Principles of Distributed Biological Computation in Non-Growing Cells

A rational architecture for chemical wires

Author: Carlos Toscano Ochoa

TESI DOCTORAL UPF / year 2021

THESIS SUPERVISOR
Jordi Garcia-Ojalvo
Department of Experimental and Health Sciences

A mis padres, sin cuyo incondicional apoyo no hubiera podido llegar a donde he llegado.

Abstract

Abstract

One of the branches of synthetic biology pursues the creation of a biological computer. Many have successfully implemented simple logic circuits rewiring cellular components. However, these circuits may aggressively compete for energy and resources from the host organism. Distributed computation proposes to split the circuit into smaller pieces and distribute them among several strains sharing information using chemical wires, usually based on quorum sensing systems. Nevertheless, most published research has been focused in the creation of proofs of concept, evading the fundamental question of the design constraints self imposed due to the use of chemical wires. Here we propose a unified and rational wire architecture consisting in three strains: one that emits a communication molecule called biobit, one that degrades and one that senses it. Our model predicts the strain ratios that allow desired behaviours, and shows how a wire is even able to work as a timer.

Resumen

Una de las ramas de la biología sintética persigue la creación de un computador biológico. Muchos han implementado con éxito circuitos lógicos simples reconectando componentes celulares. Sin embargo, estos circuitos pueden competir agresivamente por la energía y los recursos del organismo anfitrión. La computación distribuida propone dividir el circuito en partes más pequeñas y distribuirlas entre varias cepas que comparten información utilizando cables químicos, generalmente basados en sistemas de quorum sensing. Sin embargo, la mayoría de las investigaciones publicadas se han centrado en la creación de pruebas de concepto, evitando la cuestión fundamental de las limitaciones de diseño autoimpuestas debido al uso de cables químicos. Aquí proponemos una arquitectura de cables unificada y racional que consta de tres cepas: una que emite una molécula llamada biobit, otra que la degrada y otra que la detecta. Nuestro modelo predice las frecuencias relativas de cepas que permiten los comportamientos deseados y muestra cómo un cable puede funcionar como un temporizador.

Acknowledgements (Español)

Espero no olvidar a nadie. Empezando por alguien muy muy importante: mi compañera y amiga Letícia, con quien he compartido momentos tan divertidos a lo largo de los años y de los lugares, y a su Jordi, un tipo que cuando durmió lo que le tocaba se dejó conocer, te alegraba con su humor absurdo. A Marçal, un amigo fricazo que redefinió lo que significa ser el abogado del diablo y el pensamiento crítico (ojalá nunca me toques de reviewer) y a Claudia, que ya podríamos haber salido más antes más veces. A Rosa, con quien compartí momentos relajados de conversación profunda y la que puede que sea mi conversación más larga en catalán caminando por Lyon. A Elena Abad, que me haciste sentir acogido en el grupo en el que acababa de entrar. A Pablo Ruiz, porque las conversaciones contigo siempre fueron divertidas y estimulantes y disfruté mucho de tu divertida calma. A Pablo Casaní porque invocas el buen rollo en un radio de 10 metros y nos lo pegas (además de la repostería valenciana). A Keith por nuestras charlas en las que me acompañabas a fumar y discutíamos de las sutilezas del lenguaje (de cómo te molestan los bichos jeje). A Maria por su capacidad de estar siempre feliz y pegártelo. A Gang, con quien pude comunicarme finalmente al final de la tesis, poque resultaste ser interesante y divertido. A Gabriel por las cuan-

tas conversaciones sesudas que mantuvimos antes de que llegara el virus. A Leila por unas cuantas conversaciones en las que me dejaste con más ganas de te apuntaras a más quedadas (y mira que yo he ido a bien pocas desde que empezó la pandemia). A Cristina la nueva técnico porque además de ser una persona dulce, eres una increíble profesional que resolvía problemas. A Josep Mercadal por las conversaciones sobre matemáticas y música (¡hay que ver lo que tardaste en empezar a responder a los buenos días!). A Francesc por las conversaciones sobre percepción, y por lo perplejo que me quedé contigo y tu capacidad de hacer ciencia. A mi *esclavo* Nº1 Edu, porque no me he podido reír más en mi vida, y por las charlas científicas y las caminatas, y los desfases. Y a mi *esclavo* Nº2 Joan, pues casi que por lo mismo jajaj. Qué bueno fue el año que compartí con vosotros, y qué me gusta que aprendierais lo básico de un laboratorio. A mi jefe Jordi, por darme la oportunidad de ver cómo era el mundo académico y por estar siempre con semblante amable.

A mis propios compañeros originales de iGEM Sevilla 2011 (Arcanum Project). Porque es ahí donde está el germen de esta idea que no ha dejado de obsesionarme desde hace una década. Gracias por permitirme desarrollarlo y así poder llevar el modelo lo más lejos que he podido.

A Nastassia por nuestros cotilleos y nuestros paseos. A Júlia Domingo por los pitis y las birras y las charlas de ciencia. A Marc Talló por esos cafés con cuenta gota que nos hemos tomado, las charlas de proyectos creativos, y los desfases en los que me has embolado. A Dimitri, también por los momentos de cruzarnos en horas intempestivas y las conversaciones sobre biología sintética. A Raúl, con quien coincido: mucho tardamos en conocernos, y qué conversaciones. Es-

pero que su camino *preper* siga recto. A Víctor por las conversaciones de friqueo, que seguro que volverán en breve. A Aina por las conversaciones profundas, y el libro prestado que empieza a fosilizarse en mi estantería. Espero que te vaya bien por tu pueblo. A Salva Durán por ser una persona a la que jamás pensé que llegaría a untarle la calva de azul, y sin embargo mira la de vueltas que da la vida. A los chicos de iGEM 2016, si no me dio un chungo era porque me lo pasaba bien con vosotros y la experiencia vivida fue genial. De ellos, en particular, a Arianne Bercowsky, a la que llamaba la bestia porque considero que lo eres. Llegarás lejos. Y a su relajado y amigable Joan. A Antonia, con quien me encantó compartir las charlas semanales sobre cómo mejorar ese robot de Lego, y por reconciliarme con la sobrasada.

A Sergio, porque además de haberme ayudado a ponerme en forma en el final de mi doctorado, te has convertido en un buen amigo. Me temo que tendré que volver a volver a perder peso, que escribir esto me ha pasado factura.

Al Onma y la Helena, que me acogieron primero los dos cuando llegué a la ciudad, y luego ella durante un par de añitos. Y qué bien me lo pasé con vosotros.

A Jesús, que has sido un auténtico *brother from another mother* cuando estábamos los dos en aquel periodo tan turbio como delirante y que no cambio por nada del mundo. ¡Qué año, por Dios! Y a Tania cuando llegó después, que me tuvo que soportar hasta que me aceptó (con lo rarísimo que soy jeje).

Al Señor Doctor Ilustrísimo Don Jordi Piñero, que ha sido una amistad de esas imperecederas que apareció aquí en Barcelona de repente cual Pokémon mítico, y con el que he tenido, tengo y espero seguir teniedo tanto la amistad como nuestros momentos de friqueo

del bueno sobre matemáticas. Y a su señora Dolors, que a veces me coge el teléfono y mantenemos unas charlas muy salás, y que la pandemia no me ha permitido conocer en persona más de lo que me hubiera gustado.

Por supuesto, a Lorena. Has sido mi hermana mayor aquí en esta ciudad. El apoyo, los cuidados y el cariño que tú me has dado han sido absolutos. Que me habéis dado. Que ahí está Héctor también, con quien me lo he pasado genial y he compartido momentos geniales. Además, casi todo lo que sé de biología molecular te lo debo a ti. Me parte el corazón no poder verte tanto ahora con esto del virus. Qué me gustaría volver a los cigarritos fuera en el PRBB. A ir a buscarte a por un café. En estos últimos tiempos, a bajarnos a charlar y despacharnos a gusto criticando al gobierno. Lo primero que haré en cuanto se pueda, va ser coger el coche e ir a verte. A veros, que ahora sois 3. Qué me gustó participar de tu embarazo y conocer a Martina, que la próxima vez que la vea podré hablar con ella. Y gracias a tus padres, que también me acogieron con los brazos abierto. Muchas gracias, Lorena. Te quiero muchísimo.

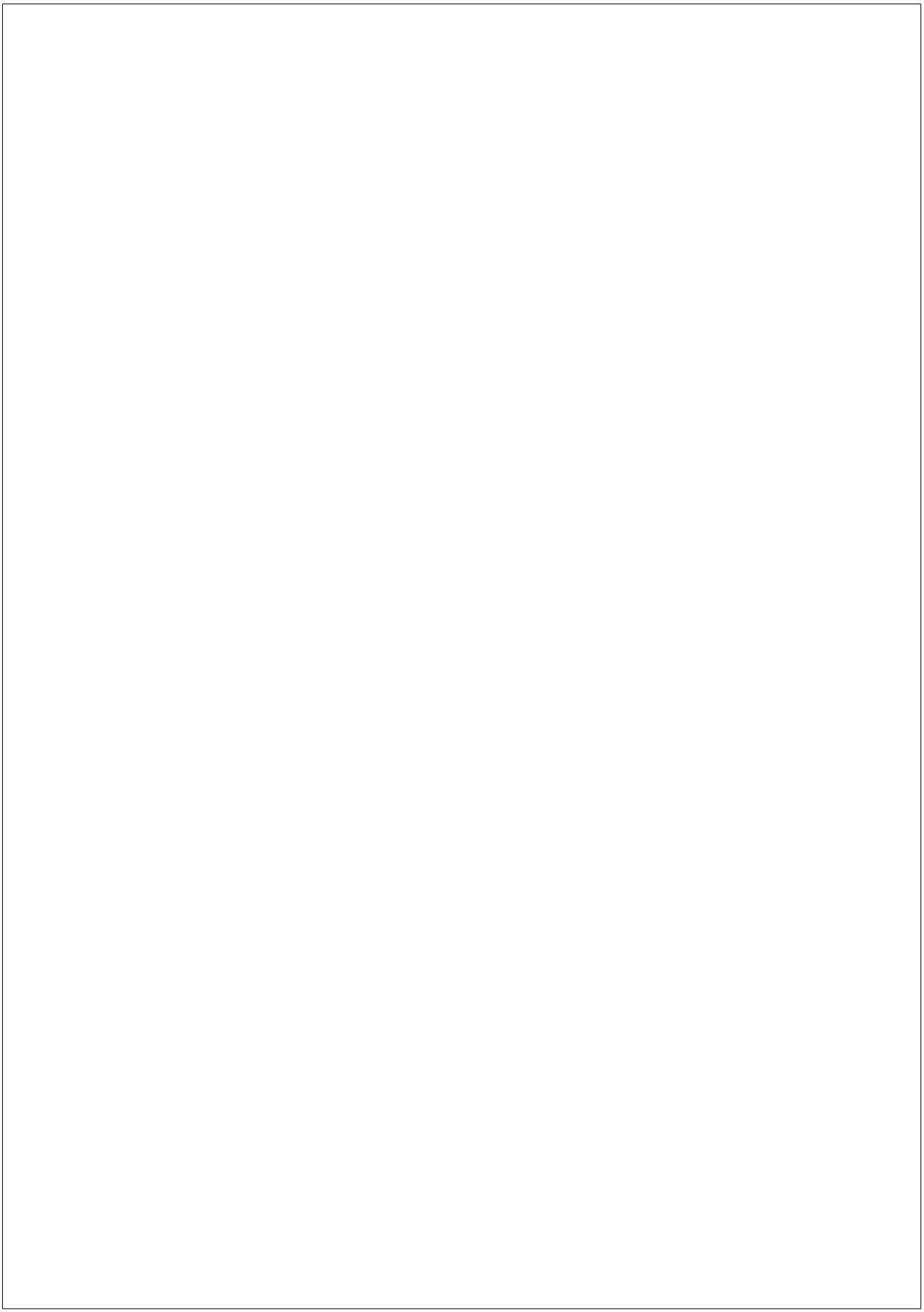
Mi familia tenía que estar por aquí también. A veces creo que no os merezco. Muchas gracias, Mamá y Papá, por el apoyo a lo largo de estos años. Vosotros sabéis bien que han sido muy complicados para mí, y habéis atravesado la península cada vez que ha hecho falta. Os he extrañado mil veces, y tengo ganas de poder veros más. Me gusta que seais mis padres, soy un afortunado. Lo mismo digo de mi hermano, que al principio de mi doctorado hablábamos de higos a brevas, y ahora estamos hablando todas las semanas varias veces. Gracias por las visitas y por estar conmigo durante los momentos buenos y los no tan buenos. Me muero de ganas de volver

a tener tiempo para pasarlo contigo. Y con mi nueva cuñada Lucía, que me cae genial y me falta el tiempo para conocerla más. Y gracias también a mi abuela Pepa, que nos llamamos todas las semanas, y todas las semanas nos decimos un te quiero desde que me fui de Sevilla. Gracias a Neska, por acompañarnos a todos, y acompañar a mis padres cuando se quedaron solos en la casa. Y gracias Lhasa por continuar acompañándolos cuando Neska nos dejó.

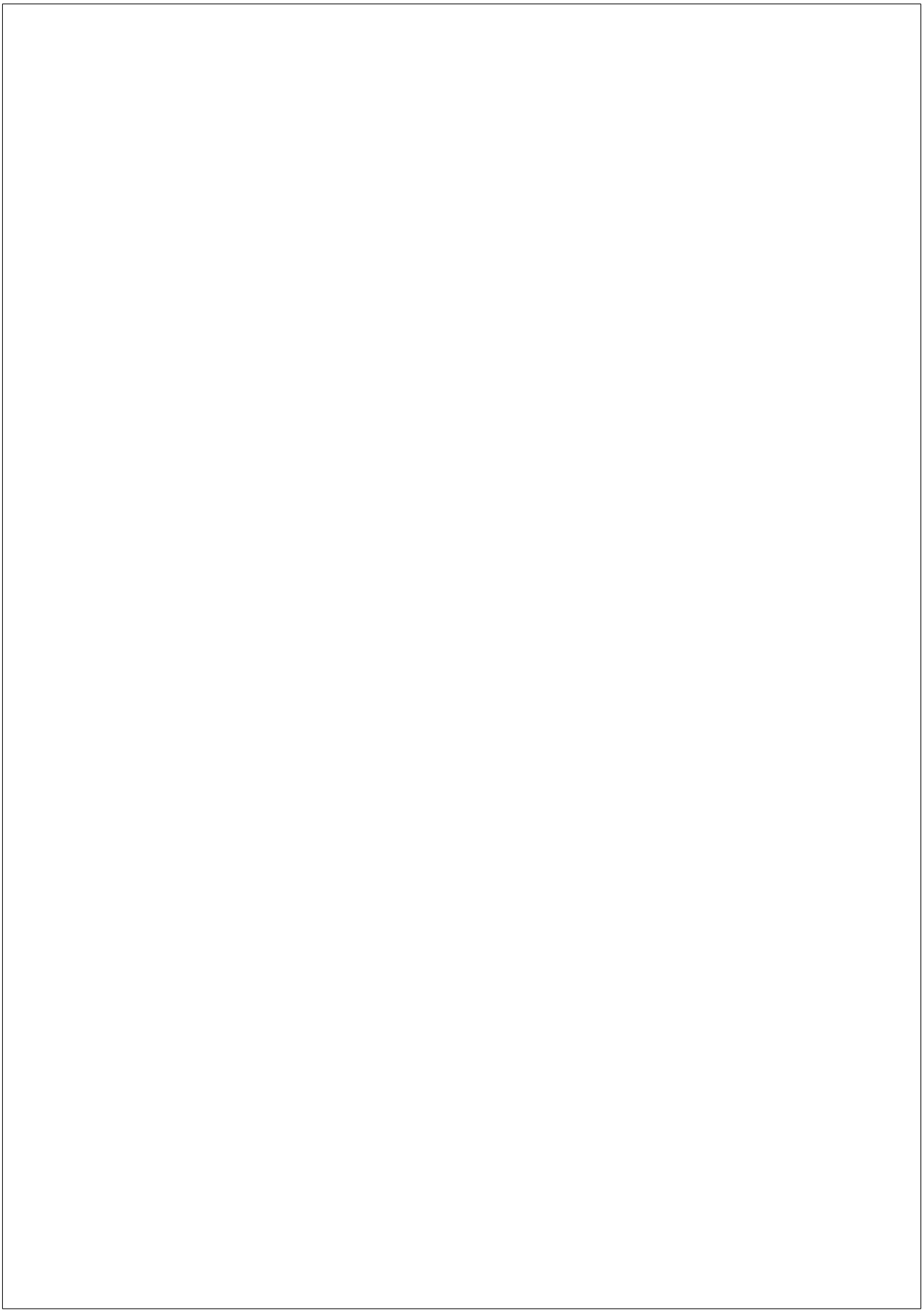
Y por supuesto gracias a ti, Manuel. Porque todavía no me creo la suerte que tengo de tenerte conmigo. La vida contigo es verdadera, es fácil, y es bonita. Y que la vida sea así no tiene precio. Me has ayudado muchísimo en los últimos años del doctorado, que han sido muy duros. Y tú lo has hecho todo más sencillo. ¡Y por supuesto me quedo contigo después de que me den el título! ¡Te quiero muchísimo, moreno! A ti y al perro llamado Alfredo con el que juego todas las noches!

A todos vosotros que de alguna forma me habéis acompañado: muchas gracias, porque lo he notado. ¡Gracias!

Contents



List of Figures



List of Tables

Chapter 1

INTRODUCTION

1.1. The dawn of synthetic biology

It is usually the case of most [*synthetic*] biology students to think that synthetic biology started with the iGEM competition[1], or with a couple of papers that appeared twenty years ago (I am talking about the *repressilator* and the toggle switch, but we will discuss these later). We will got through these when it proceeds. But if we naively approach the pure concept of synthetic biology, as any undergrad student would do, maybe the definition that we can find in Wikipedia could be a basic, but yet complete, one: *Synthetic biology is a multidisciplinary area of research that seeks to create new biological parts, devices, and systems, or to redesign systems that are already found in nature.* Note that this definition is not precisely constraining *what* it does mean when it refers to *systems*, beyond the given and obvious constrain of being a *biological* system.

[1] iGEM is the acronym for international Genetically Engineered Machine competition (www.igem.org).

Thus, we could perfectly be talking about creating new organisms from scratch, or rather we could be talking about new approaches for implementing *things* that we certainly do not identify as living beings. Whereas the idea of creating an organism from scratch is self-explanatory, and we can find teams trying to create a synthetic or minimal cell [65, 55], the other alternative is more cryptic. A good couple of examples are the implementation of systems for information storage [79, 28, 29, 83], or the creation of a RGB color perception and screen output in bacteria [46].

But I was creating some anticipation about the origins of *synthetic biology*. When did all of this start? Well, it may surprise you that the first use of the terms was in a book published in 1910 by Stéphane Leduc named *"Théorie Physico-Chimique de la vie et générations spontanées"* [2]. And the terms *synthetic biology* was not mentioned haphazardly. There was a whole chapter directly named *La biologie synthétique* . That chapter started orbiting the idea of comparing Biology with the development of other sciences such Chemistry. In 1910, Chemistry was starting to master the synthesis of several compounds. And the author, without any prior prejudice, was claiming in that chapter that, in the first place, Biology should complete its descriptive quest. And only after that, it would be possible to attack the problem of mastering what he for the very first time defined as synthetic biology. What would he meant by mastering synthetic biology? Let *Monssieur* Leduc explain himself:

> *The synthesis of life cannot be the sensational discovery that this expression evokes. If evolution has been accomplished as it is currently conceived,* ***the synthesis of life can only begin with intermediate productions between***

the mineral kingdom and living beings, having only a few rudimentary attributes of life, to which laboriously, little by little, others will be added through actions which, so to speak, will gradually make the first productions of synthetic biology evolve.

Well, how about that? Remember the date, dear reader: 1910. Technically, Leduc does not refer exactly to what we think about synthetic biology today. For us, it is a field of study lying in between science and engineering. For him it was a particular achievement totally in line with the previously mentioned project about minimal artificial cells. But let us be honest: it is fundamentally a matter of semantics. I am pretty sure that if Leduc comes back to life again, he would perfectly identify our field of study to what he defined as synthetic biology.

I am particularly fascinated with his mention to *attributes of life*, and the idea of *adding* these attributes for *summoning* life, as if they would be necessary conditions for synthetic life to emerge. And yes, most of what us *sciengineeres* do has nothing to do with creating artificial life. But just change the word life for living systems, and you get an equivalent statement.

Thus, synthetic biology seeks the rational design and implementation of systems with arbitrary desired behaviours by combining biological entities. It might be pretentious to say, but I think that this definition encapsulates whatever project or idea that you could imagine within the framework of synthetic biology.

1.2. Emulating engineering with synthetic biology

If the concept of synthetic biology appeared in the early twentieth century, why is that we place synthetic biology in the early 2000s in the collective imagination? Well, there is a big gap between conceiving the idea and being able to finely manipulate living matter. It required, as Leduc predicted, a first step consisting in describing biology. The connection between genetics and evolution with the Modern Synthesis starting with Fisher [1], Haldane [85] and Wright [147] in the early twentieth century. The structural characterisation of genetic material by Watson and Crick [142], or the first example of gene regulation with the Lac operon by Jacob and Monod [70] would happen in the mid-twentieth century. Then, it would be in the seventies when the recombinant DNA technology would be developed [69], bringing the capability of enzymatically modifying DNA and placing it inside cells. This technical achievement, together with the development of polymerase chain reaction [99] or the ability to sequence DNA [122] allowed, so to speak, the rise of genetic engineering.

And it was only then that the first attempts to fulfil Leduc's predictions with the first classical and archetypal examples of synthetic biological systems would appear in the year 2000. The *repressilator*, by Elowitz (2000) [40], a system consisting of three transcription factors that repress each other in a cyclical manner, producing oscillations in the gene expression, visible through the expression of a fluorescent reporter like GFP. Also, the toggle-switch developed by Gardner also in 2000, showed the implementation of a bi-stable system by using two transcription factors that mutually inhibit each other [52]. Other works

were focused on the understanding of the basic rules of homeostasis of genetic circuits [12].

After this, the study of gene networks and their properties [124, 96, 11], shed some light in the complex problem or gene regulation. Network motifs studies increased the understanding of the capabilities of gene regulatory networks and how these systems are able to process information [91, 92]. The standardization of biological components [41] and the iGEM contest led to the world-wide popularization of synthetic biology among students interested in genetic engineering.

At this point, we can talk about what I consider the two main branches of synthetic biology. The first branch, initiated by Drew Endy in the late 2000s, is very focused in the technical aspects, such defining and characterizing the DNA sequence that represents a minimal biological entity with function [42, 22]. Also, this branch is responsible for the development of the popular BioBrick assembly standard, a method that allows to systematically combine DNA sequences using a single protocol and four restriction enzymes [125]. The iGEM contest would be the pinnacle and the visible face of this synthetic biology branch, and its original purpose was to gradually increase the number of available BioBricks (of minimal DNA pieces), and show what can be achieved using them.

The other branch of synthetic biology is the academical one. This is what the reader will see if he/she goes to a synthetic biology congress. There, nobody is interested in standardization, and no one explicitly uses BioBrick technology. Which works could I cite here? Literally anything that you can find in Google Scholar using the key words "synthetic biology" will be categorized in this branch almost surely. The already mentioned *repressilator*, the toggle switch(es), RGB-color

perception, and any other work referenced in here will probably be a good example of this other branch. If the iGEM-branch was interested in standardization, this other branch is more interested in capabilities and proves of concept. This does not mean that we cannot see some approximation from this branch towards the standardization, and it is common to see that recent advances in academic synthetic biology are adapted and used in iGEM contest. For instance, under the name PERSIST, we can find a collection of RNA motifs and allow the bioengineer to implement behaviours directly in RNA molecules [33]. Why was it developed in the first place? Because some circuits implemented in mammals showed a decrease in its performance when they were implemented in mammal cells weeks after being transformed. The original work shows a toggle switch implemented in RNA molecules that are constitutively expressed, and the switch behaviour depends on the interactions between the RNA molecules. When compared with a toggle switch done using transcription factors, the bistable property had progressively been loosened with time due to the gene silencing experienced by DNA methylation, due to the lack of gene expression. This was solved by using genes that were constitutively expressed and with this RNA motifs encoding the behaviour. There was a problem (decrease in performance due to gene silencing), and now there is a solution that implements the same behaviour (toggle switch) using another approach (RNA). The final result is more robust behaviours. There is no explicit intention in standardization, though the RNA motifs can be easily standardized.

It will take some years (at least one decade in my opinion) to see the results of this incremental work done by the two branches. In other words: we are far away from developing something as clean and ro-

bust as electronics. On the other hand it is comprehensible, as our biological circuits have to be interpreted by organisms that we do not fully understand.

1.3. Computation meets synthetic biology

Any natural system can be seen as a computing machine: it has the ability to process internal and external information, and generate an output, a response. This is a free variation of the machine conception of the organism, proposed by Descartes in the seventeenth century. In summary: organisms are *like* machines, or organisms directly *are* machines [13]. It is not very difficult to jump from machine to computing machine. The Von Neumann architecture, which establishes the paradigm for most of the machines that surround us proposes a computer with input systems, output systems, and a central processor unit with access to memory in between. That being said, it is easy to interpret an organism as a computing machine. Any organism is able to sense its environment, process that information and alter its internal physiological state in a transient or permanent way (that could be interpreted as memory), and emit an output response. That being said, there are some philosophical works that directly critique the machine conception of the organism [103, 17, 105, 104]. One of these works, by Nicholson (2019), throws a valid criticism: machines are extrinsically purposive whereas organisms are intrinsically purposive [103]. I think that this appreciation will be valid until artificial machines reach a complexity so it will be difficult to guess if its purposiveness is extrinsic

or intrinsic [2]. Any ways, these philosophical questions do not paralyze the ubiquitous [computing] machine conception of the organism.

The consequence of not being able to dissociate the realm of computation from the realm of biology is what has pushed bioengineers to merge both realms in different ways. One of the approaches to such merge was using biological entities to explore and expand computation capabilities, with the use of DNA to perform computation or directly solve problems [73, 14, 28, 62, 3, 4]. In other approach, this time more theoretical, others preferred to explore membrane computing, a field that proposes to exploit the inner cell membrane traffic to perform computation [115, 116]. Some of these approaches have proved, at least in theory, that some extremely hard NP problems, could be solved as P problems, which is a strong motivation to keep merging computation with biology.

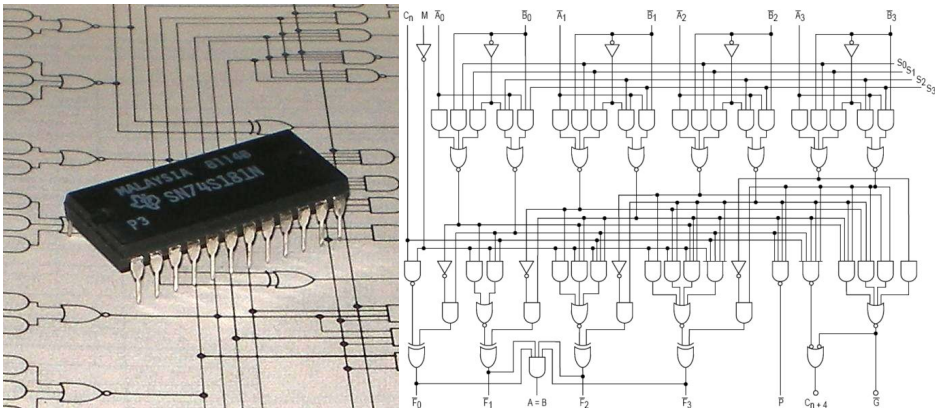


Figure 1.1: **Example of ALU 74181, showing its internal combinatorial logic.**

[2] I do not even know if such a test could exist, but it would be like a purposiveness Turing test.

And there is, of course, the obvious approach: trying to directly emulate standard computing machines using biological entities. Before I was mentioning the Von Neumann architecture, and saying that it possess a computing processor unit (CPU). One of the elements of a CPU is the arithmetic logic unit (ALU). An ALU is a unit that has implemented a combinatorial logic to give the ALU a specific purpose[3] (Figure 1.1). This internal combinatorial logic is usually a system of logic gates connected through wires. A logic gate is an object with that accepts at least one digital input and generates a digital output depending on some arbitrary internal function. For instance, the AND logic gate accepts two inputs and gill generate an output equal to 1 if and only if both inputs are equal to 1 simultaneously. And the OR logic gate only requires at least one input to be equal to 1 (not necessarily both) for emitting an output equal to 1. Just imagine that combining logic gates, we can manipulate input signals composed by strings of several bits (being each bit equal to 1 or 0), and generate an arbitrary output.

Being logic gates so simple to conceive, and yet so powerful when combined, it is natural that bioengineers have tried to implement them using biological entities. There are several approximations to implementations of logic gates [94, 144, 78, 6, 97, 15], and these planted the seed of the development of biological computing machines. One particular work by Adison Wong (2014) shows the scope of this approach by creating a half adder within a single cell [146] (Figure 1.2). A half adder is a circuit that allows to add binary ciphers (Fig. 1.2, B, left). This circuit requires that the cell must host two logic gates: an AND, and an exclusive OR (XOR). This is achieved by rewiring several

[3] Thus it is extrinsically purposive.

transcription factors with the ability to sense arabinose and ramnose, two carbohydrates. The presence or absence of both will be the representation of the two possible bit states: 1 or 0, respectively. The output was red fluorescence in the case of the sum (mod 2) of the two input bits, and green fluorescence in for representing the carry. Fluorescence levels for each fluorescent protein was measured using flow cytometry (histograms are shown in Fig. 1.2, C).

Projects such this last one (half adder implementations) might encourage the reader to think that *any* circuit might be implemented within a cell. Well, an obvious limitation relies in fact that genetic programs require a minimum space for the required DNA, and cell volume is very limited. But it is not necessary to reach such limits to encounter implementation difficulties. Researchers in the field soon noticed that any computer/machine implementation based on genetic networks within a single cell would then be unfeasible. Because our own implementations compete for energy and resources, our circuits can affect the host cell, and thus affect themselves[26, 149]. This gene load or metabolic burden is an unavoidable fundamental problem when implementing in single cells, and the best that we have been able to do is to characterize or predict it [25, 106]. Some researches advocate for pushing forward the single-cell circuit paradigm by merging circuits based on gene expression and transcription factors with metabolic networks [58], or by taking advantage of already existing endogenous circuits [100].

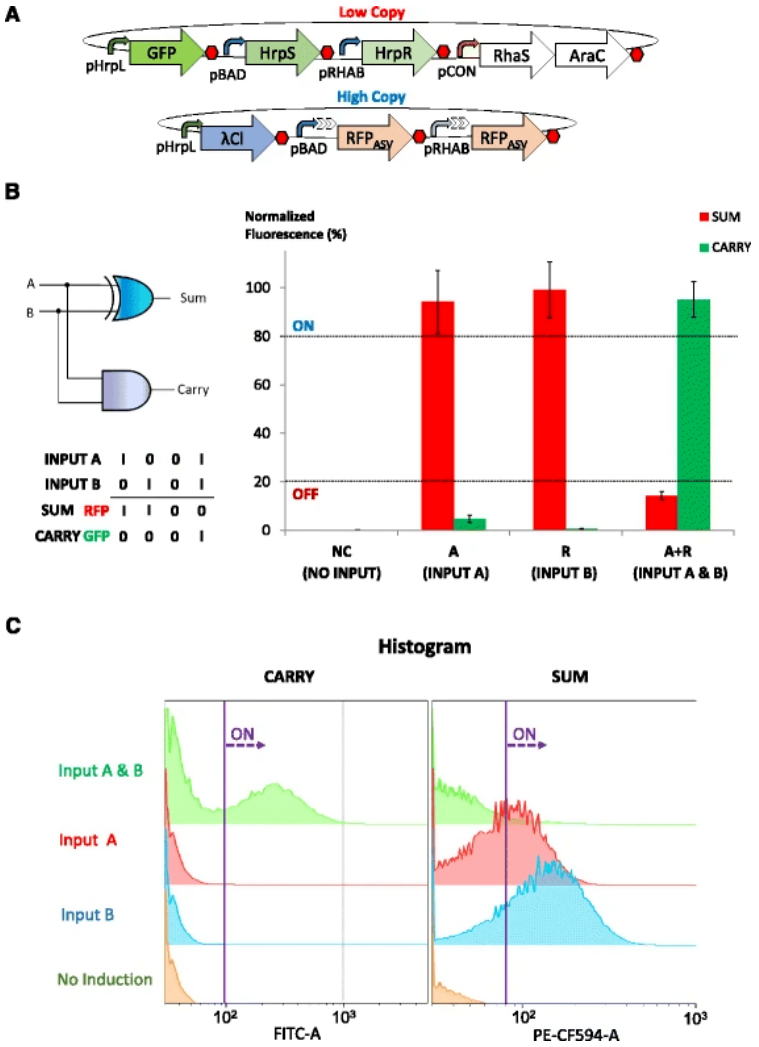


Figure 1.2: **Single cell half adder implementation and results. From Wong (2014) [146].**

1.4. Divide et impera: distributed biological computation

These observations, together with the fact that actual living organisms have a division of labour between cells, and previous works about microorganisms consortia engineering [18, 107], pushed the conceptual framework towards the development of distributed computation systems, also referred as consortia computation. The title of this section, the latin locution *divide et impera*, attributed to Philip II of Macedon, emphasizes the idea of first dividing a community to being able rule over it. A bit warlike, but certainly appropriate.

Under this approach, single pieces of a circuit, like logic gates or little networks motifs are implemented in single cells, and then connected among them by using chemical wires to transfer information [132, 119, 127, 66, 74, 88]. The use of cell-cell communication has not been restricted to the idea of mixing different strains, but also to give to one strain a single circuit that, by means of cell-cell communication using diffusible molecules, can make appear collective behaviours [31, 114, 30, 89, 151]. One of the advantages of distributed computation approaches are the maximization of the genetic parts recycling [87], and the avoidance of metabolic burden saturation, unlocking the possibility of creating scalable systems. Almost all chemical wires in these contexts are based on diffusible compounds, being quorum sensing systems from Gram-negative bacteria the most prominently used [109]. But there has been other alternative implementations such using yeast and sex hormones as chemical wires [136], or the more original approach of using bacterial conjugation for wiring [57].

But distributing the logic labour among several strains and wiring

them with chemical wires such quorum sensing is not a definitive solution. If quorum sensing Gram negative bacteria is used, and it is the most widely used, we have a couple of problems. In two works previously mentioned by Solé (2013) [128] and Macía (2014) [87], a problem is noted: in the classical implementation of an electronic ALU, wires are physically isolated, whereas in a distributed circuit, chemical wires throuw their chemical messengers to the media. This implies that, if we want our combinatorial logic circuits to work, each chemical messenger **must** be chemically unique. This imposes a first restriction: when combinatorics and *unique chemical compounds* is not something that works well when numbers get bigger and bigger. We certainly have a big catalogue of quorum sensing molecules, but it is nothing if we want to create an ALU like the one depicted in Fig. 1.1.

Is that the only problem? I am afraid this is not the case. The ubiquitous use of quorum sensing molecules from Gram negative bacteria has a problem associated: cross talk [110, 93]. This effect happens when the quorum sensing molecule from one species is able to interact and induce the receptor for a different molecule. Now remember that all these wiring molecules must be mixed in a liquid. Under that circumstance, orthogonality (the ability of one cell-cell communication system to not interact with other cell-cell communication system) is more an ideal than a real feasible path. One possible solution to the wiring problem is the spacial segregation of the different strains coexisting in the culture [86, 132].

1.5. Updating the wiring problem

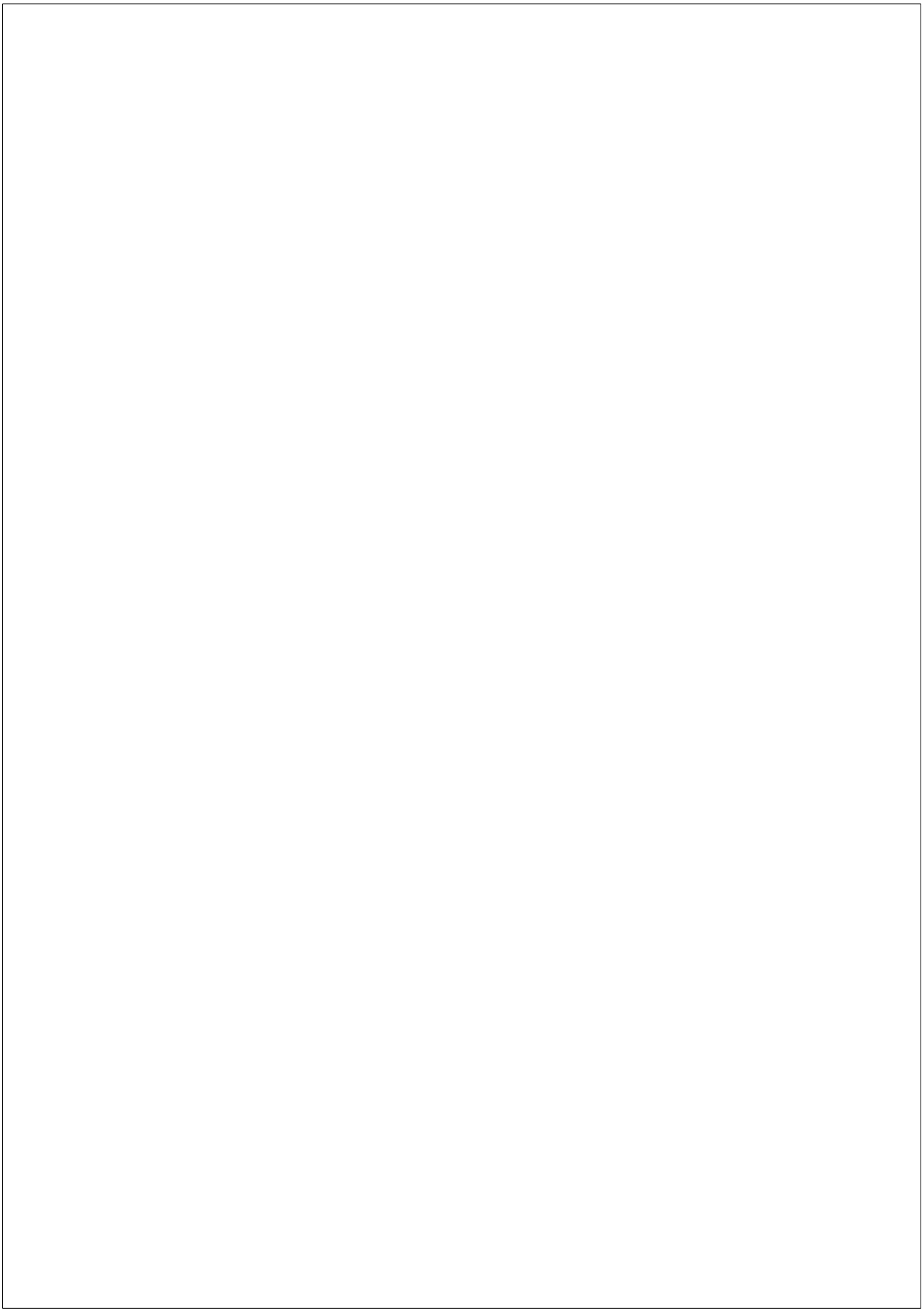
Nevertheless, even when all this effort has been paid, there is a lack of precise definition of *what* a wire is, or what a wire should be. Here we present precise definitions of a wire and its three components: the emitter component, the receptor component, but also the sink, a component responsible for the active degradation of the chemical compound that mediates the communication. Even more, we establish a theoretical and representational framework for the wire itself and provide mathematical definitions of what we call the wire space. This theoretical framework is built upon the aim to provide more a practical tool for bioengineers rather than a systems biology description of the cell-cell communication phenomena, even though some suggestions for extending the model and make it more descriptive are provided.

One fundamental detail has been missing from the beginning: we have seen some illustrative and archetypal examples of distributed computation, but it is not common to see a detailed discussion about how the ratios among the different strains affects the whole system. There are just two works that mention this particular issue of the ratio effect [136, 126]. But they lack a proper general explanation of this effect. This dissertation will cover the ratio problem in a general and theoretical way, as an extension of the wiring problem.

Also, we prove that our model predicts several behaviours that rely only in the wire itself: beyond the fact that it can transmit information, it can also work as a buffer, by remembering how much time is has been in the ON state. One main limitation of our model is that is has been elaborated assuming that cells do not divide and are sufficiently well mixed. Even when this is a strong assumption, it has been proved that *Escherichia coli* cells are able to respond to external stimuli when they

have arrested their growth because they are in stationary phase [53]. This establishes the less restrictive environment for the implementation of consortia circuits connected through chemical wires.

This introduction wants to establish the general purpose and capabilities of this wire model. But the concrete implementation proposal imposes new and deeper implications and problems that require to be extensively discussed. For that matter, the experimental part of this dissertation possesses its own experimental introduction, where some of the works mentioned here will be discussed again in a more detailed way.



Chapter 2

WIRE MATHEMATICAL MODEL: THE WIRE SPACE

In this chapter, we will establish the precise concept of what a wire is. We will start by defining the basic building blocks, followed by the description of the properties that we expect a wire should have. In this part, it is important to remark the underlying tension that exists between a desirable property and a feasible property. This is due to the obvious truth that we will face: engineering living systems has its own constrains, and we have to play with the already existing properties that living matter has when trying to impose or extract a desirable property. After establishing the mathematical model for the wire, we will explore its implications and properties, all under the umbrella of the Wire Space concept, the fundamental subject of this dissertation

The trilinear coordinate system, an original way of presenting information, will be of great help in this chapter. It will help us visualize every discussed property. Also, gaining intuition about this coordinate system will also help us to understand the bigger implications it has in

Chapter 4, where it will transcend its original representational purpose. There is no theory without its dark corners, and so we will discuss the limitations of the model, but also how and when to overcome them.

Let it be noted that in this chapter we will deal with the definition of chemical wires, in which three strains share information through a chemical diffusible messenger. But it is conceivable a wire in which the information is transmitted using light (through bioluminiscence), though we will not cover this assumption by now.

The mathematical formalisation exposed in this work is intentionally intended to be a as simple as possible. No general theory about wiring different strains for synthetic biology purposes has been proposed, and foundational propositions should be simple and elementary. Some may describe this model as a *toy model*, and I am aware that meanwhile for some readers that may be seen as an invitation to understanding, some others may experience some rejection. But I strongly believe that *toy models* are not intrinsically or essentially good or bad by themselves. They will be as good as we are good as players. After all, toys are meant to play with. And even at the risk of being pretentious, I encourage the reader to join me through the development of the model, as it may lead us to unexpected or surprising results, given its *a priori* simplicity.

2.1. Wire definition

In the context of distributed biological computation, a wire is defined as a set of three strains each of these containing one of the following three components: the emitter component, the receptor component, and the sink component. A wire has the ability to share infor-

mation by controlling the concentration of a chemical messenger in the media, which we call biobit or bbit. This bbit control will be determined by the ratio between the three strains. We will precisely define each of these elements:

The biobit

A biobit (bbit from here to the end), is any molecule that fulfils the following properties:

- Biosynthesizeble: it can be biologically synthesized by the cells.
- Secretable: it can be secreted to the extracellular medium.
- Biodetectable: it can be dynamically sensed by a biological system.
- Biodegradable: it can be actively and biologically degraded by the cells.
- Non-toxicity: it must be non toxic for the living systems that participate in the culture.
- Long half-life: the molecule must be stable by long periods of time.

It is important to notice that even when the mandatory requirements of a bbit are essentially satisfied, some bbit candidates may present partial fulfilment of the requirements. For instance, we may imagine the existence of a molecule that has a toxicity that depends on its concentration. In this case we would be limited by the range of concentrations at which the molecule is non-toxic.

The Emitter Half-Wire Component

The emitter half-wire component (or EHW) is the minimum set of genes required to sense an input signal and, in response to it, synthesize and secrete bbit to the media. Notice that with this definition we are not imposing how the secretion process occurs. A strain containing this component will be called the emitter. Figure 2.1 shows a scheme of the genetic construct implemented in the emitter.

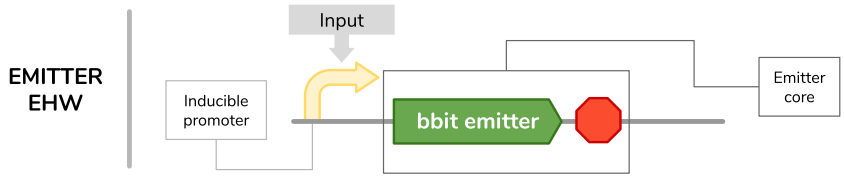


Figure 2.1: **Emitter component architecture**. Promoters are depicted as yellow arrows, terminators are depicted as red octagons, and elongated rectangles correspond to transcripts.

The Receptor Half-Wire Component

The receptor half-wire component (or RHW) is the minimum set of genes required to sense bbit concentration in the media and, in response to it, generate an output response. We assume that receptor genes are constitutively expressed, and that when bbit concentration raises above some threshold, it will induce the expression of some response gene (or genes). A strain containing this component will be called the receptor strain (Figure 2.2).

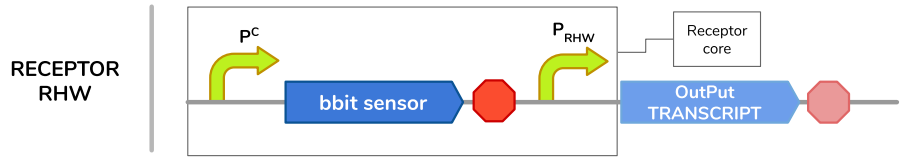


Figure 2.2: **Receptor component architecture**. Promoters are depicted as yellow arrows, terminators are depicted as red octagons, and elongated rectangles correspond to transcripts. P^C means constitutive promoter and P^{RHW} means RHW-dependent promoter.

The Sink Component

The sink component (or SINK) is the minimum set of constitutively expressed genes required to degrade bbit molecules. It does not participate in the production nor the detection of bbit concentration. A strain containing this component will be called the sink (Figure 2.3).

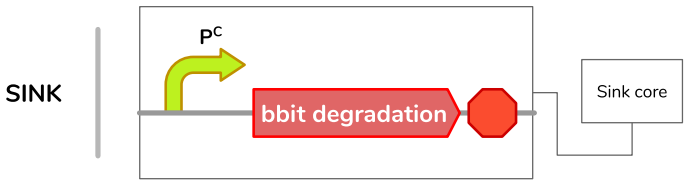


Figure 2.3: **Sink component architecture**. Promoters are depicted as yellow arrows, terminators are depicted as red octagons, and elongated rectangles correspond to transcripts. P^C means constitutive promoter.

Strain designing principles

When designing our three components for our three strains, we must take into account some relevant constrains. Notice that with these definitions we are not imposing the specific nature of the gene products (RNA or proteins), but the desired properties. So we can imagine a receptor component based in the interaction between the bbit and a RNA transcript to induce gene expression.

Another relevant aspect to take into account is that the capabilities of each strain must be physically bounded to the cells themselves. This means that if, for instance, the emitter strain can produce bbit because it secretes an enzyme that diffuses through the medium, it will not lead to the desired behaviours that we are looking for: the component must be physically contained and limited by the strain. The same applies to RHW and SINK strains. Nevertheless, it could be conceivable that the strain had its gene products attached to the membrane. In this case, the strain would addapt to the wire designing principles (as we shall see).

A relevant point, this time one more subtle, concerns the biodetectability of the bbit. Typically, this detection will be carried out by a receptor protein that will bind to the bbit. But here, the key point is that this union must be very dynamical, meaning that the binding reaction that will take place must be reversible and not shifted to the side where the ligand is bound to the receptor.

The exogenous genetic components described above, which define our distributed circuit, will demand energy and resources from the cell, so we should be conservative and minimalist when designing our circuits. As properly claimed in the component definitions, these should contain the *minimal* set of genes. It is known that metabolic

burden can alter the logic and behaviour of the genetic circuits[25, 106], and the wire components are genetic circuits in fact.

Concerning the interaction between the bbit molecule and the organism used as chassis, it will be desirable to consider a organism that is not affected by the presence of bbit molecules. Note that we are not saying here that it shouldn’t be toxic, but it should not interact at all, or at least it will interact the least possible.

2.1.1. Wires of non-growing cells

When considering the physiologic state, we propose the use of non-growing cells. Even when there are scenarios in which we may want our system to work in a state of high metabolic activity, like exponential growth or sustained growth in bioreactors, we want to restrict this work to non-growing cells for several reasons. First of all, it allows us to ignore the exponential growth in our model, thus fixing the amount of cells over time in our approach. Second, if we implement a wire in exponential growth, cells transition through several physiological states and over a limited period of time. In non-growing cells, the inner metabolic state is more stable over longer periods of time [101, 8]. It has been shown that *Escherichia coli* can induce gene expression from synthetic constructs when it is in stationary phase in a very robust manner [53]. So, *E. coli* may be a good candidate for implementing wires as a non-growing chassis.

2.2. Minimal model

Now we will focus in defining the minimal wire model. For that matter, first we will recall the prerequisites of the model and the basic assumptions defined in the previous section. We will start with the idealized framework of non-growing, divided in three different strains, the emitter, the receptor, and the sink strains. We will go through the development of the model assuming stationary phase cultures of *E. coli*.

Basic definitions

Each of the three strains defined above is idealized as a liquid culture of non-growing cells. Then, the cell concentration for each culture is fixed. Because the wire is the mix of these three strains, the first parameter we must define is the cell concentration of the wire co-culture. Let be P this general cell density. P can be measured in cells $\cdot$ mL^{-1}; alternatively optical density, OD_{600}, can also be used. For sure, due to the limited range of linear behaviour of the Beer–Lambert law when measuring cell density in a spectrophotometer, a correction must be applied to the measure to ensure linearity between corrected OD_{600} and cells $\cdot$ mL^{-1}. As this work is intended to be used as a tool for bioengineering, here I propose the usage of OD_{600} for practical reasons: it is easy to measure, and it is easy to get the corrected value.

Now, in a wire co-culture with cell density P, we can define three new parameters: the cell densities of emitter, receptor and sink strains. These will be called E for the cell density of the emitter cells, S for the cell density of the sink, and R for the cell density of the receptor strain. Thus, the overall cell density of the wire co-culture P can be expressed

as the sum of the three strain densities: $P = E + S + R$. In this work, however, we will be working with the ratios of these strain densities to the total population density:

$$\varepsilon = \frac{E}{P} \qquad \sigma = \frac{S}{P} \qquad \rho = \frac{R}{P}$$

It is trivial to see that $\varepsilon + \sigma + \rho = 1$. And through this dissertation, we will focus on the effect that altering this ratios has over the wire behaviour.

The Wire Space and the triangle plot representation

With the elementary definitions established, we can define the main concept of this dissertation, the Wire Space:

Definition 1. Let be the **Wire Space**, denoted by Δ, defined as the set of triplets composed by real numbers ε, σ and ρ, each representing sink, emitter and receptor strains ratios respectively, so their sum equals 1, and they all belong to the interval $[0, 1]$:

$$\Delta = \{\, (\varepsilon, \sigma, \rho) \mid \varepsilon + \sigma + \rho = 1, \;\; \forall\, \varepsilon, \sigma, \rho \in [0, 1] \,\}$$

At this point, and before proceeding to the description of the model itself, I will introduce to the reader the triangle plot representation (also known as ternary plot, simplex plot), a trilinear coordinate system in equilateral triangles. This representation takes advantage of Viviani’s Theorem, published in 1659 by Italian mathematician Vivenzo Viviani. The theorem claims that given an arbitrary point inside an equilateral

triangle, the sum of the three distances from that point to the sides of the triangle is constant and equals the height of the triangle [139, 76].

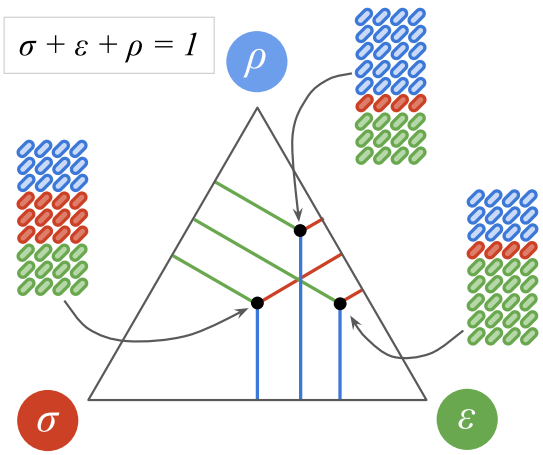


Figure 2.4: **Triangle plot representation.** Based on Viviani’s Theorem, that claims that given any point within an equilateral triangle, the sum of the distances from that point to the three sides is constant. This allows us to represent all possible combinations of emitter, receptor and sink strains as points within a triangle.

If we start with a triangle of height 1, there exists a perfect bijection between all possible combinations of wire strain co-cultures and a point inside a triangle: we can represent the relative frequencies of each strain in the wire as a distance between a point and one side of the triangle in an unique way. For that matter, we must agree which vertex represents each strain. Placing the triangle on one of its sides, the bottom left vertex will be the corresponding to σ, the Sink strain; the bottom right vertex will correspond to ε, the Emitter strain; and the top vertex will correspond to ρ, the Receptor strain. The intuition is simple: the closer the point is to a specific vertex, the more of that

strain will dominate in the culture.

The reader may have the impression that here we are presenting the information of three independent variables, ε, σ and ρ, when actually we only have two independent variables. This is because the cell density P is known, once we know the cell density of two of the three strains in the media, the cell density for the remaining strain in automatically known. The same happens when dealing with ratios because, once two of the three strain rations have been fixed, for instance ε and σ, the remaining one can be easily obtained as $\rho = 1 - \varepsilon - \sigma$.

Now, we have a tool to represent Δ, the whole Wire Space, by using the triangle plot by virtue of the Viviani's Theorem. For instance, a culture composed by the three strains at equal concentrations will be shown as a point in the middle of the triangle. More examples are shown in Figure 2.4. But also, we can imaging a culture only composed by a 50% mixture of emitters and receptors as a point lying in the midpoint of $\overline{\rho\varepsilon}$ triangle side, or a culture entirely composed by sink strain by a point lying directly on the σ vertex.

The Minimal Wire Model

Once the basic definitions have been established and also the representational framework has been detailed, we will dive in the mathematical model. Firstly, let me introduce the assumptions:

- The bbit intrinsic degradation is negligible.
- bbit precursors will never limit its production.
- The bbit production rate will be proportional to the Emitter cell density.

- Emitter cells will switch instantaneously from uninduced to induced when exposed to the input.
- The bbit degradation rate follows a Michaelis-Menten dynamics.
- Receptor cells participate neither in the synthesis nor the degradation of bbits.
- Receptor cells do not hijack bbit molecules, and the binding/unbinding is considered to be dynamical.
- Output production from receptor cells will be an instantaneous function of the bbit in the media.

This set of assumptions establishes the baseline of an idealised and minimal model for wires. We will cover what happens if there are not instantaneous processes in the next chapter. Nevertheless, besides these excluded assumptions could show a more realistic description, the minimal idealised model can work as a discard model. This means that if the minimal model predicts a desirable result, then it is worth to experiment with more detailed wire models. But if the minimal wire model predicts that the wire won't work, then it is worthless to test what would happen under more descriptive models.

Having explained this little disclaimer about the utility of the model, let me explicitly write the model. We will be using the differential equation, that captures everything we have been discussing:

$$\frac{\mathrm{d}B}{\mathrm{d}t} = K_E E - \frac{D_m B}{K_S + B} S \tag{2.1}$$

being B the bbit concentration, K_E the emission constant, E the emitter strain density, D_m the maximum bbit degradation constant, K_S the

sink constant, and S the sink strain density. All constants are assumed to be non-negative. Recall that, besides it is not essential, I would recommend measuring cell density in corrected OD_{600}. We will discuss the solutions of Eq. (2.1) but in other form, that explicitly shows its dependency from ε and σ, the relative ratios of emitter and sink strains respectively, starting from:

$$\frac{\mathrm{d}B}{\mathrm{d}t} = K_E \varepsilon P - \frac{D_m B}{K_S + B} \sigma P$$

and extracting common factor P, we get the **wire differential equation**:

$$\boxed{\frac{1}{P}\frac{\mathrm{d}B}{\mathrm{d}t} = K_E \varepsilon - \frac{D_m B}{K_S + B} \sigma} \tag{2.2}$$

We have to keep in mind that all constants will be fixed except for the emission constant, K_E, that will switch between two values: K_E^{OFF} and K_E^{ON}. Whereas K_E^{ON} will acquire an arbitrary positive value, K_E^{OFF} will have a value close to 0, as this constant will represent what happens when there is no input and bbit emission by emitter strain is negligible. I want to emphasize that this value will be expected to be close to 0, but not necessarily 0. This is due to what is called promoter leakiness, the situation in which some gene expression happens even when the gene is repressed. However, we will can assume $K_E^{OFF} = 0$ under certain circumstances. Everything explained so far is summarized in Figure 2.5. One consideration here: we will start by exploring what happens to Eq. (2.2) with an arbitrary K_E value, and later we will cover what happens with the interplay between the two K_E values: K_E^{OFF} and K_E^{ON}.

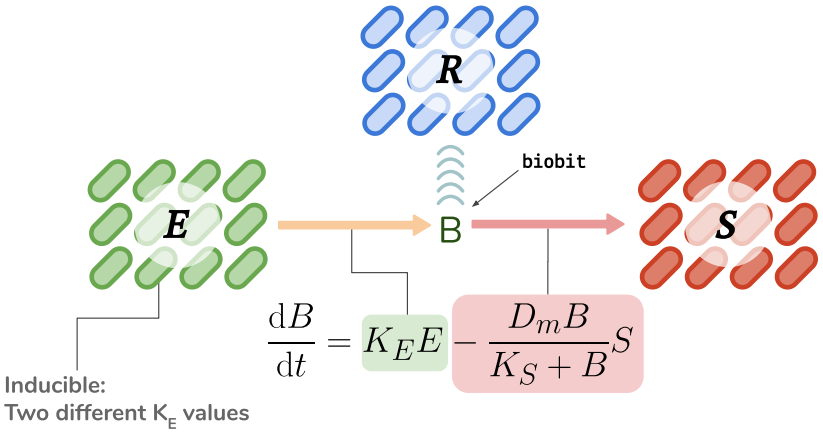


Figure 2.5: **Wire model graphical abstract**

2.2.1. Fixed points: bbit concentration at equilibrium

We will first identify the fixed points of the system. This will allow us to know the value that bbit concentration, B, reaches at equilibrium ($B = B_{eq}$). Because equilibrium is achieved when *enough time* has passed, and also it implies that there is no variation of bbit concentration in the media, we can mathematically express this scenario this way:

$$\lim_{t\to\infty} B(t) = B_{eq} \iff \frac{\mathrm{d}B}{\mathrm{d}t} = 0$$

And by applying the second part of this implication to Eq. (2.2), with some algebraic manipulation, with get the following expression for the **bbit concentration at equilibrium**:

$$\boxed{B_{eq} = \frac{K_E K_S \varepsilon}{D_m \sigma - K_E \varepsilon}} \tag{2.3}$$

We will explore more this formula in the next section, where we will associate it with specific sub-regions of the wire space.

2.2.2. Solution of wire differential equation

Even though for most of the calculations we will be dealing with it is not strictly necessary to find the solution for Eq. (2.3), we are going to develop it here. The following is the **implicit solution of the wire equation**:

$$\boxed{Ce^{\left(\frac{K_E K_S \varepsilon}{B_{eq}}\right)^2 Pt} = K_E K_S \varepsilon \; e^{-\frac{K_E K_S \varepsilon}{B_{eq}} B(t)} \left(1 - \frac{B(t)}{B_{eq}}\right)^{-K_S D_m \sigma}} \tag{2.4}$$

being C the constant of integration. This implicit solution will be helpful for obtaining relevant time measurements. Now, we will focus in the **explicit solution of the wire equation**:

$$\boxed{B(t) = B_{eq} + \frac{D_m \sigma}{K_E \varepsilon} W(D)} \tag{2.5}$$

with D being equal to:

$$D = \frac{(Ce^{(D_m\sigma - K_E\varepsilon)^2 Pt + K_E K_S \varepsilon})^{\frac{1}{-K_S D_m \sigma}}}{-K_S D_m \sigma}$$

The $W(x)$ is the Lambert W Function. The only thing we need to know for our purposes about this function is that $W(0) = 0$. Lambert W Function is bivaluated and the solution of the model requires considering the two branches: $W_0(x)$ and $W_{-1}(x)$ Also, $W(x)$ cannot be written as a composition of elementary functions, but its value can be approached numerically with the Taylor series:

$$W_0(x) = \sum_{n=1}^{\infty} \frac{(-n)^{n-1}}{n!} \, x^n$$

For the estimation of C, the integration constant, we can assume that when $t = 0$, $B = B(0)$, and then we get the following value for C:

$$C = e^{-K_E K_S \varepsilon \frac{B(0)}{B_{eq}}} \left[K_E K_S \varepsilon - K_E K_S \varepsilon \frac{B(0)}{B_{eq}}\right]^{-K_S D_m \sigma}$$

We can push forward this integration constant assuming $B(0) = 0$ to get:

$$C_0 = (K_E K_S \varepsilon)^{-K_S D_m \sigma}$$

Because our analysis of the model will consist in numerical simulations, we can use our analytical solution to validate our numerical approach. In Figure 2.6) it is shown the comparison between simulations and analytical solutions. We can appreciate that the different branches of W Lambert function represent different set of solutions. We also can see that the numerical simulations are more powerful that analytic solutions as these have some problems due to complex values appearing. But we will assume that we can trust in numerical simulations. For these simulations, a culture consisting in only emitter and sink cells was used.

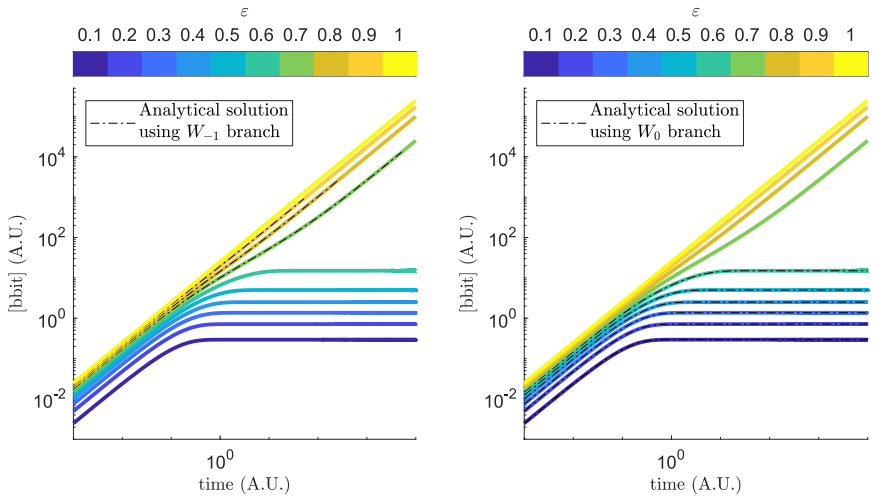


Figure 2.6: **Comparison between analytic and numerical solutions**. Parameter values from Table 2.1.

Parameter	Value
P (Population)	50
K_E (Emitter constant)	0.5
D_m (Maximum degradation)	1
K_S (Sink constant)	5

Table 2.1: Parameter values used in Fig. 2.6

Taking into account everything explained in this subsection, we will consider numerical simulations valid, and we will use then by default unless specified.

2.3. Convergence and divergence regions

Let us start our section doing a little bit of experimental simulations. Using the triangular representation, we can simulate several combinations of emitter, sink and receptor ratios homogeneously distributed throughout the wire space:

In Figure 2.6 we can appreciate that whereas for some culture-points the bbit concentration seems to reach equilibrium (green dots and lines), other culture-points correspond to divergent bbit concentration over time when simulated (purple dots and lines). To make some sense of these two qualitatively different behaviours, we shall recall on Eq. 2.2, but emphasizing the biosynthesis and degradation terms in green and red respectively:

$$\frac{1}{P}\frac{\mathrm{d}B}{\mathrm{d}t} = K_E\varepsilon - \left(\frac{D_m B}{K_S + B}\sigma\right)$$

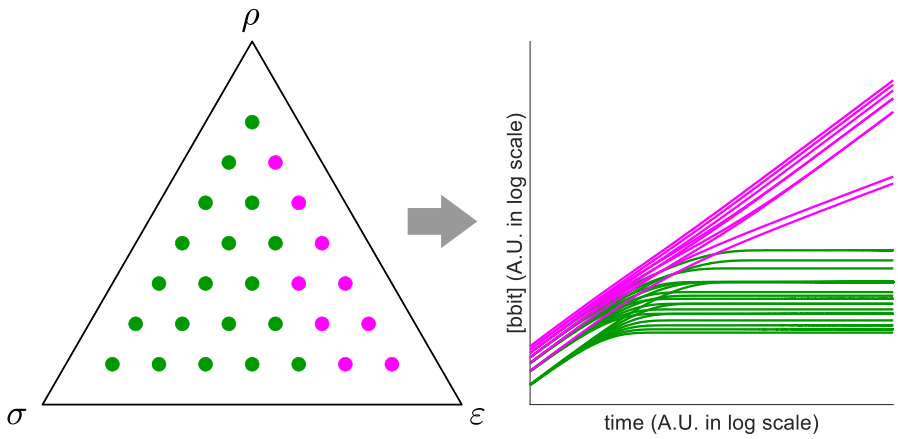


Figure 2.7: **Convergent and divergent biobit time series**. Equally distributed points in the wire space coloured in green if the associated simulation converges, or in magenta if the corresponding simulation diverges. Parameter values from Table 2.1.

When B_{eq} is reached the following is satisfied:

$$B(t) = B_{eq} \implies \frac{\mathrm{d}B}{\mathrm{d}t} = 0 \implies K_E\varepsilon = \frac{D_m B}{K_S + B}\sigma$$

And so, we can graphically summarize this last equality by plotting these two processes (biosynthesis and degradation) depending on the bbit concentration, B, in Figure 2.8:

Figure 2.8 shows the phase plane of our simple model. Notice that, in this specific case, there is one scenario in which biosynthesis and degradation curves do not intersect, at least in the time window that is explored here. Will these curves intersect beyond that time window?, or will they keep distant from each other no matter how large the bbit concentration gets? If we want to unravel what is going on here, we

must start at the very beginning.

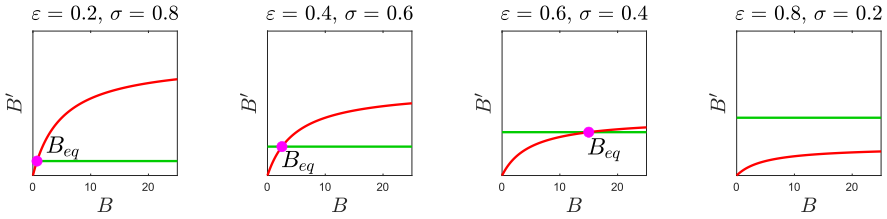


Figure 2.8: **Phase portraits for wire model**. Biobit synthesis rate is depicted in green, degradation rate is depicted in red, and biobit concentration at equilibrium B_{eq} is shown as a magenta dot when proceeds. Parameter values from Table 2.1.

2.3.1. Condition for convergence

First, let me remind the reader Eq. 2.3, or the explicit expression for B_{eq}:

$$B_{eq} = \frac{K_E K_S \varepsilon}{D_m \sigma - K_E \varepsilon}$$

B_{eq}, the bbit concentration at equilibrium, must be a positive value. And because all constants are positive, the numerator of E. 2.3 is positive. However, the denominator can be positive or negative depending on the relative ratios of ε and σ. So, if we want to impose the positiveness of B_{eq}, the following must be satisfied:

$$D_m \sigma - K_E \varepsilon > 0$$

which implies the following **condition for convergence**:

$$\boxed{B_{eq} > 0 \iff \frac{\varepsilon}{\sigma} < \frac{D_m}{K_E}}$$

So, now we know that when a culture-point is is associated with a convergent bbit concentration time-series, then the condition for convergence must be satisfied. Figure 2.9 summarizes what we have being discussing until now.

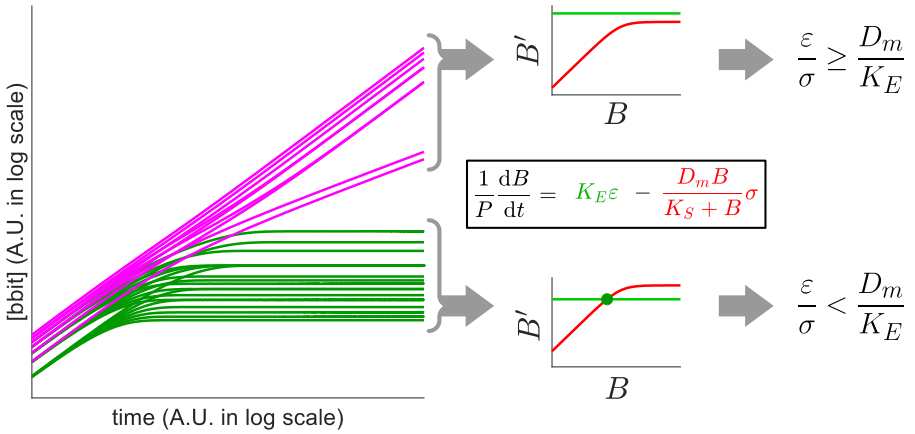


Figure 2.9: **Convergence, divergence, and phase portraits**. Convergent time-series correspond to intersecting biosynthesis and degradation rates, whereas in the case of divergent trajectories the rates do not intersect.

2.3.2. Definition of convergence and divergence regions

Recalling the Definition 1 of the Wire Space, Δ, let me proceed to define the two first sub-regions of the Wire Space:

Definition 2. Let be the **convergence region**, denoted by C, defined as the set of elements from the wire space so $\frac{\varepsilon}{\sigma} < \frac{D_m}{K_E}$:

$$\mathrm{C} = \left\{ (\varepsilon, \sigma, \rho) \in \Delta \mid \frac{\varepsilon}{\sigma} < \frac{D_m}{K_E} \right\}$$

Definition 3. Let be the **divergence region**, denoted by D, defined as the set of elements from the wire space so $\frac{\varepsilon}{\sigma} \geq \frac{D_m}{K_E}$:

$$\mathrm{D} = \left\{ (\varepsilon, \sigma, \rho) \in \Delta \mid \frac{\varepsilon}{\sigma} \geq \frac{D_m}{K_E} \right\}$$

Definition 4. Let be the **horizon of divergence**, denoted by $\hat{\mathrm{H}}$, defined as the set of elements from the wire space so $\frac{\varepsilon}{\sigma} = \frac{D_m}{K_E}$:

$$\hat{\mathrm{H}} = \left\{ (\varepsilon, \sigma, \rho) \in \Delta \mid \frac{\varepsilon}{\sigma} = \frac{D_m}{K_E} \right\}$$

Recall that in the extreme case of a culture exclusively composed by receptor and emitter strains, when $\sigma = 0$, the ratio ε/σ is not defined. To avoid this definition problem, consider the equivalent definition for divergence, $K_E\varepsilon \geq D_m\sigma$, and thus, the definition is complete.

The main corollary is that the Wire Space can be divided into two disjoint regions, and thus, using set notation:

$$\boxed{\Delta = \mathcal{C} \cup \mathcal{D}} \tag{2.6}$$

And trivially, the horizon of divergence $\hat{\mathrm{H}}$ is included in the divergence region:

$$\boxed{\hat{\mathrm{H}} \subset \mathrm{D}} \tag{2.7}$$

Now we can explore more about the explicit form of the horizon of divergence. Let us start by considering a culture composed solely by emitter and sink strains, then $\rho = 0$, and $\varepsilon + \sigma = 1$. This consideration places the culture-point in the bottom side of the triangle representation, the $\overline{\sigma\varepsilon}$ side. The horizon of divergence will occur when the denominator in Eq. 2.3 (the expression of B_{eq}) will be 0. And we are going to find the values $\varepsilon = \hat{\varepsilon}$ and $\sigma = \hat{\sigma}$ that ensure that denominator equals 0:

$$D_m\hat{\sigma} = K_E\hat{\varepsilon} \implies D_m(1-\hat{\varepsilon}) = K_E\hat{\varepsilon}$$

whose solution for $\hat{\varepsilon}$ and $\hat{\sigma} = 1 - \hat{\varepsilon}$ is

$$\hat{\varepsilon} = \frac{D_m}{K_E + D_m} \qquad \hat{\sigma} = \frac{K_E}{K_E + D_m}$$

And so, we define the **point of divergence**, $\hat{h}$, as a culture-point with the following coordinates $(\varepsilon, \sigma, \rho)$:

$$\hat{h} = (\hat{\varepsilon}, \hat{\sigma}, 0) = \left(\frac{D_m}{K_E + D_m}, \frac{K_E}{K_E + D_m}, 0\right) \tag{2.8}$$

Because the $\hat{\mathrm{H}}$ is defined as $\varepsilon/\sigma = D_m/K_E$, the geometric locus that satisfies that definition is the straight segment $\overline{\rho\hat{h}}$, that irradiates from the top vertex ρ and hits the bottom side $\overline{\sigma\varepsilon}$ at the point of convergence, $\hat{h}$. Hence, the horizon of divergence $\hat{\mathrm{H}}$ admits an alternative explicit definition:

Definition 5 (*Explicit definition for* $\hat{\mathrm{H}}$)**.**

$$\hat{\mathrm{H}} = \{\, (\varepsilon, \sigma, \rho) \in \Delta \mid \ (\varepsilon, \sigma, \rho) = (k\hat{\varepsilon}, k\hat{\sigma}, 1-k), \text{ with } k \in (0, 1]\}$$

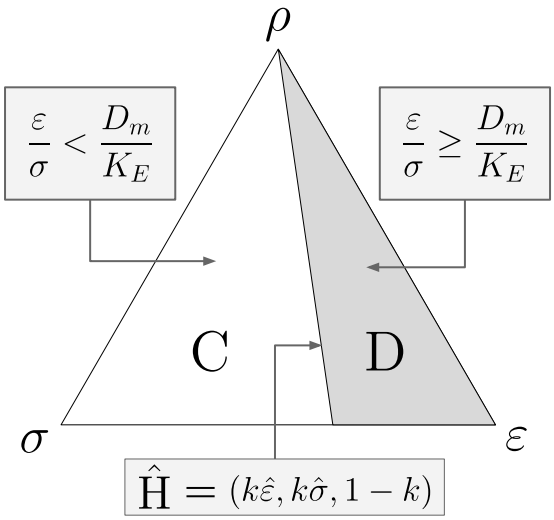


Figure 2.10: **Convergence and divergence regions, and horizon of divergence.** Analytic definitions are shown.

2.3.3. The convergence region and B-isolines Lemma

Assuming that we have two cultures $u_1 = (\varepsilon_1, \sigma_1, \rho_1)$ and $u_2 = (\varepsilon_2, \sigma_2, \rho_2)$ such that $u_1 \neq u_2$, and $u_1, u_2 \in \mathrm{C}$, can point-cultures u_1 and u_2 reach the same B_{eq}? If so, can we find a geometric place within the triangle such that if two culture-points lay in this place we can

ensure that they will reach the same B_{eq}? That geometric place will be called a B**-isoline**, and the following lemma details its mathematical description:

Lemma 1 (B**-Isolines Lemma**)**.** All points along any straight line that radiates from the top ρ vertex of the wire space triangle and hits its $\overline{\sigma\varepsilon}$ bottom side reach the same B_{eq} value.

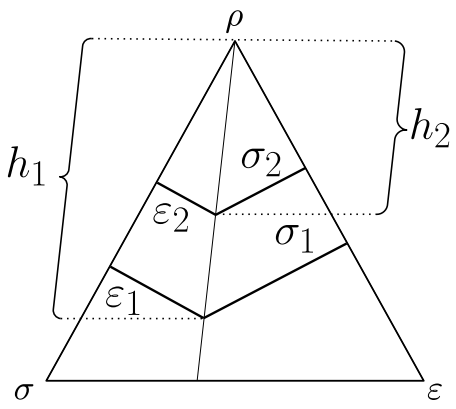


Figure 2.11: **Graphical support for B-Isolines Lemma**

Proof. Assuming that B_{eq} is reached, we begin by noting that the expression (2.3) of the bbit concentration at equilibrium can be rewritten as $B_{eq} = \frac{K_E K_S}{D_m \sigma/\varepsilon - K_E}$. Thus, B_{eq} depends on the consortium composition only via the ratio σ/ε. In turn, geometric similarity shows in a straightforward manner that the ratio σ/ε is constant for all points belonging to any straight line radiating from the ρ vertex of the triangle. To see this, consider two points along one of these radiating lines, characterized respectively their strain relative frequencies $(\varepsilon_1, \sigma_1)$ and $(\varepsilon_2, \sigma_2)$, as shown in Fig. 2.10. Let h_1 and h_2 be distances between

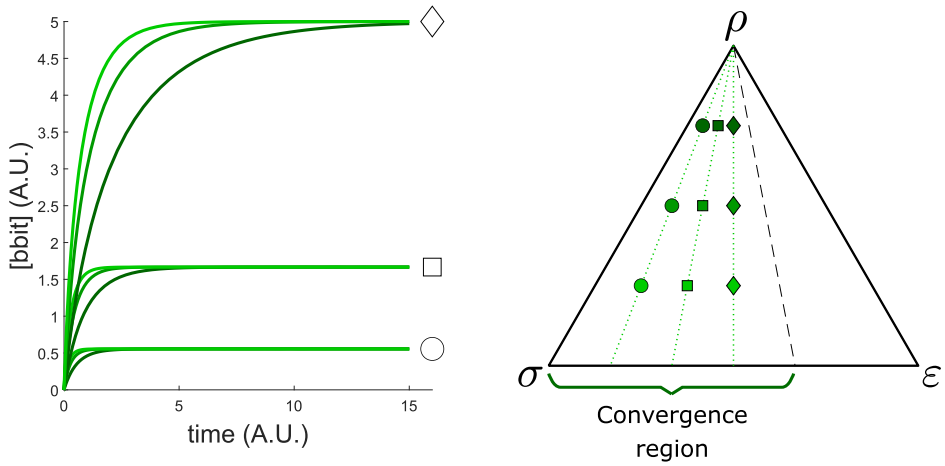


Figure 2.12: **B-Isolines behaviour**. All cultures within the convergence region lying in any segment irradiating from the top vertex ρ to the $\overline{\sigma\varepsilon}$ bottom side will reach the same biobit concentration at equilibrium. Parameter values from Table 2.1.

these two points and the ρ vertex. The similarity between the two triangles delimited by the ε and h segments in one case, and by the σ and h segments in the other, leads to:

$$\frac{\varepsilon_1}{\varepsilon_2} = \frac{h_1}{h_2} = \frac{\sigma_1}{\sigma_2} \quad \Longrightarrow \quad \frac{\sigma_1}{\varepsilon_1} = \frac{\sigma_2}{\varepsilon_2}$$

Thus $B_{eq,1} = B_{eq,2}$, which proves the lemma. □

Figure 2.12 shows simulations done over several cultures placed on top of three B-isolines. It is easy to appreciate that whenever the cultures are placed in the same isoline, they reach the same B_{eq} value. The only difference between cultures sharing B-isoline is that it seems that the closer a culture is to the top ρ vertex. the longer it will take to reach B_{eq}. We will explore in depth this temporal behaviour in the next chapter.

The last question that remains to be discussed is how B_{eq} grows as a culture approaches to $\hat{\mathrm{H}}$. To approach this question, we will use the previous lemma. The B-isolines lemma ensures that, assuming our cultures are within the convergence region, we can work with those in the $\overline{\sigma\varepsilon}$ side of the triangle (where $\rho = 0$), and extrapolate what happens there to the whole convergence region.

In the $\overline{\sigma\varepsilon}$, $\sigma + \varepsilon = 1$, and using $\sigma = 1 - \varepsilon$, Eq. 2.3 can be rewritten as:

$$B_{eq} = \frac{K_E K_S \varepsilon}{D_m - (D_m + K_E)\varepsilon}$$

In the convergence region, $\varepsilon \in [0, \hat{\varepsilon})$, and it is trivial that when $\varepsilon = 0$, then $B_{eq} = 0$. It is not difficult to see as well that when $\varepsilon \to \hat{\varepsilon} \implies B_{eq} \to \infty$.

Dividing Eq. 2.3 by ε and defining $\sigma/\varepsilon = x$ we get:

$$B_{eq} = \frac{K_E K_S}{D_m x - K_E}$$

being $x \in (0, K_E/D_m)$. And in this last equation, everything is a fixed parameter except x, which implies that, the B_{eq} experiment a hyperbolic growth as a culture approaches $\hat{h}$ (the point of divergence), being $\hat{h}$ the singularity of the hyperbola. B-Isolines lemma allows the following generalization: for every fixed $\rho \neq 1$, if a culture approaches $\hat{\mathrm{H}}$, then B_{eq} experiments a hyperbolic growth with the singularity in $\hat{\mathrm{H}}$. This last statement can also be appreciated in Figure 2.12, in which three B-isolines are shown with constant distances among them for a fixed ρ value, which nevertheless translates into a increasing and non-linear B_{eq} value.

2.3.4. The divergence region and $\hat{v}$-Isolines Lemma

We have being discussing mainly about the convergence region. But what about the divergence region? In the diverge region, we know that $D_m\sigma \leq K_E\varepsilon$, so $D_m\sigma - K_E\varepsilon \leq 0$, which implies, by Eq. 2.3, that $B_{eq} < 0$. This is impossible as B is the bbit concentration and hence it must be positive. We also know that because B diverges, $\lim_{t\to\infty} B(t) = \infty$. Let us rearrange Eq. 2.2, the bbit differential equation, as:

$$\frac{\mathrm{d}B}{\mathrm{d}t} = P\left(K_E\varepsilon - \frac{D_m}{\frac{K_S}{B}+1}\sigma\right)$$

Now we take limits to get the expression of $\hat{v}$, the **asymptotic divergence rate**:

$$\boxed{\hat{v} = \frac{\mathrm{d}B_{\hat{v}}}{\mathrm{d}t} = \lim_{B\to\infty}\frac{\mathrm{d}B}{\mathrm{d}t} = P(K_E\varepsilon - D_m\sigma)} \tag{2.9}$$

The first thing to pay attention to is that the term $K_E\varepsilon - D_m\sigma$ is positive, because if we consider we are in the diverge region, then $D_m\sigma - K_E\varepsilon \leq 0$. This asymptotic divergence rate means that when $t \to \infty$, then:

$$B' = \frac{\mathrm{d}B}{\mathrm{d}t} \approx \hat{v}$$

Or in other words,

$$\lim_{t\to\infty}\frac{B'}{\hat{v}} = 1$$

Because all terms involved in $\hat{v}$ are constants, then the solution to Eq. 2.9 is linear and has this expression:

$$\boxed{B_{\hat{v}}(t) = \hat{v}t + B(0)} \tag{2.10}$$

The final, rather trivial, corollary is that, provided that a culture is placed in the divergence region, then:

$$\lim_{t\to\infty} \frac{B(t)}{B_{\hat{v}}(t)} = 1$$

Unless B_{eq}, $\hat{v}$ is a bounded value. Knowing that when any of the three strains, ε, σ or ρ, approaches 1 then the other two must necessarily approach 0, it is rather trivial that:

$$\hat{v} \in [0, PK_E]$$

Now, we can formulate an analogous question for the asymptotic divergence rate, $\hat{v}$, as we did for the B_{eq}: what is the geometric place within the divergence region D that ensures that if two cultures lay in that place, then they both will reach the same asymptotic divergence rate? Let me introduce the following lemma:

Lemma 2 (**$\hat{v}$-Isolines Lemma**)**.** All points along any straight segment laying in the divergence region, D, and parallel to the horizon of divergence, $\hat{\mathrm{H}}$, will reach the same asymptotic divergence rate, $\hat{v}$.

Proof. Let us start by recalling the main characteristics of the points laying in the horizon of divergence, $\hat{\mathrm{H}}$. In Figure 2.13, panel A, the horizon of divergence $\hat{\mathrm{H}}$ is the segment $\overline{\rho\hat{h}}$, being $\hat{h}$ the point of divergence. The grey shaded area corresponds to the divergence region D, and left white region to the convergence region, C. Any point that belongs to $\hat{\mathrm{H}}$ has the coordinate expression $(\varepsilon, \sigma, \rho) = (k\hat{\varepsilon}, k\hat{\sigma}, 1-k)$, with $k \in (0, 1]$, being these values the distances between the point and the three sides of the triangle $\widehat{\sigma\varepsilon\rho}$.

Continuing in Figure 2.13 panel B: the segment $\overline{\rho'\hat{h}'}$ is parallel to $\hat{\mathrm{H}}$, and we will call it $\hat{\mathrm{H}}'$. That segment $\hat{\mathrm{H}}'$ hits the segment $\overline{\rho\varepsilon}$ in point

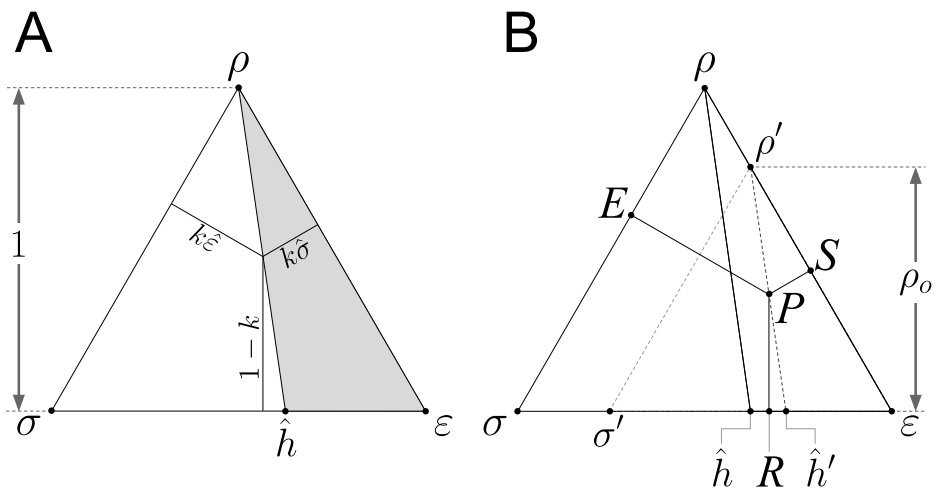


Figure 2.13: **Graphical support for v̂-Isolines Lemma**

ρ'. Starting at ρ', we can define a new segment $\overline{\rho'\sigma'}$ that is parallel to the side $\overline{\rho\sigma}$. Now, the triangle $\widehat{\sigma\varepsilon\rho}$ is geometrically similar to the triangle $\widehat{\sigma'\varepsilon\rho'}$, with the scaling origin at vertex ε. The height of triangle $\widehat{\sigma'\varepsilon\rho'}$ is ρ_o, and because the height of the original triangle $\widehat{\sigma\varepsilon\rho}$ is 1, ρ_o constitutes the scaling factor. Also, every ρ_o defines a unique $\hat{\mathrm{H}}'$ segment, so all the points that belong to $\hat{\mathrm{H}}'$ differ in their k value, with $k \in (0, 1]$. Now we select a point $P \in \hat{\mathrm{H}}'$, and the segments $\overline{PS}$ and $\overline{PR}$, both perpendicular to $\overline{\varepsilon\rho}$ and $\overline{\sigma\varepsilon}$ sides respectively. Because the length of these two segments $\overline{PS}$ and $\overline{PR}$ are the coordinates σ and ρ of point P with respect to new triangle $\widehat{\sigma'\varepsilon\rho'}$, and that triangle is scaled down by a factor ρ_o, then the values of these lengths are $|\overline{PS}| = k\hat{\sigma}\rho_o$ and $|\overline{PR}| = (1 - k)\rho_o$. But also, coordinates σ and ρ of point P are the same for triangle $\widehat{\sigma\varepsilon\rho}$.

Now we want the length of the segment $\overline{PE}$, but in virtue of Viviani’s Theorem applied to triangle $\widehat{\sigma\varepsilon\rho}$, $|\overline{PE}| + |\overline{PS}| + |\overline{PR}| = 1$, so

$|\overline{PE}| = 1 - k\hat{\sigma}\rho_o - (1-k)\rho_o$. Hence, the triangular coordinates of P are $(\varepsilon, \sigma, \rho) = (1 - k\hat{\sigma}\rho_o - (1-k)\rho_o, k\hat{\sigma}\rho_o, (1-k)\rho_o)$.

Given two arbitrary points P and P' that belong to $\hat{\mathrm{H}}'$ having asymptotic divergence rate $\hat{v}$ and $\hat{v}'$ respectively, only differing in their k values (k for P and k' for P'). If $\hat{v} = \hat{v}'$, then $\hat{v}/\hat{v}' = 1$:

$$\frac{\hat{v}}{\hat{v}'} = \frac{P\left[K_E(1 - k\hat{\sigma}\rho_o - (1-k)\rho_o) - D_m k\hat{\sigma}\rho_o\right]}{P\left[K_E(1 - k'\hat{\sigma}\rho_o - (1-k')\rho_o) - D_m k'\hat{\sigma}\rho_o\right]}$$

which simplifies into

$$\frac{\hat{v}}{\hat{v}'} = \frac{K_E - K_E\rho_o + K_E k\rho_o - k\rho_o(K_E\hat{\sigma} + D_m\hat{\sigma})}{K_E - K_E\rho_o + K_E k'\rho_o - k'\rho_o(K_E\hat{\sigma} + D_m\hat{\sigma})}$$

but because $\hat{\sigma} = K_E/(D_m + K_E)$, then $-k\rho_o(K_E\hat{\sigma} + D_m\hat{\sigma}) = -K_E k\rho_o$, cancelling all terms involving k (denominator work analogously). Then we find that

$$\frac{\hat{v}}{\hat{v}'} = \frac{K_E - K_E\rho_o}{K_E - K_E\rho_o} = 1$$

which proves the lemma. □

A consequence of the $\hat{v}$-isolines lemma is that each isoline solely depends on its *height*, ρ_o, that:

$$\hat{v} = PK_E(1 - \rho_o)$$

This alternative expression for $\hat{v}$ is not practical, as it is desirable to know the asymptotic divergence rate directly knowing ε and σ. Nevertheless, it allows us to know that $\hat{v}$ varies from 0 in $\hat{\mathrm{H}}$, to $PK_E\varepsilon$ in the ε vertex in a linear manner.

In Figure 2.14 it is summarized the behaviour of different cultures placed in the divergence region. The points are strategically placed over different isolines to emphasize the expected behaviours.

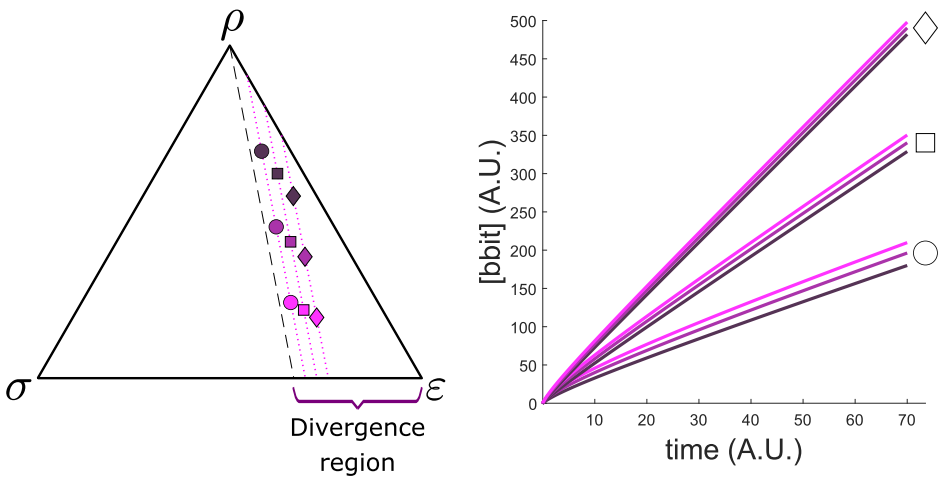


Figure 2.14: **v̂-Isolines behaviour**. All cultures within the divergence region lying in any segment parallel to the horizon of divergence $\hat{\mathrm{H}}$ will reach the same asymptotic divergence rate. Parameter values used taken from Table 2.1.

Because $\hat{v}$ is bounded in the interval $[0, PK_E]$, we can divide $\hat{v}$ by PK_E to obtain the normalized asymptotic divergence rate, $\overline{v} \in [0, 1]$:

$$\boxed{\overline{v} = \frac{\hat{v}}{PK_E}} \tag{2.11}$$

2.3.5. Intrinsic and extrinsic convergence

When discussing about the convergence/divergence regions in the past, I have had to face some specific criticism about the use of the term *divergence* (specially when talking with physicists). The problem arises from the fact that there is no such thing as divergence in the realm of biological systems. And the critics are totally right: you can-

not expect the concentration of any arbitrary molecule to increase *ad infinitum*. If concentration rises above a specific threshold, you cannot expect to have a diluted molecule, as it will precipitate. Not to mention the case in which the concentration keeps raising reaching the same density of molecules expected in a pure crystal. Keep raising from that point, and you are forcing atoms to be closer than they should be.

The point is: something will eventually happen that will prevent your concentration to keep increasing. Nevertheless, we cannot always predict what or how it will happen. It could maybe be that bbit molecule interacts in a harmful manner when its concentration is too high, becoming toxic. Maybe it tends to aggregate or self-react, or maybe it can alter the membrane potential of the host cells. The nature of the bbit and the nature of the host cell define an unique and, I hypothesize, unpredictable scenario that will prevent bbit concentration from keep increasing. This is what I call **extrinsic convergence**: the arrest of bbit concentration due to factors that are independent to the wire model. This critical concentration must be measured experimentally for each pair bbit-host if bbit concentration is expected to reach high values. Under this assumption, all divergence region is actually a region of extrinsic convergence. Even more, some sub-regions in the convergence region that are close enough to $\hat{\mathrm{H}}$, the horizon of divergence, will reach B_{eq} values above that critical concentration that promotes extrinsic convergence.

Obviously, the regions in which that critical concentration is not reached, is what I call **intrinsic convergence**, always belonging to the convergence region. I call it intrinsic, because it only depends on the model’s prediction.

I hope this section will bring understanding upon the use of the

term *divergence*, as it was never intended to evoke an actual never-ending increase of bbit concentration. In the context of this wire model, divergence means that bbit concentration will keep rising until some unpredictable event will prevent it.

2.4. The role of the receptor strain

So far, we have being discussing quantitatively the bbit concentration as an interplay between emitters and sink strains. Also, we have covered how we can understand this relations in the representational framework of the triangle representation of the Wire Space. But we have not defined any particular role for what is *not* ε of σ. Up to now, ρ could be literally interpreted as $1 - \varepsilon - \sigma$ and we wouldn't have to assume its particular behaviour or nature. Of course, ρ represents the ratio of receptor strain in the culture. This strain has all the necessary genetic equipment to sense the presence (or absence) of bbit, and respond to it. Now, we will discuss how the receptor strain reads the information stored in bbit concentration and translates it into output gene expression.

2.4.1. Receptor strains and Hill-equations

We will start by assuming our receptor strain follows a general Hill function:

$$r(B) = \frac{r_m B^n}{K_A^n + B^n} \tag{2.12}$$

in which r_m is the maximum response, K_A is the B concentration at which half of the maximum response in reached ($r(K_A) = \frac{1}{2} r_m$), and n is the Hill-coefficient governing the response steepness. There

could be some leakiness of the genes involved in output synthesis, but for practical reasons, we will consider it negligible. If we imagine the output is GFP fluorescence, leakiness would be the fluorescence in absence of bbit. Even when there exist systems with low leakiness, recent works provide implementation strategies to achieve robust negligible leakiness [21].

2.4.2. Output gene expression is measured with respect to receptor strain cell density

A really important consideration that has to made when discussing receptor strain deals with the answer to this question: what is *the response* in terms of receptor concentration? Are we counting the whole culture fluorescence? The answer is no. The above Hill-equation is referred as output per receptor cell unit. Let me explain: imagine that we have a stationary phase culture entirely composed by the receptor strain. Because cells are non-growing, $P = R$. Under that circumstance, we test the receptor strain by exposing it to different bbit concentrations to get experimental data about the relationship between bbit concentration and output gene. Let us imagine that this output gene is GFP fluorescence. Then, we are not really interested in the absolute values of fluorescence from the culture, but the amount of fluorescence per unit cell in this receptor mono-culture. Because we assume that we can measure cell density using corrected OD_{600}, we would divide the fluorescence data by the corrected OD_{600} to get the gene expression per unit cell. These values will typically be easy to fit to a Hill-function, and it is from *these data* that we will be able to measure the characteristics of our receptor's transfer function.

Another relevant detail is that Eq. 2.12 tells us the relation between bbit concentration and output gene expression when the system has reached equilibrium, but it captures anything about its temporal dynamics. A good descriptive equation for a receptor strain expressing GFP as output gene would be:

$$\frac{\mathrm{d}[\mathrm{GFP}]}{\mathrm{d}t} = \alpha \frac{B^n}{K_A^n + B^n} - \gamma[\mathrm{GFP}] \tag{2.13}$$

This last equation assumes that GFP is synthesized in response to bbit concentration B following an activating Hill-function dynamics (almost no synthesis if $B \approx 0$, and an α rate of synthesis for $B \ggg K_A$). Also, Eq. 2.13 assumes that GFP is degraded within the receptor cells at a rate proportional to its own concentration. The preceding equation allows us to capture the temporal dynamics, but we will focus on the equilibrium, when [GFP] keeps its concentration stable, and thus we can measure the Hill equation associated with this equilibrium, its transfer function [24].

We will discuss more on temporal dynamics in the next chapter, but let me recall the reader that with this simple wire model we just want to check if, all time given, it is possible to expect a desired behaviour from our hypothetical wire.

As a last detail: concerning the receptor strain response, the relevant parameters are K_A, where we find the inflexion point of our Hill-function, and n, the Hill-coefficient, that governs the steepness around the inflexion point. For that matter, we can work with the normalized response:

$$\boxed{\overline{r}(B) = \frac{r(B)}{r_m} = \frac{B^n}{K_A^n + B^n}} \tag{2.14}$$

This last expression allows us to bound response values in the interval $[0, 1]$, and we will be using this last expression for our model. However,

for what we are going to explain now, we only need to know that response can be mapped onto $[0, 1]$ interval, and follows a sigmoidal function.

2.4.3. New response regions and horizons

Receptor strain are responsible for establishing the relation between bbit concentration and the equilibrium reached by output gene in response. Last Eq. 2.14, allows us to split the response into three different regimes:

- **Uninduced receptor**: denoted by U, when receptor level of induction is less or equal than 5% of its maximum response level.
- **Induced receptor**: denoted by I, when receptor level of induction is greater or equal than 95% of its maximum level of induction.
- **Variable level of induction**: denoted by V, when receptor presents an intermediate level of induction between 5% and 95% of its maximum level of induction.

Later, we will properly define these regions with respect to the Wire Space and the triangle representation. But let us first focus in the boundaries between these three regimes. These last three receptor strain induction levels define two important bbit concentrations: T_5, the threshold that separates the uninduced and the variable induction; and T_{95}, the threshold that separates the variable induction and the fully induced states.

Now, we are going to find the geometric locus of T_5 and T_{95} in the Wire Space, to properly define our three new receptor regimes as

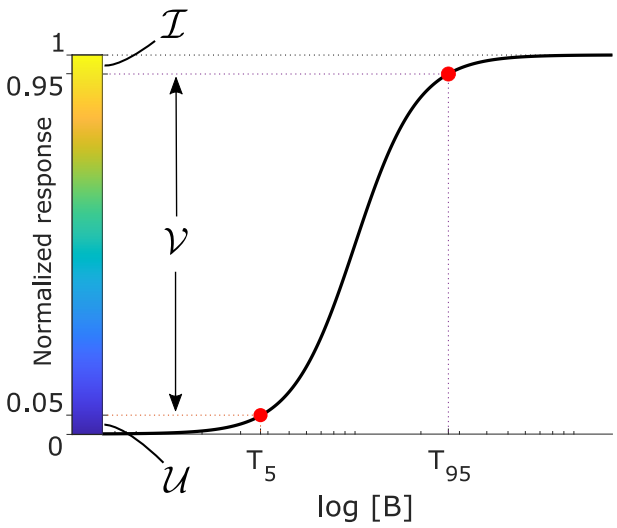


Figure 2.15: **Receptor strain transfer function defines levels of induction**. The three different levels of response induction define special B_{eq} concentrations T_5 and T_{95} that split the whole range of induction between uninduced (U), variable level of induction (V), and fully induced receptor (I).

regions. To do so, we will start by placing ourselves in the $\overline{\sigma\varepsilon}$ bottom side of our triangle, where $\rho = 0$, to find the corresponding ε and σ values:

$$B_{eq} = T_5 = \frac{K_E K_S \varepsilon}{D_m \sigma - K_E \varepsilon} = \frac{K_E K_S \varepsilon}{D_m(1-\varepsilon) - K_E \varepsilon}$$

And solving for ε, we can get σ easily as we have imposed that $\rho = 0$ and hence $\varepsilon + \sigma = 1$, and then we get the point $(\varepsilon_5, \sigma_5, 0)$, where:

$$\varepsilon_5 = \frac{D_m T_5}{K_E K_S + (D_m + K_E) T_5} \qquad \sigma_5 = \frac{K_E K_S + K_E T_5}{K_E K_S + (D_m + K_E) T_5}$$

Because these are the ε and σ values that will be associated with

$B_{eq} = T_5$ in the bottom $\overline{\sigma\varepsilon}$ side of our triangle, we can use the B-isolines lemma 1 to explicitly define a new horizon:

Definition 6 (*Explicit definition for* $\overline{T}_5$)**.** Let be the horizon T_5, denoted by $\overline{T}_5$, the set of points inside the convergence region such $B_{eq} = T_5$:

$$\overline{T}_5 = \{\, (\varepsilon, \sigma, \rho) \in \mathrm{C} \mid \ (\varepsilon, \sigma, \rho) = (k\varepsilon_5, k\sigma_5, 1-k), \text{ with } k \in (0, 1]\}$$

It is trivial to proceed analogously for defining T_{95}:

Definition 7 (*Explicit definition for* $\overline{T}_{95}$)**.** Let be the horizon T_{95}, denoted by $\overline{T}_{95}$, the set of points inside the convergence region such $B_{eq} = T_{95}$:

$$\overline{T}_{95} = \{\, (\varepsilon, \sigma, \rho) \in \mathrm{C} \mid \ (\varepsilon, \sigma, \rho) = (k\varepsilon_{95}, k\sigma_{95}, 1-k), \text{ with } k \in (0, 1]\}$$

Where ε_{95} and σ_{95} are defined as:

$$\varepsilon_{95} = \frac{D_m T_{95}}{K_E K_S + (D_m + K_E) T_{95}} \qquad \sigma_{95} = \frac{K_E K_S + K_E T_{95}}{K_E K_S + (D_m + K_E) T_{95}}$$

These two new horizons allow us to split convergence region C into three new sub-regions that we define as follows:

Definition 8 (*Uninduced region*)**.** Let be the **receptor uninduced region**, denoted by U, the set of points inside the convergence region such:

$$\mathrm{U} = \left\{ (\varepsilon, \sigma, \rho) \in \mathrm{C} \mid 0 \leq \frac{\varepsilon}{\sigma} \leq \frac{\varepsilon_5}{\sigma_5} \right\}$$

Definition 9 (*Variable region*)**.** Let be the **receptor variable induced region**, denoted by V, the set of points inside the convergence region such:

$$\mathrm{V} = \left\{ (\varepsilon, \sigma, \rho) \in \mathrm{C} \mid \frac{\varepsilon_5}{\sigma_5} < \frac{\varepsilon}{\sigma} < \frac{\varepsilon_{95}}{\sigma_{95}} \right\}$$

Definition 10 (*Induced region*)**.** Let be the **receptor induced region**, denoted by I, the set of points inside the convergence region such:

$$\mathrm{I} = \left\{ (\varepsilon, \sigma, \rho) \in \mathrm{C} \mid \frac{\varepsilon_{95}}{\sigma_{95}} \leq \frac{\varepsilon}{\sigma} < \frac{\hat{\varepsilon}}{\hat{\sigma}} \right\}$$

With respect to this last region, the receptor induced region I, because it is defined upon points belonging to the convergence region C, it is implicit that $\varepsilon/\sigma < \hat{\varepsilon}/\hat{\sigma}$, but I prefer to make it redundant and as clear as possible.

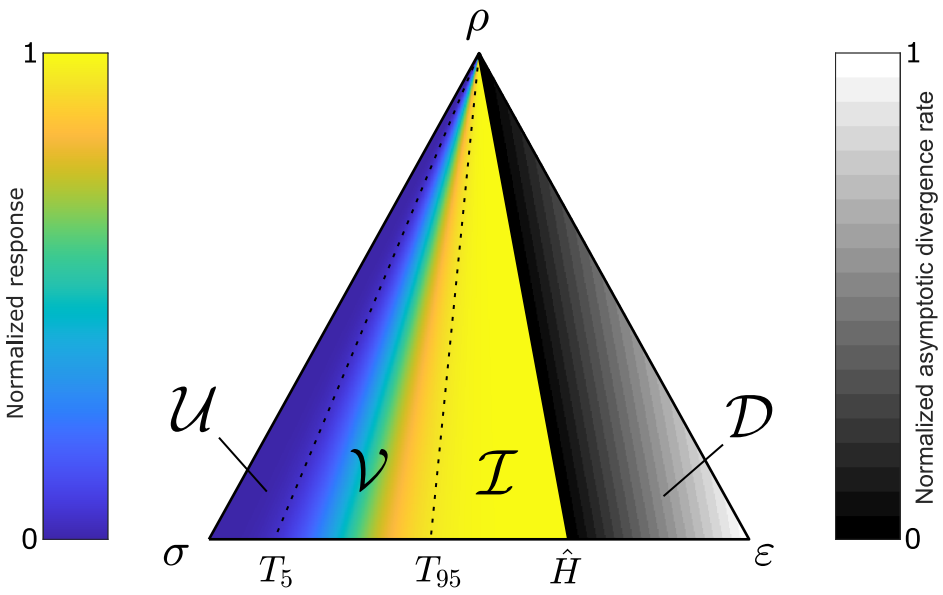


Figure 2.16: **Mapping of $\overline{r}$ and $\overline{v}$ in the triangle representation of the wire space.** Parameter values taken from Table 2.2.

Parameter	Value
P (Population)	50
K_E (Emitter constant)	0.5
D_m (Maximum degradation)	1
K_S (Sink constant)	5
K_A (Inflexion point for Hill Eq.)	1
n (Hill coefficient)	3

Table 2.2: Parameter values used in Fig. 2.16

With respect to what happens with divergence region and the receptor induction level, receptor strain will be totally induced, but here the big difference is that whereas in the convergence region $B = B_{eq}$,

no equilibrium is reached in divergence. Hence, we will be more focused in the normalized asymptotic divergence rate (Eq. 2.11), as it better captures the increasing bbit concentration.

Now, with these new definitions, we can map the different new convergence sub-regions into the triangle representation. Using the normalized response, $\overline{r} \in [0, 1]$, and the normalized asymptotic bbit divergence rate, $\overline{v} \in [0, 1]$, we can express and summarize both the different levels of induction in the convergence region, and the whole range of bbit increasing concentration rates in Figure 2.16.

2.5. Wire States

So far, we have covered what happens assuming that all constants involved in the model assume fixed values. But as I pointed at the beginning of this chapter, the emitter strain has two physiological states: uninduced or induced. What are the implications of this change? We will try answering this question by using the triangle representation. We will also present the conditions space, a *naïve* theoretical framework for dealing with different environmental variables, as a developing seed for future wire research.

2.5.1. Proper states

In this subsection, we will briefly discuss what we call the **wire proper states**. These are the minimal states that a wire is assumed to have. They are given by the fact that the emitter strain is inducible and thus it possesses two different physiological states: fully uninduced, when $K_E = K_E^{OFF} \approx 0$, and fully induced, when $K_E = K_E^{ON} > 0$. In

terms of the model, it only affects the K_E parameter. We will note this difference by writing the words ON and OFF as a super-index when pertinent.

Because there are two different sets of parameters for a wire (even when they only differ in one, K_E), there are two different wire states, which are associated with two different triangle representations (Figure 2.17)

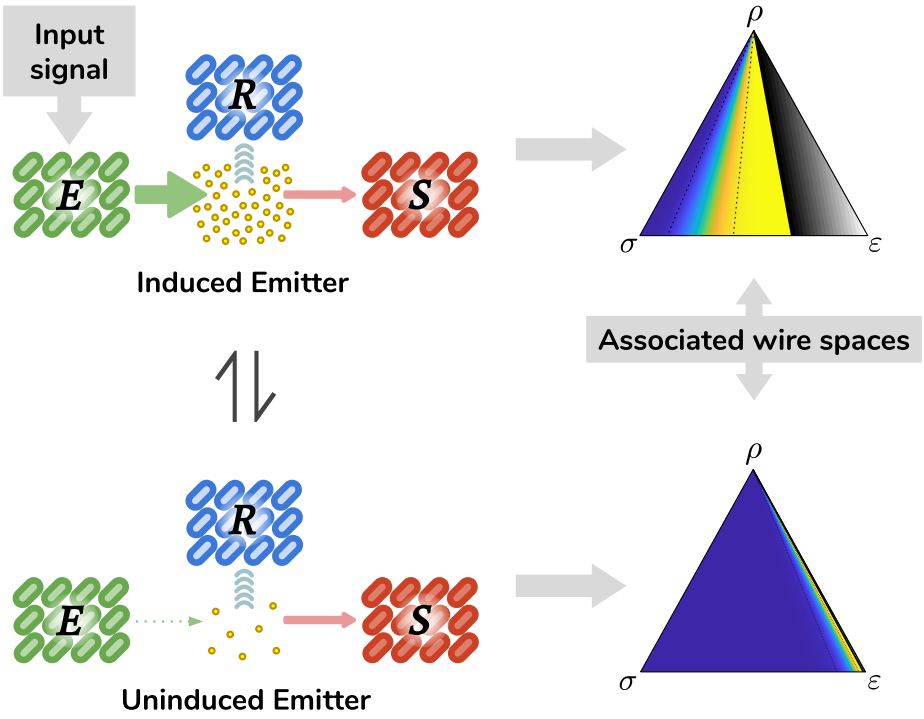


Figure 2.17: **Wire proper states** Depending on the emitter induction state, the whole wire space can switch between the two associated wire proper states, which is reflected in the triangle representation.

When emitter strain switches from uninduced to induced, $K_E^{OFF} \rightarrow$

K_E^{ON}, and all horizons (divergence and response thresholds) move from being close to the $\overline{\rho\varepsilon}$ side to being more distributed and closer to the $\overline{\rho\sigma}$ side of the triangle. Because every horizon is uniquely defined by a ratio ε/σ, this means that these ratios get smaller as $K_E^{OFF} \rightarrow K_E^{ON}$. Recall that a culture may be placed in a specific point within the wire space, and switching from uninduced to induced and *vice versa* does not affect the point position of any other parameter except from K_E.

The specific proper state of a wire will be denoted by Δ^{OFF} is the emitter is uninduced, and by Δ^{ON} if the emitter is induced, using the super-index notation. So at the end, $K_E^{OFF} \implies \Delta^{OFF}$ and $K_E^{ON} \implies \Delta^{ON}$.

2.5.2. Conditions Space

When we think about a wire, at the end it is all about measuring parameters. We should measure K_E^{OFF} and K_E^{ON} in the case of the emitter strain, K_S and D_m for the sink strain, and T_5 and T_{95} for the receptor strain. As a first approach, there is no reason for assuming that the two emitter states will affect sink or receptor strains. And as a consequence, we encounter ourselves with two proper states: Δ^{OFF} and Δ^{ON}. But in which experimental conditions did we measure our constants? Obviously, it is assumed that all parameters are measure in a fixed scenario. For instance, we may measure *Escherichia coli* MG1655 strain hosting our wire constructs in pSB1C3 plasmid in Davis media as a baseline (we will study experimental conditions in chapter 4). But what about temperature, pH, light exposure or salts concentrations? Let us imagine that we took our measurements at 37Cº with the default Davis media salts concentrations. But what if we want our wire

to work in another temperature or pH? For every single condition, wire parameters may vary, and so the wire space properties. A change in temperature, salts or pH can potentially alter at least one parameters. Hence, we need a simple theoretical setup to deal with a changing environmental conditions.

We will start by intuitively introducing the conditions space, noted by Ω, as the set of all conditions (intrinsic or experimental) that require measuring again our wire parameters. By definition, **assuming fixed experimental conditions**, a wire has at least two proper states, Δ^{OFF} and Δ^{ON}. These states constitute the intrinsic conditions of a wire because of its inducible nature. And thus, we now that at least, for some fixed experimental conditions:

$$\Omega = \{\Delta^{OFF}, \Delta^{ON}\}$$

If two sets of parameter measurements corresponding to the proper states of a wire are given to us, a detailed description in which these measurements were taken is mandatory. This implies that there is no such thing as *fundamental* proper states, as they will always depend on the experimental conditions. Thus, proper states give us a fixed picture of a wire for specific experimental conditions.

But what if we want to vary our experimental conditions? Do proper states lack meaning as they always depend on them? Definitely no. If we want to test our wire in two different temperatures such 37ºC and 30ºC (for example), we should measure Δ^{OFF} and Δ^{ON} under these two specific temperatures, which makes four sets of measurements. Note that Δ^{OFF} and Δ^{ON} are mutually exclusive states, as they are temperatures such 37ºC and 30ºC. And the same will be true for every single experimental variable (understanding experimental as environ-

mental). But the different values that the experimental variables may acquire are not exclusive among them. We can test our proper states at 30ºC, pH = 7 and blue light exposure, and test again our proper states at 30ºC but in darkness and pH = 5.

Let me introduce the omega notation for conditions: when describing the different states a variable may acquire, we will wire the variable with capital omega as a super-index. This way, the element within the wire that defines the two proper wire states, the emitter strain denoted by EHW, may be induced or uninduced. This would represented as $\mathrm{EHW}^{\Omega} = \{\mathrm{OFF}, \mathrm{ON}\}$. Because there is a direct relation between the states of the emitter strain and the proper states of the wire, we can also represent the previous description of EHW^{Ω} with the equivalent symbols for wire proper states, so $\mathrm{EHW}^{\Omega} \equiv \Delta^{\Omega} = \{\Delta^{OFF}, \Delta^{ON}\}$.

If we want to test our wire under three different temperatures, we should represent this fact by $\mathrm{T}^{\Omega} = \{23°\mathrm{C}, 30°\mathrm{C}, 37°\mathrm{C}\}$. If we want test our wire in two different pHs 5 and 7, we represent this as $\mathrm{pH}^{\Omega} = \{5, 7\}$. And for an arbitrary variable X, we denote the finite set of values it may acquire as X^{Ω}. Let us proceed now with the precise definition of the conditions space.

Definition 11 (*Conditions Space*)**.** Let $\mathrm{X}_1^{\Omega}, \mathrm{X}_2^{\Omega}, ..., \mathrm{X}_n^{\Omega}$ be a finite number of independent variables with a finite number of values each. We define the **condition space**, denoted by Ω, as the cartesian product of all the variables:

$$\Omega = \prod_{i=1}^{n} \mathrm{X}_i^{\Omega} = \mathrm{X}_1^{\Omega} \times \mathrm{X}_2^{\Omega} \times \cdots \times \mathrm{X}_n^{\Omega}$$

The requirement of the finite number of values for each variable

obeys the practical constrain that, experimentally, we are unable to measure in a truly continuous way. In actual terms, once we will have identified our independent variables with at least two different values each, we should get an estimation for K_E, D_m, K_S, T_5 and T_{95} for every condition in the condition space. For example, lets imagine that we have two variables, one being the wire proper states $\Delta^{\Omega} = \{\Delta^{OFF}, \Delta^{OFF}\}$, and another some different temperatures T^{Ω} = $\{30°\mathrm{C}, 37°\mathrm{C}\}$, because we want our wire to work at both 30ºC and 37ºC. Then, our condition space is:

$$\Omega = \{(\Delta^{OFF}, 30°\mathrm{C}), (\Delta^{OFF}, 37°\mathrm{C}), (\Delta^{ON}, 30°\mathrm{C}), (\Delta^{ON}, 37°\mathrm{C})\}$$

And so, we should get an estimation for the constants for every element of the condition space Ω. For calculating the number of different conditions in which we are required to measure our parameters, defining $|\mathrm{X}_i^{\Omega}|$ as the number of elements that variable X_i may acquire, then:

$$|\Omega| = \prod_{i=1}^{n} |\mathrm{X}_i^{\Omega}| \tag{2.15}$$

To distinguish and emphasize the main role that proper states play in the conditions space Ω, we may want to write it as Δ/i, being i some arbitrary condition set. If for instance we wan to test our wire under two different temperatures, defining the parameter vector $(K_E, K_S, D_m, T_5, T_{95})$, we can give a summary of these values under the two temperatures like this:

$$\begin{cases} \Delta^{OFF}/30°\mathrm{C} = \left(K_E^{OFF}, K_S^{OFF}, D_m^{OFF}, T_5^{OFF}, T_{95}^{OFF}\right)/30°\mathrm{C} \\ \Delta^{ON}/30°\mathrm{C} = \left(K_E^{ON}, K_S^{ON}, D_m^{ON}, T_5^{ON}, T_{95}^{ON}\right)/30°\mathrm{C} \\ \Delta^{OFF}/37°\mathrm{C} = \left(K_E^{OFF}, K_S^{OFF}, D_m^{OFF}, T_5^{OFF}, T_{95}^{OFF}\right)/37°\mathrm{C} \\ \Delta^{ON}/37°\mathrm{C} = \left(K_E^{ON}, K_S^{ON}, D_m^{ON}, T_5^{ON}, T_{95}^{ON}\right)/37°\mathrm{C} \end{cases}$$

And explicitly writing down the measured values for each parameter.

One may argue that some phenomenological law could be deduced to know the explicit relation that exists between some specific experimental variable and our parameters. One could systematically measure the wire parameters at different temperatures and try to infer the relation that ties temperature with the parameters. It would definitely be a desirable work to do, but the work increases very fast when trying to infer these relations if more experimental variables have to be tested in the same way. This is due to the fact that assuming independence between these experimental variables may hide synergistic or antagonistic effects between these variables that may get reflected upon the wire parameters.

In what follows, we will be assuming that our wire works under controlled experimental conditions, and thus, we will not make use of this last piece of theory about conditions space.

2.6. Operative regions: digital and buffer regions

The proper states may define new states. Each point within the wire space will be associated with two possible scenarios: converging to a specific B_{eq} (or to an asymptotic divergence rate) when Δ^{OFF}, or to an another specific B_{eq} (or to an asymptotic divergence rate) when Δ^{ON}. Equivalently, theses points can be grouped according to the regions they *simultaneously* belong when Δ^{OFF} or Δ^{ON}. In other words, the **intersections** between regions may allow the existence of what we call **operative regions**. These regions, whose existence is

not granted, are the basis of the wire capabilities.

Let us start by intuitively defining the operative regions. We define the digital region, denoted by greek letter Φ, as the intersection between the uninduced region when the emitter in uninduced, and the induced region when the emitter is induced. And we define the buffer region as the intersection between the uninduced region when the emitter in uninduced, and the divergence region when the emitter is induced.

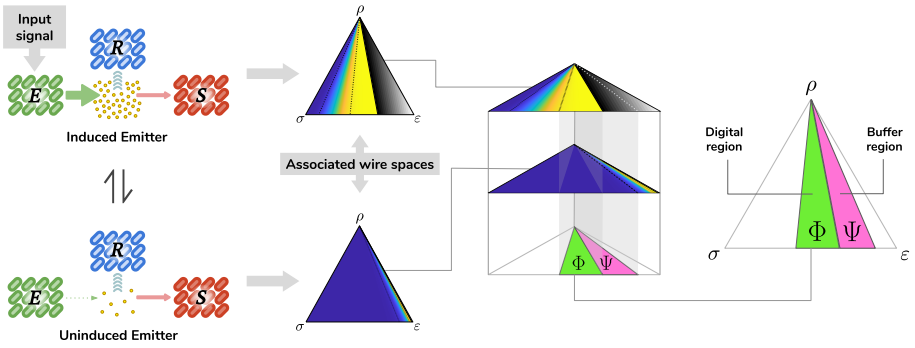


Figure 2.18: **Digital and buffer regions**. Operative regions (digital and buffer) are defined as the intersections between specific regions of the proper states.

We will be exploring and properly defining these new regions in the following subsections.

2.6.1. Digital region

As previously mentioned, digital region is the intersection between two pre-existing regions, one corresponding to the Δ^{OFF} proper state and another to the Δ^{ON} proper state. In the case of Δ^{OFF}, we are talk-

ing about the uninduced region U, and in the case of Δ^{ON}, the region will be I. To emphasize the proper state and not being always referencing the Δ^{Ω} notation, the Ω notation will be extended to the regions belonging to each proper state, so U when the proper state is Δ^{OFF} will be represented as U^{OFF}, and analogously, I^{ON} will represent the other case.

Because U^{OFF} and I^{ON} belong to their corresponding convergence regions C, any point representing a culture placed on this digital region will be associated with two different B_{eq} values. One case when the emitter is uninduced, and hence we expect a $B_{eq} \approx 0$, that will be associated with a normalized response less that 5% by definition. In the other case, B_{eq} will reach a value that is enough to generate a receptor normalized response greater than 95% again by definition.

Let us properly define the digital region:

Definition 12 (*Digital region*)**.** Let be the **digital region** denoted by Φ, defined as the set of points from the wire space so:

$$\Phi = \{w \in \Delta \mid w \in \mathrm{U}^{\mathrm{OFF}} \cap \mathrm{I}^{\mathrm{ON}}\}$$

This last definition is certainly minimalist, but lacks operability, and because we are discussing operative regions, it requires another approach. For that matter, let us recall definitions of both uninduced and induced regions (Def. 8 and Def. 10 respectively). Assuming an arbitrary point $w = (\varepsilon, \sigma, \rho)$, U is defined as all points such $0 \leq \varepsilon/\sigma \leq \varepsilon_5/\sigma_5$, and I is defined as all points such $\varepsilon_{95}/\sigma_{95} \leq \varepsilon/\sigma \leq \hat{\varepsilon}/\hat{\sigma}$.

Taking this into account, we can imagine two different cases (Figure 2.19). In one case, the ratio ε_5/σ_5 when the proper state is Δ^{OFF}

is smaller than the ratio $\hat{\varepsilon}/\hat{\sigma}$ when the proper state is Δ^{ON} (case 1 in Fig. 2.19). Equivalently, this is the same as when the $\hat{\mathrm{H}}^{\mathrm{ON}}$ is closer than T_5^{OFF} to the $\overline{\varepsilon\rho}$ side of the triangle. The other case if when the converse is true and T_5^{OFF} horizon is the closest one to the $\overline{\varepsilon\rho}$ side (case 2 in Fig. 2.19).

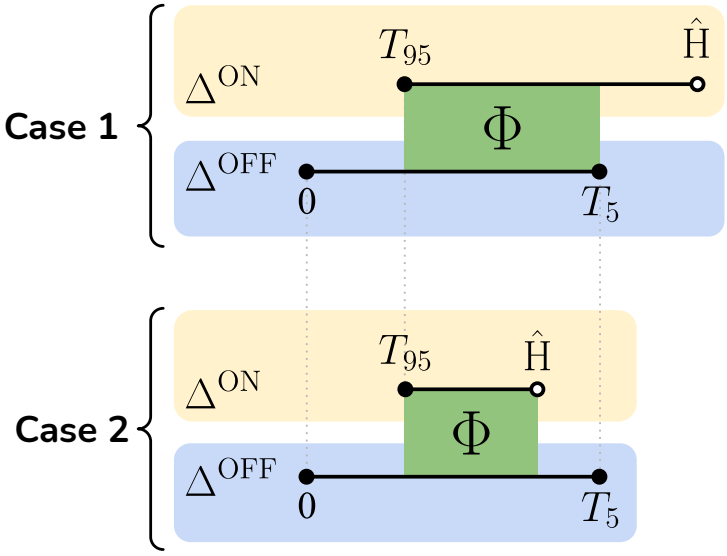


Figure 2.19: **Digital region definition**. This region is defined as the intersection between I^{ON} and $\mathrm{U}^{\mathrm{OFF}}$ regions.

This allows us to reformulate Definition 12 by using inequalities in terms of model parameters:

Definition 13 (*Explicit definition for digital region*)**.** Let be the **digital region** denoted by Φ, defined as the set of points $w = (\varepsilon, \sigma, \rho)$ from the wire space so:

$$\Phi = \left\{ w \in \Delta \mid \left(\frac{\varepsilon_{95}}{\sigma_{95}}\right)^{ON} \leq \frac{\varepsilon}{\sigma} < \min\left\{ \left(\frac{\hat{\varepsilon}}{\hat{\sigma}}\right)^{ON}, \left(\frac{\varepsilon_5}{\sigma_5}\right)^{OFF} \right\} \right\}$$

The existence of this digital region is not granted. The fact that a wire is an inducible system as described in this thesis does not allow by itself such regime. The parameters, measured under the two proper states Δ^{Ω} must fulfil what it is called as the **existence condition for digital region**:

$$\boxed{\exists \Phi \neq \varnothing \iff \left(\frac{\varepsilon_{95}}{\sigma_{95}}\right)^{ON} < \left(\frac{\varepsilon_5}{\sigma_5}\right)^{OFF}} \tag{2.16}$$

The disappearance of the ratio representing $\hat{\mathrm{H}}^{ON}$ in the existence condition is pretty trivial: if the ratio associated with T_5^{OFF} is bigger than the ratio associated with $\hat{\mathrm{H}}^{ON}$, then it must also be bigger than the ration associated with T_{95}^{ON} horizon. This reasoning is left to the reader as an exercise.

Let us stop here for a moment and recall our previous discussion about Ω, the conditions space. The previous existence condition for Φ assumes only the proper states under some fixed experimental conditions. Written in conditional notation, Φ/e, with e being the unique and fixed experimental condition in which the proper states have been measured. But what if we want our wire to work under different environmental conditions? Can we find a robust formulation of the existence condition for Φ that works under any of the desired conditions?

Labelling these arbitrary conditions as $i \in \Omega$, we can formulate the **digital region strong definition**:

Definition 14 (*Strong definition for digital region*)**.** Let be the **digital region** denoted by Φ, defined as the set of points $w = (\varepsilon, \sigma, \rho)$ from the wire space such:

$$\Phi = \bigcap_{i \in \Omega} \Phi / i$$

An explicit strong definition can be stated by taking maximum and minimum of all ratios under all conditions considered. Now, let me introduce the **strong existence condition for digital region**:

$$\exists \Phi \neq \varnothing \iff \max \left\{ \frac{\varepsilon_{95}}{\sigma_{95}} \Big/ i \right\}^{ON} < \min \left\{ \frac{\varepsilon_{5}}{\sigma_{5}} \Big/ i \right\}^{OFF} \quad \forall i \in \Omega \tag{2.17}$$

This last strong existence condition for Φ emphasizes this more restrictive scenario. Starting with the strong definition, there could be a conditioned digital region that does not present any intersection with other conditioned digital regions, and thus, there wouldn’t exist any region in the wire space that would allow cultures to work digitally under all conditions. Or maybe there is an experimental condition that does not allow the existence of any digital region.

But **why is this digital region so important**? If a culture corresponds to a strain combination that is associated with a point lying in the digital region, the time it will take to switch from $\Delta^{OFF} \rightarrow \Delta^{ON}$ and vice-versa will be **fixed in any case**. This implies that a culture in this regime would work like a light switch, in which the ON/OFF state of

the switch gets reflected instantly in the light bulb. And precisely **this** is the typically desired behaviour that most of the distributed synthetic constructs require.

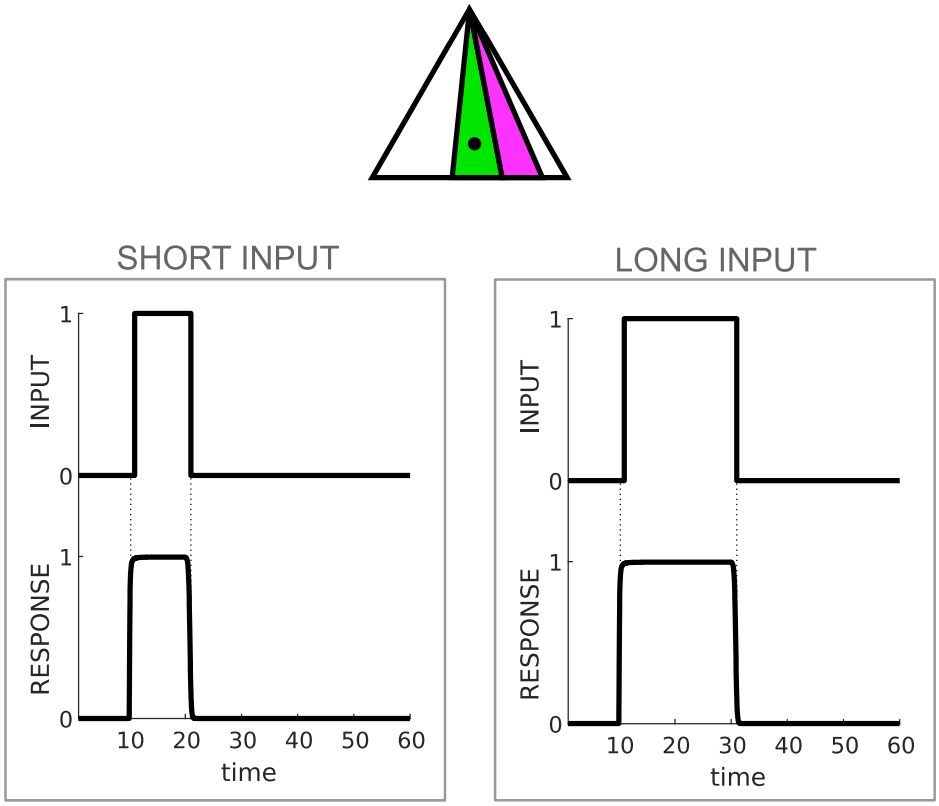


Figure 2.20: **Digital region behaviour**. Output signal reconstructs input signal if enough time is provided when the culture is in the digital region. Parameter values from Table 2.2.

In Figure 2.20, two simulations are done over a point lying in the digital region Φ (in green). It is possible to appreciate what was commented in previous paragraph. The response duration, the output, is

almost the same as input duration. This behaviour is what can be expected for any point lying in the digital region (if enough time passes by): a purely digital behaviour. In the next chapter we will explore the time dynamics in a more deeper way.

2.6.2. Buffer region

Unlike copper wires in electronics, biological wires based on the proposed architecture can extend their properties beyond the digital transmission of information. This is because biological wires can store information by themselves. We say *can store* instead of directly saying *store*, because, in the same way that digital region was not granted to exist, the same applies to this new buffer region.

Buffer region is again an intersection between two pre-existing regions, being the first one U^{OFF}, the uninduced region when the proper state is Δ^{OFF}, and D^{ON}, the divergence region when the proper state is Δ^{ON}. We will denote this new operative region with greek letter Ψ:

Definition 15 (*Buffer region*)**.** Let be the **buffer region** denoted by Ψ, defined as the set of points from the wire space so:

$$\Psi = \{w \in \Delta \mid w \in \mathrm{U}^{\mathrm{OFF}} \cap \mathrm{D}^{\mathrm{ON}}\}$$

Let us think about this last definition before proceeding to the explicit definition. By picking an arbitrary point w that belongs to buffer region, we are implying that $w \in \mathrm{U}^{OFF}$, which means that bbit concentration will reach an equilibrium that will correspond to a normalized response less or equal than 5% of the maximum response. But when

the emitter is induced (and so the proper state is Δ^{ON}), $w \in \mathrm{D}^{OFF}$, meaning that the culture will never reach equilibrium as it is placed in the divergence region. Hence, the bbit concentration will start increasing in a quasi-linear manner.

As mentioned above, cultures lying in this operative region Ψ can store information. In this case, we have to imagine what happens with input signal of variable length. Let us assume that the nature of the input signal is anything that can be *added* and *removed* at will to the culture. A good example is blue light. There are several synthetic constructs that posses this property. The clearest case in my opinion is transcription factor EL222, a protein with the ability to interact with specific DNA sequences that when placed in the right position in the promoter region of a gene can promote or inhibit transcription [71]. Under this circumstance, blue light exposure will induce emitter strain, and bbit concentration will start increasing in the media. But because our culture is placed in the buffer region, bbit concentration will not reach any equilibrium and will keep increasing. This means that at some point we can always expect the receptor strain to be fully induced and produce output signal. But here comes the main difference with digital region: the time duration of the input signal will determine the maximum bbit concentration reached in the media. And because sink strain is responsible for leading this bbit concentration to 0 again, and at bbit concentrations greater than K_S the degradation rate becomes constant, **the time it will take for bbit concentration to reach almost 0 will depend on the duration of the input signal**.

This implies that buffer region can store the duration of the input signal in the same fashion a cooking timer does: the more we spend winding the cooking timer, the more it will take to stop its mechanism.

This is actually how we introduced biological synthetic wires, as the simplest way to implement a biological timer [134]. Recall here that we are claiming that the same strains, with the same genetic programming, and the same experimental conditions and input signal, can switch from being a digital wire to become an *involuntary* system to store temporal information. And this would be achieved just by adjusting the ratios between emitter, sink and receptor strains.

Now that we have exposed how and why this buffer region works in a qualitative manner, let me present the explicit definition of buffer region. The uninduced region, U, is defined as the set of points $w = (\varepsilon, \sigma, \rho)$, such the following inequality is satisfied: $0 \leq \varepsilon/\sigma \leq \varepsilon_5/\sigma_5$. And the divergence region is the wire space subset in which the following inequality holds: $\hat{\varepsilon}/\hat{\sigma} \leq \varepsilon/\sigma$. As in the case of digital region, these inequalities can be represented as intervals, being the intersection of these intervals the origin of buffer region. This is summarized in Figure 2.21, in which intervals are represented as a segment in the case of U region, and a semisegment in the case of D region (because the ratio ε/σ can reach any value up to ∞).

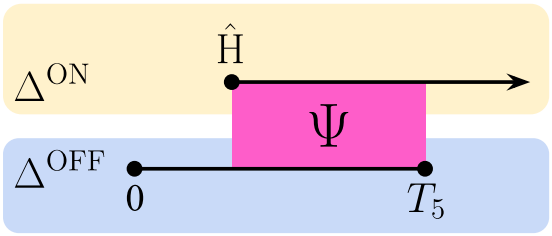


Figure 2.21: **Buffer region definition**. This region is the intersection between $\mathrm{U}^{\mathrm{OFF}}$ and D^{ON} regions.

Taking this into account, we can explicitly define our buffer region

in terms of model parameters:

Definition 16 (*Explicit definition for buffer region*)**.** Let be the **buffer region** denoted by Ψ, defined as the set of points $w = (\varepsilon, \sigma, \rho)$ from the wire space so:

$$\Psi = \left\{ w \in \Delta \mid \left(\frac{\hat{\varepsilon}}{\hat{\sigma}}\right)^{ON} \leq \frac{\varepsilon}{\sigma} < \left(\frac{\varepsilon_5}{\sigma_5}\right)^{OFF} \right\}$$

Again, this last definition allows us to define the **condition for buffer region existence**:

$$\boxed{\exists \Psi \neq \varnothing \iff \left(\frac{\hat{\varepsilon}}{\hat{\sigma}}\right)^{ON} < \left(\frac{\varepsilon_5}{\sigma_5}\right)^{OFF}} \tag{2.18}$$

And in an analogous way, we can reference to wires working in several experimental conditions and expecting our wires to work under any of these. In that case, because our parameters may vary when a wire is under a particular experimental condition. That would be associated with a different set of parameter values and hence a different buffer region Ψ/i for each condition $i \in \Omega$ the conditions space. And then, we can find a much more strong definition for buffer region that will grant to work under any experimental condition:

Definition 17 (*Strong definition for buffer region*)**.** Let be the **buffer region** denoted by Ψ, defined as the set of points $w = (\varepsilon, \sigma, \rho)$ from the wire space so:

$$\Psi = \bigcap_{i \in \Omega} \Psi / i$$

This last strong definition unlocks the **strong existence condition for buffer region**:

$$\exists \Psi \neq \varnothing \iff \max \left\{ \frac{\hat{\varepsilon}}{\hat{\sigma}} \Big/ i \right\}^{ON} < \min \left\{ \frac{\varepsilon_5}{\sigma_5} \Big/ i \right\}^{OFF} \quad \forall i \in \Omega \tag{2.19}$$

Now, with the proper tools to handle with buffer region, we can show a couple of simulations, like we did i the case of digital region, and using the same input signals, we can appreciate how the time duration of the response or output signal increases with the time duration of the input signal (Figure 2.22)

But how does response duration increases with respect to input duration? The answer is *nearly* proportionally. We say nearly because this is what can be observed if input duration is long enough, due to the fact that response in not really instantaneous with respect to input appearance. We will cover these time related topic in the following chapter.

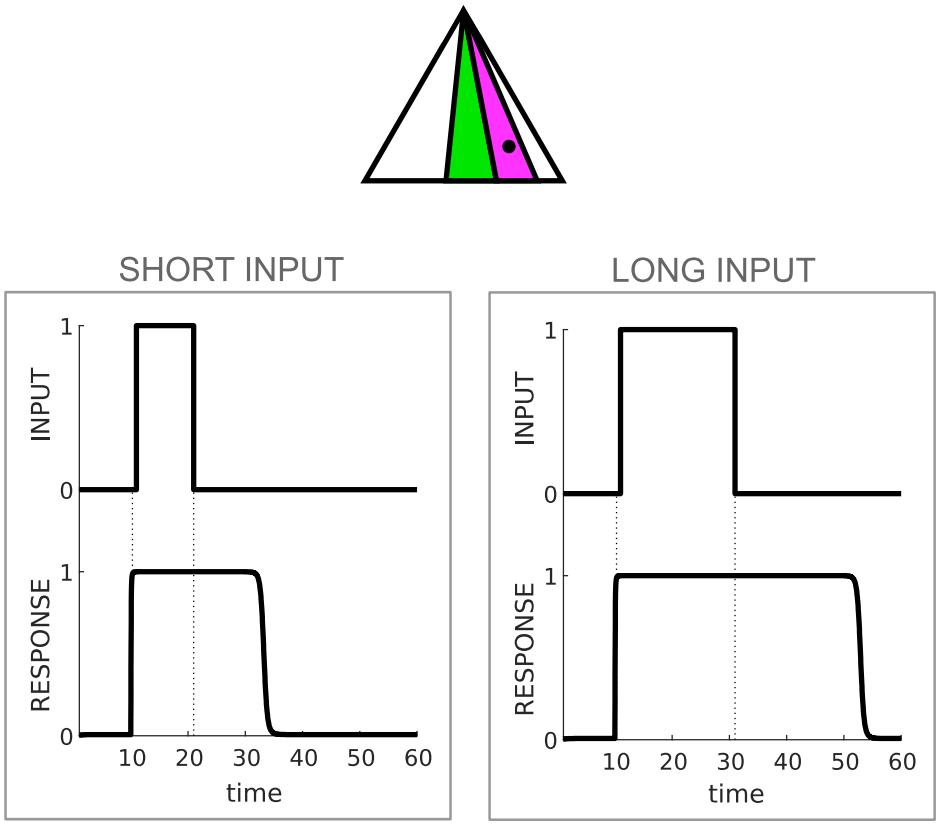
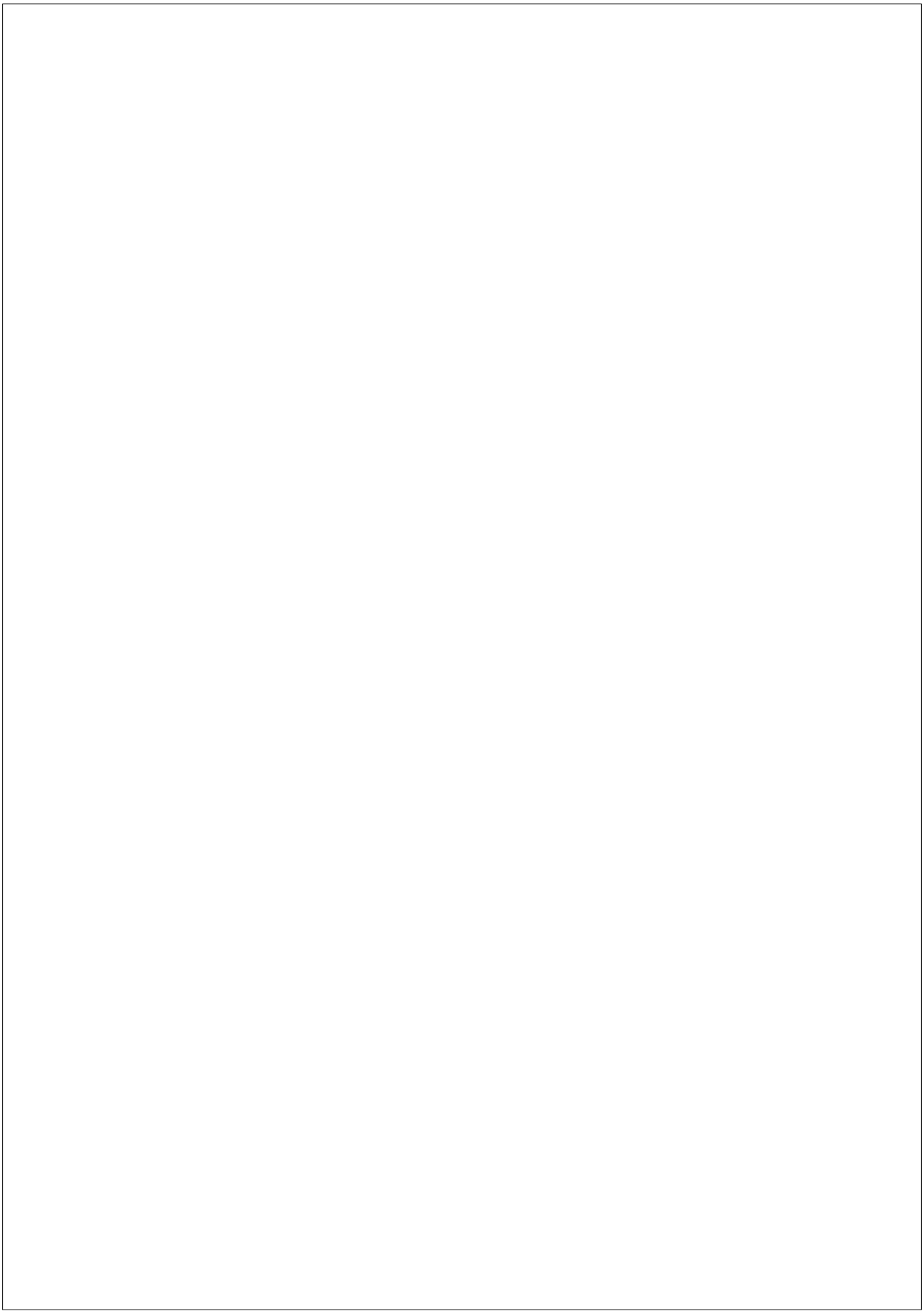


Figure 2.22: **Buffer region behaviour**. Output signal increases its duration depending on the input signal duration. Parameter values from Table 2.2.



Chapter 3

WIRE DYNAMICS

Whereas in the previous chapter we have tried to define in the most rigorous way possible an idealized wire architecture, in this chapter we will explore the consequences of this definition. First, we will address the question of the interplay between time and a wire. Second, we will explore the inferred properties by exposing our wire to digital periodic signals. Finally, we will extend our wire model to make it more descriptive by breaking the assumption of instantaneous behaviour in the emitter and in the receptor strains.

3.1. Wire temporal dynamics

Up to now, we have covered what happens with the different regions, states, and operative region. We have had a glimpse of how input signals can be modulated in digital and buffer regions. We have seen how digital region more or less preserves the duration of the input signal, whereas buffer region multiplies the output response duration.

In this chapter, we will dig into the characteristics and differences between the digital and buffer regions in terms of time.

3.1.1. Time-related definitions

Let us start by defining the basic variables that we will be using. The main reference framework in this case will start by considering a square pulse input signal. In this case, it may be helpful to imagine blue light as an input, with the protein EL222 in the emitter strain being responsible for transducing blue light exposure into biobit production and release in an instantaneous manner.

In Figure 3.1 a square input signal appears on top. From this top plot we can define t_{IN} as the of the input signal time duration. With respect to the bottom plot, we are seeing the response from the receptor, particularly the normalized response (as it is between 0 and 1). In blue, all time-points associated with a response less or equal to 0.05 (5% of the maximum response) are emphasized (blue shading shows the vertical region corresponding to this situation). Conversely, all time-points associated with a response greater or equal to 0.95 are shown in yellow (yellow shading shows the vertical region corresponding to this situation).

We define t_{OUT} as the time duration associated with normalized responses greater or equal to 0.95. But we can see other two magnitudes in Fig. 3.1. t_{ON} is the required time to switch from an uninduced to a fully induced response. In turn, t_{OFF} is a little bit trickier to define, or better, to properly explain its motivation. t_{OFF} it the time duration that ranges from the disappearance of the input signal up to the point

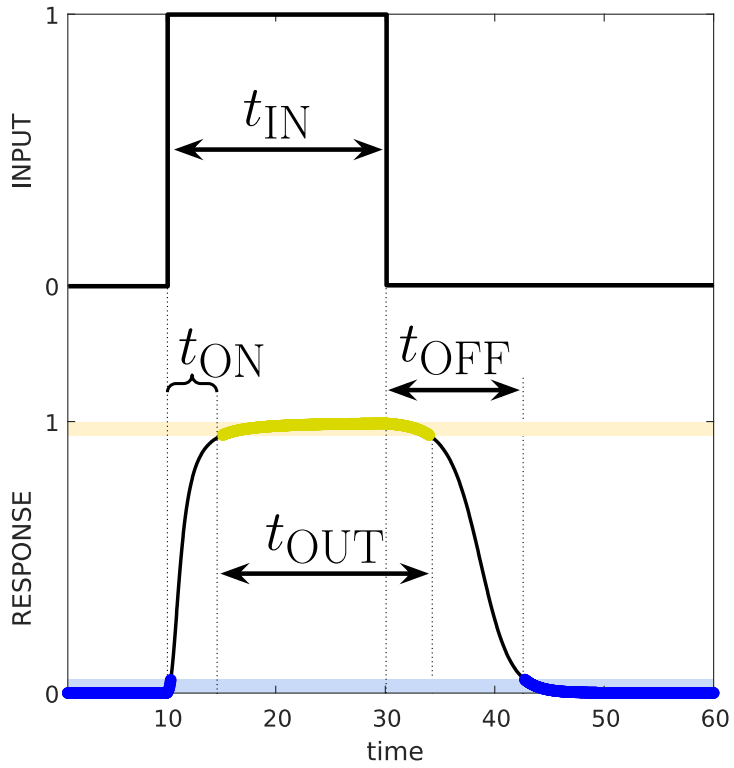


Figure 3.1: **Time-related variables**. When input is a square signal, we can define t_{IN} as its time duration. Then, the three time-related variables t_{OUT}, t_{ON}, and t_{OFF} are derived from the output signal. Parameter values for this archetypal simulation from table 3.1, but $P = 5$ in this case.

at which normalized response falls bellow 0.05. As the reader may see, the definition of t_{OFF} is fundamentally different from t_{ON}. If we define t_{OFF} in an analogous way as we have done with t_{ON}, then it should range from the moment the response drops below 0.95 up to the moment it reaches 0.05. Nevertheless, if define t_{OFF} as we have

Parameter	Value
P (Population)	50
K_E^{ON} (Induced emitter constant)	0.5
K_E^{OFF} (Uninduced emitter constant)	0.01
D_m (Maximum degradation)	1
K_S (Sink constant)	5
K_A (Inflexion point for Hill Eq.)	1
n (Hill coefficient)	3

Table 3.1: Parameter values used in Fig. 3.1

done, it will behave differently in digital and buffer regions. In digital region, an ideal wire would have $t_{\mathrm{OFF}} = 0$, whereas in buffer region, an ideal wire could never reach 0. Rather, it would approach a fixed value corresponding to the time extension promoted by the buffer region, and then the normalized response would *instantaneous* switch to 0 with extreme steepness.

3.1.2. Ideal wires and operative regions

Taking into account all constrains and definitions we have covered in this chapter, when we introduce these new time related variable: what would constitute an ideal wire? For notation purposes that will become clear in future sections, let me introduce a new definition, the Λ-region:

Definition 18 (Λ-*region*)**.** Let be the Λ**-region** denoted by Λ, defined as the set of points $w = (\varepsilon, \sigma, \rho)$ under the proper state Δ^{ON} such:

$$\Lambda = \mathrm{C}^{ON} \cup \mathrm{D}^{ON}$$

Note that any wire in the Λ-region will reach full receptor induction at some point, that is $\overline{r} = 1$. In this new region, we can talk about how an input signal is modified when it passes through the wire and generates an output. Recall that operative and buffer regions (Φ and Ψ respectively) belong to Λ-region, but there are points beyond buffer region that do not belong to neither Φ or Ψ, but they still belong to the Λ-region. The point of this new region is that for each point $w \in \Lambda$, we can define a new function, $\Lambda : w \to \Lambda(w)$ that will tell us how output signal is depending on the input signal when a wire is *ideal*. But what is an ideal wire?

If we consider a culture placed within the digital region, then we define the ideal wire as the one with both t_{ON} and t_{OFF} equal 0. In this scenario, $t_{\mathrm{IN}} = t_{\mathrm{OUT}}$ and hence the signal is perfectly transmitted. In that scenario. Because such wires cannot obviously exist, we expect our wires to minimize t_{ON} and t_{OFF} as much as possible is we pursuit the ideal wire with perfect input to output reconstruction. It is precisely in this scenario that we say that $t_{\mathrm{IN}} = \Lambda(w) t_{\mathrm{OUT}}$ with $\Lambda(w) = 1$ for every point that belongs to the digital region Φ.

In the case of a culture placed in the buffer region, t_{OFF} cannot be 0 by definition, as bbit concentration never reaches equilibrium and the more time the emitter is induced, the higher the bbit concentration, and hence the longer it will take for the sink strain to decrease bbit

concentration bellow T_5 threshold.

Considering $t_{\mathrm{ON}} = 0$, if we allow t_{OFF} to be as small as possible, then $t_{\mathrm{OUT}} = \Lambda(w)t_{\mathrm{IN}}$ with $\Lambda(w)$ being a value greater than 1. Hence, the ideal wire would have $t_{\mathrm{OFF}} = (\Lambda(w) - 1)t_{\mathrm{IN}}$.

We will properly define function $\Lambda(w)$ when studying t_{OFF}, as it does not play a relevant role in the first study case: t_{ON}. But in summary, an ideal wire is that one that minimizes both t_{ON} and t_{OFF}, being $\inf(t_{\mathrm{ON}}) = 0$ and $\inf(t_{\mathrm{OFF}}) = (\Lambda(w) - 1)t_{\mathrm{IN}}$[1].

3.1.3. The Δ^{OFF} to Δ^{ON} switching time: t_{ON}

The next question we have to answer is: where in the wire space t_{ON} is as low as possible? Dealing with this question is not particularly easy. One should start by Eq. 2.4 (the implicit solution for wire equation), and then find the time it takes for the wire to make $B_{eq} \geq T_{95}$ starting with $B(0) = 0$. We can perform numerical simulations to shed some light to the problem.

In Figure 3.2, a fine mesh grid of points was selected along the wire space. Simulations assuming $B(0) = 0$ as initial condition and then switching to Δ^{ON} were done using parameters from Table 3.1, and t_{ON} was obtained for each point. For points whose associated B_{eq} value corresponds with a response lower than 0.95 (not fully induced receptor), t_{ON} was counted as the time required to reach 95% of *its maximum achievable response*. A contour plot shows the t_{ON}-isolines. Horizons T_5, T_{95} and $\hat{\mathrm{H}}$ are shown for Δ^{ON}. We can see that there is a singularity on ρ vertex, where t_{ON} tends to infinity. Nevertheless, this change is very non-linear, and most of the wire space presents a rather

[1]The function $\inf$ is the *infimum*, that gives the highest of the lower bounds in an ordered set without necessarily belonging to it

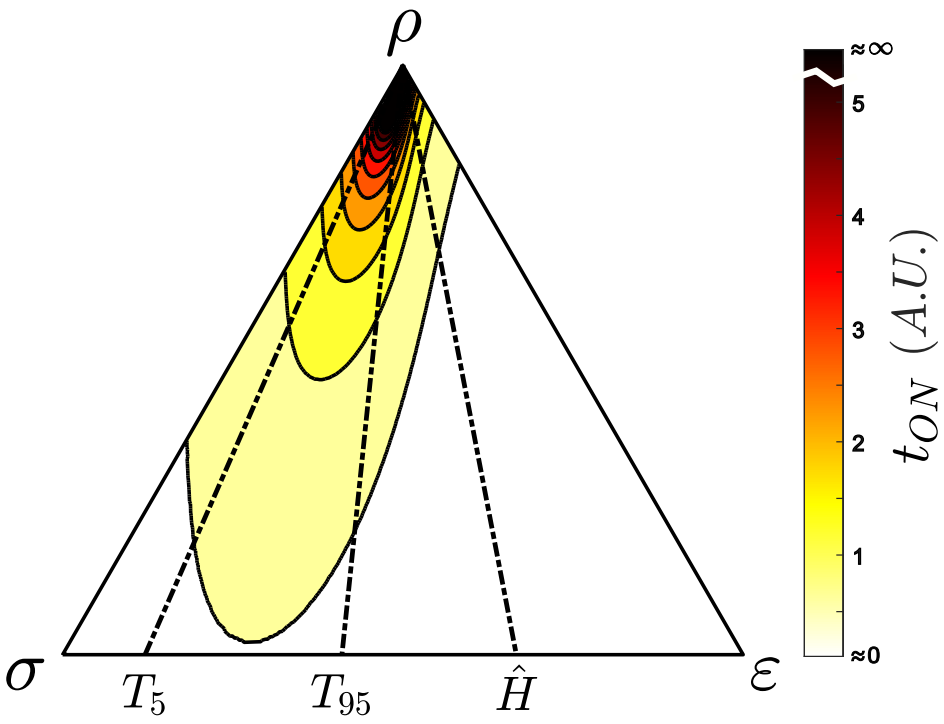


Figure 3.2: **Wire space and** t_{ON}. A contour plot for t_{ON} with the t_{ON} isolines is shown. Parameter values from Table 3.1.

gradual increase in t_{ON} as the points get closer to ρ vertex. If we focus on the induced region I^{ON}, between T_{95} and $\hat{H}^{ON}$ horizons, we can appreciate that most of the region presents low t_{ON} values, finding the minimum in the neighborhood of the point of divergence. Obviously, the points near the ε vertex show the minimum t_{ON} values, as the bbit asymptotic divergence rate finds its maximum when $\varepsilon = 1$. It is also worth noting that there seems to be a B-isoline corresponding to a bbit concentration near K_M, the inflexion point of the Hill equation for the receptor, that shows the highest t_{ON} values for any fixed ρ value.

Constraining our dissertation to what happens in digital and buffer regions, because both regions share the subset U^{OFF}, both will always start a fixed B_{eq} concentrations. And trivially, these concentrations will always be associated with normalized responses below or equal to 0.05. We can consider, as we have done in Fig. 3.2, that $B(0) = 0$ without loosing any generality (in any case, by considering $B(0) = 0$, we would be overestimating t_{ON} by a small quantity, as $B_{eq} > 0$ for any interior point in the convergence region). And fixing an arbitrary point within the digital region, a new B_{eq} will be reached when Δ^{ON} is new proper state. The same argument applies to buffer region, but in this case, what if constant is $\hat{v}$, the bbit asymptotic divergence rate. Because T_{95} is a fixed parameter, the time it will take for a wire in the buffer region to reach T_{95} will be fixed for any fixed point.

As a conclusion, t_{ON} **is always fixed for any fixed point within the operative regions**.

3.1.4. The Δ^{ON} to Δ^{OFF} switching time: t_{OFF}

In the case of t_{OFF}, the digital and buffer regions differ in a qualitative manner. Whereas in digital region Φ when B_{eq} is reached under proper state Δ^{ON}, the time it will take for a fixed culture-point to reach the new $B_{eq} \approx 0$ under proper state Δ^{OFF} is fixed, in buffer region it works differently, as we will see. But as a first remark: t_{OFF} is fixed for any point within the convergence region.

In buffer region Ψ, an ideal wire will present $t_{\mathrm{OUT}} = \Lambda(w)t_{\mathrm{IN}}$, with $\Lambda(1) \geq 1$ a value that will depend on the position of the point in the buffer region[2]. If we can estimate $\Lambda(w)$, we can find a lower bound to

[2]Technically, $\Lambda(w) > 1$ almost everywhere in the buffer region, because the

t_{OFF} it terms of t_{IN}. We will start by assuming t_{IN} arbitrarily large. In that scenario, Eq. 2.2 (the wire differential equation) asymptotically becomes Eq. 2.9 (the asymptotic divergence wire differential equation):

$$\hat{v} = \frac{\mathrm{d}B_{\hat{v}}}{\mathrm{d}t} = P(K_E\varepsilon - D_m\sigma) \tag{3.1}$$

whose solution is Eq. 2.10:

$$B_{\hat{v}}(t) = P(K_E\varepsilon - D_m\sigma)t + B(0) \tag{3.2}$$

If we assume $B(0) = 0$, for t_{IN} large enough, $B(t_{\mathrm{IN}}) \approx B_{\hat{v}}(t_{\mathrm{IN}})$, and thus, the reached bbit concentration after t_{IN} in the Δ^{ON} proper state will be:

$$\boxed{B(t_{\mathrm{IN}}) \approx B_{\hat{v}}^{ON}(t_{\mathrm{IN}}) = P(K_E^{ON}\varepsilon - D_m\sigma)t_{\mathrm{IN}}} \tag{3.3}$$

Also, at some point, the wire proper state Δ^{ON} will switch to Δ^{OFF} (at which $K_E^{OFF} = 0$ as an approximation). And hence, sink strain is the only one operating here, becoming this:

$$\hat{v}^{OFF} = \frac{\mathrm{d}B_{\hat{v}}^{OFF}}{\mathrm{d}t} = -D_m\sigma P \tag{3.4}$$

whose solution is:

$$B_{\hat{v}}^{OFF}(t) = -D_m\sigma Pt + B(0) \tag{3.5}$$

In this last equation, we want $B_{\hat{v}}^{OFF}(t) = 0$, which will only occur when $t \approx t_{\mathrm{OFF}}$. Also, in this last equation, $B(0) = B(t_{\mathrm{IN}})$, all together becoming:

$$0 = -D_m\sigma Pt_{\mathrm{OFF}} + P(K_E^{ON}\varepsilon - D_m\sigma)t_{\mathrm{IN}} \tag{3.6}$$

equality only holds in the horizon of divergence

Solving for t_{OFF} we get:

$$\boxed{t_{\mathrm{OFF}} \approx \left(\frac{K_E^{ON} \varepsilon}{D_m \sigma} - 1 \right) t_{\mathrm{IN}}} \tag{3.7}$$

Recall that this last formula is an approximation for t_{OFF}, as we are using linear approximations for our differential equations. Now it is turn to remember:

$$\inf(t_{\mathrm{OFF}}) = (\Lambda(w) - 1) t_{\mathrm{IN}} \tag{3.8}$$

With $w \in \mathrm{D}^{ON}$ an arbitrary point. And from this identity, we can obtain an analytic approximation for $\Lambda(w)$, the scaling factor that verifies $t_{\mathrm{OUT}} = \Lambda(w) t_{\mathrm{IN}}$ in the divergence region D when the proper state is Δ^{ON}:

$$\boxed{\Lambda_{\mathrm{D}}(w) = \frac{K_E^{ON} \varepsilon}{D_m \sigma}} \tag{3.9}$$

From this point to the end, we will write Λ_Φ, Λ_Ψ or Λ_{D} to inform the reader that we are constraining our function $\Lambda(w)$ to digital, buffer of divergence region respectively (when the proper state is Δ^{ON}) when convenient. And that will happen depending on where in the Λ-region the point w is placed. It is at this point, that we can properly define function $\Lambda(w)$, the **rescaling factor function**:

$$\boxed{\Lambda(w) = \begin{cases} 1 & w \in \Phi \\ \frac{K_E^{ON} \varepsilon}{D_m \sigma} & w \in \mathrm{D}^{ON} \end{cases}} \tag{3.10}$$

with $w = (\varepsilon, \sigma, \rho)$. And thus, the following identity expresses the time rescaling in the divergence region D^{ON} (and hence in the buffer region) when the emitter is induced:

$$\boxed{t_{\mathrm{OUT}} = \frac{K_E^{ON} \varepsilon}{D_m \sigma} t_{\mathrm{IN}}} \tag{3.11}$$

Note that this last equation assumes an ideal wire, but the greater t_{IN}, the closer the scaling factor to $\Lambda(w)$, as the deviations in bbit time series from the linear asymptotic solution will be negligible for bigger B values. The scaling factor $\Lambda(w)$ is adimensional, and ranges from 1 when $K_E\varepsilon = D_m\sigma$ (the horizon of divergence $\hat{\mathrm{H}}$) to ∞ for points placed near $\overline{\rho\varepsilon}$ side, where $\varepsilon/\sigma \to \infty$.

Also, explicitly representing $\Lambda(w) = K_E^{ON}/D_m \cdot \varepsilon/\sigma$, it is more evident that $\Lambda(w)$ solely depends in the ε/σ ratio, meaning that $\Lambda(w)$ presents Λ-isolines irradiating from the top ρ vertex towards the $\overline{\sigma\varepsilon}$ side of the triangle. A demonstration for such statement is analogous as such for B-isolines Lemma 2.3.3.

We know that all points lying in the divergence horizon $\hat{\mathrm{H}}$ will be ideally associated to $\Lambda(w) = 1$ (also, here B_{eq} diverges, constituting a singular phenomenon). And from Eq. 3.11 we know that $\Lambda(w) \in [1, \infty)$. But now we will constrain our cultures within the buffer region Ψ, in which cultures must verify the inequality $(\hat{\varepsilon}/\hat{\sigma})^{ON} \leq \varepsilon/\sigma \leq (\varepsilon_5/\sigma_5)^{OFF}$. We already know what happens to those cultures placed in the horizon of divergence, but analysing the associated $\Lambda(w)$ for cultures lying on top of T_5^{OFF} horizon, we can analytically obtain the boundaries for $\Lambda(w)$ constrained to the buffer region Ψ.

Cultures placed in the T_5^{OFF} horizon will present a ratio $\varepsilon/\sigma = (\varepsilon_5/\sigma_5)^{OFF}$, so in this region, the maximum value $\Lambda(w)$ can reach is:

$$\boxed{\max(\Lambda_\Psi) = \frac{K_E^{ON}}{D_m}\left(\frac{\varepsilon_5}{\sigma_5}\right)^{OFF} = \frac{K_E^{ON}}{K_E^{OFF}}\frac{T_5}{T_5 + K_S}} \tag{3.12}$$

Hence, we can analytically determine the boundaries of $\Lambda(w)$ for cultures placed within the buffer region Ψ ($\forall w = (\varepsilon, \sigma, \rho) \in \Psi$) as:

$$\boxed{\Lambda_\Psi(w) \in \left(1, \frac{K_E^{ON}}{K_E^{OFF}}\frac{T_5}{T_5 + K_S}\right)} \tag{3.13}$$

Note how $\max(\Lambda_\Psi)$ depends on the ratio between emission constants when the emitter strain is induced or uninduced. If an emitter presents negligible leakage in the genes responsible for the bbit biosynthesis, $K_E^{OFF} \approx 0$, and hence Λ_Ψ will range from 1 to infinity. So, **the leakiness of the emitter strain's genes determines the maximum $\Lambda(w)$ value that can be reached in buffer region**. Parameter T_5 determines the lower threshold for bbit detection by receptor strain. If receptor strain is very sensitive (low T_5), K_S (the parameter that determines the affinity of the enzymes responsible for bbit degradation[3]), will contribute to decrease $\max(\Lambda_\Psi)$ for low affinities (high K_S). On the other hand, if receptor strain presents low sensitivity to bbit molecules (high T_5), then K_S may have negligible contribution.

I have to admit that I like variable $\Lambda(w)$ because in its analytical expression is condensed some of the key features of our three strains: the inducible nature of emitter strain, the bbit affinity of sink strain, and the sensitivity of receptor strain.

Last thing I would like to share with the reader is what I consider a beautiful and clear comparison between digital and buffer regions. In Figure 3.3, two different points were selected from digital region Φ (in green) and buffer region Ψ (in magenta), both with the same amount of receptor strain (same ρ value). Simulations for each point were done using as input a square signal with time duration t_{IN}, and then t_{OFF} was calculated for each case. This example illustrates that whereas the culture in digital region shows a rather constant t_{OFF} values (green dots), the culture in buffer region responds in a linear manner with t_{IN}, being the $\Lambda(w) - 1$ the slope to such line.

[3]Even when K_S determines the affinity of the enzyme for its substrate, this correlation is indirect, meaning that the lower K_S, the higher the affinity.

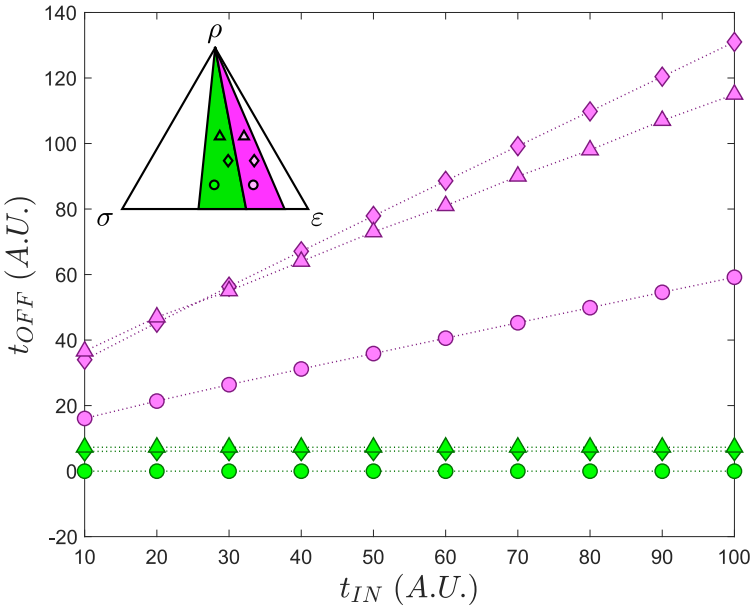


Figure 3.3: t_{OFF} **as a function of** t_{IN} **in digital and buffer region.** Three points were selected for each region (digital or buffer). Several simulation were done with different t_{IN} values, and t_{OFF} values were calculated for each case. Parameter values from Table 3.1.

This constitutes a demonstration that biological wires under the proposed architecture can store information about t_{IN} in a transient manner. I illustrate this as follows: imagine entering a room with the light on. By changing the switch to the OFF position, the light will take to turn off an amount of time that is proportional to the time duration the switch was in the ON position, allowing to know when the switch was turned on in first place. This is why I call t_{OFF} a **buffer time**. The reader might argue then that it also takes time for a culture in the digital region to *turn off*. Yes, but this amount of time is independent of the input duration, is characteristic of the position of the culture within the

wire space. And this is the reason why I call this characteristic time a **lag time**, and cannot carry any information beyond how ideal or not is a wire.

3.1.5. The time-response interplay

The last subject I would like to point at is the interplay that exists between time (t_{ON} and t_{OFF}) and response. When we discussed about receptor strain and output response, we mentioned that response (normalized or not) would be measured as response per receptor strain unit. Nevertheless, in a co-culture with our three strains mixed, it might be difficult to specifically measure response per unit of receptor strain. In turn, it would be easier to measure response per unit of culture[4], and then extrapolate the obtained value to the receptor strain ratio, which at the end is just multiplying by ρ. This is what is shown in Figure 3.4.

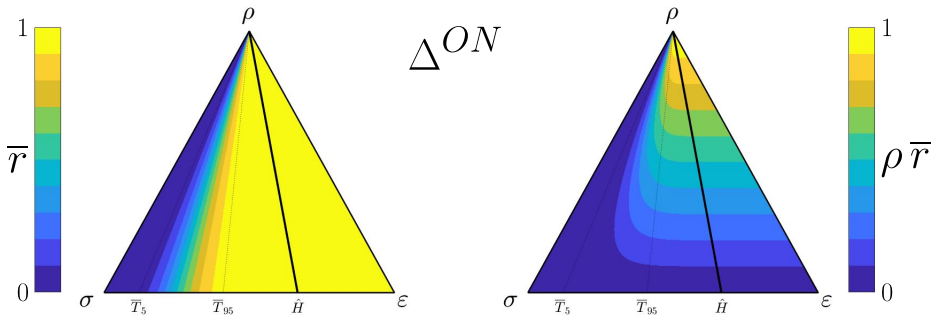


Figure 3.4: **Response per unit receptor and per unit culture**. Parameters from Table 3.1.

In Fig. 3.4, a wire in proper state Δ^{ON} is shown. In the left panel, the normalized response per receptor strain unit, shows how all cul-

[4]By unit of culture I mean dividing by the corrected $\mathbf{OD_{600}}$.

tures placed in the convergence and divergence region present fully induced receptor strain. However, receptor strain is less and less represented when cultures a nearer of $\overline{\sigma\varepsilon}$ triangle side. And even when these cells are fully induced, this has two consequences:

- **Response detection threshold**: if we are measuring our receptor's output signal, we may encounter sensitivity issues for low ρ values. Imagine we want to measure fluorescence, but we cannot sense fluorescence below some arbitrary threshold. Hence, experiments using fully induced receptor strain mixed with a non-fluorescent equivalent strain at different and known ratios must be done to find the response detection threshold.

- **Ineffective information flow when combining wires**: if the receptor's output is not meant to be measured (like GFP fluorescence), but to synthesize another different biobit as part of another new wire, we may encounter some problems, because the low ratio of receptor strain is equivalent to low emission constant for the next wire. Note that here this strain is both receptor of one biobit, and emitter of another different biobit.

And it precisely with this last second option that I spot a dilemma: the closer to the ρ vertex a culture is, the less of a problem constitutes the response detection threshold and the ineffective information flow for connected wires. But there is a big *but*: near the ρ vertex, both t_{ON} and t_{OFF} increase, moving away our wire from the idea wire. In summary: the wire gets slower, and input signals get distorted.

So at the end, a bioengineer has to choose somewhere between two undesirable design extremes: slower and less ideal wires for high

receptor strain ratios, or faster and more ideal wires at expense of low receptor strain ratios.

3.2. Digital periodic signals

In the previous sections we have focused on the properties of a single wire. And the reader may think that we first need to know how single wires work before connecting them and doing something useful. But this is not true. Thanks to the buffer region, a single wire can present interesting properties that operate over digital periodic signals. In this section, after explaining the basic definitions, we will explore how a digital periodic signal (DPS) is modified when a wire processes it. Also, we will show that wires can filter DPSs if all requirements are fulfilled, and how we can *program* a wire to filter different DPSs by just adjusting the relative ratios.

3.2.1. Basic definitions

A digital periodic signal, DPS, is a series of pulses with a period τ. The first part of the period represents presence of emitter strain's inducer or input, and the second part a lack of it. This division in the period can be described with the duty cycle, denoted by d_c, which describes the percentage of the period that corresponds to presence of input (Figure 3.5).

We will study how a DPS is modified when it acts as an input signal and how the resulting output is, all in terms of its duty cycle. In this sense, divergence and buffer regions play the major role as we shall see.

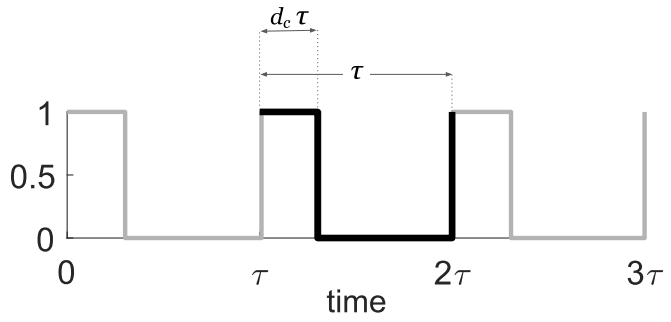


Figure 3.5: **Duty cycle definition**.

3.2.2. Duty cycle modification

We are going to start this subsection directly doing simulations to get some intuition about the alteration of DPSs piped through a wire as inputs. In Figure 3.6 we can see a DPS in the top left panel. In that case, $\tau = 10$ and $d_c = 30\%$. Then 200 points lying in the red segment $\overline{\mathrm{AB}}$, with $\rho = 0.5$ were selected to do simulations (bottom panel showing the wire space with operative states). For each point, simulations were done using parameters from Table 3.1. The result of each simulation was normalized to get $\overline{r}$, the normalized response, and then each resulting data is shown in a squared color map (top right panel). The vertical axis corresponds to time in arbitrary units, the horizontal axis to position with respect to segment $\overline{\mathrm{AB}}$, and the color scale shows the normalized response. In the color map, the digital and buffer regions, denoted as Φ and Ψ respectively, are delimited with vertical grey dashed lines. The continuous black lines delimit where T_5 concentration was reached (black line delimiting blue area), and where T_{95} was reached (black line delimiting yellow area).

We can see that all simulations to the left of the digital region show

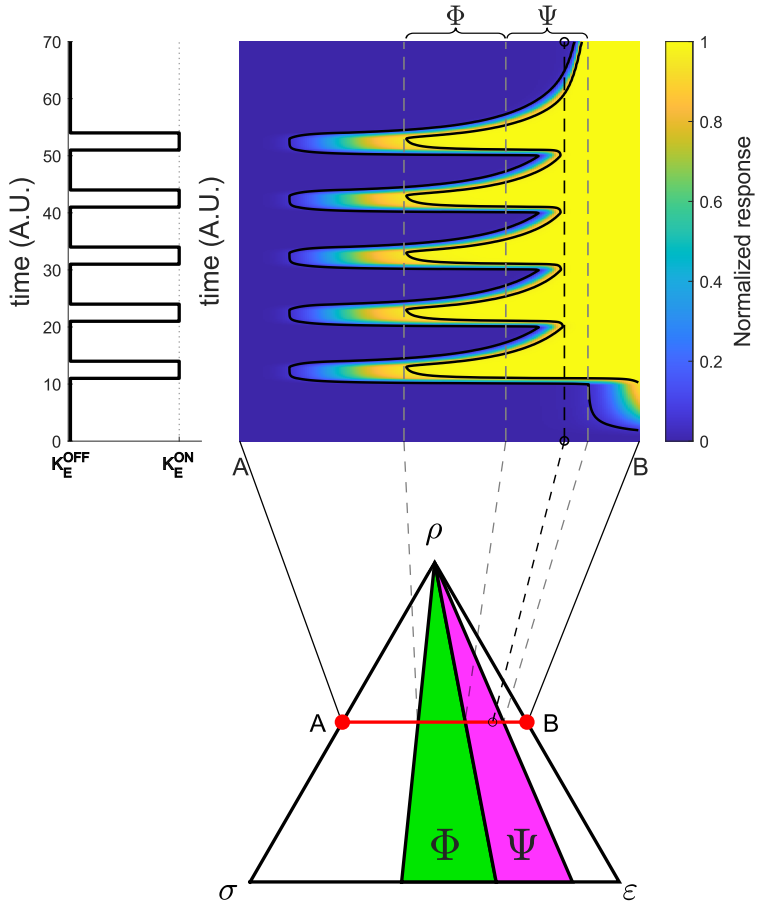


Figure 3.6: **DPS piped through a subset of the wire space**. Parameter values from Table 3.1.

no response or a variable level of receptor induction, whereas all simulations in the digital region Φ show full level of receptor induction. It is precisely in this digital region that we can appreciate our simulated wire is not ideal, as for every simulation, $t_{\mathrm{ON}} > 0$ and $t_{\mathrm{OFF}} > \Lambda(w) - 1$. This is evident because there is not an instantaneous jump from fully

uninduced receptor to fully induced receptor and vice versa. Because we already know that our wire is not ideal, this implies that we can expect the same effect in the buffer region.

In the case of the buffer region, the output signal increases its duty cycle with respect to the DPS by at least $\Lambda(w)$. Note that this *at least* is very important, because $t_{\mathrm{OFF}} > \Lambda(w) - 1$. But at least, we can find a lower bound for this increase in duty cycle with $\Lambda(w)$. But let us no anticipate what is about to come. Because something relevant happens here in buffer region: there is a point that makes this increase in the output d_c of the output signal reach 100%. And from that point to the right, the output signal is no longer a periodic signal. Receptor strain is fully induced always. In other words: there is an specific point that **filters the input DPS**. This is denoted with a vertical dashed black line. And from that point to the right, receptor cells get fully induced. Moving to the right beyond the buffer region (the remaining D^{ON} region), our wire starts to show some level of induction even when there is no inducer in the media.

The questions that we shall answer here are: Can we find the relation between the duty cycle of the input and that of the output? And, can we estimate where is the filtering threshold? let us find out.

When dealing with DPSs, d_c is a percentage, but we will treat the duty cycle d_c as a number between 0 and 1. For properly comparing duty cycles, we will denote d_c^{IN} for the duty cycle of the input signal, and d_c^{OUT} for the duty cycle of the output signal.

The first formalism is that, focusing on a single period τ of the input signal, the part of the period that comprises presence of inducer has a time duration of $d_c^{\mathrm{IN}}\tau$. This implies that $t_{\mathrm{IN}} = d_c^{\mathrm{IN}}\tau$. Now, if we assume that our DPS is fed as input into a culture lying within the buffer region

when the proper state is Δ^{ON}, and recalling Eq. 3.9, that describes how an output signal is scaled when the wire is ideal, we get:

$$t_{\mathrm{OUT}} = \Lambda(w) t_{\mathrm{IN}} = \Lambda(w) d_c^{\mathrm{IN}} \tau \tag{3.14}$$

If we divide this last expression by the period τ at both sides, we obtain the duty cycle of the output region in terms of the input's duty cycle and the scaling factor $\Lambda(w)$:

$$\boxed{d_c^{\mathrm{OUT}} = d_c^{\mathrm{IN}} \Lambda(w)} \tag{3.15}$$

And thanks to the definition of function $\Lambda(w)$ (Eq. 3.10, we can properly unpack last equation and precisely define d_c^{OUT} in terms of d_c^{IN} for cultures lying in the Λ-region:

$$\boxed{d_c^{\mathrm{OUT}} = \begin{cases} d_c^{\mathrm{IN}} & w \in \Phi \\ d_c^{\mathrm{IN}} \frac{K_E^{ON}}{D_m} \frac{\varepsilon}{\sigma} & w \in \mathrm{D}^{ON} \end{cases}} \tag{3.16}$$

This last expression confirms that when a culture is lying in the divergence horizon $\hat{\mathrm{H}} \subset \mathrm{D}$, where $K_E \varepsilon = D_m \sigma$, there is a perfect match between input and output duty cycles ($d_c^{\mathrm{OUT}} = d_c^{\mathrm{IN}}$)[5]. Nevertheless, the ε/σ ratio can be arbitrarily big when cultures are close enough to the $\overline{\rho\varepsilon}$ triangle side. So now we want to constrain our analysis to what happens within the buffer region. To do so, we need to impose that $\Lambda(w) = \max(\Lambda_\Psi)$, to find the maximum output duty cycle that we can expect in the buffer region. Using Eq. 3.12, we obtain the **maximum output duty cycle in buffer region**:

$$\boxed{\max(d_{c,\Psi}^{\mathrm{OUT}}) = d_c^{\mathrm{IN}} \max(\Lambda_\Psi) = d_c^{\mathrm{IN}} \frac{K_E^{ON}}{K_E^{OFF}} \frac{T_5}{T_5 + K_S}} \tag{3.17}$$

[5]Recall that in $\hat{\mathrm{H}}$ the scaling factor Λ is 1.

This last formula tells us that there is a fundamental limit to the expected duty cycle of the output DPS given an input duty cycle. Nevertheless, if the emitter has almost zero leakiness, we may expect to get arbitrarily big $\max(d_{c,\Psi}^{\mathrm{OUT}})$ values. The reader may notice that $\max(d_{c,\Psi}^{\mathrm{OUT}})$ can even be bigger than 1, a case that seems paradoxical because d_c should never be greater than 1. Later we will see that there is some scenarios in which $\max(d_{c,\Psi}^{\mathrm{OUT}}) > 1$ has some useful meaning.

3.2.3. Filter horizons

In the last subsection, we have set the rules to know how an input DPS is modified when passing through a wire. We have seen that the input duty cycle is modified by the scaling function $\Lambda(w)$, which directly depends on the ε/σ ratio of a given duty cycle d_c^{IN} (Eq. 3.16). And we finished by pointing out that $\max(d_c^{\mathrm{OUT}})$ may equal 1. When a duty cycle equals 1, we are not in front of a DPS anymore, because 100% of the period τ, we have full induction of the receptor cells. In other words: there is no valid period because the input DPS has been filtered.

Before digging into the mathematical formalism, let us do some simulations to get some intuition about what is going on. In Figure 3.7, a fine mesh of points were distributed across the whole wire space. For each point, several simulations were done using parameter values from Table 3.1, and using as input DPSs with $\tau = 20$, and input duty cycle ranging from 10% to 90% in intervals of 10. Then, when a specific point had the property of filtering an input DPS, the position of the point together with the input duty cycle was stored. Figure 3.7 shows where the filtering points for each input duty cycle are located as continuous colored lines. Pink shaded region shows where the

buffer region is located.

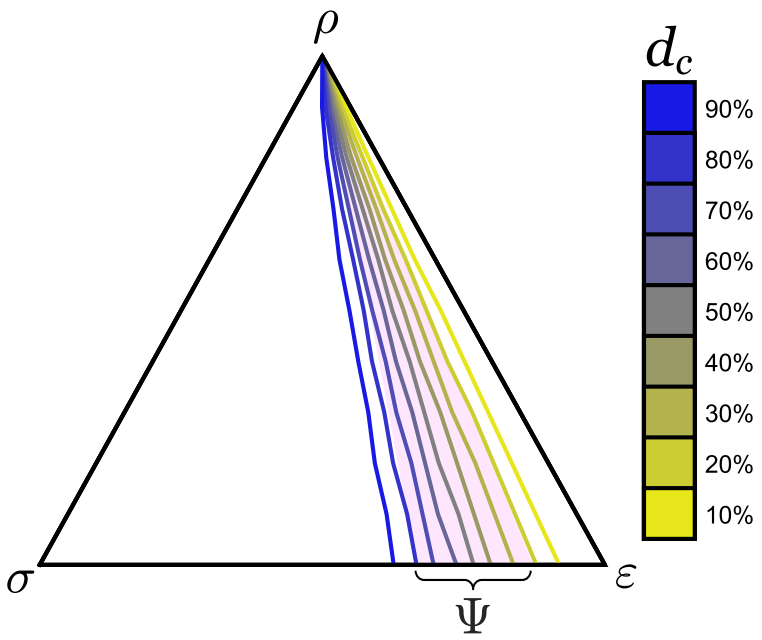


Figure 3.7: **Filter horizons for different duty cycles**. Parameter values from Table 3.1.

The first thing I would like the reader to notice, is that the simulated filter horizons seem to irradiate from the top ρ vertex to the bottom $\overline{\sigma\varepsilon}$ side. I say *seem* because I am pretty sure that everyone can see that these horizon seem to be straight, but they are not. This is because this is not an ideal wire, and some minor deviations are normal here. Should we expect these horizons to be straight lines? According to Eq. 3.16, the output duty cycle solely depends on the ε/σ ratio, which is constant across any segment irradiating from the top ρ vertex to the bottom $\overline{\sigma\varepsilon}$ side. Hence, our analytical prediction proposes a reasonable scenario not far from simulations. So far, so good. Actually, the

fact that a our wire is non-ideal is the reason why the reader may point out the fact that some input DPSs are filtered in the digital regions (those with high duty cycle). This happens because of the lag time, that is enough to filter the signal.

But let us go back to our discussion about Figure 3.7. Given the fact that every simulated horizon is not a straight segment, but rather a curvy one, that means that every horizon is bounded by two different ε/σ ratios. If we collect these boundaries for each input duty cycles, we can get some intuition on how the filtering horizons are related with the ε/σ ratios. And this is exactly what is shown in Figure 3.8:

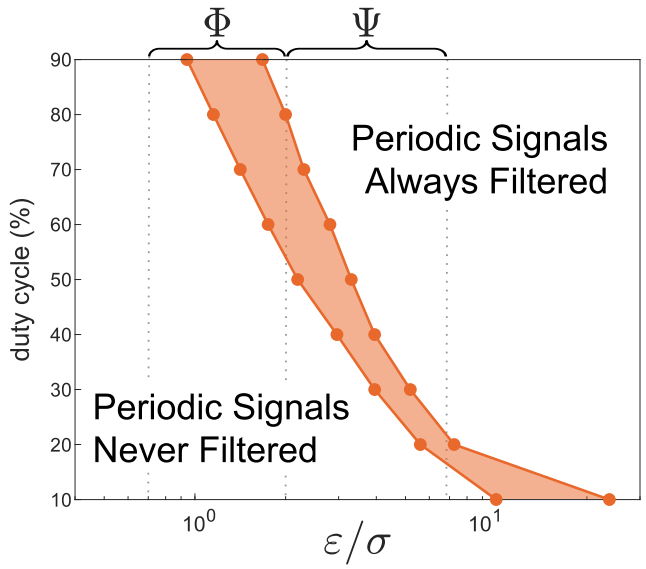


Figure 3.8: **Filter phase plane**. Parameter values from Table 3.1.

Recall that in Figure 3.8 we are using the simulations used in Figure 3.7. What we can see is that given an input duty cycle, we can always find a ε/σ ratio what will ensure that our input DPS will be fil-

tered. The orange shaded region depicts ratios in which our input DPS may or may not be filtered precisely depending on the specific location of the point within the wire space. The next question is: can we analytically attack this problem and establish some general laws for filtering DPSs?

If we want to filter the input DPS, then we have to impose $t_{\mathrm{OUT}} = \tau$. Hence, Eq. 3.15 becomes $\tau = \Lambda_{\mathrm{D}} d_c^{\mathrm{IN}} \tau$. Note that the solution to this equation is independent from period τ, meaning that we just care about the duty cycle d_c when dealing with signal filtering. At the end, the solution becomes:

$$\boxed{\Lambda_{\mathrm{D}} = \frac{1}{d_c^{\mathrm{IN}}}} \tag{3.18}$$

Remember that by using Λ, we are implicitly assuming that our wire is ideal. If we want to solve the ratio ε/σ that defines the filter threshold, we can *unpack* Λ_{D} and solve for the ratio. Then, it is possible to find the specific ratio ε/σ that filters a given duty cycle:

$$\boxed{\left(\frac{\varepsilon}{\sigma}\right)_{d_c^{\mathrm{IN}}} = \frac{D_m}{d_c^{\mathrm{IN}} K_E^{ON}}} \tag{3.19}$$

I would like to share a relevant warning about filter horizons: a filter horizon guarantees that an input DPS will be filtered if the ratio ε/σ is *higher* that the value in Eq. 3.19. That does not mean that an input DPS cannot be filtered if that ratio is smaller. In that case, it will simply depend on the specific period τ and the non-ideality of the wire. Note that if we consider an ideal wire, and a culture placed in the convergence region, then $\Lambda = 1$, and there cannot exist any filtering. Nevertheless, as mentioned before, the lag time $t_{\mathrm{OFF}} > 0$ may induce some filtering if d_c^{IN} is big enough. This last formula relates ratios with filtered duty cycles id the ratios correspond to points within the diver-

gence region D^{ON}. And it is easy to see that the bigger d_c^{IN}, the smaller the ratio, and so the closer the points will be to $\hat{\mathrm{H}}^{ON}$, the divergence horizon when the proper state is Δ^{ON}. Conversely, if d_c^{IN} gets smaller, the ratio will increase, corresponding to points near the $\overline{\rho\varepsilon}$ side of the triangle. But then a question arises: what is the smallest input duty cycle that the buffer region can filter?

To answer this question, we will impose the ratio to be $\varepsilon/\sigma = (\varepsilon_5/\sigma_5)^{OFF}$, and solve for d_c^{IN}, resulting in:

$$\boxed{\min d_{c,\Psi}^{\mathrm{IN}} = \frac{K_E^{OFF}}{K_E^{ON}} \frac{T_5 + K_S}{T_5} = \frac{1}{\max \Lambda_{Psi}}} \tag{3.20}$$

This last equation defines the interval of input duty cycles that buffer region Ψ can filter:

$$\boxed{\text{If } d_{c,\Psi}^{\mathrm{IN}} \in \left(\frac{1}{\max \Lambda_{\Psi}}, 1 \right), \text{ then } d_{c,\Psi}^{\mathrm{IN}} \text{ can be filtered.}} \tag{3.21}$$

Equation 3.20 shows again that the smaller the leakiness of the promoter, the less limited a wire will be to filter digital periodic signals (DPSs).

3.2.4. A study case: how to make a wire aware of weather seasons?

Thinking about this wire's new ability to produce differential responses depending upon the specific input duty cycle, I came with this problem: could a wire be be aware of weather seasons? And when I say a wire, I am not denying the possibility of a system that could be found in nature, as long as it somehow fulfils the requirements to

be analogous to a wire. And also: the only variable that is allowed to be sensed is light. It would not be allowed to measure temperature at all, just light. If we allow the nature of the input signal to be day light, we can imagine an emitter strain that will produce bbit in response to light if the light intensity is above to characteristic threshold. Such systems exist and before we have referred to the particular example of transcription factor EL222, that possesses the ability to sense the presence of blue light and promote gene expression from a specific promoter.

Every day of a year, and more evidently every month, has a number of daylight hours. Depending on where in the World you are right now, these daylight hours vary according to the season of the year. And because each year is a periodic event, the daylight hours per month follow an oscillatory pattern in duty cycle with respect to a full 24 hours day (the period). In my case, I am writing this dissertation from Barcelona. A quick search on the Internet allows to find the data shown in Table 3.2.

Now, the question is: can we program a wire to discriminate between warm months and cold months? And doing it solely according to the daily circadian cycle? The answer is yes: we can reuse the same parameters that we have been using during this chapter, and just by adjusting the ratios we can accomplish our task.

If we want to find the filter ratio, for a given input duty cycle d_c^{IN}, Eq. 3.19 (the filter ε/σ ratio) allows us to find the explicit solution for the filtering horizon. By placing the culture in the $\overline{\sigma\varepsilon}$ bottom side, where $\rho = 0$ and hence $\sigma = 1 - \varepsilon$, we can find the filter horizon explicit

Month	Daylight hours	d_c
January	9.58	0.40
February	10.62	0.44
March	11.94	0.50
April	13.33	0.56
May	14.51	0.60
June	15.12	0.63
July	14.83	0.62
August	13.80	0.57
September	12.48	0.52
October	11.10	0.46
November	9.89	0.41
December	9.25	0.39

Table 3.2: Monthly daylight hours in Barcelona

expression with a free parameter $k \in (0, 1)$:

$$(\varepsilon, \sigma, \rho)_{d_c^{IN}} = \left(\frac{D_m}{d_c^{\mathrm{IN}} K_E^{ON} + D_m} k, \frac{d_c^{\mathrm{IN}} K_E^{ON}}{d_c^{\mathrm{IN}} K_E^{ON} + D_m} k, 1 - k \right) \quad (3.22)$$

So any culture whose ε/σ ratio verifies $\varepsilon/\sigma \geq (\varepsilon/\sigma)_{d_c^{\mathrm{IN}}}$ will filter the input duty cycle d_c^{IN} in the receptor's output.

According to Table 3.2, if we want to filter May, June and July, setting a filter $d_c^{\mathrm{IN}} = 0.59$ we will able to filter *at least* these three months. And if we want to know if that input duty cycle can be filtered by cultures placed in the divergence region, we can use Eq. 3.21, the filtering interval for our buffer Ψ region. Using the parameter values from Table 3.1, we can calculate the lower bound for our filter interval and check if our input duty cycle can be filtered in the buffer region:

$$d_c^{\mathrm{IN}} = 0.59 \in (0.29, 1)$$

And hence our input duty cycle can be filtered by the buffer region. Substituting the parameter values in Eq. 3.22, we can find the filter horizon for our input duty cycle:

$$(\varepsilon, \sigma, \rho)_{d_c^{IN}=0.59} = (0.7722\,k, 0.2278\,k, 1-k)$$

By adding a tiny amount δ to ε and subtracting the same amount to σ when $k = 1$ so $((0.7722 + \delta)k, (0.2278 - \delta)k, 1 - k)$, we can be sure our culture will not be place on top of the filter horizon. In this case, we have choose $\delta = 0.011$. And by selecting k big enough (ρ small), we are forcing our wire to behave in a more ideal way, so $k = 0.95$ will be the value used for our simulations. Results are shown in Figure 3.9.

Although a month has between 28 and 31 days, for visualization purposes we will imagine the every month has just 5 days, more than enough to test the filtering capabilities of our wire. We can appreciate how the response get effectively filtered for months May, June and July. Nevertheless, an extra month has also been filtered: August. This is because given a duty cycle, bbit concentration is bounded even in the buffer region for cultures placed in the non-filtered region ($\varepsilon/\sigma < (\varepsilon/\sigma)_{d_c^{\mathrm{IN}}}$). Needless to say that we a constraining our analysis to DPSs, not to single pulses with arbitrary t_{IN} duration. What happens here is that an input DPS gets transformed into an output DPS with an increased duty cycle. And this can only happen, as I was saying, if bbit concentration in bounded *even* if the culture is placed in the divergence region. This can only occur if the bbit concentration falls bellow T_5 at some point.

But if the culture is in the filtered region, or when the duty cycle changes and now it can be filtered, which is the case of Fig. 3.9, then bbit concentration never has the chance to reach again T_5 before another period τ hits the system, increasing more and more its average

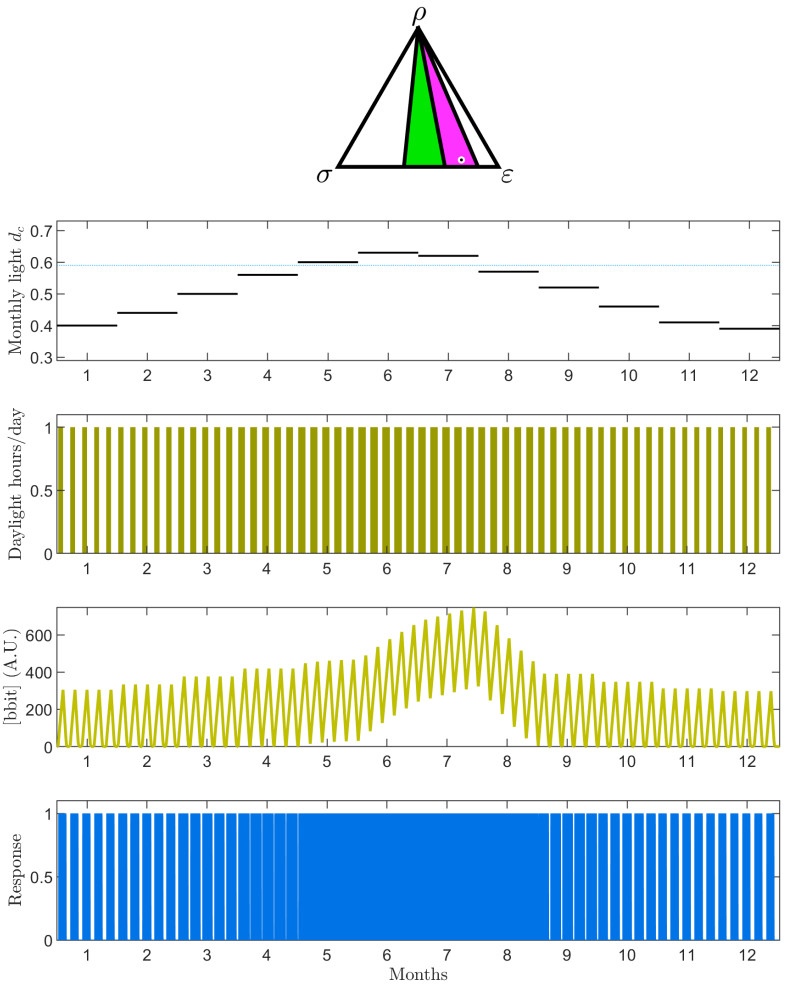


Figure 3.9: **Detecting warm months with a wire**.

concentration with each period. We can appreciate this effect in the

warm months of Fig. 3.9, in the bbit concentration panel (continuous yellow line).

At some point August arrives, and the input duty cycle is now smaller than the filter duty cycle. But now the avegare bbit concentration is high after three months of growth. Now the scenario is different: the decrease in bbit concentration each period τ is bigger than the accumulation that precedes it in the same period. The consequence here is that the average bbit concentration decreases. And the other consequence is that August, a month that presents a daylight duty cycle that should not be filtered, gets filtered as a consequence.

In other words, this filter horizons establish regions analogous to operative regions, but they require DPS to act like this. And in the same way t_{OFF} can tell us the value of t_{IN} when our culture is in the buffer region, the time it takes for a wire to restore its periodic output in the receptor after a filtering event can tell us the time extent of such event.

3.3. Non-instantaneous wire behaviour

According to our own wire model, K_E makes this switch between K_E^{OFF} and K_E^{ON} instantaneously. But we already know that this is a convenient simplification to gain knowledge about the ideal capabilities of a wire. Nevertheless, if we want to precisely define K_E^{ON} we can say that this is the maximum value that K_E can reach, and thus is a limit for input *intensity*. That is: for increasing values of input (considering the input is a chemical substance, then this would be a concentration), we can not increase K_E anymore.

But we may need to use a more descriptive model for extracting

information from time series. We know that the response, the output, will be proportional to a Hill-function depending on biobit concentration B. But that circumstance constrained to equilibrium, but what if want to describe what happens dynamically? Then the **non-instantaneous receptor equation** arises:

$$\boxed{\frac{\mathrm{d}[\mathrm{Out}]}{\mathrm{d}t} = L_R + \alpha_R \frac{B^n}{K_R^n + B^n} - \gamma_R[\mathrm{Out}]} \tag{3.23}$$

We have added a leakiness term L_R. Emitter strain is not so different from a receptor strain. Emitting biobit depends on the presence of an arbitrary input. And that input sensing is analogous to a receptor associated process. From that point of view, bbit emission is just a convenient output responding to an external input, and thus we can write an analogous descriptive differential equation for emitter strain:

$$\frac{\mathrm{d}K_E}{\mathrm{d}t} = L_E + \alpha_E \frac{[\mathrm{In}]^m}{K_E^m + [\mathrm{In}]^m} - \gamma_E K_E \tag{3.24}$$

This last equation present an important dissonance that it is not commonly found in biological modelling: the entity that varies is not a physical one, but a number that aggregates all the plausible proteins and molecules required for emitting an arbitrary biobit. Notice how in both previous equations, a degradation term is added. This is because output gene products required to be actively degraded to keep its concentration constant. The same applies for the gene products required to synthesize and emit biobit to the media.

If we solve Eq. 3.24 considering negligible leakiness ($L_E = 0$), we get the following solution:

$$K_E([\mathrm{In}], t) = \frac{\alpha_E}{\gamma_E} \frac{[\mathrm{In}]^m}{K_E^m + [\mathrm{In}]^m}(1 - e^{-\gamma_E t}) + e^{-\gamma_E t} K_E([\mathrm{In}], 0) \tag{3.25}$$

Applying limits for $t \to \infty$ in the absence of input ($[\mathrm{In}] = 0$) give us the value K_E^{OFF}. On the other hand, in input increases arbitrarily, then the Hill factor becomes 1; if we then apply limits for $t \to \infty$, then we get the value for K_E^{ON}:

$$\begin{cases} K_E^{OFF} = \lim_{t\to\infty} K_E(0,t) = 0 \\ K_E^{ON} = \lim\limits_{\substack{y\to\infty \\ [\mathrm{In}]\to\infty}} K_E([\mathrm{In}],t) = \frac{\alpha_E}{\gamma_E} \end{cases} \tag{3.26}$$

Now, we are going to define the input parameter $\{\mathrm{In}\}$ that can take two values, 0 or 1, depending if there is no input, or if there is input at concentration high enough so $K_E^{ON} \approx \frac{\alpha_E}{\gamma_E}$. By introducing this new parameter, we are implicitly removing the Hill factor, which allows us to write the non-instantaneous wire equation:

$$\boxed{K_E(t) = \frac{\alpha_E}{\gamma_E}\{\mathrm{In}\}(1 - e^{-\gamma_E t}) + e^{-\gamma_E t} K_E(0)} \tag{3.27}$$

This last equation, the **non-instantaneous wire differential equation**, captures in the most simple way how adding input in the media does not translate into a instantaneous switch from K_E^{OFF} to K_E^{ON}. Together with the analogous differential equation for the receptor strain (Eq. 3.23), we can obtain the **non-instantaneous wire system**:

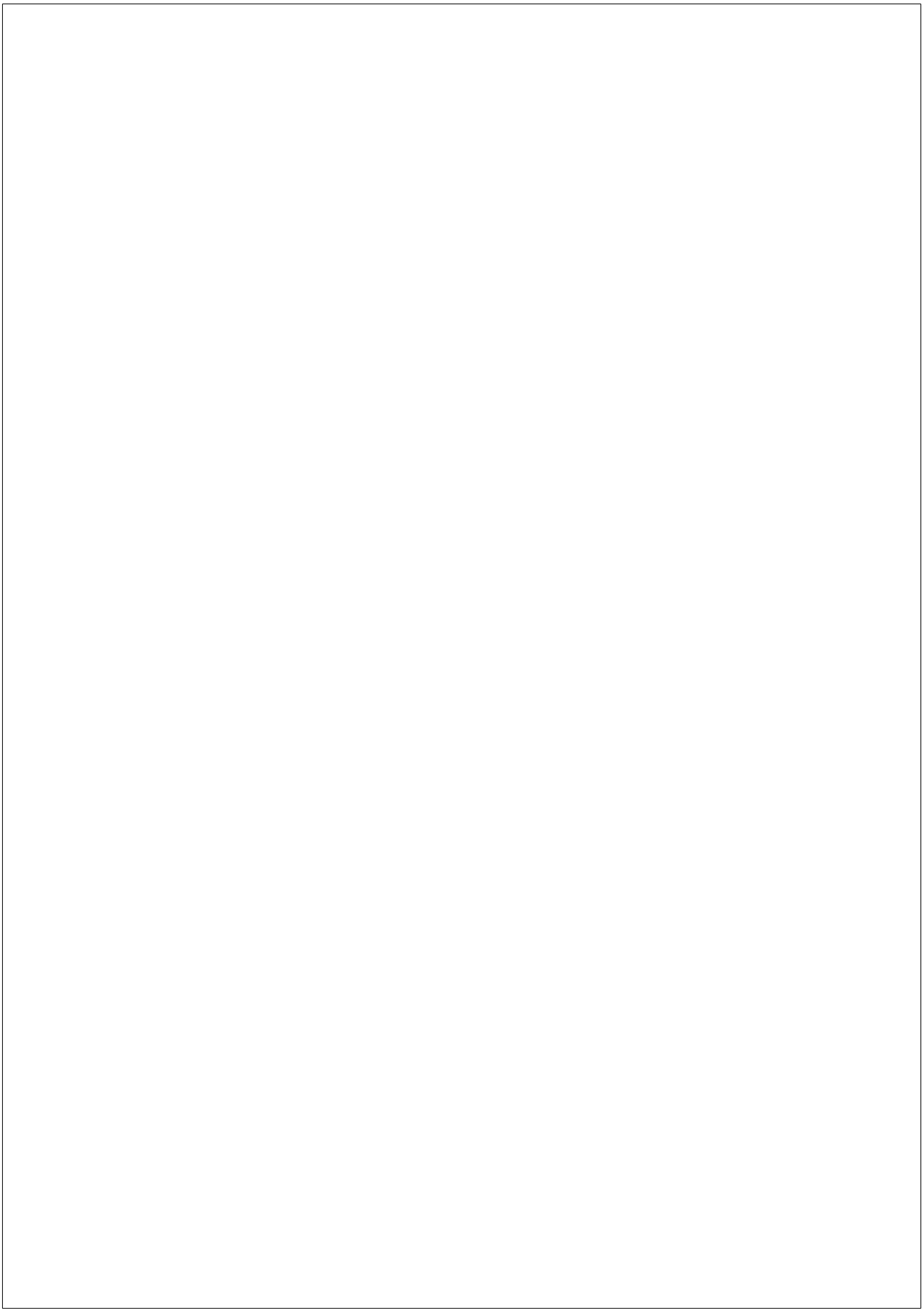
$$\boxed{\begin{cases} \dfrac{\mathrm{d}B}{\mathrm{d}t} = K_E(t) P\varepsilon - \dfrac{D_m B}{K_S + B} P\sigma \\ \\ \dfrac{\mathrm{d}[\mathrm{Out}]}{\mathrm{d}t} = \alpha_R \dfrac{B^n}{K_R^n + B^n} - \gamma_R[\mathrm{Out}] \end{cases}} \tag{3.28}$$

This system of differential equations has the advantage of being modular in its non-instantaneous property. This means that if for any reason we consider that emitter constant reaches equilibrium much more

faster than output does, then we can substitute the non-instantaneous wire equation ruling biobit dynamics by its original instantaneous version. Conversely, we can keep the non-instantaneous wire differential equation and assume the output is instantaneous. This last assumption is equivalent to evaluate $B(t)$ with the Hill function associated with the receptor strain to obtain the response level.

An extremely important detail: the receptor differential equation informs about how response varies per receptor strain cell unit. This means that if our culture is in a place where receptor cells are underrepresented, if the receptor response is maximum, then we have multiply it by ρ to obtain the expected output per unit of culture (and not per unit of receptor strain cell).

This system and its possible instantaneous alternatives can be used as more descriptive alternatives to test once the original wire equation has confirmed that the wire possesses operative regions. Also, it is possible to solve the system analytically to obtain a function that we can fit to experimental data.



Chapter 4

EXPERIMENTAL IMPLEMENTATION

In the previous chapter we have been discussing and developing the theory and the conceptual framework of what a wire should be, and the exploring the consequences and corollaries that follow from its definition. In this chapter I will present to the reader a possible approach to implementing a wire. Needless to say that the experimental design that I will present here does not pretend to be anywhere close to *canonical*. I mean, it is simplistic to pack a whole experimental approach as good or bad. Some of my decisions could be regarded as canonical, yes (the less risky ones), and some decisions were directly bad.

I have been thinking a lot about the right way to tell the story of an experimental path towards a goal. In a paper, it is only shown what *presumably* worked. But it does not tell you the hidden details, or the dead ends that the authors had to face when debating about possible experiments. And without denying that what appears in a

paper *is* science, I think it is my duty to inform the reader about the whole picture. After all, many of my bad decisions were not thought as bad in the beginning. And the same could probably have happened to you as well.

An important warning to the reader: this chapter, I am afraid, is the story of an unfinished story. The combination of decisions that turned to be bad at the end, together with the hit of the coronavirus pandemics, resulted in an interruption of the experiments. I know that at some point I will finish my experiments, because I need to see an experimental validation of my model. But it won’t happen here. Does it mean that this chapter is worthless? Not at all! Precisely because a lot of the decisions that I took I think they were actually good, and others turned to be bad, and I want to expose them here, this has value. Because it might be case that you are also interested in this particular problem of wiring distributed synthetic biological circuits. Because you may even try implementing the wire architecture that I propose here, and **I don’t want you to make the same mistakes** that I did. And because I was working on the possible fixes of some of my initial mistakes.

At the end, want I what you, the reader, to experience while reading this chapter is why I took the decisions I took, how I discovered if they were good or bad, and the possible improvements that I thought while doing experiments. In essence, the experimental motivation for every step towards validating my model. And I strongly believe that it is precisely *there* where the scientific thinking inhabits.

4.1. Introduction to the experimental part

In this section we will discuss the implementation generalities: which biobit molecule and genes will be the selected ones and why? How about the host organism? And what about the experimental conditions? Let’s start with the discussion.

4.1.1. Quorum sensing as a source for wire components

The search for a biobit candidate is necessarily linked to the search of candidate genes. A biobit candidate requires the minimum set of genes so it can be synthesized and secreted to the external media. Above all small molecule systems out there, there is one that has been widely used in synthetic biology: LuxI/LuxR quorum sensing system from *Vibrio fischeri*. But what is quorum sensing? Firstly described in 1970 by Nealson, Platt and Hastings [102], quorum sensing (QS from here to the end) is a collection of biological solutions present in bacteria that allow the cells to sense their own cellular concentration and generate a particular behaviour in response to it [98]. These QS systems can be divided in two groups depending on the bacterial group: Gram positive bacteria use little peptides calles autoinducer peptides (AIP) as signaling molecules, and Gram negative bacteria use homoserine lactones (HSL) [95]. Figure 4.1 shows the archetypal *LuxI/LuxR* system from Gram negative *V. fischeri* (A, left), and the as well archetypal *agrBDCA* from Gram positive *Staphylococcus aureus* (B, right).

The underlying physiological mechanism is the same for both quorum sensing systems: a subset of genes is responsible for synthesiz-

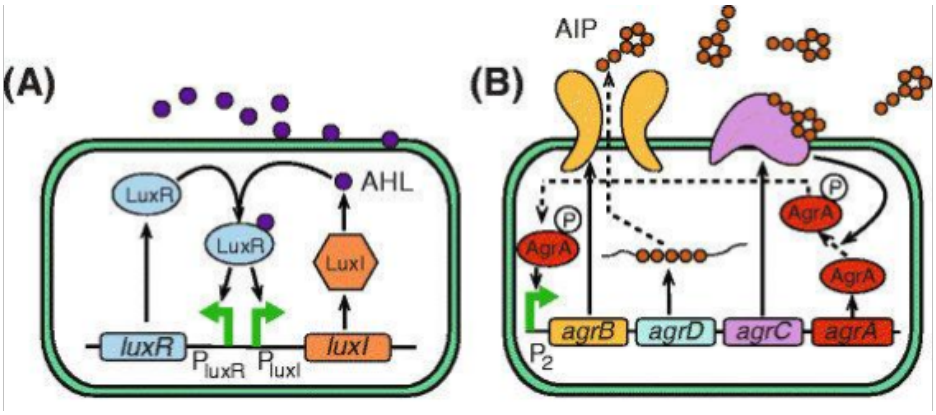


Figure 4.1: **Quorum sensing examples from Gram negative and Gram positive bacteria.** From Kong, 2014 [77].

ing and exporting the small diffusible molecule (autoinducer from here to the end). When bacterial cell density is low, assuming a constant rate of synthesis, the concentration that can be reached in the media of such autoinducer is limited. Nevertheless, if for any reason cell density increases, the overall rate of synthesis will increase as a consequence. And it turns out that the same bacterial cells have another subset of genes that are sensing the autoinducer concentration. It could happen that this concentration raises above some specific concentration threshold. It is then when the sensing genes get induced and promote the expression of surrogate genes [95, 98]. It is also a common feature to find the genes responsible for autoinducer production among these surrogate genes, inducing a positive feedback loop that promotes a rapid switch in gene expression pattern.

QS systems have been used for synthetic biology purposes many times. The idea is simple: split the set of genes described above in two halves. One half containing the genes responsible for autoinducer

biosynthesis, and another half containing the genes responsible for sensing autoinducer concentration. Then, place both components in two different host strains, *et voilà*, you have created a classical chemical wire [77].

As the reader may also assume, QS systems found in Gram negative bacteria are much more convenient for synthetic biology purposes for a simple reason: once the system is divided in two components, you can have a emitter component and a receptor component solely composed by one gene each. On top of that, many acyl-homoserine lactones can passively diffuse through the membrane, not requiring transporters of any kind [98].

And among all Gram negative quorum sensing systems, there is one that has raises above them all for being widely used: *LuxI/LuxR* from *Vibrio fischeri*. Not only that, it has been proven to perfectly work in other famous chassis apart from *V. fischeri*: *Escherichia coli*. Actually, all synthetic biology works cited in this section are done using *E. coli* as host organism.

4.1.2. LuxI, LuxR and its biobit 3-OXO-C_6-HSL

LuxI is an acyl-homoserine synthase, an acyltransferase enzyme that that catalyzes the following reaction: acyl-[ACP] + S-adenosyl-L-methionine $\rightarrow$ [ACP] + S-methyl-5'-thioadenosine + N-acyl-L-homoserine lactone. Being ACP the Acyl-Carrier Protein, a protein in the middle of lipid metabolism that provides acyl groups to several metabolic pathways. The product that constitutes our biobit is the N-acyl-L-homoserine lactone specifically produced by *LuxI*, 3-OXO-C_6-HSL (Figure 4.2).

According to our own restrictions about what a bbit should be, this

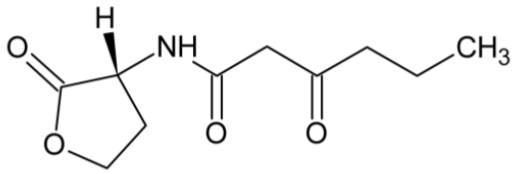


Figure 4.2: **3-OXO-C_6-HSL structure.**

molecule can definitely be synthesized in a living system by LuxI protein, and it is *de facto* secretable as it naturally diffuses through the membrane. It also is biodetectable, because is can be sensed by another gene product, in this case *LuxR*. We can also assume it is non-toxic at least in the concentrations used in all the mentioned projects: *E. coli* can perfectly live in the presence of this lactone. What about its degradation rate? Well, in every informal talk about this particular issue, researchers tend to agree on its low degradation rate, which is actually what made me choose this biobit candidate for testing my model. But which is its precise degradation rate? And more importantly, what does *low* means in this specific context? What makes a degradation rate be qualified as *low*?

The first question can be easy answered by taking a look at the existing literature. Kaufmann et al. directly measured in 2005 this degradation rate as $3.07 \times 10^{-5} s^{-1}$ [75], or reported as 3×10^{-4}min^{-1} by Marc Weber and Javier Buceta [143] from the same source. In other works, this value is taken as reference, so this is the primary source for several works [108, 67]. In one of these works, the degradation rate was measured again in treated seawater ($0.102 \pm 0.003\ h^{-1}$) [67]. I know: we will not be doing our experiments in seawater, but as a first approach, I think it is good to know that the degradation rate of

the molecule we are interested in is low across different conditions. Other work by Fekete (2010) indirectly measured this abiotic degradation rate by fitting a model to experimental data shed other value, such as $0.005545\ h^{-1}$ [45], a value used again in other modelling works [126]. Lactone half-life has been measured also by using biosensors and the results confirms half-live values of several days for the closely related lactone C_6-HSL under mildly acidic conditions [32]. In that work, 3-OXO-C_6-HSL was measured again over several hours up to 24h, showing almost no noticeable degradation in the first 6 hours, and a decrease in roughly 20% after 24 hours since the beginning of the experiment.

All these works orbit around the idea that degradation rate of 3-OXO-C_6-HSL is associated with a half-live of several days, in accordance with the informal commentaries about its stability. Now, the relevant question though is if that half live is enough, or equivalently, if the degradation rate is *low*. The answer is simple: it all depends on the emission rate and the time it takes to generate a response. In other words, it depends on the average dynamics of wire genes expression rate. If our constructs express their genetic programming in a few hours, then a 3-OXO-C_6-HSL half-live of days would qualify the bbit abiotic degradation rate as low or negligible. If, on the other hand, our wire requires days to fully express its genes, then the time scales of abiotic degradation and accumulation of bbit are comparable. This would imply that the abiotic degradation rate cannot be regarded as small enough to be negligible.

4.1.3. AiiA as the Sink Component

We have covered all requirements about our bbit candidate 3-OXO-C_6-HSL except one in particular: it has to be biodegradable. In other words, a functional biotic degradation is required. We need an enzyme that actively degrade our lactone. And it turns out that answer to our problem resides in another natural process: quorum quenching, or the ability of one specie to disrupt another's quorum sensing [153, 35, 27, 111][1]. Among all these QS disrupting strategies, one of them, by *Bacillus thuringiensis* and other *Bacillus* species happens to be very convenient for synthetic biology. This Gram positive bacteria has a protein, AiiA, with AHL-lactonase activity [118, 60, 34]. In a work by Wang et al. in 2004 [141], the enzymatic activity of AiiA was tested for several AHL (Acyl-Homoserine Lactone) molecules. Setting the activity for 3-OXO-C_8-HSL as 100%, the activity of AiiA towards our bbit 3-OXO-C_6-HSL was measured to be 91.8% in comparison.

This enzyme has been used in many synthetic biology projects. For instance, and starting in the purely theoretical realm, a circuit proposed by Borek in 2016 was analyzed through simulations for testing the ability to generate Turing patterns [16]. In that circuit, H_2O_2 and an AHL would intimately regulate the expression of each other in a way that would promote the appearance of Turing Patterns when cells would be placed in 2D cultures.

And in the realm of actual implementations, several oscillator architectures have been proposed, with the ability to synchronize the oscillations among all cells within the culture [114, 31]., and with the

[1] In one of the referenced works (Pietschke, 2017 [111]) another measurement of our bbit, 3-OXO-C_6-HSL, indirectly points that no abiotic degradation was observed in the first 24 hours (Fig 1A from the paper).

extra ability to even tune these oscillations [121].

These works, however, put all required genes within the same strain, the same cell. What about distributing the labor among several strains? In the specific case of AiiA enzyme, I have found a couple of examples. The first one, a synthetic amplifier composed by three strains one containing *AiiA* gene [152]. And the second one, a wire. Yes, you have read it right: there is one paper using almost the same architecture of what I define as a wire.

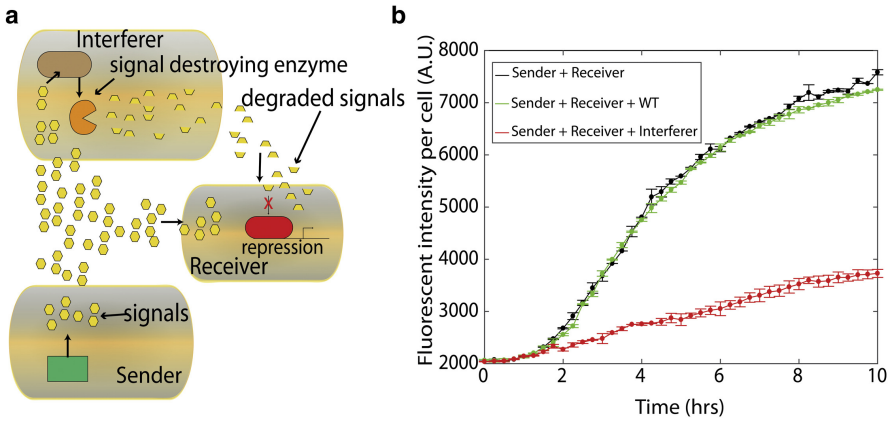


Figure 4.3: **Almost a wire architecture**. From Silva (2017) [126]

In Figure 4.3 from Silva (2017) [126], the authors propose mixing three different strains. One strain acting as sender (analogous to our emitter), one acting as receiver (analogous to our receptor strain), and one labeled as interferer (analogous to our sink strain). The only difference that this wire presents with respect to ours would be that the *signal destroying enzyme* gene is induced in the presence of bbit within the interferer strain[2].

[2]A question to the reader: which name do you prefer, Interferer or Sink? I definitely prefer Sink.

It turns out that this paper from 2017 uses the same genes that we are discussing right now: *LuxI* as the emitter component, *LuxR* for the receptor component, and finally *AiiA* for the sink component. The same genes that I would use to test the wire postulates. And in Figure 4.3 panel b we can appreciate the effect of mixing *receivers* and *sender* only, or mixing these two strains **plus** the *interferer* strain: the output produced by receiver strain is drastically reduced because the bbit concentration raises slower now that there is an AHL-lactonase enzyme in the media. In that work, researchers wanted to prove that distributing a mix of strains one of them being a signal *interferer*, you can create enzymatic walls that will prevent communication molecule to freely diffuse through the agar. In other words, they showed how you can limit or *pipe* the biobit effective flow in two dimensions.

Needless to say that our biobit, 3-OXO-C_6-HSL, can diffuse into the cell in the same fashion it diffuses out of the cell (an advantage of using a molecule that can passively diffuse through the membrane). I think this claim is pretty obvious, but I prefer to emphasize it: an emitter strain is a source of bbit, and a sink strain is... well, it *is* a biobit sink. And even when AiiA protein cannot go out of the cell, bbit will go inside the cell, specially because of the induced lack of bbit molecules that will be locally produced by AiiA enzyme.

4.1.4. Playing with ratios

So far we know that lactones have been used for synthetic biology purposes many times. We also know that *LuxI*, *LuxR* and *AiiA* genes are a popular selection when developing synthetic biology projects[3].

[3]Not to mention the iGEM projects that accumulate each year using these three particular genes.

We have also even seen a paper that describes a wire architecture with a regulatory feedback implemented. But what about the ratios?

In the previous chapter, the ratios among the three strains participating in the wire were critical to predict the big regions that define qualitatively different behaviours. Has any research groups noticed that ratios were something relevant? Actually, the last paper that we have been discussing has a section specifically dedicated to what happens if the ratio of emitter, sink and receptor strains vary. They placed a colony in the center of an agar plate composed by emitter strain, and a bacterial lawn composed by sink and receptor strain (recall that the original paper calls these strains senders, receivers and interferer strains, and that biobit induces AiiA expression in sink strain). They define normalized interference as the relative ratio of sink strain with respect to receptor strain, hence ranging from 0 to 1. In Figure 4.4, from Silva’s 2017 paper, shows the activation ratio (the ratio at which output fluorescence was observable) with respect to normalized interference for different amounts of emitter cells in the colony at the center of the agar plate. They noticed that by increasing or decreasing the amount of emitter by a factor of 5, a shift in the activation radius curve could be observable, as predicted by their mathematical model, and supported by the experimental results. Note that in that work, emitter strain is constitutively induced and therefore it does not respond to the presence of any input signal.

This last observation was summarized by the authors with following sentence: *"the ratio of sender to interference strains dictates the spatial extent of activation"*. Translation into *wire language*: the ratio of emitter versus sink strain determines the biobit concentration and thus the level of induction of receptor strain. Of course, in this experimental

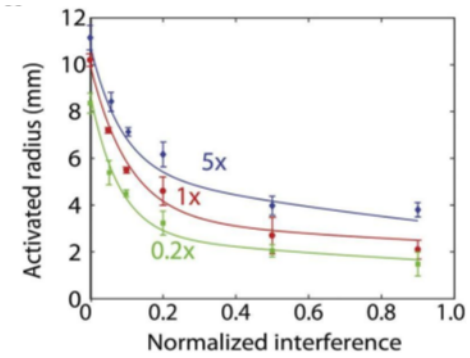


Figure 4.4: **Effect of varying the ratio of emitter strain with respect to sink strain**, from Silva (2017) [126].

setup there is the space variable (which is reflected in the mathematical modelling), and diffusion through the space plays a relevant role here, but at least, the major role of strain ratios is emphasized as deserved in this paper.

Is this the only paper that explores the ratio issue? Not at all. There is another paper by Urrios (2018) [136] that explores biological distributed computation using chemical wires, but this time in yeast. In this case, there exists again an almost perfect match with our proposed wire architecture. They use as biobit the α-factor pheromone from *Saccharomyces cerevisiae* (αSc), and the emitter strain is in this case HTX1:αSc, a yeast strain with the ability to secrete αSc in response to external glucose (the input). The receptor strain, labeled αSc:INS has the ability to sense the presence of biobit (αSc in this case) and secrete insulin as output. A *sink* was used in this setup in two different ways: one directly adding an enzyme that degrades αSc, so playing with its concentration would be somehow equivalent to altering the ratio of a phantom sink strain. And in another alternative setup,

a proper sink strain would produce the degradation enzyme, but in response to a second biobit, the α-factor from *Candida albicans* (αCa). This second biobit would be produced by a second emitter strain that would sense the presence of glucose as input. Even when the authors played with altering the ratio of the different strains in the co-culture, at the end, they were not trying to prove the capabilities or properties of a wire by itself, but going beyond and showing how distributed consortia allows modifying the behaviour of a circuit without genetic engineering. And by when they introduced a proper sink strain, they did it to illustrate a prove-of-concept incoherent feed-forward loop, introducing non-linear dynamics.

But there is a critical detail about Urrios' paper that I have not mentioned yet and it changes everything: cells were growing exponentially. That means that ratios may be very unstable over time, because a tiny difference between strain ratios, will probably be amplified in a future. And a varying ratio, in wire space, corresponds to a point with the ability to move across the wire space. If you remember previous chapter, sure you know that this moving point may escape from operative regions to unoperative regions, ending the desired behaviour[4].

4.1.5. Strain ratios in growing cultures

In 2011, I also imagined that a wire would be composed by exponentially growing strains. I had a minor doubt: would there be enough time for a culture in a flask to perform all the desired computations before reaching stationary phase? And I say it was a minor issue be-

[4] In an informal talk at BDebate at Barcelona 2019, I had the chance to talk with Urrios and Canadell, and they literally told me that the wire model would had been the solution to their problems, as they had difficulties dealing with strain ratios.

cause you can also switch to a bioreactor, so your cells will always be in exponential phase (I always conceived these circuits to be implemented in *E. coli*). But in a bioreactor, the bbit equation would require a dilution term, that in this case would be totally unavoidable... Well, minor problems compared with the big elephant in the room. Ten years ago, when I was thinking in this project, I thought that another problem that would exist would affect communities composed by more than one strain. These strains would compete for available nutrients in the media, and because of their different genetic programs, they would have different growth rates, which in turn would cause one strain to dominate the media forcing the others to decrease its concentration. This could affect the proper performance of the whole system, so it would be desirable for the strains in the culture to have a *stoichiometric stabilization system* that would bound their ratios in intervals in which the culture would respond as desired.

At that time, I used to think that exponential phase was the solution and that a *stoichiometric stabilization system* for strains would do the trick. I was thinking in adapting the ideas from the paper *A synthetic Escherichia coli predator-prey ecosystem* by Balagadde (2008) [10]. In that paper, an intelligent coupling between two quorum sensing systems (LuxI/LuxR and LasI/LasR) with a kill-switch (*ccdB* gene) strategically distributed between two *E.coli* strains, would surrogate one strain' survival upon the relative ratio with the other strain. This would produce in turn oscillations, hence the title of the paper.

A less lethal approach was used by Imperial College of London in iGEM 2016, under the *brand* Ecolibrium[5]. Their idea had something similar to Balagadde's paper: the use of cross diffusible chemical sig-

[5]http://2016.igem.org/Team:Imperial_College

nals between two strains. The main difference is that instead of killing the host, in response it would express *Gp2* gene from *E. coli* bacteriophage T7, a tiny protein with the ability to bind to RNA polymerase and arrest growth. To finely tune this activity, an ingenious use of complementary microRNAs would measure the differences in growth between both strains, constituting what I think is an early and beautiful example of a double reciprocal antithetic integral feedback, popularized by Mustafa Khammash in the late 2010s [19].

If the reader has some experience with wet-lab synthetic biology, then he/she may imagine how painful might be to implement all these *stoichiometric stabilization systems*, and on top of that, adding the logic gates, the wire components... While one of the points of distributed circuits was not to over-saturate a cell with genes. Too many thing just to keep the promise of circuits in exponential growth phase. Not to mention the difficulties that the model could show by adding exponentially growing strains and *stoichiometric stabilization systems* terms (in 2011 I was thinking in adapting Lotka-Volterra equation to the model to see the point moving in the triangle).

But if exponentially growing cells may result in problems, as it unavoidably provides a source of non-linear dynamics to the conundrum, what could we do then?

4.1.6. Stationary phase to the rescue

Hopefully for us, a paper appeared in 2014 by Gefen by the title *Direct observation of single stationary-phase bacteria reveals a surprisingly long period of constant protein production activity* [53]. The premise of the paper was so simple that it is astonishing that no one had done this before: let’s induce synthetic constructs in *E. coli* while

the culture is in stationary phase. Me and my wet-lab colleague were surprised that a paper like this could appear in 2014, because this should have been done at least in the early 2000s. We were wrong.

Using a plate-reader with a robotic arm to dispense inducer at different time points, and an *E.coli* culture that would express fluorescent protein mCherry in response to aTc (anhydrotetracycline), they were able to measure the expression rate when cells were induced at different time points with respect to the change from exponential to stationary phase. In Figure 4.5 (from Gefen 2014), this setup is shown in panel A. In panel B, the optical density measured in arbitrary units of a representative well is shown, with a grey area emphasizing the exponential growth zone.

In Fig. 4.5C, mCherry fluorescence is shown for every well, depending on the time point the well was induced. It is crystal clear that in stationary phase, mCherry fluorescence increases linearly, implying a **constant rate of synthesis**. Finally, in panel D, another critical detail is shown: promoter activity falls a factor of 10 when cells enter in stationary phase, but then remains **constant** for hours, defining what the authors have named as CASP (constant activity stationary phase). Other experiments were done to measure the cost of expressing these exogenous proteins, and the results were straight forward: almost no observable cost.

This CASP region is what we need for our wire. At the cost of promoter activity being 10 times smaller than exponential phase’s, we have gained the following features:

- Cells do not grow, and if they do, they do it in a negligible manner. From Gefen’s paper: *"extremely slow growth (>20 h per division) was observed in only a small percentage (less than 5%) of bac-*

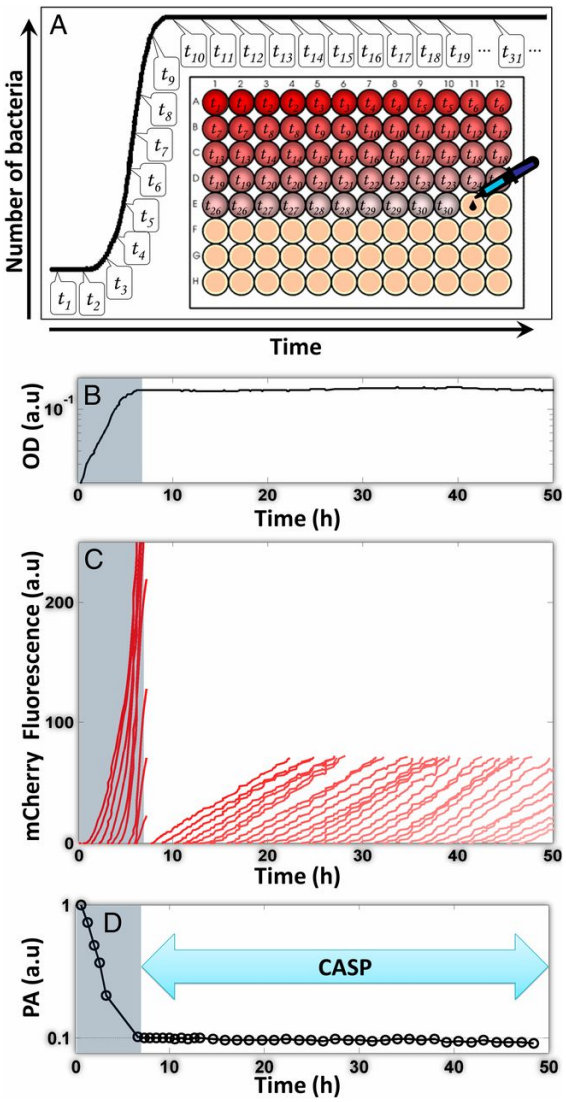


Figure 4.5: ***E. coli* can induce genes in stationary phase for a long period of time**. From Gefen (2014) [53].

teria". This is perfect for our wires, as the strain ratios will remain constant over time. In other words: where before there was an exponential growth term, now there is a constant, and hence no requirement for orthopedic *stoichiometric stabilization systems*.

- Promoter activity is constant. In other words, once a gene has been induced, it reaches equilibrium and then the expression rate remains fixed over time.

- Promoter activity remains constant for a long period of time, allowing the implementation of ambitious circuits that could require more time that the one that offers a culture growing exponentially in a flask.

- There is no need for bioreactors, and we can do our experiments in a flask with shaking. This adds the advantage from the modelling point of view of not requiring a dilution term, mandatory in the case of a bioreactor.

In summary, our experiments will be done in stationary phase cultures of *E.coli*, but which *E.coli* strain is best suited for the role?

4.1.7. *E.coli* △SdiA: the AHL-blind strain

In biology it is often very difficult to ensure that an arbitrary biological system will not be affected by an external factor. When an external factor and a biological system show this property, the lack of any cross effect between both, we say they are orthogonal[6]. I would calmly bet

[6]The only geometric concept or operation that fits in here is the dot product. At least I've always thought in dot product when talking about orthogonality.

that there is no such thing as pure orthogonality in biological systems, but rather negligible cross effects that we can consider as orthogonal.

When an interaction is non-orthogonal in the context of synthetic biology, we usually say there is some significant interaction or crosstalk. Specifically, crosstalk can be observed in quorum sensing when the lactone synthase from one specie interacts with the proper lactone receptor from another receptor from another specie, that is, a receptor that shows a preference for a different lactone. This lack of specificity found in lactone receptors, this crosstalk, is very ubiquous among Gram negative and positive quorum sensing systems [64, 7, 120], and has been used in synthetic biology [148, 140, 133]. In Figure 4.6, from Tekel (2019) [133], several AHL synthase enzymes from different bacterial species were expressed from a host *E.coli* grown in liquid LB media. Then the liquid cultures were centrifuged and filtered to collect the produced AHL (each enzyme produces a different lactone). This cell-free supernatant with lactones was mixed with several reporter strains carrying a unique AHL-sensing transcription factor (including our LuxR protein), coupled to GFP expression. Normalized GFP fluorescence after 8 hours is shown for each combination as a number in the matrix instances, with a color code proportional to its value.

In that experiment, we can appreciate that precisely our LuxR protein is not particularly specific to our bbit 3-OXO-C_6-HSL, and actually is activated by many other AHL synthases, with the exception of SinI, that catalyzes the synthesis of long chain AHL [51]. On the other hand, QS signal receptor TraR, from *Agrobacterium tumefaciens*, turns out to be very specific of its ligand, the lactone 3-OXO-C_8-HSL [138].

The point is: orthogonality and quorum sensing systems are two natural antagonists, with very few exceptions. And what is the prob-

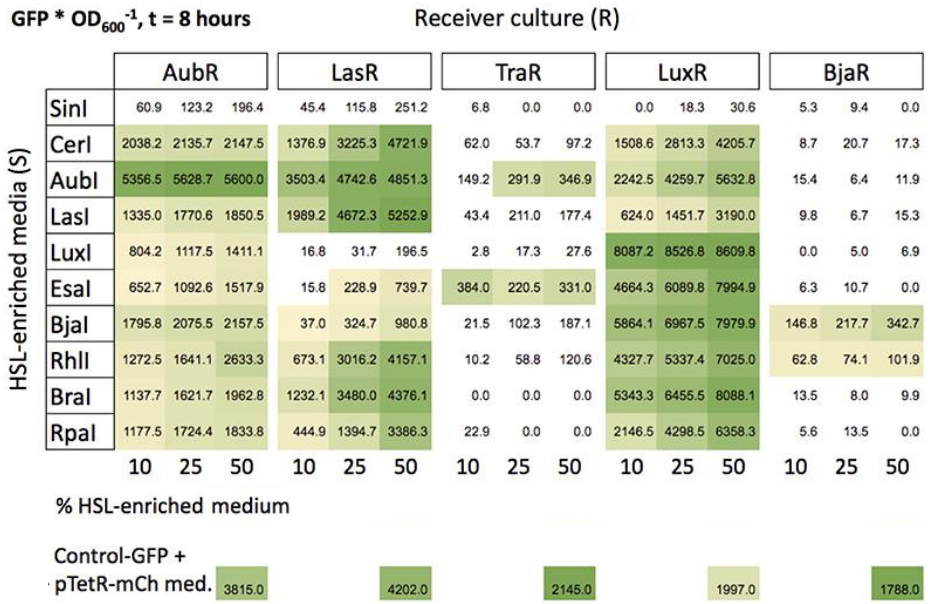


Figure 4.6: **Crosstalk between several AHL sender/receiver quorum sensing systems**. Adapted from Tekel (2019) [133]

lem, you may ask, if the wire implementation that we are discussing here is based on just *one* QS system? Well, the problem is that we are going to be producing lactones from plasmids within a host organism that happens to be a Gram negative bacteria, *E. coli*. And yes, even when the biobit that we want to use is the proper AHL from *V. fischeri*, it could produce some crosstalk with *E. coli*’s metabolism. But how? After all, LuxR protein will control the expression of some reporter gene, how could LuxR influence *E. coli* metabolism? Well, it doesn’t. Actually, even though *E. coli* cannot synthesize AHL because it lacks an AHL synthase encoding gene, it do has a gene from the LuxR family, *SdiA*, that can sense the presence of lactones and alter

the expression pattern of some of its genes in response (Van Houdt, 2006) [137]. And hence, orthogonality is again at risk here.

In Van Houdt's work [137], an *E. coli* promoter library expressing GFP as reporter was tested in the presence of C_6-HSL. The used *E. coli* strain was MG1655, considered as the *wild type* in the lab. They showed that 6 genes were upregulated and 9 were downregulated in the presence of C_6-HSL. Their research also shows that these gene expression alterations were eliminated when gene *SdiA* was knocked-out, and that this AHL responsiveness was temperature dependent (at 30°C, but negligible at 37°C). And yes, they also reported that the same effects described for C_6-HSL were extensible to our biobit candidate, 3-OXO-C_6-HSL. And this work was done using LB media, but MG1655 $\triangle$*SdiA* strain has been reported to grow in minimal media such as M9 medium with 1% glycerol [72] and MOPS medium with 0.4% glucose [9].

4.1.8. General experimental baseline

According to everything exposed in this section, we will be doing our experiments using *E. coli*'s $\triangle$*SdiA* strain. The host strain will be transformed with *LuxI* gene for the emitter component, *LuxR* for the receptor component, and *AiiA* for the sink component. The biobit molecule will be 3-OXO-C_6-HSL, which is stable enough to do experiments lasting a couple of days. The cells will be in stationary phase, in particular in the CASP regime, to take advantage of the constant rate of protein expression. And we can do our experiments in minimal media, because two reasons: $\triangle$*SdiA* knock-out is perfectly viable in such conditions, and rich media such LB presents autofluorescence that could mess with fluorescence if we select our output gene to be

GFP fluorescence.

4.2. Plasmid design and construction

All the required genetic constructs for our three wire components will be built in plasmids. These plasmids will always be pSB1C3 by default unless other way specified. pCB1C3 is a plasmid from the Parts Registry, the data base and repository from famous iGEM competition. In Figure 4.7, a scheme is shown. This plasmid possesses Chloramphenicol resistance and a origin of replication modified from pMB1 ori site, which confers the property of being a high copy number plasmid (between 100 and 300 copies per cell according to its official parts registry entry[7]).

The multiple cloning site (MCS) of iGEM plasmids is composed solely by four restriction sites distributed in two sequences called biobrick prefix (containing EcoRI and XbaI restriction sites), and biobrick suffix (containing SpeI and PstI restriction sites). The purpose of these sequences is to supply an standardized method for classical restriction/ligation cloning that allows to concatenate fragments in any desired way. pSB1C3 plasmid has two terminator flanking the MCS to transcriptionally isolate the insert. And finally, two sequences flanking these two terminators, VF2 (upstream) and VR (downstream), allow standard PCR amplification for fragment size analysis in colony PCRs, or for Sanger sequencing.

What, according to iGEM paradigm, is strictly called a *BioBrick*, is the insert that in located in between the prefix and the suffix sequences, *including both*. These BioBricks sequences can be what-

[7] http://parts.igem.org/Part:pSB1C3

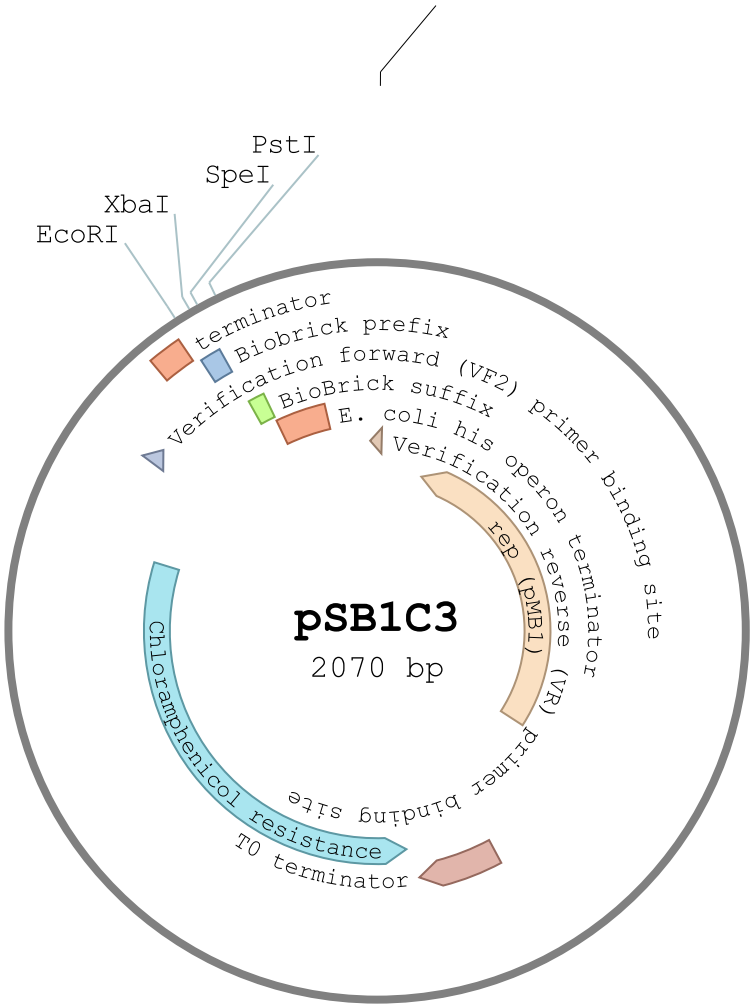


Figure 4.7: **pSB1C3 plasmid scheme.**

ever we could find among the constituent elements of a gene, such promoters, ribosome binding sites, coding sequences, terminator or even protein domains. But also any combination of these elements. By using the BioBrick assembly method which, remember, it is just a

clever variation of restriction and ligation method, we ensure that the concatenation of two BioBricks will produce a new BriBrick. That is, a new sequence bounded by a BioBrick prefix and a BioBrick suffix. An in the middle, a minor scar will remain in between the two original BioBrick predecessors.

Nevertheless, we will not be using the standard biobrick assembly method, so we will not discuss here how it works. For joining DNA pieces we will be using Gibson assembly [56]. This cloning method allows to join DNA fragments that have been previously amplified by PCR using primers that add homology regions to the extremes of at least one of the fragments to be joined (between 20 and 40 bp of homology are recomended). Gibson reaction consists in a clever mix of 5’-exonuclease, a DNA polymerase, and a ligase. The exonuclease degrades each dsDNA strand by its 5’ extremes, exposing a single stranded DNA at the ends of the dsDNA fragments. Because these extremes have homology thanks to the previous PCR amplification, these ssDNA extremes can anneal, joining the extremes. The DNA polymerase fills the ssDNA gaps that remain, and finally the ligase reconstructs the dsDNA (Figure 4.8).

This method requires previous purification of the fragments amplified by PCR, and is an isothermal reaction (50ºC). Even when commercial products claim that reaction can take place in around 15 minutes, it is a common practice to leave the isothermal reaction for at least 1 hour. More hours is perfectly fine as well, because once the fragments have been joined together, the reaction cannot chemically modify the DNA in the mix by any means. Actually, because Gibson assembly can handle the joining of more than one insert, it is also common to leave the reaction overnight if more than more than three

fragments have to be joined, of if the homology sequences are less than 20 bp.

The standard procedure after Gibson cloning, if we want to transform by electroporation, will consist in a desalination process using Millipore dyalisis membranes. These circular disks are placed floating in type I water (MilliQ water), and then the Gibson reaction product $(20\mu L)$ is pipetted in the center of the circular membrane. After 15 minutes, the liquid drop in the middle of the disk (the Gibson reaction) is pipetted out again and stored for later use. After this procedure, the reaction product is ready for electroporation. If the transformation process is thermal shock (with chemically competent cells), then the desalination is not required at all.

4.2.1. Design and construction of Receptor strain

The emitter strain will consist in the *E. coli* BW25113 $\triangle$SdiA strain transformed with the pRHW plasmid (consisting in the receptor component as an insert in the pSB1C3 plasmid). The proposed scheme for receptor component is depicted in Figure 4.9.

The receptor core will have coding sequence *LuxR* with BioBrick code BBa_C0062, expressed from the constitutive promoter P(Lac)IQ, which provides a high transcription rate (BioBrick code: BBa_I14032). *LuxR* coding sequence will be followed by two terminators, whose BioBrick codes are BBa_B0010 and BBa_B0012 respectively. *LuxR* dependent promoter Plux, with BioBrick code BBa_R0062, finishes the receptor core. The output part of the receptor component is composed by GFP coding sequence with degradation tag LVA (BioBrick code: BBa_J04031) followed by terminator sequence BBa_B0012. All ribosome binding sites in this construct are the same, with BioBrick

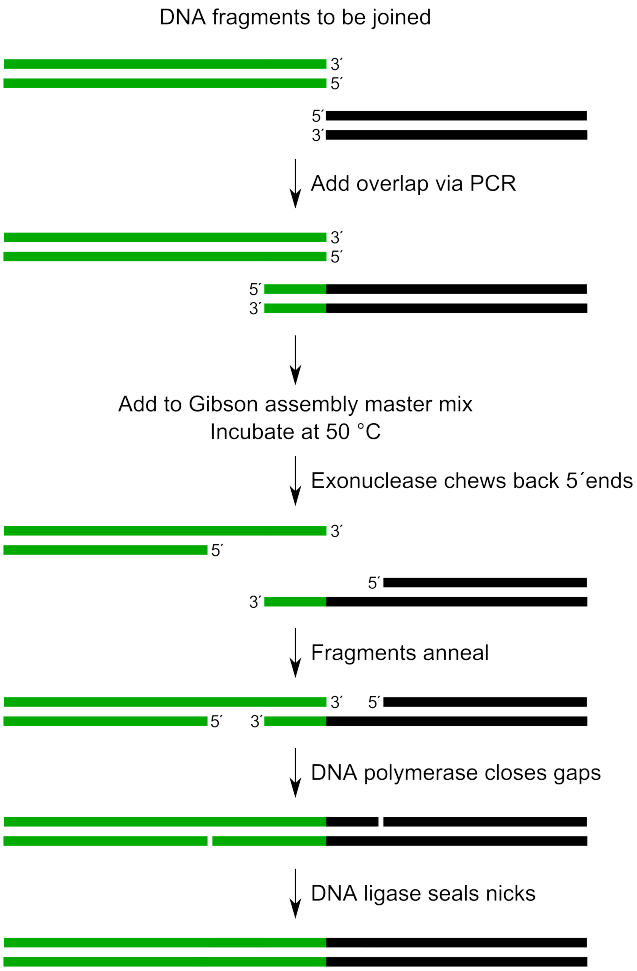


Figure 4.8: **Gibson assembly overview**.

code BBa_B0034.

I found this sequence of BioBricks divided in two different BioBricks:

- **BBa_J09855**: a BioBrick composed by the BioBrick sequence

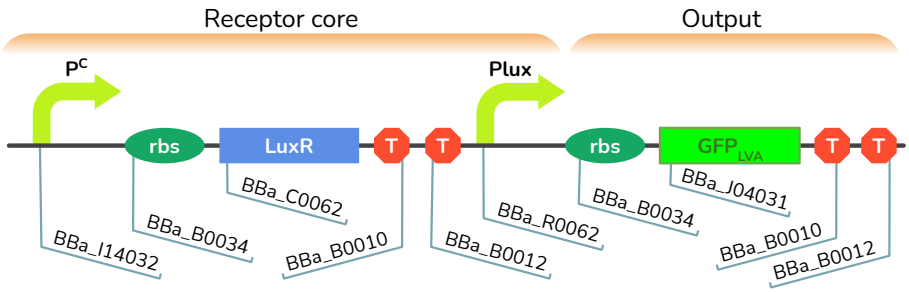


Figure 4.9: **Receptor component BioBrick outline**.

BBa_I14032, BBa_B0034, BBa_C0062, BBa_B0010, BBa_B0012, BBa_R0062, which covers the whole receptor core.

- **BBa_I763020**: a BioBrick composed by the BioBrick sequence BBa_B0034, BBa_J04031, BBa_B0010, BBa_B0012, which constitutes the output part of the receptor component.

Recall that choosing GFP coding sequence as the output transcript is an arbitrary choice, uniquely motivated in our case because it is a standard and useful approach to measure gene expression.

These two BioBricks (in pSB1C3 plasmids) were available in the 2016 iGEM BioBrick distribution, where they are found as red-dyed dried DNA that has to be resuspended in 10μL before transformation. From these 10μL, 3 were used to perform chemical transformation using NZY5α cells from Nzytech[8] following its standard procedure, plated in LB with Chloramphenicol (LB +Cm from here to the end), and cultured overnight at 37°C to select transformed clones.

[8]Beware! These commercial cells are useful to store plasmids, but avoid using them for experiments in minimal media, as they possess *thi*-1 mutation, requiring thiamine in the media to grow.

Colonies were tested by colony PCR using VF2 and VR primers, and positive colonies were cultured overnight in 5mL of LB +Cm at 37°C with orbital shaking overnight. Next day, 1mL of saturated culture was mixed with 1mL Glycerol 50% and stored at -80°C, while the remaining 4mL were used to do miniprep protocol using Sigma miniprep kit (following their standard protocol, but elluting in 50μL at the end). Purified plasmid was then analyzed using Nanodrop Lite to measure its concentration and purity. From this pure plasmid solution, a couple of aliquots were sent for Sanger sequencing with primers VF2 and VR (less than 5μL per sequencing reaction). The remaining pure plasmid was then stored a -20°C for future uses. Sequencing results (chromatogram ab1 files) were analyzed against the expected sequence in Benchling. Bad sequencing results were always discarded. A plasmid was considered to be sequence-confirmed when there was a match between sequencing results from VF2 (forward) and VR (reverse) primers.

BioBrick BBa_I763020 containing GFP had a mutation at residue 238 where a Lysine (K) had been changed by a glutamic acid (E). The mutation implies the change from a positively charged amino acid to a negatively charged amino acid. Even though this mutation K238E is placed just in the locus where the GFP sequence finishes and the LVA degradation tag starts, we will clone mutant with K and mutant with E residue. To revert the K238E mutation, we will amplify BBa_I763020 with two pairs of primers introducing a point mutation that will reverse the mutation to the original amino acid K. The strain with the original mutation will be amplified without using the inner primers.

The primers used for the two Gibson assemblies are summarized in Figure 4.10, and the expected clones will be these:

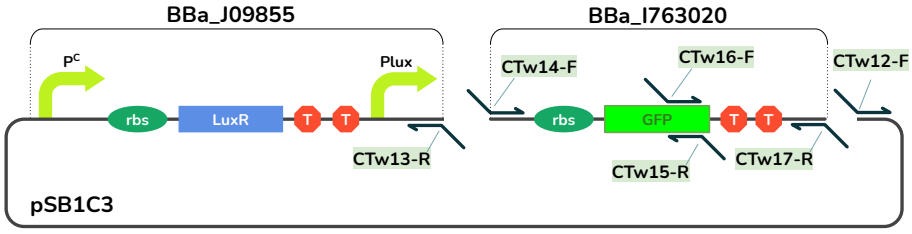


Figure 4.10: **Receptor component Gibson primers**.

- **pRHW$_K$** will require the following PCR amplifications with the corresponding primers:
  - **Backbone** (pSB1C3 with BBa_J09855) will be amplified by primer pair **CTw12-F/CTw13-R**.
  - **Insert 1** (from pSB1C3 with BBa_I763020) will be amplified with primer pair **CTw14-F/CTw15-R**. Primer CTw15-R contains a nucleotide change to reverse K238E mutation back to K (Lysine).
  - **Insert 2** (from pSB1C3 with BBa_I763020) will be amplified with primer pair **CTw16-F/CTw17-R**. Primer CTw16-F contains a nucleotide change to reverse K238E mutation back to K (Lysine).
- **pRHW$_E$** will require the following PCR amplifications with the corresponding primers:
  - **Backbone** (pSB1C3 with BBa_J09855) will be amplified by primer pair **CTw12-F/CTw13-R**.
  - **Insert** (from pSB1C3 with BBa_I763020) will be amplified with primer pair **CTw14-F/CTw17-R**.

After PCR amplification, the size of the PCR products was checked with agarose electrophoresis, and gel purified, their concentrations were measured using Nanodrop Lite, and two Gibson assemblies were done. Gibson assembly products were directly transformed into NZY5α cells. Colony PCR was done using VF2 and VR primer pair over the transformant candidates to check insert size. Positive transformants were cultured in 5mL of LB +Cm overnight at 37^{o}C shaking. 1mL of the saturated culture was mixed with glycerol 50% and stored in -80^o. A miniprep was done with the remaining 4 mL, and the DNA concentration was measured again using Nanodrop Lite, and an aliquote was sent for Sanger sequencing using VF2 and VR primers.

For the pRHW$_K$ plasmid, two clones that passed the process described above were selected, labeled pRHW$_K$1 and pRHW$_K$2, and for pRHW$_E$ plasmid. another two clones passed the above process and were labeled as pRHW$_E$1 and pRHW$_E$2.

These plasmids were transformed again by electroporation in the *E.coli* BW25113 Δ*SdiA* strain from Keio Knock-out collection [9]. The sequences of all primers described here can be found in Table 4.1.

4.2.2. Design and construction of Emitter strain

The emitter strain will consist in the *E. coli* BW25113 ΔSdiA strain transformed with the pEHW plasmid (consisting in the emitter component as an insert in the pSB1C3 plasmid). The process for implementing the emitter component will follow exactly the same plan as we did for the receptor component, so this time we will skip the minor details and will focus only in the particularities of this component. The proposed scheme for emitter component is depicted in Figure 4.11.

Primer name	Sequence (5’ to 3’)
CTw12-F	TATATACTAGTAGCGGCCGCT...
	GCAGTCCGGCAAAAAAGGG
CTw13-R	TTTCTCCTCTTTCTCTAGTAT...
	TTATTCGACTATAACAAACCATTTTCTTG
CTw14-F	TAGTCGAATAAATACTAGAGA...
	AAGAGGAGAAATACTAGATGCG
CTw15-R	GTCGTTTGCAGCAGGCCTTTT...
	GTATAGTTCATCCATGCCA
CTw16-F	TGGCATGGATGAACTATACAA...
	AAGGCCTGCTGCAAACGAC
CTw17-R	CCCTTTTTTGCCGGACTGCAG...
	CGGCCGCTACTAGTATATAAA
VF2	CCACCTGACGTCTAAGAAAC
VR	GTATTACCGCCTTTGAGTGA

Table 4.1: Primers used for pRHW$_{E/K}$ plasmid construction.

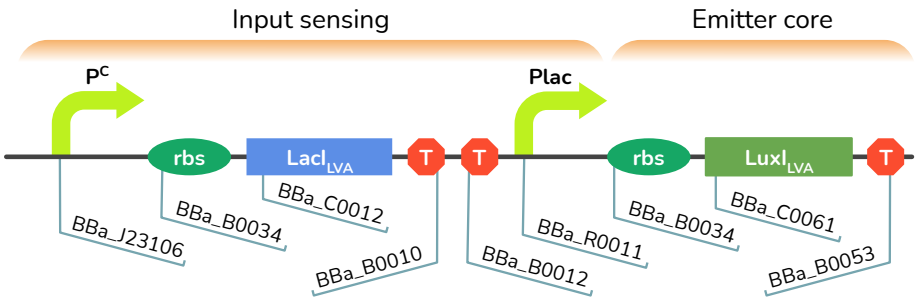


Figure 4.11: **Emitter component BioBrick outline**.

As the reader may notice, we need configure our emitter core to respond to an external input. In this case the input will be presence

of IPTG[9], an analogous to lactose, that can bind to transcription factor LacI. This protein is normally bound to its promoter Plac repressing transcription, and it releases itself from the promoter when it is bound to IPTG, allowing RNA polymerase to initiate transcription [130]. Using IPTG and LacI repressor for inducing gene expression is a very common practice in synthetic biology.

The sequence of BioBricks shown in Fig. 4.11 was found in the following BioBricks:

- **BBa_K735051**: a BioBrick composed by the BioBrick sequence BBa_J23106, BBa_B0034, BBa_C0012, BBa_B0010, BBa_B0012, and BBa_R0011, which covers the input sensor.
- **pSB1AK3 plasmid**[10], containing a BioBrick composed by the BioBrick sequence BBa_B0034, and BBa_C0061, which constitutes core emitter component.

In Figure 4.11, a terminator with BioBrick code BBa_B0053 is shown after *LuxI* coding sequence, but it does not appear in the source plasmid in previous list. This is because that terminator, BBa_B0053, is placed *after* the BioBrick suffix, that is, one of the default terminators flanking the prefix/suffix region.

These plasmids were transformed in NZY5α strain. A colony PCR using VR2 and VF primers was done to find positive colonies according to insert size. Positive colonies were culture overnight in 5mL LB+Cm 37ºC shaking overnight. From that culture, 1 mL was store

[9]When IPTG 2mM is bought in liquid format, it has to be stored at -20ºC. And it remains liquid at that temperature. This is normal and yes: it is just IPTG and water, nothing else.

[10]Courtesy of Blai Vidiella, from Complex Systems Lab at UPF

at -80°C as glycerol stock, and the remaining 4mL were used to do a miniprep. Purified plasmids were stored at 4°C and Sanger sequenced confirmed. Then, the following primers were design to perform a Gibson assembly.

pEHW will require the following PCR-amplified primers with the corresponding primers:

- **Backbone** (pSB1C3 with BBa_K735051) will be amplified by primer pair **CTw6-F/CTw7-R**.
- **Insert** (from pSB1AK3 plasmid) will be amplified with primer pair **CTw4-F/CTw5-R**.

Figure 4.12 summarizes the backbone, insert and primers for Gibson assembly.

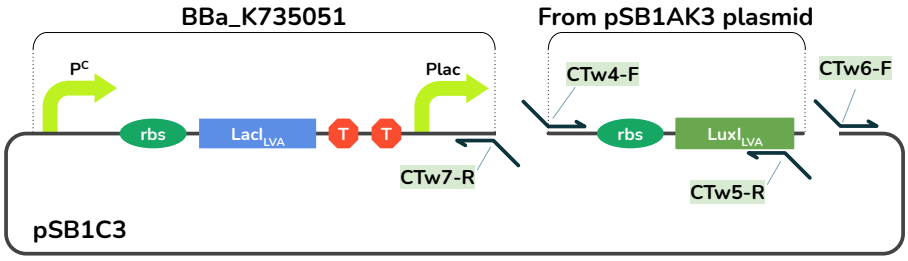


Figure 4.12: **Emitter component Gibson primers**.

After PCRs were done to the source plasmids, desired fragments were size verified and gel-purified. Their concentrations were measured and a Gibson assembly was done. Then the Gibson assembly product was transformed into NZY5α cells. Colonies were selected according to size with a colony PCR, and size-positive colonies were cultured in 5mL LB+Cm 37°C shaking overnight. From that culture,

glycerol stock was stored and from the remaining culture, a miniprep was done. The plasmid was sent for Sanger sequencing, and the sequence of the new plasmid pEHW was confirmed to be positive. Then, the plasmid was transformed by electroporation in the *E. coli* BW25113 $\triangle$SdiA strain, and positive colonies were glycerol stored.

The sequences of all primers required to amplify by PCR the fragments detailed here can be found in Table 4.2.

Primer name	Sequence (5’ to 3’)
CTw4-F	ATACTGAGCACATACTAGAG...
	AAAGAGGAGAAATACTAGATGACT
CTw5-R	TGCAGCGGCCGCTACTAGTA...
	GAGATCTACACTAGCACTATCA
CTw6-F	ATAGTGCTAGTGTAGATCTC...
	TACTAGTAGCGGCCGCTGCA
CTw7-R	ATCTAGTATTTCTCCTCTTT...
	CTCTAGTATGTGCTCAGTATCTTGTTATCC
VF2	CCACCTGACGTCTAAGAAAC
VR	GTATTACCGCCTTTGAGTGA

Table 4.2: Primers used for pEHW plasmid construction.

4.2.3. Design and construction of Sink strain

The sink strain will consist in the *E. coli* BW25113 $\triangle$SdiA strain transformed with the pSINK plasmid (consisting in the sink component as an insert in the pSB1C3 plasmid). The sink component will consist in the BioBrick sequence depicted in Figure 4.13. It is designed to constitutively express AiiA protein, the HSL lactonase enzyme.

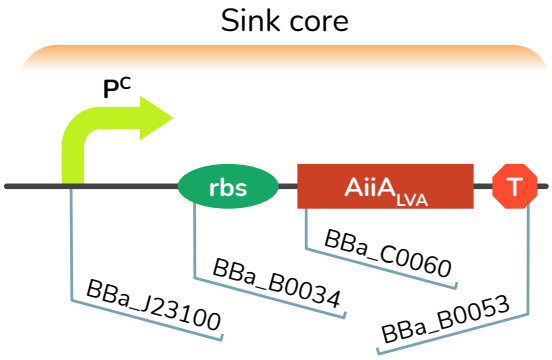


Figure 4.13: **Sink component BioBrick outline**.

The following list shows BioBricks from iGEM 2016 distribution that contain the required elementary parts to build the sink component:

- **BBa_K880005**: a BioBrick composed by the BioBrick sequence BBa_J23100, and BBa_B0034.
- **BBa_K516022**, containing a BioBrick composed by the BioBrick sequence BBa_B0032, BBa_C0060, BBa_B0010, and BBa_B0012. From this BioBrick we are only interested in BBa_C0060, AiiA gene with LVA degradation tag.

This two BioBricks, stored in pSB1C3 plasmid, were treated in the same way as the other two plasmids. They were transformed in NZY5α, glycerol stock stored, and size and sequence confirmed by Sanger sequencing. Then, primers for Gibson assembly were designed according to Figure 4.14.

The DNA insert and backbone to create **pSINK** plasmid by Gibson assembly will require the following primers:

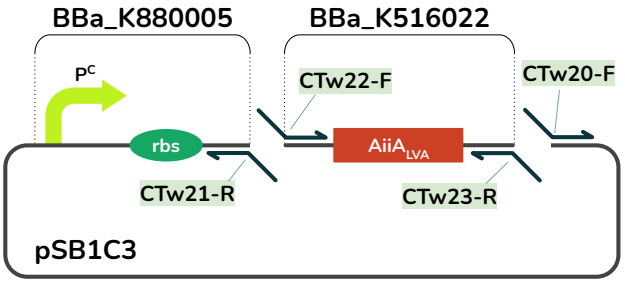


Figure 4.14: **Sink component Gibson primers**.

- **Backbone** (pSB1C3 with BBa_K880005) will be amplified by primer pair **CTw20-F/CTw21-R**.

- **Insert** (pSB1C3 with BBa_K516022) will be amplified with primer pair **CTw22-F/CTw23-R**.

After PCR amplification, size verification by electrophoresis and gel-purification of these two fragments, a Gibson assembly was done. The product was transformed in NZY5α cells by heat shock. Colonies were size verified using VF2 and VR primers. Positive colonies were cultured in 5mL LB+Cm 37ºC shaking overnight. From that culture, 1mL was stored as glycerol stock, and from the remaining 4mL, a miniprep was done. The plasmid was sent for Sanger sequencing, and the sequence of the new plasmid pSINK was confirmed to be positive.

The sequences of all primers required to amplify by PCR the fragments detailed here can be found in Table 4.3.

Primer name	Sequence (5’ to 3’)
CTw20-F	ATAATACTAGTAGCGGCCGCTGCAGTCCGGC... AAAAAAGGG
CTw21-R	CTTCTTTACTGTCATCTAGTATTTCTCCTCT... TTCTCTAGTAGCTAGC
CTw22-F	CTAGAGAAAGAGGAGAAATACTAGATGACAG... TAAAGAAGCT
CTw23-R	CCCTTTTTTGCCGGACTGCAGCGGCCGCTAC... TAGTATTATTAAGCTACTAAAGCGTAGTTT
VF2	CCACCTGACGTCTAAGAAAC
VR	GTATTACCGCCTTTGAGTGA

Table 4.3: Primers used for pSINK plasmid construction.

4.2.4. The motivation for degradation tags

As you may have seen, almost all protein coding sequences contain an LVA degradation tag. This degradation tag (or *ssrA*-dependent tag) consists in a C-terminal peptide tag with AANDENYALVA sequence of amino acids which is recognized by the proteolytic machinery of *E. coli* [47, 80, 59], and its use is ubiquitous in synthetic biology [5, 31, 113, 129].

These tags allow circuits to act faster upon inputs, which is useful for dynamical circuits. But mind the fact that when bacterial cells are exponentially growing, all cellular proteins undergo a dilution process due to the constant increase in cytoplasm volume before each cell division. This dilution rate makes unnecessary the existence of a very relevant degradation machinery. But when cells arrest their growth in stationary phase, the dilution disappears and proteolytic degradation

that regulates the amount of proteins within the cell. The main reasoning in our case is not that our circuit requires very fast responses. But because our proteins are expressed from strong promoters with high transcription rates (the same can be argued for ribosome binding sites), the lack of dilution might increase protein concentration above toxic levels. It is for that matters that adding a degradation tag like LVA *ssrA* tag we ensure that toxic concentrations will not be reached.

4.3. Experimental results

Now that we have our three strains, it is turn for the experiments. However, and as I said at the beginning, we were not able to prove nor disprove the model. For that matter, I will present the results in chronological order, so the reader will follow the reasoning after each experiment and the motivation for the next one. Each experiment needs to address a specific question that will be properly motivated. It also demands an experimental design that will be sketched. And finally a brief discussion of the results. I hope it will help you understand the line of reasoning after the sequence of experiments.

4.3.1. Receptor strain experiments

First, we are going to analyse the receptor strain, carrying the $pRHW_E$ or the $pRHW_K$ plasmid. There is one main final objective for this strain: to find the transfer function the relates biobit concentration with the output at equilibrium. In our case, the output will be GFP flu-

orescence. Recall that this function will be defined by the Hill-function

$$r(B) = \frac{r_m B^n}{K_A^n + B^n} \tag{4.1}$$

whose parameter values we need to characterize. But also, we have to check if there is any difference between the two version of pRHW plasmids, how these plasmids affect growth. Or if the solvent of our biobit, 3-OXO-C_6-HSL, affects *E. coli*. The consequences of culturing *E.coli* in LB media or in minimal media, etc. Many variables that may or may not affect the receptor strain behaviour and our ability to measure our output, GFP fluorescence.

Testing receptor strain growth in LB media

For this first experiment, $pRHW_E$ and $pRHW_K$ plasmids were transformed in *E. coli* MG1655 strain (wild type). Two precultures from the corresponding glycerol stocks were done in LB+Cm media at 37°C shaking overnight. From these two precultures, new cultures were placed with an initial OD_{600} = 0.01 in LB+Cm and 200μL were dispensed in a 96-well plate. A gradient of biobit 3-OXO-C_6-HSL was prepared ranging from 100μM and decreasing by halfs. To do so, a 100 times more concentrated gradient was first done in DMSO, the solvent for stocking 3-OXO-C_6-HSL, and 2μL were added to each well. A well with just LB was placed as blank control, and a well with only DMSO (no biobit in the media) was used to control the culture is absence of biobit. The 96-well plate was measured in a Tecan plate-reader for OD_{600} and GFP fluorescence shaking at 37°C. Three replicas were done, and representative cases for each plasmid are shown in Figure 4.15.

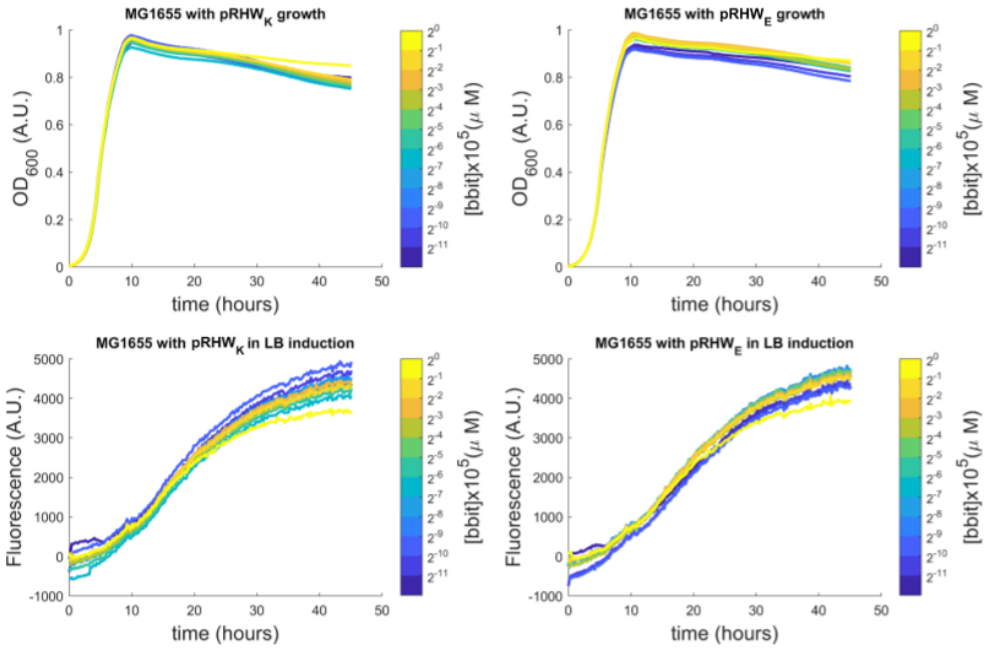


Figure 4.15: **Receptor strain growth and induction.**

The obvious thing that can be noticed is that the biobit range does not cover the whole induction range, because all cultures have been fully induced, except the case for the lack of inducer (data not shown). That could be seen as a mistake, but it allows us to verify that output gene can be induced when cells reach stationary phase, as fluorescence increases when cells have arrested their growth. Growth arrest happens around 10 hours after the beginning of the experiment. Also, it is possible to appreciate that between 10 and 20 hours, fluorescence was increasing in a linear manner, but after 20 hours, fluorescence decreases its increasing rate like it is approaching equilibrium. This is supporting evidence for Gefen results on CASP phase [53], but also for the effect of LVA degradation tag.

Stationary phase and induction in LB or Davis Media

For the next experiment, we will be using our strain *E. coli* BW25113 $\triangle$SdiA transformed with pRHW$_E$ plasmid. We will place precultures in LB+Cm overnight at 37^{o}C with shaking. From these precultures, two new cultures were prepared with initial OD$_{600}$ = 0.01 in 10mL of LB media or David media (supplemented with casamino acids and 0.4% glucose), and then cultured with shaking at 37^{o}C for 48 hours, so we are sure that our cultures will have reached stationary phase.

After that time, half of each culture was centrifuged and filtered to collect the filtered stationary phase media (fST). This fSP (from LB and from Davis media) was used together with the remaining culture to set the initial OD$_{600}$ = 1 in a 96-well plate, with a range of biobit starting at 100μM and decreasing by a factor of 10 (instead to 2 like the previous experiment). Also, culture with only DMSO (the solvant for 3-OXO-C$_6$-HSL) was used to control the absence of biobit, but also take into account the effect of the solvant. Also, filtered media for each media tested was place in the 96-well plate with the biobit-inducer range as a control. And another row of the plate containing the exact same setup (cell + biobit-inducer range) was placed with host strain without plasmid. Antibiotic (Cm 10μg/mL) was added depending on the cultured strain.

Figure 4.16 shows a comparison between *E. coli* BW25113 $\triangle$SdiA strain with or without pRHW$_E$ plasmid in LB media or David media. The first relevant detail is that we can observe an initial growth even when cells were supposedly in stationary phase both in LB or in Davis media, though Davis media shows a smaller growth rate. My only thought about this effect is that altering or decreasing the reached OD$_{600}$ may trigger growth again [23]. But this, again, is something positive in this

specific context. Because we can appreciate how there is less growth in the strain containing pRHW$_E$ plasmid when biobit concentration is high, compared with the case of not having the plasmid, were no relevant change is observed. This fits with the hypothesis that neither biobit nor the biobit solvent might affect growth.

If we compare between growth (OD_{600}) against receptor induction (fluorescence) only for cells containing pRHW$_E$ plasmid, we can see that that the underlying cause of this slower growth rate (and maximum reached cell density) may actually be due to GFP expression, only induced at high biobit concentration.

When looking at fluorescence emitter by cells with the plasmid in Davis media, it is clear that GFP expression is responsible for the negative growth effect. And also, we can appreciate how GFP expression is more stable in Davis media, as in LB media it seemed to generate a pulse and then decrease, whereas is Davis media fluorescence keeps at high values for hours.

Stationary phase cells without OD_{600} correction in Davis media or LB induced 24 or 48 hours after inoculation.

In the previous experiment, we have found that once cells have arrested their growth by reaching stationary phase, if we artificially modify cell density, then more growth is induced. Actually, it is known that *E. coli* arrest its growth in the stationary phase onset not when there is nutrient deprivation, but when these nutrients fall below some threshold [50], or when cell density surpasses some critical value [23]. In this experiment, precultures of the *E. coli* BW25113 $\triangle$SdiA strain with and without pRHW$_E$ in LB (with appropriate antibiotic). Then, both strains were culture in Davis media (just with Glucose 0.4%) or

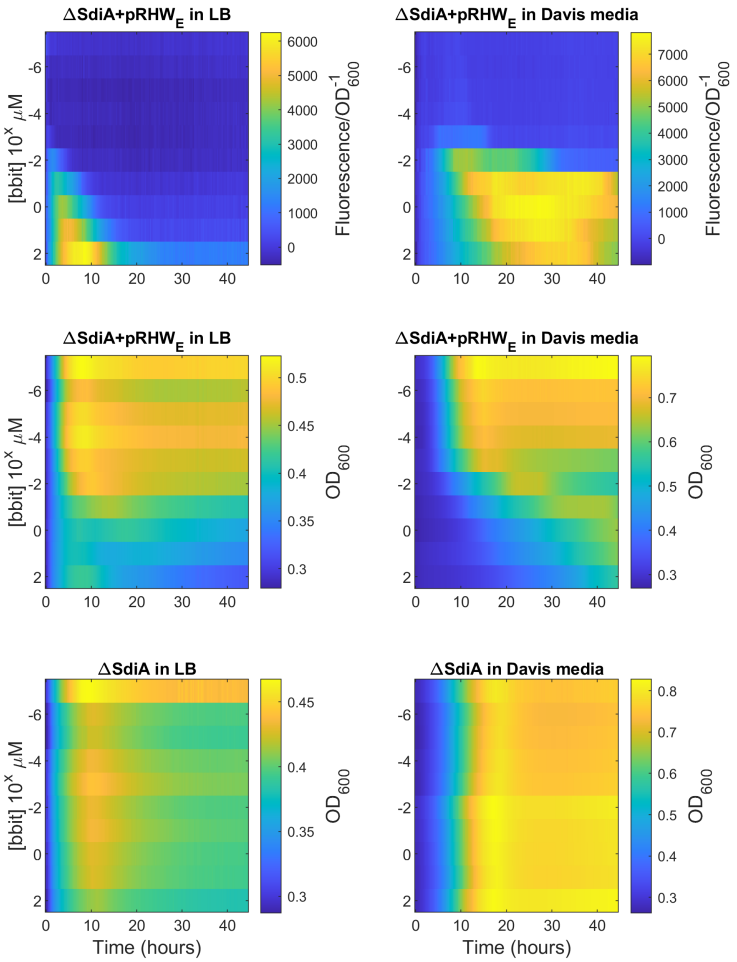


Figure 4.16: **Testing growth and induction in different media.** Host strain growth with or without pRHW$_E$ plasmid in LB or Davis media plus biobit inducer range.

LB for 24 or 48 hours (again, with proper antibiotics). After this, 200μL were directly placed, without altering cell densities, in 96-wells plate and a range of biobit (3-OXO-C_6-HSL in DMSO) was added to the four combinations. OD_{600} and GFP fluorescence was measured in plate-reader. All culture processes here were done at 37ºC with shaking.

Figure 4.17 shows OD_{600} and GFP fluorescence across all media and hours combinations. It is clear that there is some regrowth in the first hours at least when the culture was induced 24 hours after inoculation. That seems to be also the case for inducing 48 hours after inoculation in LB. Only in the case of Davis media induced 48 hours after inoculation, cell density (OD_{600}) seems to be relatively stable over time.

With respect to receptor strain induction, GFP fluorescence tends to show a pulse in the first hours, and then decreases in a more or less pronounced way. This pulse is extremely evident in the case of LB media induced 24 hours after inoculation, whereas the pulse disappears when cells are cultured in Davis media and induced 48 hours after inoculation. This results were successfully replicated.

This last experimental condition (cultures induced 48 hours in Davis media) presents ideal characteristics:

- Cell density remains constant for tens of hours.
- Induction curves show a behaviour compatible with GFP constant rate of biosynthesis and a degradation rate proportional with the GFP concentration:

$$\frac{\mathrm{d}[\mathrm{GFP}]}{\mathrm{d}t} = \alpha - \gamma[\mathrm{GFP}] \tag{4.2}$$

  which is always determined to asymptotically reach equilibrium.

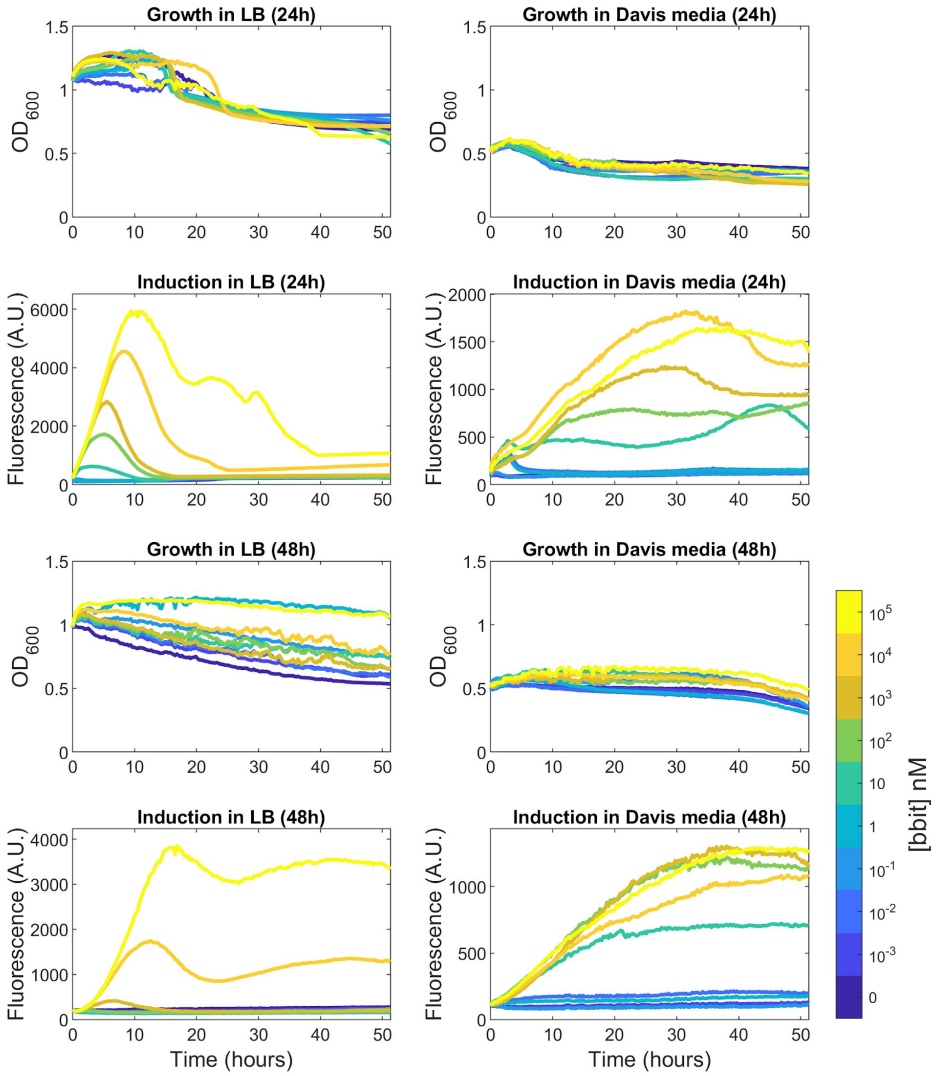


Figure 4.17: **Inducing receptor strain in stationary phase**. Receptor strain hosting pRHW$_E$ plasmid in LB or Davis media induced 24 or 48 hours after inoculum.

- Fluorescence increases linearly at least for the 10 first hours, agreeing with Gefen (2014) [53].
- Induction seems to show a good sigmoidal relation (Hill-function curve) with respect to biobit concentration, allowing the refinement of previous Eq. 4.2 into:

$$\frac{\mathrm{d}[\mathrm{GFP}]}{\mathrm{d}t} = \alpha \frac{B^n}{K^n + B^n} - \gamma[\mathrm{GFP}] \tag{4.3}$$

Finding receptor strain parameters.

Once we have found good and stable experimental conditions for proper output reading, we are going to proceed to measure receptor parameters. Now, I am going to declare the protocol in the most detailed way possible, so anyone will be able to replicate the experimental conditions. The consensus protocol is composed by the following steps:

1. Place precultures in 5mL LB adding appropriate antibiotics (Cm 10µg/mL for the pRHW$_E$ and Kn 25µg/mL for ΔSdiA[11]) from glycerol stock tips and culture at 37oC shaking at 225rmp overnight (24 hours).
2. Next day, add 10µL from this culture to 10 mL of Davis Media with 0.4% Glucose with antibiotic depending on the cultured strain (initial OD$_{600} \approx 6 \cdot 10^{-3}$ measured with Eppendorf BioPhotometer Model #6131). Then proceed to culture at 37ºC shaking at 225rmp overnight for 48 hours.

[11] Mutants from the Keio collection have a Kanamycin resistance cassette, and adding this antibiotic prevents contamination.

3. Fill a 96-well plate will 198µL of culture and add 2µL of biobit. Our biobit, 3-OXO-C_6-HSL, is stocked in DMSO at 100mM. Each well of the plate-reader contained 200µL (198µL of culture and 2µL of DMSO with desired 3-OXO-C_6-HSL concentration). The original inducer was diluted in DMSO to get a full range of inducer to add to the wells. Biobit must be diluted in factors of 10 starting at 100µM. Filter some milliliters of the culture by centrifuging 10 minutes at 10000 rpm and then pass supernatant through 0.2µm pore filter. Use this filtered stationary phase media for blank measurement. Use electronic pipette if possible to minimize temperature changes, and fill the plate placed over a warm surface (37°C), as temperature changes may trigger undesired growth. Seal the plate before culturing in the plate-reader, because it prevents evaporation.

4. Culture in plate-reader (TECAN in our case) with orbital shaking at 37°C. Measure optical density at 600nm (9nm bandwidth) and GFP fluorescence with excitation wavelength 483nm (9nm bandwidth) and emission 540nm (bandwidth 20 nm) and gain 80. I did one measurement each 15 minutes.

This pipeline was followed doing induction curves with the *E. coli* BW25113 ΔSdiA strain with pRHW$_E$, with 6 biological replicas and 2 experimental replicas. A Hill function (Eq. 4.1) was fitted to each culture and for each time-point. Hill functions were normalized and then Hill-coefficient and K value distributions were analyzes (Figure 4.18).

Figure 4.18 shows the parameter distributions for Hill coefficient and K value. The two colors represent the two experimental replicas. K parameter showed log-normal like distribution, so a normal distribution was fitter to the logarithm of the K values. This distributions

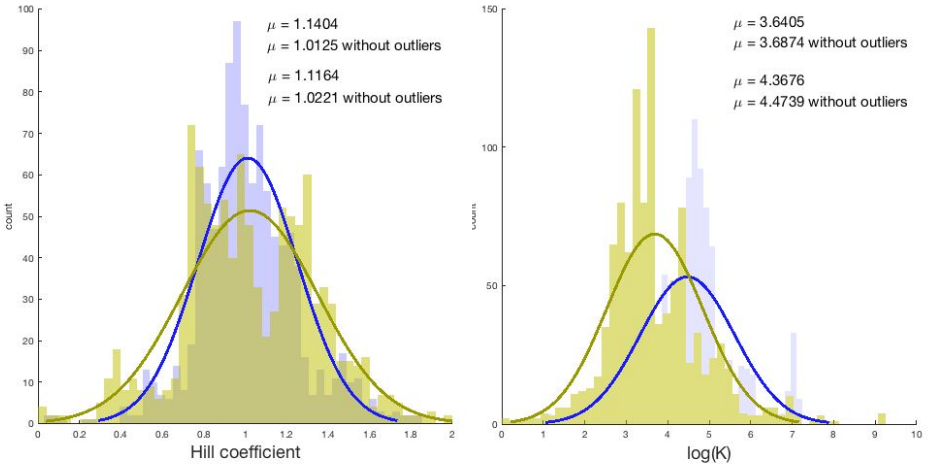


Figure 4.18: **Parameter distribution for Hill function fitting**. Histograms for Hill coefficient and K value for two independent replicas. A normal distribution was fitted in each case.

seem to define our receptor parameters in a simple way: for Hill parameter, $n \approx 1$, which implies that there is no apparent cooperativity in the biobit-LuxR binding process, and so the Hill function becomes a Michaelis-Menten equation. On the other hand, $K = e^4$ nM ≈ 54.5982 nM.

With these two parameters we can, assuming that our receptor follows a Hill function, find the receptor thresholds: $T_5 \approx 2.87$ nM and $T_{95} \approx 1.037$ µM.

4.3.2. Emitter strain experiments

When dealing with emitter strain, we only have to characterize one value, the emitter parameter K_E when the emitter is fully induced, which in our case happens when IPTG concentration is 100µM.

I conceived this experiment in a very straightforward manner: mixing emitter strain in stationary phase with IPTG (100µM), then filtering the media, and measuring 3-OXO-C_6-HSL concentration with small molecule mass spectrometry. Nevertheless, economic restrictions forced us to be more creative, so we had to come with an alternative approach. We have a characterized receptor strain that could measure biobit concentration (though in a very limited range). So, what we are going to mix in known ratios receptor and emitter strain and then add IPTG to induce emitter strain. Then, from the receptor induction curve and its temporal dynamics, we should be able to deduce K_E.

But whereas with the direct approach using mass spectrometry we could find K_E using the wire differential equation, if we want to characterize K_E parameter by fitting temporal series to some model, that model must be more descriptive.

Using the non-instantaneous wire system model

We already have shown the more descriptive non-instantaneous receptor strain differential equation at the end of previous chapter. Adapted for our wire implementation, our non-instantaneous receptor equation would look likes this:

$$\frac{\mathrm{d}F_{OD}}{\mathrm{d}t} = \frac{\mathrm{d}[\mathrm{GFP}]}{\mathrm{d}t} = \alpha_R \frac{B^n}{K_R^n + B^n} - \gamma_R[\mathrm{GFP}] \tag{4.4}$$

which is the same as equation 3.23. Recall that we are writing this equation with [GFP] as output, and by doing that we are implicitly assuming that there exist a linear correlation between GFP concentration and measured fluorescence per receptor strain cell unit, F_{OD}. If we

solve this equation, we obtain the following expression:

$$F_{OD}(t) = \frac{\alpha_R}{\gamma_R}\frac{B^n}{K_R^n + B^n}(1 - e^{-\gamma_R t}) + e^{-\gamma_R t}F(0) \tag{4.5}$$

Assuming that $F_{OD}(0) = 0$, we finally obtain the explicit solution for non-instantaneous expected fluorescence per receptor strain cell unit:

$$\boxed{F_{OD}(t) = \frac{\alpha_R}{\gamma_R}\frac{B^n}{K_R^n + B^n}(1 - e^{-\gamma_R t})} \tag{4.6}$$

Because we will be measuring OD_{600} and GFP fluorescence in a plate-reader in experiments mixing emitter and receptor strains, we need analytical tools to predict how F_{OD} will depend on time and strain ratios. For that matter, the next question that we need to address is: should we consider that the emitter strain switches from K_E^{OFF} to K_E^{ON} instantaneously or not?

Assuming K_E^{OFF} switches to K_E^{ON} instantaneously.

The first assumption will be to consider the switch $K_E^{OFF} \rightarrow K_E^{ON}$ to be instantaneous. Under that assumption, and taking into account that in the following experiments we will not be adding sink strain to the culture ($\sigma = 0$), so the wire equation simply becomes

$$\frac{\mathrm{d}B}{\mathrm{d}t} = K_E P \varepsilon \tag{4.7}$$

whose solution is $B(t) = K_E P \varepsilon t + B(0)$, and assuming $B(0) = 0$ and $K_E^{OFF} = 0$, and $K_E = K_E^{ON}$, we can substitute $B(t)$ in Eq. 4.6 to obtain the expression that predicts how fluorescence will evolve assuming instantaneous $K_E^{OFF} \rightarrow K_E^{ON}$ switch:

$$\boxed{F_{OD}(t) = \frac{\alpha_R}{\gamma_R}\frac{(K_E^{ON} P \varepsilon t)^n}{K_R^n + (K_E^{ON} P \varepsilon t)^n}(1 - e^{-\gamma_R t})} \tag{4.8}$$

From this last expression, if we can estimate all parameters related with receptor strain by fitting Eq. 4.6 when receptor cells are induced under a biobit concentration range as we did in the previous section. Then, we can use these estimated parameters in Eq. 4.8 and fit again the obtained time series for cultures composed by known ratios of emitter and receptor strains to estimate K_E^{ON} parameter.

Assuming K_E^{OFF} switches to K_E^{ON} non-instantaneously.

Conversely, we want to consider the switch $K_E^{OFF} \rightarrow K_E^{ON}$ to be non-instantaneous. Then, $K_E(t)$ is a function that varies over time when input is present. This function was described in the previous chapter:

$$K_E(t) = \frac{\alpha_E}{\gamma_E}\{\mathrm{In}\}(1 - e^{-\gamma_E t}) + e^{-\gamma_E t}K_E(0) \tag{4.9}$$

Because we will assume $K_E^{OFF} = 0$ and the switch will be from K_E^{OFF} to K_E^{ON} ($\{\mathrm{In}\} = 1$ and $K_E(0) = K_E^{OFF} = 0$), $K_E(t)$ becomes:

$$K_E(t) = \frac{\alpha_E}{\gamma_E}(1 - e^{-\gamma_E t}) \tag{4.10}$$

Substituting in Eq. 4.7, we obtain:

$$\frac{\mathrm{d}B}{\mathrm{d}t} = K_E(t)P\varepsilon = \frac{\alpha_E}{\gamma_E}(1 - e^{-\gamma_E t})P\varepsilon \tag{4.11}$$

whose solution is:

$$B(t) = \frac{\alpha_E}{\gamma_E^2}(e^{-\gamma_E t} + \gamma_E t - 1)P\varepsilon + B(0) \tag{4.12}$$

Assuming $B(0) = 0$ and substituting in $F_{OD}(t)$ function (Eq. 4.6), we obtain the final expression:

$$\boxed{F_{OD}(t) = \frac{\alpha_R}{\gamma_R}\frac{\left[\frac{\alpha_E}{\gamma_E^2}(e^{-\gamma_E t} + \gamma_E t - 1)P\varepsilon\right]^n}{K_R^n + \left[\frac{\alpha_E}{\gamma_E^2}(e^{-\gamma_E t} + \gamma_E t - 1)P\varepsilon\right]^n}(1 - e^{-\gamma_R t})} \tag{4.13}$$

A expression that, again, we can use in an analogous way to the previous situation by first fitting a biobit induction curve with receptor strain and obtain the characteristic receptor parameter. Then, we use this last expression to estimate the two characteristic parameters for emitter strain, α_E and γ_E, being $K_E^{ON} = \alpha_E/\gamma_E$.

Defining the emitter + receptor co-culture experiment

Now that we have the two fitting tools for the two possible scenarios that we can find, let me define the experimental setup. This experiment will start exactly as we did for receptor experiments in previous section, but this time, we will also include emitter strain (*E.coli* BW25113 ΔSdiA + pEHW) in our precultures. From these precultures, new cultures in Davis media were placed but this time **without** antibiotics. We have to do this because we will be mixing host strain with plasmid and without plasmids. These strains were then cultured at 37°C with shaking for 48 hours.

	E 0 IPTG	R 0 bbit	R 1E-2 bbit	R 1E-1 bbit	R 1 bbit	R 10 bbit	R 1E2 bbit	R 1E3 bbit	R 1E4 bbit	R 1E5 bbit	
	R:E 10:0 +IPTG	R:E 9:1 +IPTG	R:E 8:2 +IPTG	R:E 7:3 +IPTG	R:E 6:4 +IPTG	R:E 5:5 +IPTG	R:E 4:6 +IPTG	R:E 3:7 +IPTG	R:E 2:8 +IPTG	R:E 1:9 +IPTG	
	R:E 10:0 0 IPTG	R:E 9:1 0 IPTG	R:E 8:2 0 IPTG	R:E 7:3 0 IPTG	R:E 6:4 0 IPTG	R:E 5:5 0 IPTG	R:E 4:6 0 IPTG	R:E 3:7 0 IPTG	R:E 2:8 0 IPTG	R:E 1:9 0 IPTG	
fSP	R:Δ 10:0 0 IPTG	R:Δ 9:1 0 IPTG	R:Δ 8:2 0 IPTG	R:Δ 7:3 0 IPTG	R:Δ 6:4 0 IPTG	R:Δ 5:5 0 IPTG	R:Δ 4:6 0 IPTG	R:Δ 3:7 0 IPTG	R:Δ 2:8 0 IPTG	R:Δ 1:9 0 IPTG	Δ

Figure 4.19: **96-well plate distribution scheme**.

These three cultures (emitter strain, receptor strain and host strain without plasmid) were mixed or not in different combinations according to the scheme shown in Figure 4.19. In that scheme, some wells

appear with just one letter **R**, **E** or Δ (receptor, emitter and host strain without plasmid respectively). If in a position in the table two letters appear, then it represents a co-culture of the two depicted strains in the ratios shown below. The last row in each position shows the specific inducer added to the well: biobit 3-OXO-C_6-HSL with the final concentration in nM, or the input molecule IPTG at 100µM (0 IPTG represents no IPTG was not added). fSP means filtered stationary phase media. 200µL of final volumen was added to each well and the empty wells were filled with water. There is enough space in the 96-wells plate to place 2 replicas. The plate was sealed and cultured in Tecan plate-reader with the settings following the protocol described in previous section.

Normalization and background subtraction

This topic is never the main section of any paper[12]. Nevertheless, the formulas defined in this chapter to fit experimental data require to obtain F_{OD} values (fluorescence per receptor strain cells unit). And taking into account how the experiment was designed, this case requires a fine and well motivated explanation about background substraction and normalization for complex co-cultures.

Let us start by dissecting what each OD_{600} and fluorescence measured value means for each well. In Figure 4.20 we can see two values that appear in each well: OD_{bg} and F_{bg}. These two values are what the media (M) contributes by itself in each well, and so, it appears below every well.

We can also see the values OD_R, OD_E and OD_Δ, the contribu-

[12]Gefen's paper from 2014 [53], done by Balaban's group, does not even mention if their data have been normalized.

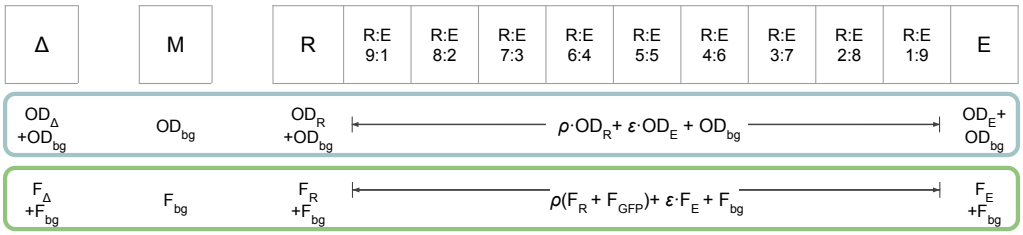


Figure 4.20: **OD_{600} and fluorescence in each well**.

tion to the optical density measurement that directly depends on cells in the media (receptor, emitter and host strains respectively). Analogously, F_R, F_E and F_Δ correspond to their contribution to fluorescence measurement, but only with the autofluorescence (not the one that depends on any GFP).

Finally, there is another contribution to fluorescence, F_{GFP}, the fluorescence that only depends on GFP. As the reader may have seen, this value only appears in wells that have emitter and receptor cells simultaneously, because it will only be there where biobit will be both produced and sensed when IPTG will be added.

We can see that in the wells that contain both emitter and receptor cells, the contribution to the optical density and fluorescence from receptor and emitter cells is scaled according to the relative ratios ρ and ε. Because the same culture volume was added to each well, we can subtract OD_{bg} and F_{bg} to every well. I am afraid it is not so trivial for the other magnitudes. Why? Because the following magnitudes depend on cell density, and cell density may present small variations between wells due to the small pipetting volume (and hence the not so small associated error).

We have defined F_{OD} magnitude from the analytical point of view.

Now it is turn to define it from the experimental one. When talking about a specific strain X, we can define the cell density of that strain as OD_X - OD_{bg}, being OD_X the measured optical density (which includes OD_{bg}). Exactly the same can be said about fluorescence: the autofluorescence of a culture that only depends on cells is F_X - F_{bg}. Note that we are not assuming that our strain X is expressing any fluorescent protein, and if that is the case, then the last expression will account for both the cell autofluorescence and fluorescence from the fluorescent protein.

Now, we can define the fluorescence per X cell density unit as:

$$F_{OD,X} = \frac{F_X - F_{bg}}{OD_X - OD_{bg}} \tag{4.14}$$

We can then calculate $F_{OD,R}$ and $F_{OD,E}$ using the value that we have in the wells where there is only emitter or receptor strain. This magnitude allows us to guess the expected fluorescence that we should observe when other cell density is observed. If we assume that no GFP fluorescence will be observed, then the expected fluorescence in wells containing co-cultures of receptor and emitter strain, $F^{exp}_{R:E}$, will be:

$$F^{exp}_{R:E} = F^{obs}_{OD,R}(OD_R - OD_{bg})\rho + F^{obs}_{OD,E}(OD_E - OD_{bg})\varepsilon \tag{4.15}$$

This is linear combination of the expected fluorescence from each strain, but calculated for each well (in which we know the strain ratio) using the observed F^{obs}_{OD} values for each strain: $F^{exp}_{R:E} = F^{exp}_E + F^{exp}_R$. Because we have assumed that the fluorescence measurement in each well never is due to GFP, any fluorescence surplus can be considered to be dependent on GFP expression, thereby a simple substation will give us the desired quantity:

$$F_{GFP} = F^{obs}_{R:E} - F^{exp}_{R:E} - F_{bg} \tag{4.16}$$

With this last quantity, we can calculate $F_{OD,GFP}$, the expected GFP fluorescence per receptor strain cell unit, which we can then fit and compare with our previous analytical formulas:

$$\boxed{F_{OD,GFP} = \frac{F_{GFP} - F_{bg}}{(OD_{R:E} - OD_{bg})\rho}} \tag{4.17}$$

In Figure 4.21 we can appreciate the effect of this treatment when analysing plate-reader data from the row that contains the emitter and receptor strains co-cultured with IPTG. The bottom plot shows the effect of our data processing. The smoothest curves are plotted with continuous lines, whereas the ones that show unexpected peaks are plotted with dots. Speaking about peaks: all curves seem to reach equilibrium *after* a peak followed by a soft fluorescent decrease. For that matter, we will only use the continuous lines before the peak to fit our data to our two models.

Parameter estimation for emitter strain using dynamical data

Even when we want to measure emitter strain parameters, we are first goint to measure receptor strain parameters first. We will be using the formulas describing output fluorescence and fitting data corresponding to the wells containing receptor strain with a biobit induction range. For that matter, we will use Eq. 4.6 ($F_{OD}(B,t)$). The previous fitting method adjust a Hill function to each time point and estimates parameters from the parameter distribution across the whole time series. The main difference with this new approch is that in this case, we will use the whole data set (across different biobit concentrations and time points) to fit a surface assuming that output is non-instantaneous. Figure 4.22 shows the fittest $F_{OD}(B,t)$ surface together with the experimental data.

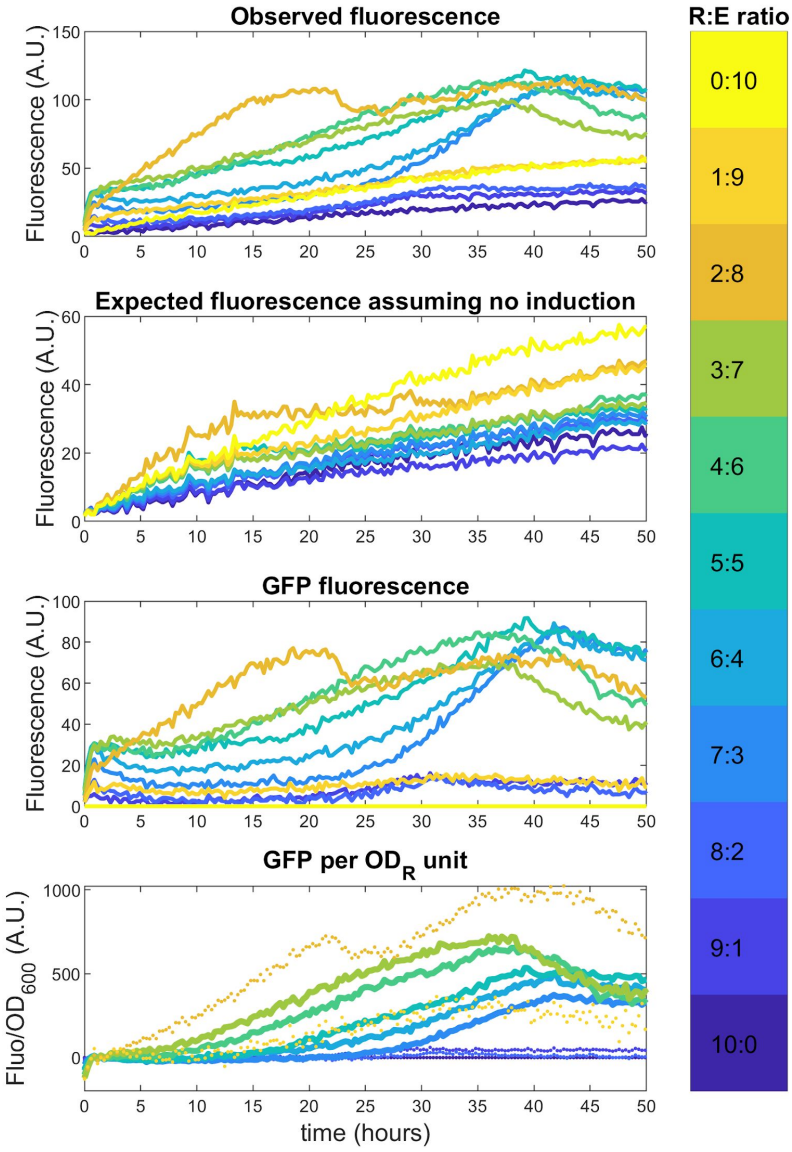


Figure 4.21: **Successive steps for obtaining GFP per unit of receptor cell.**

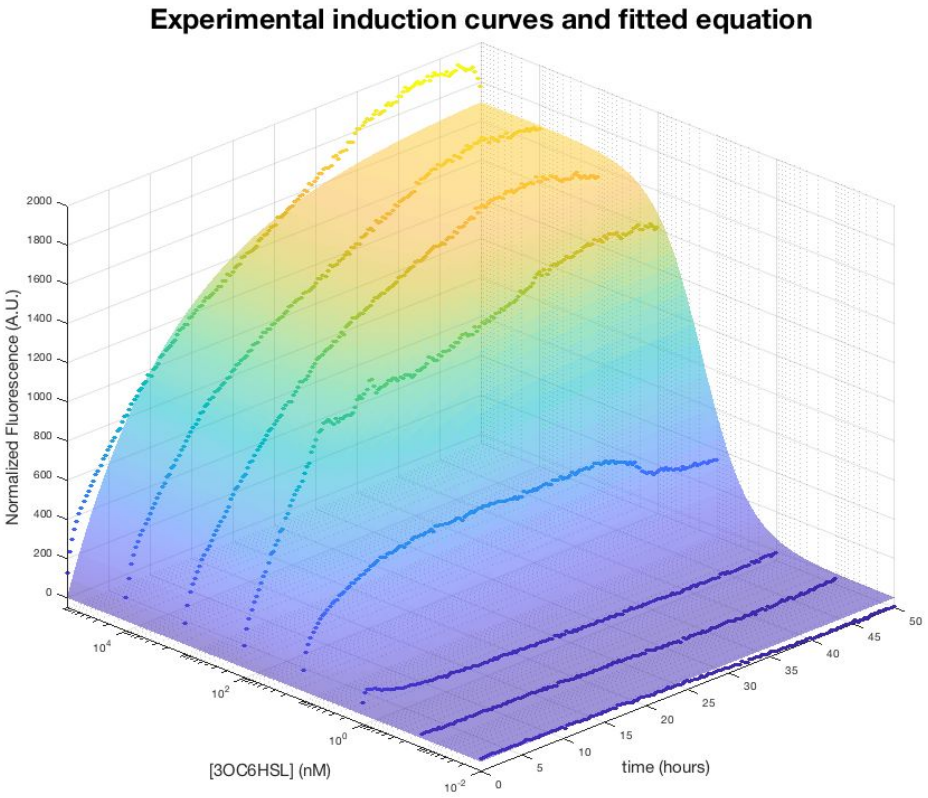


Figure 4.22: **Fitting $F_{OD}(B,t)$ to receptor strain induction data.**

Recall that $F_{OD}(B,t)$ has the following expression:

$$\boxed{F_{OD}(t) = \frac{\alpha_R}{\gamma_R}\frac{B^n}{K_R^n + B^n}(1 - e^{-\gamma_R t})} \tag{4.18}$$

After fitting, we obtained the parameter estimation ($R^2 = 0.98288$) described in Table 4.4. The only parameter that is different from previous receptor strain measurements is K_R, which in the parameter distribution approach gave a value of $K_R \approx 54.6$ nM. Nevertheless, it still shares the same order of magnitude.

Parameter	Value	95%C.I.
K_R	18.254	(17.5649, 18.9432)
n	1.1322	(1.0911, 1.1733)
α_R	145.8505	(142.6988, 149.0022)
γ_R	0.084835	(0.084835, 0.087168)

Table 4.4: Parameter estimation for receptor strain.

Once we have measured the parameters corresponding to the receptor strain, we can proceed to adjust our two $F_{OD}(\varepsilon, t)$ formulas (one assuming instantaneous emitter and another one being non-instantaneous). We will use the previously estimated parameters, so we only need to find the emitter specific parameters.

If we assume that the emitter strain switches from $K_E^{OFF} \rightarrow K_E^{ON}$ non-instantaneously, the equation that we must adjust is Eq. 4.13:

$$F_{OD}(t) = \frac{\alpha_R}{\gamma_R} \frac{\left[\frac{\alpha_E}{\gamma_E^2}(e^{-\gamma_E t} + \gamma_E t - 1)P\varepsilon\right]^n}{K_R^n + \left[\frac{\alpha_E}{\gamma_E^2}(e^{-\gamma_E t} + \gamma_E t - 1)P\varepsilon\right]^n}(1 - e^{-\gamma_R t}) \quad (4.19)$$

We can use the parameters from Table 4.4 and fit the surface to $F_{OD,GFP}$ inferred from the row with input ITPG plus emitter and receptor strains mixed in known ratios. Figure 4.23 shows the resulting fitting ($R^2 = 0.851$). For this fitting, we are only using (as we previously mentioned), the time series showing smooth responses and avoiding the final peaks. An analogous fitting process was done using Eq. 4.8 (response with instantaneous $K_E^{OFF} \rightarrow K_E^{ON}$ switch):

$$F_{OD}(t) = \frac{\alpha_R}{\gamma_R} \frac{(K_E^{ON} P\varepsilon t)^n}{K_R^n + (K_E^{ON} P\varepsilon t)^n}(1 - e^{-\gamma_R t}) \quad (4.20)$$

For this approach, similar results to those shown in Fig. 4.23 were obtained ($R^2 = 0.8164$ in this case). From this last expression Eq. 4.20,

we can directly obtain an estimation for K_E^{ON}.

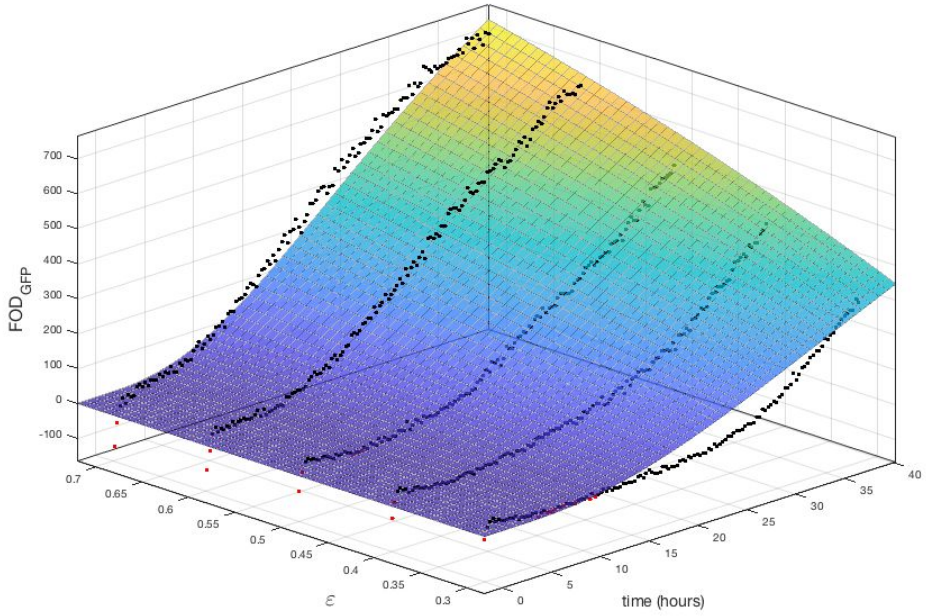


Figure 4.23: **Fitting $F_{OD,GFP}(\varepsilon, t)$ to receptor+emitter strains co-culture data.**

Parameter	Value	95%C.I.
α_E (non-instantaneous emitter)	0.1037	(0.08384, 0.1235)
γ_E (non-instantaneous emitter)	0.05919	(0.03365, 0.08473)
K_E^{ON} (instantaneous emitter)	0.8941	(0.8664, 0.9218)

Table 4.5: Parameter estimation for emitter strain.

Is we want to obtain K_E^{ON} when $K_E^{OFF} \rightarrow K_E^{ON}$ transition is non-instantaneous, the recall that $K_E^{ON} = \alpha_E/\gamma_E = 1.752$. We can appreciate how there exists a difference of barely one unit between the two

K_E^{ON} estimations, being the estimation assuming non-instantaneous wire to be slightly more precise (higher R^2 value). But we may agree that $K_E^{ON} \approx 1$.

Sadly for us, this method was extremely difficult to replicate. I think that this may be due to the extremely complex experimental setup, that accumulates a lot of steps and thus it tends to accumulate error. Then, these errors probably are increased with the data preparation and analysis. I believe that even when the theory is, in my humble opinion, essentially correct, a simpler or more direct approach should be used for estimating K_E^{ON}. I still think that direct 3-OXO-C_6-HSL measurement by mass spectrometry might be the right way to characterize both emitter and sink strains.

4.3.3. Discussion about the experiments

As I anticipated in the beginning, it was not possible to conclude the experimental validation. This was due to the pandemics, but also because measuring the wire parameters using indirect methods sequentially adds experimental noise. I advocate for directly measuring our wire parameters by quantifying HSL in supernatants using mass spectrometry over different time points. Also, the only strain that can be properly characterized is the receptor strain. Nevertheless, even though Hill coefficient and K measurements were consistent over the experiments, it was not the case for the maximum response. I suspect that this might be due to the combined effect of the degradation tags and doing our experiments in stationary phase. As I will explain in the last chapter of this dissertation, proteolytic degradation increases in stationary phase specially for proteins tagged with LVA-degradation tag.

Finally, we have found and established a useful protocol. We have tested how *E. coli* BW25113 $\triangle$Sdi) is virtually blind to HSL molecules, widely used in synthetic biology. Also, we have proved that Davis media is allows stable growth and induction of our host strain. GFP fluorescence was observable without the media showing any significant autofluorescence. And also, induction after 48 hours ensures constant gene expression with residual growth if any. All these reasons define a good and suitable experimental setup for synthetic biology.

Chapter 5

CONCLUSION

We have covered what synthetic biology is, and the relation that it has with computation. In particular we have discussed the branch that seeks the emulation of combinatorial logic circuits using biological entities to represent logic gates. And having restricted ourselves to this concrete topic, we have again focused in a very specific path: distributed biological circuits.

Under this paradigm, there is a distribution of logic labour among the different strains that comprise the culture. These strains will share information using chemical wires. These wires, frequently (but not exclusively), will be based on quorum sensing systems from Gram negative bacteria. Splitting sender elements from receiver elements, and placing them in two different strains, we can define the emitter and the receptor components of the wire. Thus, the two strains will be connected by sharing a diffusible molecule in the media. In the case of quorum sensing from Gram negative bacteria, these small diffusible molecules are acyl-homoserine lactones (HSL). We have discussed how this wire paradigm posses its own problems.

Initially formulated under the name *wiring problem* [128], it reflects the fact that due to the lack of physical isolation of the wire molecules, it is required to have unique molecules for each wire between two strains, and combinatorial logic circuits tend to increase the number of wires very fast. I honestly think that we could talk about the ***wiring problems*** (in plural), to capture all the problems that appear because chemical wires are being used:

- The number of different required wires increases with the complexity of the circuit [128].
- When using some cell-cell communication systems, such quorum sensing, cross-talk limits the number of compatible systems that present orthogonality [93, 148, 110].
- A wire might not being suitable for computation purposes, because it must present operative regions: digital region and buffer region. Also, in a distributed circuit with several wires, the ratios of all strains coexisting in the culture must be so every wire is in its operative regions. This last requirement is proposed in this work.

In this work we also propose a specific wire architecture, with a third strain called sink, with the purpose of actively degrade the signaling molecule, the biobit. We have introduced a representational framework that allows to visualize the Wire Space and gain intuition about the general wire behaviour. We have explored the properties and capabilities of this architecture. And finally we have tried to implement it. Even though it was not possible to successfully finish the experimental implementation, we have defined specific experimental

conditions that allow to test synthetic constructs in stationary phase. And now, let me discuss each milestone.

5.1. A coherent wire architecture

It is difficult to compare the proposed wire architecture against other wire architectures, as it has never been proposed nor published an official wire architecture. Needless to say that this should not constitute a problem, because wires have been definitely used in several works. And even when the purpose of such works was not to test the properties of their wires, implicitly they were using some specific wire architecture. Up to date, there is no work that specifically proposes a wire architecture. All works analyzed in this thesis make use of wires because it is required to test another hypothesis.

One of the main advantages of distributed circuits is that it is possible to reprogram them by replacing strains in the culture [86], or by directly adjusting the relative ratios of the strains comprising the culture, which I personally consider to be more interesting and less obvious. These two new possibilities are opposing with the alternative of genetically modifying the circuit. There are two works that explore this last possibility, the ratio modification. The first work by Silva (2017) [126] directly uses strain ratio modification to test the capabilities of its initial hypothesis about the signal (biobit) propagation along spatially distributed circuits. The second work, by Urrios (2018) [136], analyses the idea of modularity and plays with the strain ratio modification, but always in the context of a specific circuit (an incoherent feed forward loop). Thus, the focus is more on the circuit itself than in the general principles that may underline the use of chemical wires.

Our wire architecture is a conceptual framework to rationally implement chemical wires and predict the effect of strain ratio modification. Thus, these two works could have been done using tools that would have allowed to predict the regions in the wire space with the most interesting properties for each case for each wire.

There is a limitation though: my wire architecture is intended to be used in non-growing cells. We have confirmed what Gefen proved in 2014: *E. coli* can induce gene circuits in stationary phase [53]. But the work done by Urrios (2017) uses yeast as host organism. The idea of using yeast while the cells have arrested their growth has been investigated mainly for biotechnology purposes [145]. And there is no work focusing on the possibility of inducing synthetic circuits in yeast with arrested growth. Nevertheless, I do not see a big challenge in implementing the circuit that they propose in *E. coli*, which would have the advantage of testing their circuit in non-growing cells.

5.2. The buffer region allows wires to act as multicellular timers

Cells live in environments that evolve or change over time, thus being able to process this temporal dimension might be advantageous. [131, 49]. An important requirement of temporal information processing is the quantification of the time interval that a given input signal has been acting upon a cell. Complementarily, cells sending signals might need to program the amount of time that the signal is active. In technological settings, both these functions are performed by timers. It is thus necessary to establish how timers are implemented in cells. Cell-

intrinsic molecular timers have been identified in recent years, having been reported to regulate a wide variety of cellular processes, including apoptosis [68, 90, 54, 48], cellular proliferation [150], cell-fate specification [36, 117], and infection response [37]. These timers usually depend on intricate molecular processes [84] that are difficult to tune and are sensitive to noise.

From a systems-level perspective, feedback-based timer circuits have been proposed in synthetic biology applications [38, 112], and have also been found to operate naturally through pulses [123, 82] and oscillations [20], and even to be induced by noise [135]. These intracellular timer circuits usually depend on an accumulating signal surpassing a threshold [81], and are thus limited by the amount of signal molecules that cells can produce and store in their interior without changing their basal metabolic state. This limitation disappears if the factor controlling the timer is exported outside the cell and stored in the extracellular medium. Cells in such multicellular timers would operate collectively, and consequently their function would be less affected by noise [43].

Our wire architecture constitutes a minimal distributed circuit by itself, which potentially can present what we have called the buffer region. Cultures placed within this region posses the ability to act as a timer, storing the information about the input time extent. It is possible to tune the timer properties by adjusting the strains ratios.

This timer property can be used to process digital periodic signals (DPS). Thus, a wire can modify the duty cycle from the input signal to the output signal. In the extreme case, the input signal can be filtered in the output and the receptor strain would generate a continuous response. A simple study case is shown to illustrate how, theoretically, a

wire could distinguish the warm months by just processing the amount of daylight hours per day.

5.3. A promising experimental setup

Even though an experimental confirmation of the ideas explained in this dissertation could not be achieved, we have tested several promising experimental conditions for testing distributed circuits. One the problems that I suspect I might have had, is the addition of degradation tags to relevant proteins. This is a common practice that limits the maximum concentration that a protein can reach, and it is useful when high renewal rate is required, for example, in dynamical circuits such synthetic oscillators [39]. Nevertheless, special care must be taken when the circuit is meant to work in stationary phase.

The *ssrA* degradation tag consists in the AANDENYALAA amino acid sequence (LVA tag) added in the C-terminal. But there are several variation to that sequence usually in the three last amino acids that correspond to different degradation rates [47, 5]. One problem is that, whereas LVA tag is one of the most commonly used in synthetic biology for its high degradation rate, it might not be suitable when the experiments are done in stationary phase. This is because the enzymes responsible for proteolytic degradation increase its concentration in stationary phase [44], and because the LVA tag is the most aggressive degradation tag [63] [1].

But beyond these setbacks, I can conclude that the experimental conditions in which experiments were done are very promising for test-

[1] Before lockdown due to the pandemics, I was creating plasmids without degradation tags and with the ASV tag, which seems to be the mildest one[63].

ing circuits in stationary phase. In summary:

- The *Escherichia coli* strain BW25113 ΔSdiA cannot sense the presence of HSL and there is no growth impairment or requirements for growing in minimal media.
- Davis media shows negligible autofluorescence and allows gene induction for tens of hours. [2]
- Gene induction can be done 48 hours after inoculum, with no major impact in OD_{600}.

5.4. Future prospects

From the theoretical point of view, there are two important extensions that would give the model a bigger scope. The first one is the ability to connect wires. This implies that the strain that acts as the receptor for one specific biobit, will produce as output a different biobit. Hence, that strain is, simultaneously, the receptor of one biobit and the emitter of another. Figure 5.1 shows a system of coupled wired connected in series to emphasize the double nature of the receptor/emitter strains. Every biobit is labeled as b_k with k ranging from 1 to 4.

In the context of the instantaneous wire model, this could be achieved by multiplying the normalized response dependent on one biobit by the emitter constant of another biobit:

$$\frac{1}{P}\frac{\mathrm{d}B_i}{\mathrm{d}t} = \overline{r}(B_{i-1})K_{E,i}\varepsilon - \frac{D_{m,i}B_i}{K_{S,i}+B_i}\sigma$$

[2] My experiments were done with 0.4% glucose, but beyond 1% glucose, OD_{600} does not linearly increase with glucose concentration anymore (unpublished data).

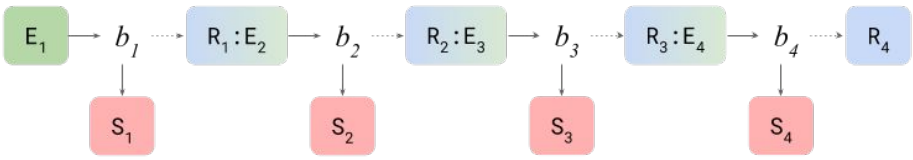


Figure 5.1: Four connected wires in series.

where $\overline{r}(B_{i-1})$ is the normalized response (Eq. 2.14), the Hill function that relates output dependent from B_{i-1}, the concentration of the previous biobit in the coupled wire system.

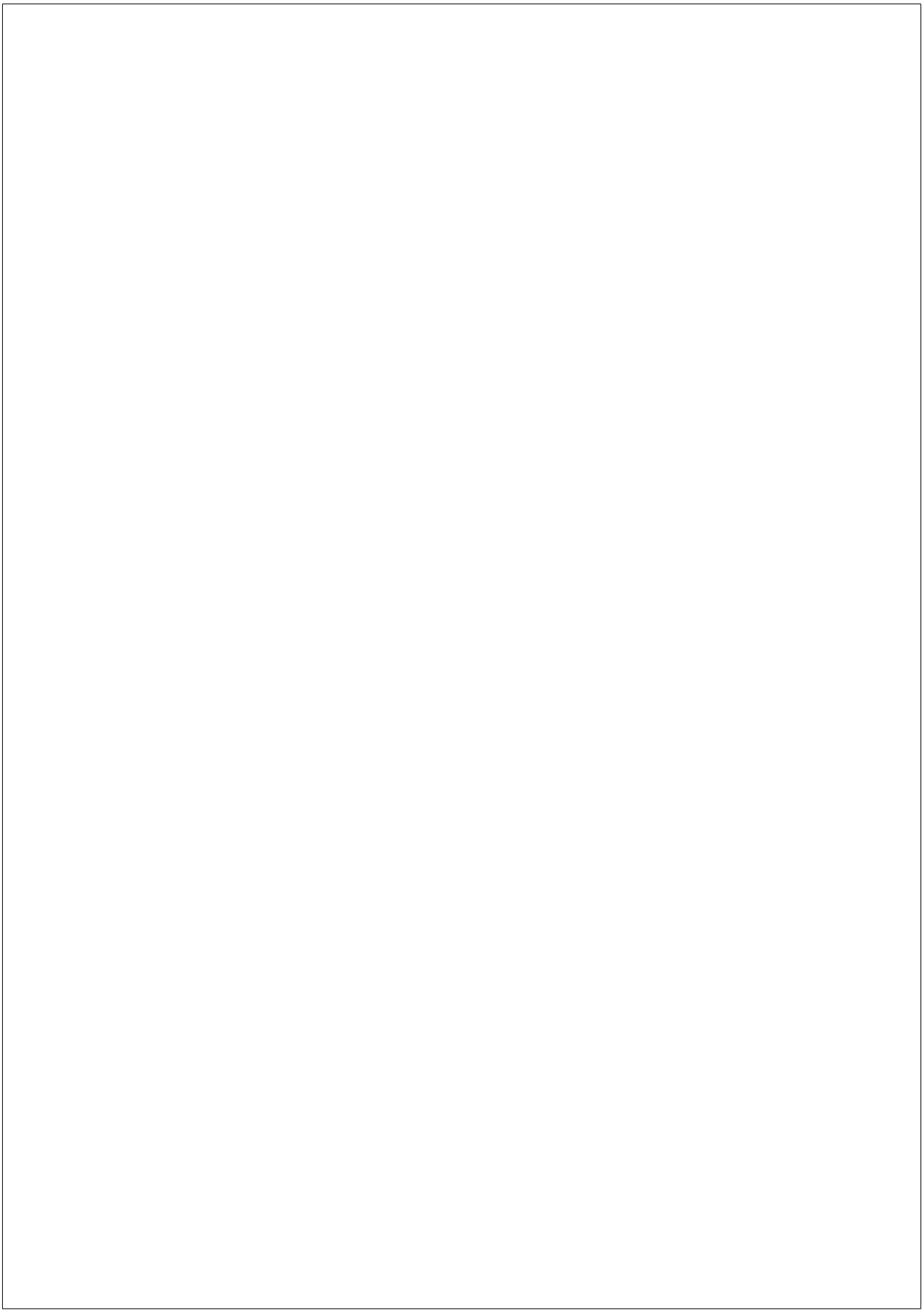
The equations presented in section 3.3 about the non-instantaneous wire models might be helpful for improving the descriptive capabilities of systems of coupled wires modelling. And the idea here is that the system of coupled wires will work if every wire is in its operative regions (buffer of digital, depending on the specific design). In this case, the triangle representation would not represent the proportion of emitter (ε), sink (σ) and receptor strains, but the proportion of emitter (ε), sink (σ), and the rest of the strains participating in the culture ($1 - \varepsilon - \sigma$). Hence, the culture would be a point within an n-simplex, a triangle of higher dimensions. And the circuit would be feasible if every wire is in its operative region.

Another possible model extension would be adding the space variable:

$$\frac{1}{P}\frac{\partial B}{\partial t} = K_E\varepsilon - \frac{D_m B}{K_S + B}\sigma + D_B\frac{\partial^2 B}{\partial x^2}$$

With this equation, the concept of feasibility depends on the tension between biobit biosynthesis, degradation, and diffusion. The idea here would be to spatially distribute the different strains on a surface, and

manipulate the information flow by creating sources and sinks regions for the different biobits in the system. This would be helpful particularly in spatially distributed networks [126].



Chapter 6

THE FINAL QUESTION

[AN EPILOGUE]

Over the last years, I have had the opportunity to develop the ideas I had 10 years ago. I know it is not common to do a PhD with your own idea. And actually, I did not come to Barcelona to do a PhD. I was not really interested in the first place. I just wanted to develop my ideas with professionals that could help me. At the end, I ended doing a PhD, but not about my ideas. I passed the two first years tracking single cells with time-lapse microscopy and microfluidics to unravel some gene expression temporal patterns. It did not work, and after these two years I had to change my research. This time to study my own project. But I had the chance to learn microscopy and microfluidics. And to study stationary phase, which would be the growth phase that I would use later. And of course improving my programming skills in MATLAB.

I have had to present several times my projects in public. I always feel very confident when I have to present, and because I have had to publicly defend two different projects (stationary phase project, and

my wire project), I know how it feels when the questions are about something you have or have not conceived. In the case of my first project, I used to be more aware of tricky questions, because it could be the case that I did not know an answer that I was supposed to know. But in the case of my own project, the feeling was totally different: I wanted to hear tricky questions. I used to know the answer to most of them, so it was kind of satisfying. But at some point the satisfaction mutated. What is the point of presenting something to some scientists if you already know the questions that they are going to ask, and hence the answers? There is no point, nor virtue in that. And I am not saying by any means that the audience was not capable of coming with a good and genuine question. The thing is that there is a fundamental difference between presenting a scientific question about Nature, and presenting a theoretical idea more related to engineering. In the first case, you are presenting your own path towards the Unknown, and in the second you are introducing your own child.

So yes, when I have presented my wiring project I have heard the same questions formulated in different ways. With time, a new question would appear, and that was like finding gold[1], because I would come back and think about the implications of the answer. It might be the case that the reader is thinking right now that *the big question* is going to be one of those. I am afraid this will not be the case. What I am pretty confident is that if the reader happens to be one of the members of the thesis jury, this paragraph is perfect for summoning tricky questions. But this is also written for the PhD student that will come after me, so I will take the risk. Any ways, *the big question*. From time to time, a person from the public, usually not interested in the intersec-

[1] The author has never found any gold.

tion between synthetic biology and computation, would formulate the innocent question:

But, what is the final purpose of this project?

And then I would say how there is no systematic and rational analysis of the properties of the chemical wires that some groups are already using. And after that, the person who just formulated the question would show me *that* face. Emptiness, like he or she is not hearing the answer. Which, it turns out, happens to be the case.

The answer that I give works very well when the person who asks is interested in computation and synthetic biology. But the question *what is the final purpose of this project?* is asking another thing. It fundamentally means: what is the purpose of creating biological computers? This question is extremely hard to answer. We have electronic machines that compute extremely fast. And now, quantum computing has joined the party and promises to be even faster. So, again: what is the point?

I have been thinking in this question for years, and I cannot figure out an answer. The question is not trivial, and Lewis Grozinger *et al* wrote a paper in 2019 in which they asked the same question [61]. They even put a fancy name to the moment a biological computer proves to be more capable at something than its silicon siblings: the cellular supremacy (analogous to quantum supremacy). In their paper, several pathways are offered, based on the properties of living organisms: they can self-organise, self-repair, they are resilient, they form distributed networks, and they can adapt and evolve. They claim that cellular supremacy will necessarily be achieved using at least one of these properties.

I have been thinking in another alternative path towards cellular supremacy based on the concept of biobit. If a bit can be 0 or 1, and a qbit can be 0, 1, or 0 and 1 simultaneously, a bbit can be whatever number between 0 and 1 (which is not really paradigm changing), but it can also store temporal information if the culture is within the buffer region. And the idea would be to think about the problems that can be solved using a biobit with its rules. Because implementation is an engineering problem. But the fundamental answer here is: which are the problems that can be solved using the biobit abstraction? So my answer is a question again, because I do not have the tools to attack such a question. But at least now there is a new *bit* model with its own rules. It is a place to start.

And "it is a place to start" is a magnificent and poetic way to finish this thesis. I will keep working in this ideas. Because now that these years during my PhD have provided me with new tools and knowledge, ***it is a place to start***.

Bibliography

[1] Fisher, R. (1930). The Genetical Theory of Natural Selection - Google Académico.

[2] Théorie physico-chimique de la vie et générations spontanées : Le Duc, Stéphane, 1853- : Free Download, Borrow, and Streaming : Internet Archive.

[3] L. M. Adleman. Molecular computation of solutions to combinatorial problems. *Science*, 266(5187):1021–1024, 1994.

[4] L. M. Adleman. Computing with DNA. *Scientific American*, 279(2):54–61, 8 1998.

[5] J. B. Andersen, C. Sternberg, L. K. Poulsen, S. P. Bjørn, M. Givskov, and S. Molin. New unstable variants of green fluorescent protein for studies of transient gene expression in bacteria. *Applied and Environmental Microbiology*, 64(6):2240–2246, 6 1998.

[6] E. Andrianantoandro, S. Basu, D. K. Karig, and R. Weiss. Synthetic biology: new engineering rules for an emerging discipline. *Molecular Systems Biology*, 2006.

[7] M. Ansaldi, D. Marolt, T. Stebe, I. Mandic-Mulec, and D. Dubnau. Specific activation of the Bacillus quorum-sensing systems by isoprenylated pheromone variants. *Molecular Microbiology*, 44(6):1561–1573, 6 2002.

[8] K. Arunasri, M. Adil, P. A. A. Khan, and S. Shivaji. Global gene expression analysis of long-term stationary phase effects in E. coli K12 MG1655. *PLoS ONE*, 9(5), 2014.

[9] T. Baba, T. Ara, M. Hasegawa, Y. Takai, Y. Okumura, M. Baba, K. A. Datsenko, M. Tomita, B. L. Wanner, and H. Mori. Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: The Keio collection. *Molecular Systems Biology*, 2, 5 2006.

[10] F. K. Balagaddé, H. Song, J. Ozaki, C. H. Collins, M. Barnet, F. H. Arnold, S. R. Quake, and L. You. A synthetic Escherichia coli predator-prey ecosystem. *Molecular Systems Biology*, 4(187):1–8, 2008.

[11] A. L. Barabási and Z. N. Oltvai. Network biology: Understanding the cell’s functional organization, 2 2004.

[12] A. Becskei and L. Serrano. Engineering stability in gene networks by autoregulation. *Nature*, 405(6786):590–593, 6 2000.

[13] R. L. I. a. t. m. biology and u. 1996. Evolution as engineering. *books.google.com*.

[14] D. Boneh, C. Dunworth, R. J. Lipton, and J. Sgall. On the computational power of DNA. *Discrete Applied Mathematics*, 71(1-3):79–94, 12 1996.

[15] J. Bonnet, P. Yin, M. E. Ortiz, P. Subsoontorn, and D. Endy. Amplifying genetic logic gates. *Science*, 340(6132):599–603, 5 2013.

[16] B. Borek, J. Hasty, and L. Tsimring. Turing patterning using gene circuits with gas-induced degradation of quorum sensing molecules. *PLoS ONE*, 11(5):17–20, 2016.

[17] M. Boudry and M. Pigliucci. The mismeasure of machine: Synthetic biology and the trouble with engineering metaphors. *Studies in History and Philosophy of Science Part C :Studies in History and Philosophy of Biological and Biomedical Sciences*, 44(4):660–668, 12 2013.

[18] K. Brenner, L. You, and F. H. Arnold. Engineering microbial consortia: a new frontier in synthetic biology. *Trends in biotechnology*, 26(9):483–9, 9 2008.

[19] C. Briat, A. Gupta, and M. Khammash. Antithetic Integral Feedback Ensures Robust Perfect Adaptation in Noisy Bimolecular Networks. *Cell Systems*, 2(1):15–26, 1 2016.

[20] H. Cai, M. Katoh-Kurasawa, T. Muramoto, B. Santhanam, Y. Long, L. Li, M. Ueda, P. A. Iglesias, G. Shaulsky, and P. N. Devreotes. Nucleocytoplasmic shuttling of a GATA transcription factor functions as a development timer. *Science*, 343(6177), 2014.

[21] B. Calles, A. Goni Moreno, and V. Lorenzo. Digitalizing heterologous gene expression in Gram-negative bacteria with a portable ON/OFF module. *Molecular Systems Biology*, 15(12), 12 2019.

[22] B. Canton, A. Labno, and D. Endy. Refinement and standardization of synthetic biological parts and devices, 7 2008.

[23] X. Carbonell, J. L. Corchero, R. Cubarsí, P. Vila, and A. Villaverde. Control of Escherichia coli growth rate through cell density. *Microbiological Research*, 157(4):257–265, 1 2002.

[24] M. Carbonell-Ballestero, S. Duran-Nebreda, R. Montañez, R. Solé, J. Macía, and C. Rodríguez-Caso. A bottom-up characterization of transfer functions for synthetic biology designs: Lessons from enzymology. *Nucleic Acids Research*, 42(22):14060–14069, 12 2014.

[25] M. Carbonell-Ballestero, E. Garcia-Ramallo, R. Montañez, C. Rodriguez-Caso, and J. Macía. Dealing with the genetic load in bacterial synthetic biology circuits: Convergences with the Ohm’s law. *Nucleic Acids Research*, 44(1):496–507, 1 2016.

[26] S. Cardinale and A. P. Arkin. Contextualizing context for synthetic biology - identifying causes of failure of synthetic biological systems. *Biotechnology Journal*, 7(7):856–866, 7 2012.

[27] K. G. Chan, W. F. Yin, C. K. Sam, and C. L. Koh. A novel medium for the isolation of N-acylhomoserine lactone-degrading bacteria. *Journal of Industrial Microbiology and Biotechnology*, 36(2):247–251, 2009.

[28] G. M. Church, Y. Gao, and S. Kosuri. Next-generation digital information storage in DNA, 9 2012.

[29] J. P. Cox. Long-term data storage in DNA, 7 2001.

[30] Y. Dang, D. A. Grundel, and H. Youk. Cellular Dialogues: Cell-Cell Communication through Diffusible Molecules Yields Dynamic Spatial Patterns. *Cell Systems*, 10(1):82–98, 2020.

[31] T. Danino, O. Mondragón-Palomino, L. Tsimring, and J. Hasty. A synchronized quorum of genetic clocks. *Nature*, 463(7279):326–330, 2010.

[32] L. Delalande, D. Faure, A. Raffoux, S. Uroz, C. D'Angelo-Picard, M. Elasri, A. Carlier, R. Berruyer, A. Petit, P. Williams, and Y. Dessaux. N-hexanoyl-L-homoserine lactone, a mediator of bacterial quorum-sensing regulation, exhibits plant-dependent stability and may be inactivated by germinating Lotus corniculatus seedlings. *FEMS Microbiology Ecology*, 52(1):13–20, 2005.

[33] B. DiAndreth, N. Wauford, E. Hu, S. Palacios, and R. Weiss. PERSIST: A programmable RNA regulation platform using CRISPR endoRNases, 12 2019.

[34] Y.-H. Dong, L.-H. Wang, and L.-H. Zhang. Quorum-quenching microbial infections: mechanisms and implications. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 362(1483):1201–1211, 7 2007.

[35] Y.-h. Dong and L.-h. Zhang. Quorum sensing and quorum-quenching. *Journal of Microbiology*, 43(March 2005):101–109, 2016.

[36] B. Durand and M. Raff. A cell-intrinsic timer that operates during oligodendrocyte development, 1 2000.

[37] B. Eckert, A. Martin, J. Balbach, and F. X. Schmid. Prolyl isomerization as a molecular timer in phage infection. *Nature Structural and Molecular Biology*, 12(7):619–623, 7 2005.

[38] T. Ellis, X. Wang, and J. J. Collins. Diversity-based, model-guided construction of synthetic gene networks with predicted functions. *Nature Biotechnology*, 27(5):465–471, 5 2009.

[39] M. B. Elowitz and S. Leibier. A synthetic oscillatory network of transcriptional regulators. *Nature*, 403(6767):335–338, 1 2000.

[40] M. B. Elowitz and S. Leibler. A synthetic oscillatory network of transcriptional regulators. *Nature*, 403(6767):335–8, 1 2000.

[41] D. Endy. Foundations for engineering biology. *Nature*, 438(7067):449–53, 2005.

[42] D. Endy. Foundations for engineering biology, 11 2005.

[43] J. T. Enright. Temporal precision in circadian systems: A reliable neuronal clock from unreliable components? *Science*, 209(4464):1541–1545, 1980.

[44] C. M. Farrell, A. D. Grossman, and R. T. Sauer. Cytoplasmic degradation of ssrA-tagged proteins. *Molecular Microbiology*, 57(6):1750–1761, 2005.

[45] A. Fekete, C. Kuttler, M. Rothballer, B. A. Hense, D. Fischer, K. Buddrus-Schiemann, M. Lucio, J. Müller, P. Schmitt-Kopplin, and A. Hartmann. Dynamic regulation of N-acyl-homoserine lactone production and degradation in Pseudomonas putida IsoF. *FEMS Microbiology Ecology*, 72(1):22–34, 2010.

[46] J. Fernandez-Rodriguez, F. Moser, M. Song, and C. A. Voigt. Engineering RGB color vision into Escherichia coli. *Nature Chemical Biology*, 13(7):706–708, 7 2017.

[47] J. M. Flynn, I. Levchenko, M. Seidel, S. H. Wickner, R. T. Sauer, and T. A. Baker. Overlapping recognition determinants within the ssrA degradation tag allow modulation of proteolysis. *Proceedings of the National Academy of Sciences of the United States of America*, 98(19):10584–10589, 9 2001.

[48] G. Fullstone, T. L. Bauer, C. Guttà, M. Salvucci, J. H. Prehn, and M. Rehm. The apoptosome molecular timer synergises with XIAP to suppress apoptosis execution and contributes to prognosticating survival in colorectal cancer. *Cell Death and Differentiation*, 27(10):2828–2842, 10 2020.

[49] M. Gabalda-Sagarra, L. B. Carey, and J. Garcia-Ojalvo. Recurrence-based information processing in gene regulatory networks. *Chaos*, 28(10), 10 2018.

[50] H. Gaimster and D. Summers. Regulation of Indole Signalling during the Transition of E . coli from Exponential to Stationary Phase. *PloS one*, pages 4–5, 2015.

[51] M. Gao, H. Chen, A. Eberhard, M. R. Gronquist, J. B. Robinson, B. G. Rolfe, and W. D. Bauer. sinI- and expR-Dependent quorum sensing in Sinorhizobium meliloti. *Journal of Bacteriology*, 187(23):7931–7944, 12 2005.

[52] T. S. Gardner, C. R. Cantor, and J. J. Collins. Construction of a genetic toggle switch in Escherichia coli. *Nature*, 403(6767):339–342, 1 2000.

[53] O. Gefen, O. Fridman, I. Ronin, and N. Q. Balaban. Direct observation of single stationary-phase bacteria reveals a surprisingly long period of constant protein production activity. *Proceedings of the National Academy of Sciences of the United States of America*, 111(1):556–61, 2014.

[54] K. Gerecht, S. Margiola, and M. M. Müller. p53 Deamidation as a Molecular Timer for Cell Death. *Biophysical Journal*, 118(3):485a, 2 2020.

[55] D. G. Gibson, J. I. Glass, C. Lartigue, V. N. Noskov, R. Y. Chuang, M. A. Algire, G. A. Benders, M. G. Montague, L. Ma, M. M. Moodie, C. Merryman, S. Vashee, R. Krishnakumar, N. Assad-Garcia, C. Andrews-Pfannkoch, E. A. Denisova, L. Young, Z. N. Qi, T. H. Segall-Shapiro, C. H. Calvey, P. P. Parmar, C. A. Hutchison, H. O. Smith, and J. C. Venter. Creation of a bacterial cell controlled by a chemically synthesized genome. *Science*, 329(5987):52–56, 7 2010.

[56] D. G. Gibson, L. Young, R. Y. Chuang, J. C. Venter, C. A. Hutchison, and H. O. Smith. Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nature Methods*, 6(5):343–345, 4 2009.

[57] A. Goñi-Moreno, M. Amos, and F. de la Cruz. Multicellular Computing Using Conjugation for Wiring. *PLoS ONE*, 8(6):65986, 6 2013.

[58] A. Goñi-Moreno and P. I. Nikel. High-Performance Biocomputing in Synthetic Biology–Integrated Transcriptional and

Metabolic Circuits. *Frontiers in Bioengineering and Biotechnology*, 7(MAR):40, 3 2019.

[59] S. Gottesman. PROTEASES AND THEIR TARGETS IN ESCHERICHIA COLI. *Annual Review of Genetics*, 30(1):465–506, 12 1996.

[60] C. Grandclément, M. Tannières, S. Moréra, Y. Dessaux, and D. Faure. Quorum quenching: Role in nature and applied developments, 9 2015.

[61] L. Grozinger, M. Amos, T. E. Gorochowski, P. Carbonell, D. A. Oyarzún, R. Stoof, H. Fellermann, P. Zuliani, H. Tas, and A. Goñi-Moreno. Pathways to cellular supremacy in biocomputing. *Nature Communications*, 10(1):1–11, 2019.

[62] F. Guarnieri, M. Fliss, and C. Bancroft. Making DNA add. *Science*, 273(5272):220–223, 7 1996.

[63] S. Guiziou, V. Sauveplane, H. J. Chang, C. Clerté, N. Declerck, M. Jules, and J. Bonnet. A part toolbox to tune genetic expression in Bacillus subtilis. *Nucleic Acids Research*, 44(15):7495–7508, 9 2016.

[64] S. J. Hagen, P. D. Pérez, and J. T. Weiss. Noise and Crosstalk in the Two Quorum Sensing Channels of Vibrio Fischeri. *Biophysical Journal*, 100(3):165a, 2011.

[65] D. A. Hammer and N. P. Kamat. Towards an artificial cell. In *FEBS Letters*, volume 586, pages 2882–2890. No longer published by Elsevier, 8 2012.

[66] S. Hennig, G. Rödel, and K. Ostermann. Artificial cell-cell communication as an emerging tool in synthetic biology applications. *Journal of Biological Engineering*, 9(1):1–12, 2015.

[67] L. Hmelo and B. A. Van Mooy. Kinetic constraints on acylated homoserine lactone-based quorum sensing in marine environments. *Aquatic Microbial Ecology*, 54(2):127–133, 2009.

[68] W. S. Hou and L. Van Parijs. A Bcl-2-dependent molecular timer regulates the lifespan and immunogenicity of dendritic cells. *Nature Immunology*, 5(6):583–589, 6 2004.

[69] D. A. Jackson, R. H. Symons, and P. Berg. Biochemical method for inserting new genetic information into DNA of Simian Virus 40: circular SV40 DNA molecules containing lambda phage genes and the galactose operon of Escherichia coli. *Proceedings of the National Academy of Sciences of the United States of America*, 69(10):2904–2909, 1972.

[70] F. Jacob and J. Monod. Genetic regulatory mechanisms in the synthesis of proteins, 1961.

[71] P. Jayaraman, K. Devarajan, T. K. Chua, H. Zhang, E. Gunawan, and C. L. Poh. Blue light-mediated transcriptional activation and repression of gene expression in bacteria. *Nucleic Acids Research*, 44(14):6994–7005, 8 2016.

[72] A. R. Joyce, J. L. Reed, A. White, R. Edwards, A. Osterman, T. Baba, H. Mori, S. A. Lesely, B. Palsson, and S. Agarwalla. Experimental and computational assessment of conditionally essential genes in Escherichia coli. *Journal of Bacteriology*, 188(23):8259–8271, 12 2006.

[73] L. Kari, G. Gloor, and S. Yu. Using DNA to solve the Bounded Post Correspondence Problem. *Theoretical Computer Science*, 231(2):193–203, 1 2000.

[74] B. D. Karkaria, N. J. Treloar, C. P. Barnes, and A. J. Fedorec. From Microbial Communities to Distributed Computing Systems, 7 2020.

[75] G. F. Kaufmann, R. Sartorio, S. H. Lee, C. J. Rogers, M. M. Meijler, J. A. Moss, B. Clapham, A. P. Brogan, T. J. Dickerson, and K. D. Janda. Revisiting quorum sensing: Discovery of additional chemical and biological functions for 3-oxo-N-acylhomoserine lactones. *Proceedings of the National Academy of Sciences of the United States of America*, 102(2):309–314, 2005.

[76] K.-I. Kawasaki. Proof Without Words: Viviani’s Theorem. *Mathematics Magazine*, 78(3):213–213, 6 2005.

[77] W. Kong, V. Celik, C. Liao, Q. Hua, and T. Lu. Programming the group behaviors of bacterial communities with synthetic cellular communication, 12 2014.

[78] B. P. Kramer, A. U. Viretta, M. D. E. Baba, D. Aubel, W. Weber, and M. Fussenegger. An engineered epigenetic transgene switch in mammalian cells. *Nature Biotechnology*, 22(7):867–870, 7 2004.

[79] H. H. Lee, R. Kalhor, N. Goela, J. Bolot, and G. M. Church. Terminator-free template-independent enzymatic DNA synthesis for digital information storage. *Nature Communications*, 10(1), 12 2019.

[80] I. Levchenko, M. Seidel, R. T. Sauer, and T. A. Baker. A specificity-enhancing factor for the clpXP degradation machine. *Science*, 289(5488):2354–2356, 9 2000.

[81] J. H. Levine and M. B. Elowitz. Polyphasic feedback enables tunable cellular timers, 10 2014.

[82] J. H. Levine, M. E. Fontes, J. Dworkin, and M. B. Elowitz. Pulsed feedback defers cellular differentiation. *PLoS Biology*, 10(1), 1 2012.

[83] K. N. Lin, K. Volkel, J. M. Tuck, and A. J. Keung. Dynamic and scalable DNA-based information storage. *Nature Communications*, 11(1):1–12, 12 2020.

[84] K. P. Lu, G. Finn, T. H. Lee, and L. K. Nicholson. Prolyl cis-trans isomerization as a molecular timer, 10 2007.

[85] C. L. T. Lucas. The Causes of Evolution. By J. B. S. Haldane London: Longmans, Green & Co., Ltd., 1932. Pp. 235. Price 7s. 6d. . *Journal of Mental Science*, 79(327):771–772, 10 1933.

[86] J. Macía, F. Posas, and R. V. Solé. Distributed computation: The new wave of synthetic biology devices. *Trends in Biotechnology*, 30(6):342–349, 2012.

[87] J. Macia and R. Sole. How to make a synthetic multicellular computer. *PLoS ONE*, 9(2), 2014.

[88] J. Macia, B. Vidiella, and R. V. Solé. Synthetic associative learning in engineered multicellular consortia. *Journal of the Royal Society, Interface*, 14(129), 4 2017.

[89] T. Maire and H. Youk. Molecular-Level Tuning of Cellular Autonomy Controls the Collective Behaviors of Cell Populations. *Cell Systems*, 1(5):349–360, 2015.

[90] S. Malladi, M. Challa-Malladi, H. O. Fearnhead, and S. B. Bratton. The Apaf-1procaspase-9 apoptosome complex functions as a proteolytic-based molecular timer. *EMBO Journal*, 28(13):1916–1925, 7 2009.

[91] S. Mangan and U. Alon. Structure and function of the feed-forward loop network motif. *Proceedings of the National Academy of Sciences of the United States of America*, 100(21):11980–11985, 10 2003.

[92] S. Mangan, A. Zaslaver, and U. Alon. The coherent feedforward loop serves as a sign-sensitive delay element in transcription networks. *Journal of Molecular Biology*, 334(2):197–204, 11 2003.

[93] J. C. March and W. E. Bentley. Quorum sensing and bacterial cross-talk in biotechnology. *Current Opinion in Biotechnology*, 15(5):495–502, 2004.

[94] N. W. Maung and C. D. Smolke. Higher-order cellular information processing with synthetic RNA devices. *Science*, 322(5900):456–460, 10 2008.

[95] M. B. Miller and B. L. Bassler. Quorum Sensing in Bacteria. *Annual Review of Microbiology*, 55(1):165–199, 10 2001.

[96] R. Milo, S. Shen-Orr, S. Itzkovitz, N. Kashtan, D. Chklovskii, and U. Alon. Network motifs: Simple building blocks of complex networks. *Science*, 298(5594):824–827, 10 2002.

[97] T. S. Moon, C. Lou, A. Tamsir, B. C. Stanton, and C. A. Voigt. Genetic programs constructed from layered logic gates in single cells. *Nature*, 491(7423):249–253, 11 2012.

[98] S. Mukherjee and B. L. Bassler. Bacterial quorum sensing in complex and dynamically changing environments, 6 2019.

[99] K. Mullis, F. Faloona, S. Scharf, R. Saiki, G. Horn, and H. Erlich. Specific enzymatic amplification of DNA in vitro: The polymerase chain reaction. *Cold Spring Harbor Symposia on Quantitative Biology*, 51(1):263–273, 1986.

[100] N. Nandagopal and M. B. Elowitz. Synthetic biology: Integrated gene circuits, 9 2011.

[101] J. M. Navarro Llorens, A. Tormo, and E. Martínez-García. Stationary phase in gram-negative bacteria, 7 2010.

[102] K. H. Nealson, T. Platt, and J. W. Hastings. Cellular control of the synthesis and activity of the bacterial luminescent system. *Journal of Bacteriology*, 104(1):313–322, 10 1970.

[103] D. J. Nicholson. Organisms≠Machines. *Studies in History and Philosophy of Science Part C :Studies in History and Philosophy of Biological and Biomedical Sciences*, 44(4):669–678, 12 2013.

[104] D. J. Nicholson. The machine conception of the organism in development andevolution: A critical analysis. *Studies in History and Philosophy of Science Part C :Studies in History and*

Philosophy of Biological and Biomedical Sciences, 48(PB):162–174, 12 2014.

[105] D. J. Nicholson. Is the cell really a machine? *Journal of Theoretical Biology*, 477:108–126, 2019.

[106] E. M. Nikolados, A. Y. Weiße, and D. A. Oyarzún. Prediction of cellular burden with host-circuit models, 4 2020.

[107] A. Pai, Y. Tanouchi, C. H. Collins, and L. You. Engineering multicellular systems by cell-cell communication. *Current Opinion in Biotechnology*, 20(4):461–470, 2009.

[108] A. Pai and L. You. Optimal tuning of bacterial sensing potential. *Molecular Systems Biology*, 5(286):1–11, 2009.

[109] K. Papenfort and B. L. Bassler. Quorum sensing signal-response systems in Gram-negative bacteria. *Nature Reviews Microbiology*, 14(9):576–588, 2016.

[110] P. D. Pérez, J. T. Weiss, and S. J. Hagen. Noise and crosstalk in two quorum-sensing inputs of Vibrio fischeri. *BMC Systems Biology*, 5(1):153, 9 2011.

[111] C. Pietschke, C. Treitz, S. Forêt, A. Schultze, S. Künzel, A. Tholey, T. C. Bosch, and S. Fraune. Host modification of a bacterial quorum-sensing signal induces a phenotypic switch in bacterial symbionts. *Proceedings of the National Academy of Sciences of the United States of America*, 114(40):E8488–E8497, 2017.

[112] D. Pinto, S. Vecchione, H. Wu, M. Mauri, T. Mascher, and G. Fritz. Engineering orthogonal synthetic timer circuits based

on extracytoplasmic function factors. *Nucleic Acids Research*, 46(14):7450–7464, 8 2018.

[113] L. Potvin-Trottier, N. D. Lord, G. Vinnicombe, and J. Paulsson. Synchronous long-term oscillations in a synthetic gene circuit. *Nature*, 538(7626):514–517, 2016.

[114] A. Prindle, P. Samayoa, I. Razinkov, T. Danino, L. S. Tsimring, and J. Hasty. A sensing array of radically coupled genetic ’biopixels’. *Nature*, 481(7379):39–44, 1 2012.

[115] G. Păun. A quick introduction to membrane computing. *Journal of Logic and Algebraic Programming*, 79(6):291–294, 8 2010.

[116] G. Păun, M. J. Pérez-Jiménez, and G. Ştefănescu. Membrane computing and programming, 8 2010.

[117] M. Raff. Intracellular developmental timers. In *Cold Spring Harbor Symposia on Quantitative Biology*, volume 72, pages 431–435. Cold Spring Harb Symp Quant Biol, 2007.

[118] G. Rampioni, L. Leoni, and P. Williams. The art of antibacterial warfare: Deception through interference with quorum sensing-mediated communication. *Bioorganic Chemistry*, 55:60–68, 8 2014.

[119] S. Regot, J. Macia, N. Conde, K. Furukawa, J. Kjellén, T. Peeters, S. Hohmann, E. de Nadal, F. Posas, and R. Solé. Distributed biological computation with multicellular engineered networks. *Nature*, pages 1–5, 12 2010.

[120] K. Riedel, M. Hentzer, O. Geisenberger, B. Huber, A. Steidle, H. Wu, N. Høiby, M. Givskov, S. Molin, and L. Eberl. N-acylhomoserine-lactone-mediated communication between Pseudomonas aeruginosa and Burkholderia cepacia in mixed biofilms. *Microbiology*, 147(12):3249–3262, 2001.

[121] B. M. Ryback, D. I. Odoni, R. G. van Heck, Y. van Nuland, M. C. Hesselman, V. A. Martins dos Santos, M. W. van Passel, and F. Hugenholtz. Design and analysis of a tunable synchronized oscillator. *Journal of Biological Engineering*, 7(1):1–10, 2013.

[122] F. Sanger, S. Nicklen, and A. R. Coulson. DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Sciences of the United States of America*, 74(12):5463–5467, 1977.

[123] S. Sen, J. Garcia-Ojalvo, and M. B. Elowitz. Dynamical consequences of bandpass feedback loops in a bacterial phosphorelay. *PLoS ONE*, 6(9), 9 2011.

[124] S. S. Shen-Orr, R. Milo, S. Mangan, and U. Alon. Network motifs in the transcriptional regulation network of Escherichia coli. *Nature Genetics*, 31(1):64–68, 2002.

[125] R. P. Shetty, D. Endy, and T. F. Knight. Engineering BioBrick vectors from BioBrick parts. *Journal of Biological Engineering*, 2, 4 2008.

[126] K. P. Silva, P. Chellamuthu, and J. Q. Boedicker. Signal Destruction Tunes the Zone of Activation in Spatially Distributed Signaling Networks. *Biophysical Journal*, 112(5):1037–1044, 2017.

[127] R. V. Solé and J. Macia. Expanding the landscape of biological computation with synthetic multicellular consortia. *Natural Computing*, 12(4):485–497, 2013.

[128] R. V. Solé and J. Macia. Expanding the landscape of biological computation with synthetic multicellular consortia. *Natural Computing*, 12(4):485–497, 12 2013.

[129] J. Stricker, S. Cookson, M. R. Bennett, W. H. Mather, L. S. Tsimring, and J. Hasty. A fast, robust and tunable synthetic gene oscillator. *Nature*, 456(7221):516–519, 11 2008.

[130] L. Swint-Kruse and K. S. Matthews. lac Operon. In *Encyclopedia of Biological Chemistry: Second Edition*, pages 694–700. Elsevier Inc., 2 2013.

[131] I. Tagkopoulos, Y. C. Liu, and S. Tavazoie. Predictive behavior within microbial genetic networks. *Science*, 320(5881):1313–1317, 6 2008.

[132] A. Tamsir, J. J. Tabor, and C. a. Voigt. Robust multicellular computing using genetically encoded NOR gates and chemical ’wires’. *Nature*, 469(7329):212–5, 1 2011.

[133] S. J. Tekel, C. L. Smith, B. Lopez, A. Mani, C. Connot, X. Livingstone, and K. A. Haynes. Engineered orthogonal quorum sensing systems for synthetic gene regulation in escherichia coli. *Frontiers in Bioengineering and Biotechnology*, 7(MAR):1–12, 2019.

[134] C. Toscano-Ochoa and J. Garcia-Ojalvo. A tunable multicellular timer in bacterial consortia. 11 2020.

[135] M. Turcotte, J. Garcia-Ojalvo, and G. M. Süel. A genetic timer through noise-induced stabilization of an unstable state. *Proceedings of the National Academy of Sciences of the United States of America*, 105(41):15732–15737, 10 2008.

[136] A. Urrios, E. Gonzalez-Flo, D. Canadell, E. De Nadal, J. Macia, and F. Posas. Plug-and-Play Multicellular Circuits with Time-Dependent Dynamic Responses. *ACS Synthetic Biology*, 7(4):1095–1104, 2018.

[137] R. Van Houdt, A. Aertsen, P. Moons, K. Vanoirbeek, and C. W. Michiels. N-acyl-L-homoserine lactone signal interception by Escherichia coli. *FEMS Microbiology Letters*, 256(1):83–89, 2006.

[138] A. Vannini, C. Volpari, C. Gargioli, E. Muraglia, R. Cortese, R. De Francesco, P. Neddermann, and S. Di Marco. The crystal structure of the quorum sensing protein TraR bound to its autoinducer and target DNA. *EMBO Journal*, 21(17):4393–4401, 9 2002.

[139] V. Viviani. *De maximis et minimis, geometrica divinatio : in qvintvm Conicorvm Apollonii Pergaei*. 1659.

[140] B. Wang, R. I. Kitney, N. Joly, and M. Buck. Engineering modular and orthogonal genetic logic gates for robust digital-like synthetic biology. *Nature Communications*, 2(1), 2011.

[141] L. H. Wang, L. X. Weng, Y. H. Dong, and L. H. Zhang. Specificity and Enzyme Kinetics of the Quorum-quenching N-Acyl Homoserine Lactone Lactonase (AHL-lactonase). *Journal of Biological Chemistry*, 279(14):13645–13651, 2004.

[142] J. D. Watson and F. H. Crick. Molecular structure of nucleic acids: A structure for deoxyribose nucleic acid. *Nature*, 171(4356):737–738, 1953.

[143] M. Weber and J. Buceta. *Stochastic Effects in Quorum Sensing*. 2015.

[144] R. Weiss and S. Basu. The Device Physics of Cellular Logic Gates. Technical report.

[145] T. C. Williams, B. Peng, C. E. Vickers, and L. K. Nielsen. The Saccharomyces cerevisiae pheromone-response is a metabolically active stationary phase for bio-production. *Metabolic Engineering Communications*, 3:142–152, 12 2016.

[146] A. Wong, H. Wang, C. L. Poh, and R. I. Kitney. Layering genetic circuits to build a single cell, bacterial half adder. *BMC Biology*, 13(1):40, 6 2015.

[147] S. Wright. EVOLUTION IN MENDELIAN POPULATIONS. *Genetics*, 16(2), 1931.

[148] F. Wu, D. J. Menn, and X. Wang. Quorum-sensing crosstalk-driven synthetic circuits: From Unimodality to trimodality. *Chemistry and Biology*, 21(12):1629–1638, 2014.

[149] G. Wu, Q. Yan, J. A. Jones, Y. J. Tang, S. S. Fong, and M. A. Koffas. Metabolic Burden: Cornerstones in Synthetic Biology and Metabolic Engineering Applications. *Trends in Biotechnology*, 34(8):652–664, 2016.

[150] E. S. Yeh and A. R. Means. PIN1, the cell cycle and cancer, 5 2007.

[151] H. Youk and W. A. Lim. Secreting and sensing the same molecule allows cells to achieve versatile social behaviors. *Science*, 343(6171), 2 2014.

[152] J. Zeng, J. Teo, A. Banerjee, T. W. Chapman, J. Kim, and R. Sarpeshkar. A Synthetic Microbial Operational Amplifier. *ACS Synthetic Biology*, 7(9):2007–2013, 2018.

[153] L. H. Zhang and Y. H. Dong. Quorum sensing and signal interference: Diverse implications. *Molecular Microbiology*, 53(6):1563–1571, 2004.