H. DOWNS, and P. PHARR. "Inference for an age-dependent, multitype branching-process model of mast cells". In: *Journal of Mathematical Biology* 25.2 (1987), pp. 203–226.

[92] O. HYRIEN. "Pseudo-likelihood estimation for discretely observed multitype Bellman–Harris branching processes". In: *Journal of Statistical Planning and Inference* 137.4 (2007), pp. 1375–1388.

[93] H. W. ENGL, C. FLAMM, P. KÜGLER, J. LU, S. MÜLLER, and P. SCHUSTER. "Inverse problems in systems biology". In: *Inverse Problems* 25.12 (2009), pp. 1–52.

[94] T. NOZOE, E. KUSSELL, and Y. WAKAMOTO. "Inferring fitness landscapes and selection on phenotypic states from single-cell genealogical data". In: *PLOS Genetics* 13.3 (2017), pp. 1–25.

[95] I. KUZMANOVSKA, A. MILIAS-ARGEITIS, J. MIKELSON, C. ZECHNER, and M. KHAMMASH. "Parameter inference for stochastic single-cell dynamics from lineage tree data". In: *BMC Systems Biology* 11.1 (2017).

[96] A. LLAMOSI, A. M. GONZALEZ-VARGAS, C. VERSARI, E. CINQUEMANI, G. FERRARI-TRECATE, P. HERSEN, and G. BATT. "What Population Reveals about Individual Cell Identity: Single-Cell Parameter Estimation of Models of Gene Expression in Yeast". In: *PLOS Computational Biology* 12.2 (2016), pp. 1–18.

[97] G. AMSELEM, C. GUERMONPREZ, B. DROGUE, S. MICHELIN, and C. N. BAROUD. "Universal microfluidic platform for bioassays in anchored droplets". In: *Lab on a Chip* 16.21 (2016), pp. 4200–4211.

[98] G. AMSELEM, P. BRUN, F. GALLAIRE, and C. N. BAROUD. "Breaking Anchored Droplets in a Microfluidic Hele-Shaw Cell". In: *Physical Review Applied* 3.5 (2015), pp. 1–5.

[99] P. ABBYAD, R. DANGLA, A. ALEXANDROU, and C. N. BAROUD. "Rails and anchors: guiding and trapping droplet microreactors in two dimensions". In: *Lab Chip* 11.5 (2011), pp. 813–821.

[100] R. DANGLA, S. LEE, and C. N. BAROUD. "Trapping Microfluidic Drops in Wells of Surface Energy". In: *Physical Review Letters* 107.12 (2011), pp. 1–4.

[101] D. J. COLLINS, A. NEILD, A. DEMELLO, A.-Q. LIU, and Y. AI. "The Poisson distribution and beyond: methods for microfluidic droplet production and single cell encapsulation". In: *Lab on a Chip* 15.17 (2015), pp. 3439–3459.

[102] L. LE CAM. "An approximation theorem for the Poisson binomial distribution". In: *Pacific Journal of Mathematics* 10.4 (1960), pp. 1181–1197.

[103] G. SAPORTA. *Probabilités, analyse des données et statistique*. Paris: Editions Technip, 2008.

[104] Q. ZHANG, T. WANG, Q. ZHOU, P. ZHANG, Y. GONG, H. GOU, J. XU, and B. MA. "Development of a facile droplet-based single-cell isolation platform for cultivation and genomic analysis in microorganisms". In: *Scientific Reports* 7.1 (2017), pp. 1–11.

[105] N. DHAR and J. D. MCKINNEY. "Microbial phenotypic heterogeneity and antibiotic tolerance". In: *Current Opinion in Microbiology* 10.1 (2007), pp. 30–38.

[106] O. PATANGE, C. SCHWALL, M. JONES, C. VILLAVA, D. A. GRIFFITH, A. PHILLIPS, and J. C. W. LOCKE. "Escherichia coli can survive stress by noisy growth modulation". In: *Nature Communications* 9.1 (2018), pp. 1–11.

[107] G. MANINA, N. DHAR, and J. MCKINNEY. "Stress and Host Immunity Amplify Mycobacterium tuberculosis Phenotypic Heterogeneity and Induce Nongrowing Metabolically Active Forms". In: *Cell Host & Microbe* 17.1 (2015), pp. 32–46.

[108] R. SAMUDRALA. "The Promise and Challenge of Digital Biology". In: *Journal of Bioengineering & Biomedical Science* 3.3 (2013), pp. 1–3.

[109] B. J. HINDSON, K. D. NESS, D. A. MASQUELIER, P. BELGRADER, N. J. HEREDIA, A. J. MAKAREWICZ, I. J. BRIGHT, M. Y. LUCERO, A. L. HIDDESSEN, T. C. LEGLER, T. K. KITANO, M. R. HODEL, J. F. PETERSEN, P. W. WYATT, E. R. STEENBLOCK, P. H. SHAH, L. J. BOUSSE, C. B. TROUP, J. C. MELLEN, D. K. WITTMANN, N. G. ERNDT, T. H. CAULEY, R. T. KOEHLER, A. P. SO, S. DUBE, K. A. ROSE, L. MONTESCLAROS, S. WANG, D. P. STUMBO, S. P. HODGES, S. ROMINE, F. P. MILANOVICH, H. E. WHITE, J. F. REGAN, G. A. KARLIN-NEUMANN, C. M. HINDSON, S. SAXONOV, and B. W. COLSTON. "High-Throughput Droplet Digital PCR System for Absolute Quantitation of DNA Copy Number". In: *Analytical Chemistry* 83.22 (2011), pp. 8604–8610.

[110] C. GUERMONPREZ. "Droplet-based Microfluidic Platform for Quantitative Microbiolgy". Thèse de doctorat. Université Paris-Saclay, 2016.

[111] H. SALMAN, N. BRENNER, C.-K. TUNG, N. ELYAHU, E. STOLOVICKI, L. MOORE, A. LIBCHABER, and E. BRAUN. "Universal Protein Fluctuations in Populations of Microorganisms". In: *Physical Review Letters* 108.23 (2012), pp. 1–17.

[112] RAINA M. MAIER. "Chapter 3 -Bacterial Growth". In: *Environmental Microbiology (Second Edition)*. Academic Press. Raina M. Maier, Ian L. Pepper, and Charles P. Gerba, 2009, pp. 37 –54.

[113] J. Q. BOEDICKER, L. LI, T. R. KLINE, and R. F. ISMAGILOV. "Detecting bacteria and determining their susceptibility to antibiotics by stochastic confinement in nanoliter droplets using plug-based microfluidics". In: *Lab on a Chip* 8.8 (2008), pp. 1265–1272.

[114] L. BARABAN, F. BERTHOLLE, M. L. M. SALVERDA, N. BREMOND, P. PANIZZA, J. BAUDRY, J. A. G. M. DE VISSER, and J. BIBETTE. "Millifluidic droplet analyser for microbiology". In: *Lab on a Chip* 11.23 (2011), pp. 4057–4062.

[115] G. SEZONOV, D. JOSELEAU-PETIT, and R. D'ARI. "Escherichia coli Physiology in Luria-Bertani Broth". In: *Journal of Bacteriology* 189.23 (2007), pp. 8746–8749.

[116] D. O'BEIRNE and G. HAMER. "Oxygen availability and the growth of Escherichia coli W3110: A problem exacerbated by scale-up". In: *Bioprocess Engineering* 23.5 (2000), pp. 487–494.

[117] M. E. COX and B. DUNN. "Oxygen diffusion in poly(dimethyl siloxane) using fluorescence quenching. I. Measurement technique and analysis". In: *Journal of Polymer Science: Part A: Polymer Chemistry* 24.4 (1986), pp. 621–636.

[118] M. A. HAMZA, G. SERRATRICE, M. J. STEBE, and J. J. DELPUECH. "Solute-solvent interactions in perfluorocarbon solutions of oxygen. An NMR study". In: *Journal of the American Chemical Society* 103.13 (1981), pp. 3733–3738.

[119] L. MAHLER, M. TOVAR, T. WEBER, S. BRANDES, M. M. RUDOLPH, J. EHGARTNER, T. MAYR, M. T. FIGGE, M. ROTH, and E. ZANG. "Enhanced and homogeneous oxygen availability during incubation of microfluidic droplets". In: *RSC Advances* 5.123 (2015), pp. 101871–101878.

[120] J. BARANYI and C. PIN. "Estimating Bacterial Growth Parameters by Means of Detection Times". In: *Applied and Environmental Microbiology* 65.2 (1999), pp. 732–736.

[121] M. J. WISER and R. E. LENSKI. "A Comparison of Methods to Measure Fitness in Escherichia coli". In: *PLOS ONE* 10.5 (2015), pp. 1–11.

[122] S. TAHERI-ARAGHI, S. D. BROWN, J. T. SAULS, D. B. MCINTOSH, and S. JUN. "Single-Cell Physiology". In: *Annual Review of Biophysics* 44.1 (2015), pp. 123–142.

[123] K. A. BRAKKE. "The Surface Evolver". In: *Experimental Mathematics* 1.2 (1992), pp. 141–165.

[124] T. C. MERKEL, V. I. BONDAR, K. NAGAI, B. D. FREEMAN, and I. PINNAU. "Gas sorption, diffusion, and permeation in poly(dimethylsiloxane)". In: *Journal of Polymer Science Part B: Polymer Physics* 38.3 (2000), pp. 415–434.

[125] R. DULBECCO. "Plaque Formation and Isolation of Pure Lines with Poliomyelitis Viruses". In: *Journal of Experimental Medicine* 99.2 (1954), pp. 167–182.

[126] A. VIGOUROUX, E. OLDEWURTEL, L. CUI, D. BIKARD, and S. VAN TEEFFELEN. "Tuning dCas9's ability to block transcription enables robust, noiseless knockdown of bacterial genes". In: *Molecular Systems Biology* 14.3 (2018), pp. 1–14.

[127] I. ARGANDA-CARRERAS, V. KAYNIG, C. RUEDEN, K. W. ELICEIRI, J. SCHINDELIN, A. CARDONA, and H SEBASTIAN SEUNG. "Trainable Weka Segmentation: a machine learning tool for microscopy pixel classification". In: *Bioinformatics* 33.15 (2017), pp. 2424–2426.

[128] J. SCHINDELIN, I. ARGANDA-CARRERAS, E. FRISE, V. KAYNIG, M. LONGAIR, T. PIETZSCH, S. PREIBISCH, C. RUEDEN, S. SAALFELD, B. SCHMID, J.-Y. TINEVEZ, D. J. WHITE, V. HARTENSTEIN, K. ELICEIRI, P. TOMANCAK, and A. CARDONA. "Fiji: an open-source platform for biological-image analysis". In: *Nature Methods* 9.7 (2012), pp. 676–682.

[129] W. D. DONACHIE and G. W. BLAKELY. "Coupling the initiation of chromosome replication to cell size in Escherichia coli". In: *Current Opinion in Microbiology* 6.2 (2003), pp. 146–150.

[130] L. ROBERT, M. HOFFMANN, N. KRELL, S. AYMERICH, J. ROBERT, and M. DOUMIC. "Division in Escherichia coli is triggered by a size-sensing rather than a timing mechanism". In: *BMC Biology* 12.1 (2014), p. 17.

[131] S. V. AVERY. "Microbial cell individuality and the underlying sources of heterogeneity". In: *Nature Reviews Microbiology* 4.8 (2006), pp. 577–587.

[132] A. AMIR. "Cell Size Regulation in Bacteria". In: *Physical Review Letters* 112.20 (2014), pp. 1–5.

[133] F. J. MASSEY. "The Kolmogorov-Smirnov Test for Goodness of Fit". In: *Journal of the American Statistical Association* 46.253 (1951), pp. 68–78.

[134] A. METRIS, S. M. GEORGES, M. W. PECK, and J. BARANYI. "Distribution of turbidity detection times produced by single cell-generated bacterial populations". In: *Journal of Microbiological Methods* 55.3 (2003), pp. 821–827.

[135] J. BARANYI and C. PIN. "A Parallel Study on Bacterial Growth and Inactivation". In: *Journal of Theoretical Biology* 210.3 (2001), pp. 327–336.

[136] J. KLEIN, S. LEUPOLD, I. BIEGLER, R. BIEDENDIECK, R. MUNCH, and D. JAHN. "TLM-Tracker: software for cell segmentation, tracking and lineage analysis in time-lapse microscopy movies". In: *Bioinformatics* 28.17 (2012), pp. 2276–2277.

[137] A. METRIS, Y. LEMARC, A. ELFWING, A. BALLAGI, and J. BARANYI. "Modelling the variability of lag times and the first generation times of single cells of". In: *International Journal of Food Microbiology* 100.1 (2005), pp. 13–19.

[138] J. M. VAN DIJL and M. HECKER. "Bacillus subtilis: from soil bacterium to super-secreting cell factory". In: *Microbial Cell Factories* 12.1 (2013), p. 3.

[139] C. E. HELMSTETTER, C. EENHUIS, P. THEISEN, J. GRIMWADE, and A. C. LEONARD. "Improved bacterial baby machine: application to Escherichia coli K-12." In: *Journal of Bacteriology* 174.11 (1992), pp. 3445–3449.

[140] J. ALDRICH. "R.A. Fisher and the making of maximum likelihood 1912-1922". In: *Statistical Science* 12.3 (1997), pp. 162–176.

[141] C. C. HEYDE and Y.-X. LIN. "On Quasi-Likelihood Methods and Estimation for Bunching Processes and Heteroscedastic Regression Models". In: *Australian Journal of Statistics* 34.2 (1992), pp. 199–206.

[142] C. OH, N. C. SEVERO, and J. SLIVKA. "Approximations to the Maximum Likelihood Estimate in Some Pure Birth Processes". In: *Biometrika* 78.2 (1991), pp. 295–299.

[143] S. MELEARD. *Modeles aleatoires en ecologie et evolution*. Springer. Collection: Mathematiques et Applications. 2016.

[144] K. B. ATHREYA and N. KAPLAN. "Convergence of the Age Distribution in the one-dimensional supercritical age-dependent branching process". In: *The Annals of Probability* 4.1 (1976), pp. 38–50.

[145] A. BREN, Y. HART, E. DEKEL, D. KOSTER, and U. ALON. "The last generation of bacterial growth in limiting nutrient". In: *BMC Systems Biology* 7.27 (2013), pp. 1–9.

[146] E. ZANG, S. BRANDES, M. TOVAR, K. MARTIN, F. MECH, P. HORBERT, T. HENKEL, M. T. FIGGE, and M. ROTH. "Real-time image processing for label-free enrichment of Actinobacteria cultivated in picolitre droplets". In: *Lab on a Chip* 13.18 (2013), pp. 3707–3713.

[147] J. CARLET, V. JARLIER, S. HARBARTH, A. VOSS, H. GOOSSENS, D. PITTET, and THE PARTICIPANTS OF THE 3RD WORLD HEALTHCARE-ASSOCIATED INFECTIONS FORUM. "Ready for a world without antibiotics? The Pensières Antibiotic Resistance Call to Action". In: *Antimicrobial Resistance and Infection Control* 1.11 (2012), pp. 1–13.

[148] J. B. LYNCH. "Multidrug-resistant Tuberculosis". In: *Medical Clinics of North America* 97.4 (2013), pp. 553–579.

[149] D. A. TADESSE, S. ZHAO, E. TONG, S. AYERS, A. SINGH, M. J. BARTHOLOMEW, and P. F. MCDERMOTT. "Antimicrobial Drug Resistance in Escherichia coli from Humans and Food Animals, United States, 1950–2002". In: *Emerging Infectious Diseases* 18.5 (2012), pp. 741–749.

[150] J. CONLY and B. JOHNSTON. "Where are all the new antibiotics? The new antibiotic paradox". In: *Canadian Journal of Infectious Diseases and Medical Microbiology* 16.3 (2005), pp. 159–160.

[151] A. RODRÍGUEZ-ROJAS, J. RODRÍGUEZ-BELTRÁN, A. COUCE, and J. BLÁZQUEZ. "Antibiotics and antibiotic resistance: A bitter fight against evolution". In: *International Journal of Medical Microbiology* 303.6 (2013), pp. 293–297.

[152] J. J. KERREMANS, P. VERBOOM, T. STIJNEN, L. HAKKAART-VAN ROIJEN, W. GOESSENS, H. A. VERBRUGH, and M. C. VOS. "Rapid identification and antimicrobial susceptibility testing reduce antibiotic use and accelerate pathogen-directed antibiotic use". In: *Journal of Antimicrobial Chemotherapy* 61.2 (2007), pp. 428–435.

[153] J. JORGENSEN and M. FERRARO. "Antimicrobial Susceptibility Testing: A Review of General Principles and Contemporary Practices". In: *Clinical Infectious Diseases* 49.11 (2009), pp. 1749–1755.

[154] S. POPPERT, A. ESSIG, B. STOEHR, A. STEINGRUBER, B. WIRTHS, S. JURETSCHKO, U. REISCHL, and N. WELLINGHAUSEN. "Rapid Diagnosis of Bacterial Meningitis by Real-Time PCR and Fluorescence In Situ Hybridization". In: *Journal of Clinical Microbiology* 43.7 (2005), pp. 3390–3397.

[155] W. STAMM and S. NORRBY. "Urinary Tract Infections: Disease Panorama and Challenges". In: *The Journal of Infectious Diseases* 183.s1 (2001), S1–S4.

[156] L. JIANG, L. BOITARD, P. BROYER, A.-C. CHAREIRE, P. BOURNE-BRANCHU, P. MAHÉ, M. TOURNOUD, C. FRANCESCHI, G. ZAMBARDI, J. BAUDRY, and J. BIBETTE. "Digital antimicrobial susceptibility testing using the MilliDrop technology". In: *European Journal of Clinical Microbiology & Infectious Diseases* 35.3 (2016), pp. 415–422.

[157] L. DERZSI, T. S. KAMINSKI, and P. GARSTECKI. "Antibiograms in five pipetting steps: precise dilution assays in sub-microliter volumes with a conventional pipette". In: *Lab on a Chip* 16.5 (2016), pp. 893–901.

[158] Y. WAKAMOTO, N. DHAR, R. CHAIT, K. SCHNEIDER, F. SIGNORINO-GELO, S. LEIBLER, and J. D. MCKINNEY. "Dynamic Persistence of Antibiotic-Stressed Mycobacteria". In: *Science* 339.6115 (2013), pp. 91–95.

[159] K. LEWIS. "Persister cells, dormancy and infectious disease". In: *Nature Reviews Microbiology* 5.1 (2007), pp. 48–56.

[160] N. COHEN, M. LOBRITZ, and J. COLLINS. "Microbial Persistence and the Road to Drug Resistance". In: *Cell Host & Microbe* 13.6 (2013), pp. 632–642.

[161] T. DÖRR, K. LEWIS, and M. VULIĆ. "SOS Response Induces Persistence to Fluoroquinolones in Escherichia coli". In: *PLoS Genetics* 5.12 (2009), pp. 1–9.

[162] J. M. ANDREWS. "Determination of minimum inhibitory concentrations". In: *Journal of Antimicrobial Chemotherapy* 48.s1 (2001), pp. 5–16.

[163] I. STOCK and B. WIEDEMANN. "Natural antibiotic susceptibility of Escherichia coli, Shigella, E. vulneris, and E. hermannii strains". In: *Diagnostic Microbiology and Infectious Disease* 33.3 (1999), pp. 187–199.

[164] A. ALIKHANI, F. BABAMAHMOODI, L. ALIZADEGAN, A. SHOJAEEFAR, and A. BABAMAHMOODI. "Minimal inhibitory concentration of microorganisms causing surgical site infection in referral hospitals in North of Iran, 2011-2012". In: *Caspian Journal of Intern Medicine* 6.1 (2014), pp. 34–39.

[165] E. F. MARLEY, C. MOHLA, and J. M. CAMPOS. "Evaluation of E-Test for Determination of Antimicrobial MICs for Pseudomonas aeruginosa Isolates from Cystic Fibrosis Patients". In: *Journal of Clinical Microbiology* 33.12 (1995), pp. 3191–3193.

[166] WORLD HEALTH ORGANISATION (WHO). *Model List of Essential Medicines*. 2017.

[167] C. TAN, R. PHILLIP SMITH, J. K. SRIMANI, K. A. RICCIONE, S. PRASADA, M. KUEHN, and L. YOU. "The inoculum effect and band-pass bacterial response to periodic antibiotic treatment". In: *Molecular Systems Biology* 8 (2012), pp. 1–11.

[168] J. B. DERIS, M. KIM, Z. ZHANG, H. OKANO, R. HERMSEN, A. GROISMAN, and T. HWA. "The Innate Growth Bistability and Fitness Landscapes of Antibiotic-Resistant Bacteria". In: *Science* 342 (2013), pp. 1068–1078.

[169] M. MUELLER, A. DE LA PENA, and H. DERENDORF. "Issues in Pharmacokinetics and Pharmacodynamics of Anti-Infective Agents: Kill Curves versus MIC". In: *Antimicrobial Agents and Chemotherapy* 48.2 (2004), pp. 369–377.

[170] S. FOERSTER, M. UNEMO, L. J. HATHAWAY, N. LOW, and C. L. ALTHAUS. "Time-kill curve analysis and pharmacodynamic modelling for in vitro evaluation of antimicrobials against Neisseria gonorrhoeae". In: *BMC Microbiology* 16.1 (2016), pp. 1–11.

[171] B. R. LEVIN and K. I. UDEKWU. "Population Dynamics of Antibiotic Treatment: a Mathematical Model and Hypotheses for Time-Kill and Continuous-Culture Experiments". In: *Antimicrobial Agents and Chemotherapy* 54.8 (2010), pp. 3414–3426.

[172] B. E. FERRO, J. VAN INGEN, M. WATTENBERG, D. VAN SOOLINGEN, and J. W. MOUTON. "Time-kill kinetics of antibiotics active against rapidly growing mycobacteria". In: *Journal of Antimicrobial Chemotherapy* 70.3 (2015), pp. 811–817.

[173] E. GULLBERG, S. CAO, O. G. BERG, C. ILBÄCK, L. SANDEGREN, D. HUGHES, and D. I. ANDERSSON. "Selection of Resistant Bacteria at Very Low Antibiotic Concentrations". In: *PLoS Pathogens* 7.7 (2011), pp. 1–9.

[174] E. TUOMANEN, R. COZENS, W. TOSCH, O. ZAK, and A. TOMASZ. "The Rate of Killing of Escherichia coli by P-Lactam Antibiotics Is Strictly Proportional to the Rate of Bacterial Growth". In: *Journal of General Microbiology* 132 (1986), pp. 1297–1304.

[175] N. SINGH and P. J. YEH. "Suppressive drug combinations and their potential to combat antibiotic resistance". In: *The Journal of Antibiotics* 70.11 (2017), pp. 1033–1042.

[176] M. BAYM, L. K. STONE, and R. KISHONY. "Multidrug evolutionary strategies to reverse antibiotic resistance". In: *Science* 351.6268 (2016), pp. 1–10.

[177] M. SCOTT, S. KLUMPP, E. M. MATEESCU, and T. HWA. "Emergence of robust growth laws from optimal regulation of ribosome synthesis". In: *Molecular Systems Biology* 10.8 (2014), pp. 747–747.

[178] H. R. MEREDITH, A. J. LOPATKIN, D. J. ANDERSON, and L. YOU. "Bacterial Temporal Dynamics Enable Optimal Design of Antibiotic Treatment". In: *PLOS Computational Biology* 11.4 (2015), pp. 1–22.

[179] P. GREULICH, J. DOLEŽAL, M. SCOTT, M. R. EVANS, and R. J. ALLEN. "Predicting the dynamics of bacterial growth inhibition by ribosome-targeting antibiotics". In: *Physical Biology* 14.6 (2017), pp. 1–17.

[180] E. TOPRAK, A. VERES, S. YILDIZ, J. M. PEDRAZA, R. CHAIT, J. PAULSSON, and R. KISHONY. "Building a morbidostat: an automated continuous-culture device for studying bacterial drug resistance under dynamically sustained drug inhibition". In: *Nature Protocols* 8.3 (2013), pp. 555–567.

[181] C. GUERMONPREZ, S. MICHELIN, and C. N. BAROUD. "Flow distribution in parallel microfluidic networks and its effect on concentration gradient". In: *Biomicrofluidics* 9.5 (2015), pp. 1–13.

[182] M. RADMAN. "SOS Repair Hypothesis: Phenomenology of an Inducible DNA Repair Which is Accompanied by Mutagenesis". In: *Molecular Mechanisms for Repair of DNA*. Basic Life Sciences. Boston, MA: Springer US, 1975, pp. 355–367.

[183] Z. BAHAROGLU and D. MAZEL. "SOS, the formidable strategy of bacteria against aggressions". In: *FEMS Microbiology Reviews* 38.6 (2014), pp. 1126–1145.

[184] D. G. ANDERSON and S. C. KOWALCZYKOWSKI. "Reconstitution of an SOS Response Pathway: Derepression of Transcription in Response to DNA Breaks". In: *Cell* 95.7 (1998), pp. 975–979.

[185] B. MICHEL. "After 30 Years of Study, the Bacterial SOS Response Still Surprises Us". In: *PLOS Biology* 3.7 (2005), pp. 1–3.

[186] J. COURCELLE, A. KHODURSKY, B. PETER, P. O. BROWN, and P. C. HANAWALT. "Comparative gene expression profiles following UV exposure in wild-type and SOS-deficient Escherichia coli." In: *Genetics* 158.1 (2001), pp. 41–64.

[187] S. FUJII, A. ISOGAWA, and R. P. FUCHS. "RecFOR proteins are essential for Pol V-mediated translesion synthesis and mutagenesis". In: *The EMBO Journal* 25.24 (2006), pp. 5754–5763.

[188] T. KATO, M. WATANABE, and T. OHTA. "Induction of the SOS response and mutations by reactive oxygen-generating compounds in various Escherichia coli mutants defective in the mutM, mutY or soxRS loci". In: *Mutagenesis* 9.3 (1994), pp. 245–251.

[189] A. AERTSEN, R. VAN HOUDT, K. VANOIRBEEK, and C. W. MICHIELS. "An SOS Response Induced by High Pressure in Escherichia coli". In: *Journal of Bacteriology* 186.18 (2004), pp. 6133–6141.

[190] T.-T. QIN, H.-Q. KANG, P. MA, P.-P. LI, L.-Y. HUANG, and B. GU. "SOS response and its regulation on the fluoroquinolone resistance". In: *Annals of Translational Medicine* 3.22 (2015), pp. 1–17.

[191] C. MILLER, L. E. THOMSEN, C. GAGGERO, R. MOSSERI, H. INGMER, and S. N. COHEN. "SOS Reponse induction by Beta-Lactams and Bacterial Defense Against Antibiotic Lethality". In: *Science* 305.5690 (2004), pp. 1629–1631.

[192] G. YIM, J. MCCLURE, M. G. SURETTE, and J. E. DAVIES. "Modulation of Salmonella gene expression by subinhibitory concentrations of quinolones". In: *The Journal of Antibiotics* 64.1 (2011), pp. 73–78.

[193] J. M. SCHOEMAKER, R. C. GAYDA, and A MARKOVITZ. "Regulation of cell division in Escherichia coli: SOS induction and cellular location of the sulA protein, a key to lon-associated filamentation and death." In: *Journal of Bacteriology* 158.2 (1984), pp. 551–561.

[194] J. WAGNER, P. GRUZ, S.-R. KIM, M. YAMADA, K. MATSUI, R. P. P. FUCHS, and T. NOHMI. "The dinB Gene Encodes a Novel E. coli DNA Polymerase, DNA Pol IV, Involved in Mutagenesis". In: *Molecular Cell* 4.2 (1999), pp. 281–286.

[195] R. T. CIRZ, B. M. O'NEILL, J. A. HAMMOND, S. R. HEAD, and F. E. ROMESBERG. "Defining the Pseudomonas aeruginosa SOS Response and Its Role in the Global Response to the Antibiotic Ciprofloxacin". In: *Journal of Bacteriology* 188.20 (2006), pp. 7101–7110.

[196] S. DA RE, F. GARNIER, E. GUÉRIN, S. CAMPOY, F. DENIS, and M.-C. PLOY. "The SOS response promotes qnrB quinolone-resistance determinant expression". In: *EMBO Reports* 10.8 (2009), pp. 929–933.

[197] H. GOTOH, N. KASARANENI, N. DEVINENI, S. F. DALLO, and T. WEITAO. "SOS involvement in stress-inducible biofilm formation". In: *Biofouling* 26.5 (2010), pp. 603–611.

[198] D. DAVIES. "Understanding biofilm resistance to antibacterial agents". In: *Nature Reviews. Drug Discovery* 2.2 (2003), pp. 114–122.

[199] J. W. BEABER, B. HOCHHUT, and M. K. WALDOR. "SOS response promotes horizontal dissemination of antibiotic resistance genes". In: *Nature* 427 (2004), pp. 72–74.

[200] M. ALAM, A. ALHHAZMI, J. DECOTEAU, Y. LUO, and C. GEYER. "RecA Inhibitors Potentiate Antibiotic Activity and Block Evolution of Antibiotic Resistance". In: *Cell Chemical Biology* 23.3 (2016), pp. 381–391.

[201] N. FRIEDMAN, S. VARDI, M. RONEN, U. ALON, and J. STAVANS. "Precise Temporal Modulation in the Response of the SOS DNA Repair Network in Individual Bacteria". In: *PLoS Biology* 3.7 (2005), pp. 1–8.

[202] Y. SHIMONI, S. ALTUVIA, H. MARGALIT, and O. BIHAM. "Stochastic Analysis of the SOS Response in Escherichia coli". In: *PLoS ONE* 4.5 (2009), pp. 1–7.

[203] P. WANG, X. ZHANG, L. WANG, Z. ZHEN, M. TANG, and J. LI. "Subinhibitory concentrations of ciprofloxacin induce SOS response and mutations of antibiotic resistance in bacteria". In: *Annals of Microbiology* 60.3 (2010), pp. 511–517.

[204] F. C. NEIDHARDT, P. L. BLOCH, and D. F. SMITH. "Culture Medium for Enterobacteria". In: *Journal of Bacteriology* 119.3 (1974), pp. 736–747.

[205] F. ST-PIERRE, L. CUI, D. G. PRIEST, D. ENDY, I. B. DODD, and K. E. SHEARWIN. "One-Step Cloning and Chromosomal Integration of DNA". In: *ACS Synthetic Biology* 2.9 (2013), pp. 537–541.

[206] S. C. CORDELL, E. J. H. ROBINSON, and J. LÖWE. "Crystal structure of the SOS cell division inhibitor SulA and in complex with FtsZ". In: *Proceedings of the National Academy of Sciences* 100.13 (2003), pp. 7889–7894.

[207] J. D. MCCOOL, E. LONG, J. F. PETROSINO, H. A. SANDLER, S. M. ROSENBERG, and S. J. SANDLER. "Measurement of SOS expression in individual Escherichia coli K-12 cells using fluorescence microscopy". In: *Molecular Microbiology* 53.5 (2004), pp. 1343–1357.

[208] C. F. ERROL, C. W. GRAHAM, S. WOLFRAM, D. W. RICHARD, A. S. ROGER, and E. TOM. *DNA Repair and Mutagenesis, Second Edition*. American Society of Microbiology, 1995.

[209] A. BATTESTI, N. MAJDALANI, and S. GOTTESMAN. "The RpoS-mediated general stress response in Escherichia coli". In: *Annual Review of Microbiology* 65 (2011), pp. 189–213.

[210] V. MOSBACH, L. POGGI, D. VITERBO, M. CHARPENTIER, and G.-F. RICHARD. "TALEN-Induced Double-Strand Break Repair of CTG Trinucleotide Repeats". In: *Cell Reports* 22.8 (2018), pp. 2146–2159.

[211] M. T. VITOR, S. SART, A. BARIZIEN, L. G. D. L. TORRE, and C. N. BAROUD. "Tracking the Evolution of Transiently Transfected Individual Cells in a Microfluidic Platform". In: *Scientific Reports* 8 (2018), pp. 1–9.

[212] S. SART, R. F. X. TOMASI, A. BARIZIEN, G. AMSELEM, A. CUMANO, and C. N. BAROUD. "Mapping Structure and Biological Functions within Mesenchymal Bodies using Microfluidics". In: *bioRxiv* (2019).

[213] Q. ZHANG and R. H. AUSTIN. "Physics of Cancer: The Impact of Heterogeneity". In: *Annual Review of Condensed Matter Physics* 3.1 (2012), pp. 363–382.

[214] Q. ZHANG, G. LAMBERT, D. LIAO, H. KIM, K. ROBIN, C.-K. TUNG, N. POURMAND, and R. H. AUSTIN. "Acceleration of Emergence of Bacterial Antibiotic Resistance in Connected Microenvironments". In: *Science* 333.6050 (2011), pp. 1764–1767.

[215] MARKETS AND MARKETS. *Microfluidics Market worth 27.91 Billion USD by 2023*. 2018.

[216] D. S. KONG, T. A. THORSEN, J. BABB, S. T. WICK, J. J. GAM, R. WEISS, and P. A. CARR. "Open-source, community-driven microfluidics with Metafluidics". In: *Nature Biotechnology* 35.6 (2017), pp. 523–529.

[217] M. LIPSITCH, R. S. SINGER, and B. R. LEVIN. "Antibiotics in agriculture: When is it time to close the barn door?" In: *Proceedings of the National Academy of Sciences* 99.9 (2002), pp. 5752–5754.

[218] D. L. SMITH, A. D. HARRIS, J. A. JOHNSON, E. K. SILBERGELD, and J. G. MORRIS. "Animal antibiotic use has an early but important impact on the emergence of antibiotic resistance in human commensal bacteria". In: *Proceedings of the National Academy of Sciences* 99.9 (2002), pp. 6434–6439.

[219] T. W. HERTEL and U. L. C. BALDOS. *Global Change and the Challenges of Sustainably Feeding a Growing Planet*. Springer International Publishing, 2016.

[220] L. ALEKSANDROWICZ, R. GREEN, E. J. M. JOY, P. SMITH, and A. HAINES. "The Impacts of Dietary Change on Greenhouse Gas Emissions, Land Use, Water Use, and Health: A Systematic Review". In: *PLOS ONE* 11.11 (2016), pp. 1–16.

université PARIS-SACLAY
ÉCOLE DOCTORALE
INTERFACES
Approches Interdisciplinaires : Fondements, Applications et Innovation

Titre : Etude de la variabilité de la croissance de bactéries en gouttes microfluidiques

Mots clés : Microfluidique, Bactéries, Croissance, Variabilité

Résumé : Cette thèse porte sur l'étude de la variabilité de la croissance de bactéries en gouttes microfluidiques. Dans un premier temps, la puce microfluidique utilisée au cours de la thèse est présentée. Elle permet d'encapsuler des bactéries individuelles dans 1500 gouttes d'un nano litre, et de suivre leur croissance en parallèle grâce à la mesure de leur fluorescence par microscopie. La relation entre fluorescence mesurée et nombre de bactérie est discutée, et plusieurs questions techniques, comme la variabilité de taille des gouttes, l'hétérogénéité de fluorescence des bactéries, sont mesurées et leurs conséquences sur les mesures de croissance quantifiées. Dans un second temps, nous développons un modèle probabiliste qui permet, à partir de la variabilité des temps de divisions des bactéries, de prédire la variabilité de croissance entre les gouttes. Pour ce faire, nous adaptons le modèle classique de Bellman-Harris. La distribution aléatoire du nombre initial de bactérie par gouttes, ainsi que les temps de divisions différents des premières générations de bactéries sont ajoutées au modèle pour l'adapter à notre système expérimental. Les contributions de ces différentes sources de variabilité à la variabilité inter-gouttes de croissance des populations de bactéries sont quantifiées, et le modèle permet bien d'expliquer la variabilité de la croissance entre les gouttes. Dans un troisième temps, nous proposons un schéma d'inférence pour résoudre le problème inverse, qui est de retrouver, à partir des courbes de croissance, la variabilité des temps de division des bactéries individuelles. Le modèle précédent ne peut être utilisé à cause des sources externes de variabilité, nous proposons donc un schéma d'inférence basé sur le suivi dans le temps de chacune des trajectoires des gouttes. Grâce à des simulations reproduisant les conditions expérimentales, nous prouvons que l'inférence est possible. Elle ne peut être appliquée à nos expériences en raison de la précision insuffisante de notre mesure de fluorescence. Enfin, la même puce micro-fluidique est utilisée pour quantifier l'action d'antibiotiques sur des bactéries, notamment la réponse Sos qui est induite lorsque l'ADN de la bactérie est endommagé. La technologie d'encapsulation en goutte est utilisée pour mesurer l'hétérogénéité de réponse des bactéries et la relier à leur capacité à survivre au stress dû à l'antibiotique, et à reformer une colonie.

Title : Studying the varibility of bacterial growth in microfluidic droplets

Keywords : Microfluidics, Bacterial Growth, Variability

Abstract : This thesis presents some results about the variability of the growth of bacteria in microfluidics droplets. In the first chapter, the microfluidic chip used throughout the PhD is presented. It allows to encapsulate bacteria in an array of 1.500 nano-liter sized droplets, and to follow their growth in each droplet in parallel through fluorescence microscopy. The link between the measured fluorescence and the number of bacteria in a droplet is discussed, and other technical questions are addressed, such as the variability in droplet size and the cell-to-cell fluorescence variability. Next, we develop a stochastic model to account for the observed variability of population size in the droplet during the exponential phase of growth. A well-known stochastic model, the Bellman-Harris model, is adapted to take into account the external sources of randomness due to our experimental system (initial distribution of bacteria per droplet, different division time of the first generations). They are taken into account, along with the effects of the cell-to-cell variability of division times in our model, which is successful to predict the variability observed in the microfluidics experiments. Then we tackle the inverse problem, which is to recover the cell-to-cell variability from the observation of the growth in droplets. We propose an inference scheme based on following each droplet in time. The deviation from pure exponential growth is linked back to the cell-to-cell variability, and this inference scheme is proven to be successful on simulations that mimic the experimental constrains. However, we cannot completely apply it to our experiments because of a lack of accuracy in our fluorescence measurements. Finally, we demonstrate how our chip can represent a gain of space and time to quantify the effect of antibiotics on a bacterial strain compared to classical susceptibility measurement methods. We also show how it can be used to study the variability of the SOS response of bacteria, which is a bacterial stress response induced when the DNA of the cell is damaged, and relate it to the ability to survive an antibiotic treatment.

Université Paris-Saclay
Espace Technologique / Immeuble Discovery
Route de l'Orme aux Merisiers RD 128 / 91190 Saint-Aubin, France