UNIVERSITÀ DEGLI STUDI DI NAPOLI
FEDERICO II

UNIVERSITÀ DEGLI STUDI DI NAPOLI FEDERICO II

PH.D.THESIS IN

INFORMATION TECHNOLOGY AND ELECTRICAL ENGINEERING

FEEDBACK CONTROL OF GENE EXPRESSION IN MAMMALIAN CELLS

LORENA POSTIGLIONE

TUTOR: PROF. DIEGO DI BERNARDO

XXIX CICLO

SCUOLA POLITECNICA E DELLE SCIENZE DI BASE
DIPARTIMENTO DI INGEGNERIA ELETTRICA E TECNOLOGIE DELL'INFORMAZIONE

Abstract

Mammalian cells are dynamical systems. They detect, adapt and respond to time-varying inputs such as environmental cues, secreted molecules, and mechanical stimuli. These processes are controlled by networks of genes, proteins, small molecules, and their mutual interactions, the so-called gene regulatory networks, showing complex topologies. Understanding how these networks work is essential to identify triggering events both in common disease as well as in rare genetic disorders.

Control Theory makes available several tools that can be applied to explore the mechanisms driving gene networks.

Although recently several successful attempts to apply the Control Theory to steer gene expression from inducible promoters have been obtained in lesser eukaryotes, the application of control engineering to mammalian cells is still in its infancy because of complexity and the limited knowledge of their transcriptional networks.

This Thesis is involved in this challenging topic. I propose the study and the *in vitro* implementation of feedback control strategies based on microfluidic platform aimed to precisely regulate the level of expression of a protein from the tetracycline inducible promoter in mammalian cells. This control approach allows to express a protein of interest at different levels or in a time-varying fashion from the same promoter and can be a unique tool for several applications, including studying the effects of gene dosage in disease, probing the function of endogenous regulatory networks, and for synthetic biology applications.

Contents

List of Figures

CHAPTER 1

Introduction

1.1 Background and motivation

Mammalian cells are dynamical systems. They detect, adapt and respond to time-varying inputs such as environmental cues, secreted molecules, and mechanical stimuli. These processes are controlled by networks of genes, proteins, small molecules, and their mutual interactions, the so-called gene regulatory networks, showing complex topologies. Understanding how these networks work is essential to identify triggering events both in common disease as well as in rare genetic disorders.

Control Theory makes available several tools that can be apllied to explore the mechanisms driving gene networks [1, 2]. The application of Control Theory to biological systems has a long history. In 50-60s, researchers started to describe a biological function, i.e. the regulation of CO_2 by the respiratory system, in terms of a feedback-control system by drawing an analogy to an electric circuit [3, 4, 5, 6]. However, the study of molecular biology from a control theory perspective has been mostly theoretical for the last 60 years, mainly because of a lack of quantitative measurements of molecular processes. The advent of new technologies to quantitatively measure molecular species and manipulate cells has now enabled practical applications of control engineering to living systems.

Recently, several successful attempts to apply the Control Theory to steer gene expression from inducible promoters have been reported in the literature, but only in lesser eukaryotes [7, 8, 9, 10, 11, 12].

Indeed transcriptional networks in mammalian cells are very complex and their knowledge is very limited. A typical mammalian cell can express more than 10^4 genes at any given time, which are regulated by tens or hundreds of different transcription factors. For these reasons, the transcription in mammalian cells is up to 1 order of magnitude slower than in bacteria or simple eukaryotes, therefore, the time scale involved in the control of gene is expression is much longer [13]. Moreover, mammalian cells have much more stringent requirements for *in vitro* experiments. For all these reasons the application of control engineering to mammalian cells is still in its infancy.

My Thesis is involved in this challenging topic. I propose the study and the *in vitro* implementation of feedback control strategies based on microfluidic platform aimed to precisely regulate the level of expression of a protein from the tetracycline inducible promoter in mammalian cells. This control approach allows to express a protein of interest at different levels or in a time-varying fashion from the same promoter and can be a unique tool for several applications, including studying the effects of gene dosage in disease, probing the function of endogenous regulatory networks, and for synthetic biology applications.

1.2 Thesis outline

In **Chapter 2** I give an overview of the application of Control Theory in Molecular Biology introducing some biological concepts, such as gene expression and inducible promoter, and I present the negative feedback control strategies used in this work.

In **Chapter 3** I provide details of the experimental platform used for controlling of gene expression in a population of mammalian cells.

In **Chapter 4** I propose the design and the *in vitro* implementation of Relay and Proportional-Integral control strategies to regulate in real-time gene expression from endogenous promoters by providing a discrete control input.

In **Chapter 5** I describe the construction of an tetracycline inducible synthetic gene network exhibiting fast dynamics. I discuss the use of the experimental platform presented in Chapter 3 to infer quantitative dynamical models of transcriptional processes from measured input and output data. Moreover I present a Model Predictive Control strategy to regulate the gene expression of this tetracycline inducible system with a discrete control input.

In **Chapter 6** I propose the implementation of an active control strategy for regulating tetracycline concentration (or any other small molecule) within a microfluidic device.

In **Chapter 7** I present the implementation of Model Predictive Control strategy for gene expression with a continuous actuation by enabling any desired tetracycline concentration to be provided to the cells in the microfluidic device. I present the preliminary experimental results of the *in vitro* set point control experiment.

All this study has been completed in the Systems and Synthetic Biology Laboratory at TIGEM under the supervision of Dr. Diego di Bernardo (Principal Investigator). The experimental platform showed in Chapter 3 was implemented by Dr. Filippo Menolascina (Institute for Bioengineering, School of Engineering, The University of Edinburgh, Scotland, UK) and Dr. Gianfranco Fiore (Department of Engineering Mathematics, University of Bristol, UK) to control gene expression in yeast cell. In collaboration with Dr. Chiara Fracassi (LIFEWARE group INRIA, Palaiseau, France), I adapted the set-up for mammalian cells and I perform the pilot study to assess the feasibility of controlling gene expression from inducible promoter

in mammalian cells described in Chapter 4. The data and the discussion presented in Chapters 5, 6 and 7 have been generated by my own work.

CHAPTER 2

Molecular Biology and Control Theory

In this Chapter, I first introduce biological concepts such as gene expression and inducible system. I then propose a brief overview of current advances in the application of Control Theory to biological systems. Finally, I present the negative feedback control strategies that I implemented in my PhD work to control gene expression in mammalian cells.

2.1 Gene expression and inducible promoter

In living cells, the complex process by which genetic information gets ultimately transformed into working proteins is called *gene expression*. The main steps of gene expression are: a) the synthesis of RNA from a DNA template known as transcription and b) the protein synthesis from an RNA template known as translation; several intermediate editing steps usually take place as well. The sequences of the DNA that encode instructions for proteins' production are called genes. In the transcription process the enzyme RNA polymerase (RNAp) binds a defined site at the promoter, a regulatory region of DNA that precedes the gene, and produces mRNA that corresponds to gene's coding sequence. The mRNA is then translated into a protein [14]. A scheme of gene expression process is shown in Figure **2-1**.

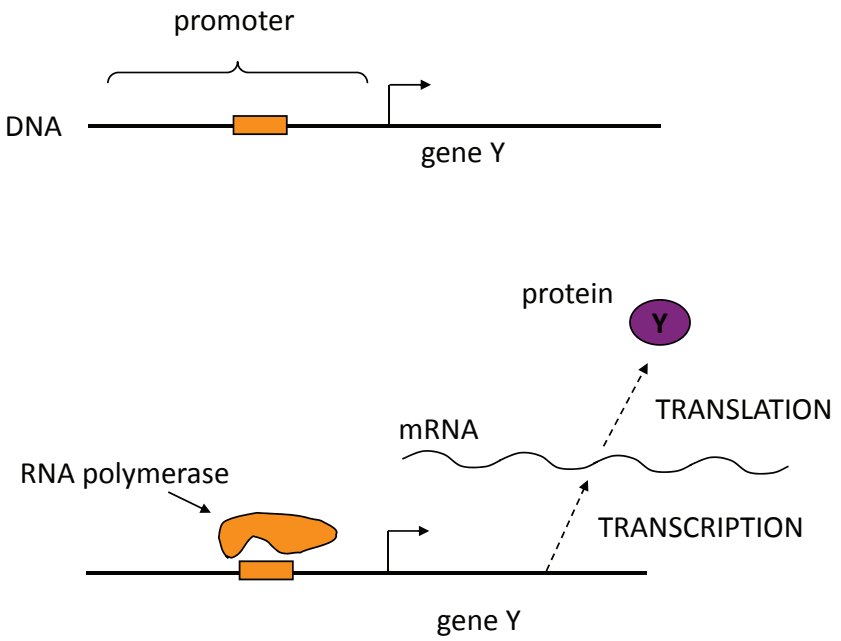


Figure **2-1**: **Scheme of gene expression process.** The main steps of gene expression are the transcritpion and tranlation processes. Transcription of a gene is the process by which RNAp (RNA polymerase) produces the mRNA corresponding to the gene coding sequence by binding a specific site of the promoter. The mRNA is then translated into a protein. Adopted from [15].

Gene expression however is a complex process regulated at several stages in the synthesis of proteins. Some proteins are structural and will accumulate at the cell-wall or within the cell to give it particular properties. Other proteins can be enzymes that catalyze certain reactions. A large group of proteins have an important role in the regulation of the genes, these are known as transcription factors [15].

Gene regulation by transcription factors can be negative or positive. In negative regulation, an inhibitor protein binds to the promoter and decreases the mRNA transcription of the gene. In positive regulation, a transcription factor is required to bind at the promoter in order to increase the mRNA transcription rate. Transcription factor proteins are themselves encoded by the genes, which are regulated by other transcription factors regulated yet

by other transcription factors, and so on. This set of interactions form a transcriptional network describing all of the regulatory transcriptional interaction in the cell [15].

The network, actually the cell, is a dynamical system: following an input signal (i.e. a molecule, a post-translational modification, a change in the physical properties of the environment), transcription factor activities change causing the variation in the production rate of proteins. The ability to apply precise inputs to the transcriptionals network in living cells is essential for interrogating and understanding complex cell signaling systems [16].

In molecular biology, the most common approach to induce a perturbation in the expression of a transcription factor in mammalian cells is the use of inducible promoters responding to small molecule. Typically, a small molecule is added to (or removed from) the cell growing medium and the expression of the transcription factor is induced.

The tetracycline-responsive promoter developed by Bujard and collegues in [17, 18] is a paradigm of this class of inducible systems.

The first critical feature of the tet-system is the regulatory protein. For example, in the Tet-OFF system shown in Figure **2-2**, the regulatory protein is the tetracycline-transactivator (tTA) protein. The tTa is a synthetic transcription factor created by fusing TetR with the C-terminal domain of VP16 [18]. The second critical component is the response plasmid which expresses a gene of interest under control of the *CMVTET* promoter which consists of seven *tet*-responsive operator motifs (*tetO7*) located just upstream of the minimal *CMV* promoter (P_{minCMV}). In the Tet-OFF system, tTA protein is able to bind the *CMVTET* thus driving the expression of the downstream gene of interest (e.g. the transcription factor to be perturbed) in absence of tetracycline (or doxycycline). Upon addition of tetracycline (or doxycycline), tTA is no longer able to bind the *CMVTET* promoter, thus preventing downstream gene expression (Figure **2-2**). Several variations of this inducible system exist, but all are based on the same principle [19].

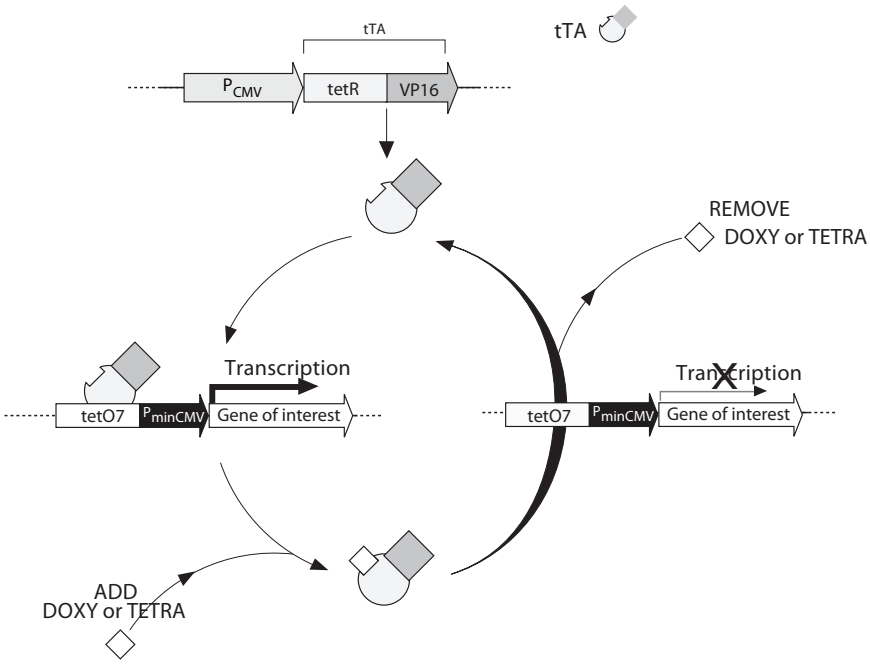


Figure **2-2**: **Scheme of gene regulation in the Tet-Off system.** The *tetO7* element is located upstream of the minimal *CMV* promoter (P_{minCMV}). tTA binds the *tetO7* and thereby activates transcription of gene of interest in the absence of tetracycline (or doxycycline). In presence of tetracycline, tTa is not able to bind the *CMVTET* promoter preventing the downstream gene expression. From `https://www.clontech.com/`

The limitation of these systems is that the downstream gene can be either fully expressed or repressed and it is very difficult to achieve intermediate expression levels by dosing the small molecule concentration, because of the highly nonlinear dose-response curve of the promoter. Hence, inducible promoters only allow "binary" expression levels. Moreover, there is no control on the expression dynamics. However, by applying feedback control strategies these limitations may be overcome and the expression level from the inducible promoter varied at will in a desired manner.

2.2 Approaches to control gene expression

In the last years, successful attempts to control gene expression in bacteria and yeast have been described in the literature. They mainly differ in the control input, the control strategy adopted and in the actuation strategies that are mainly based on optogenetics and microfluidics.

In 2011, Milias-Argeitis et al. exploited an optogenetics-based light inducible system to control gene expression from the $GAL1$ promoter in yeast cells [8]. By using light stimuli as control input, they developed an Model Predictive Control (MPC) strategy to achieve set-point regulation of the mean fluorescence of a reporter protein over the entire cell population measured via flow cytometry.

In 2014, Olson and collegues developed an optogenetic based method for programming tailor-made gene expression signals in live bacterial cells [20]. They designed an open-loop control algorithm that uses the dynamical models of light-switchable two-component systems to simulate the gene expression response and interatively optimizes the light control program until the error between the reference and the simulation is sufficiently small.

More recently, Milias-Argenteis and collegues presented a completely automatic experimental platform for the robust and precise long-term optical feedback control of gene expression driven by a light-switchable two-component system in liquid bacterial cultures [12]. They demonstrated the

ability of the optogenetic platform to satisfy set-point and tracking objectives when controlling gene expression with Proportional-Integral (PI) or MPC controllers and they performed a comparative analysis of this two control strategies.

On the other hand, microfluidic devices, allowing a tight control of cellular environment and the administration of inducer small-molecules, have been successfully exploited by Uhlendorf et al. to control the expression of a reporter protein from the Hog1-responsive promoter in yeast cells by using changes in osmotic pressure as control input [9]. Authors demonstrated that a Model Predictive Control strategy was effective in achieving for both time-constant and time-varying amount of fluorescence at the population and even the single-cell levels.

A completely automated microfluidic platform was developed to successful control in real-time gene expression in yeast cells from the *GAL1* inducible promoter as descibed in [21, 10, 11]. The authors demonstrated the ability of the platform to satisfy set-point and tracking objectives when controlling gene expression either with PI or MPC controllers, where the control inputs were galactose and glucose, while the output was the average fluorescence level of a reporter protein downstream of the GAL1 promoter across yeast cells.

Although control of gene expression in lesser eukaryotes has been successful achieved, the application of control engineering to mammalian systems is still in its infancy. The transcriptional dynamics of mammalian cells are slower than bacteria or yeast and therefore the time scale involved in the control of gene expression is much longer; moreover mammalian cells have much more stringent requirements for culture conditions. These constraints have hindered the application of automatic control to mammalian systems.

A pioneering work in mammalian cells was performed in 2011 by Toettcher and collegues who described an optogenetic approach to control the post-translational activation of the PI3K cascade, with a PI controller through automatic adjustment of the light input to single cells [16].

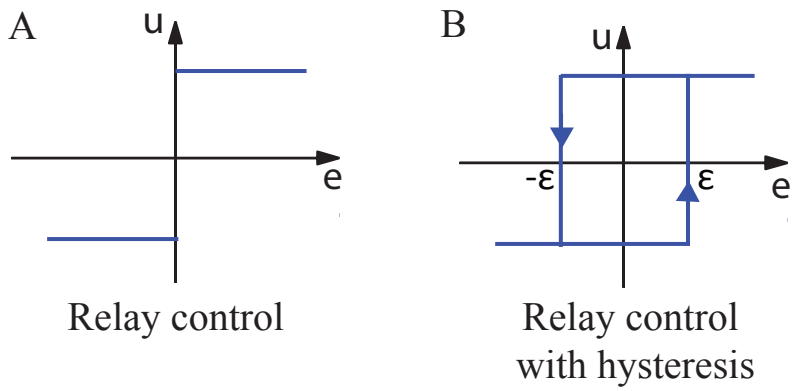


Figure **2-3**: **Relay controller.** Relay controller's characteristics with feedback error e on the x axis and the control input u on the y axis. The ideal controller is depicted in A while a version with hysteresis is shown in B.

However building on the results achieved in yeast cells by microfluidics [21, 10, 11], I decided to control gene expression from the tetracycline inducible promoter in mammalian cells with a microfluidics-based control strategy.

Next sections deal with the explanation of the control strategies I designed and implemented during my PhD work.

2.2.1 Relay

This simple control strategy can be expressed as follows:

$$u(t) = \begin{cases} u_{max} & if \quad e(t) > 0 \\ u_{min} & if \quad e(t) < 0 \end{cases} \tag{2-1}$$

where the control error $e(t) = r(t) - y(t)$ is the difference between the reference signal r and the system output y. u is the control input. This control strategy, although being simple, succeeds in keeping the system output close to the desired reference. Typically the controlled variable oscillates around the reference and it is acceptable if the oscillation is sufficiently small [22]. In addition, when the output value is very close to the reference, the control

error can rapidly change sign, thus causing the control input to continuously switch (chattering phenomenon [23]). To reduce the chattering, which can increase in presence of measurement noise, hysteresis ϵ can be added to the controller modifying the control law as follows:

$$u(t) = \begin{cases} u_{max} & if \quad e(t) \geq \epsilon \\ u_{min} & if \quad e(t) < -\epsilon \end{cases} \tag{2-2}$$

Thus in the Relay with hysteresis, the output depends on the value of past inputs. The drawback of this controller is that the amplitude of the oscillations around the set-point increases. The relation between error e and control input u is shown in Figure **2-3** for both relay controller and relay with hysteresis.

2.2.2 PID Control

Proportional-Integral-Derivative (PID) control is by far the most common controller in industry. Moreover there are also evidence that PID control is encoded in biological systems [24]. PID controllers appear in many different forms: as stand-alone controllers, as part of hierarchical, distributed control systems and built into embedded components. Block diagrams of PID controller is shown in Figure **2-4**.

The control action for a PID controller is described as:

$$u(t) = k_p e(t) + k_i \int_0^t e(\tau) d\tau + k_d \frac{de(t)}{dt} \tag{2-3}$$

where u is the control input and $e = r - y$ is the control error. y is the measured process variable and r is the reference. The control action is thus a sum of three terms: the proportional term which is proportional to the error, the integral term which is proportional to the integral of the error, and the derivative term which is proportional to the derivative of the error. The controller parameters are the proportional gain k_p, the integral gain k_i and the derivative gain k_d. The integral, proportional and derivative term can be interpreted as control actions based on the past, the present and the future represented by a linear extrapolation of the error (Figure **2-5**).

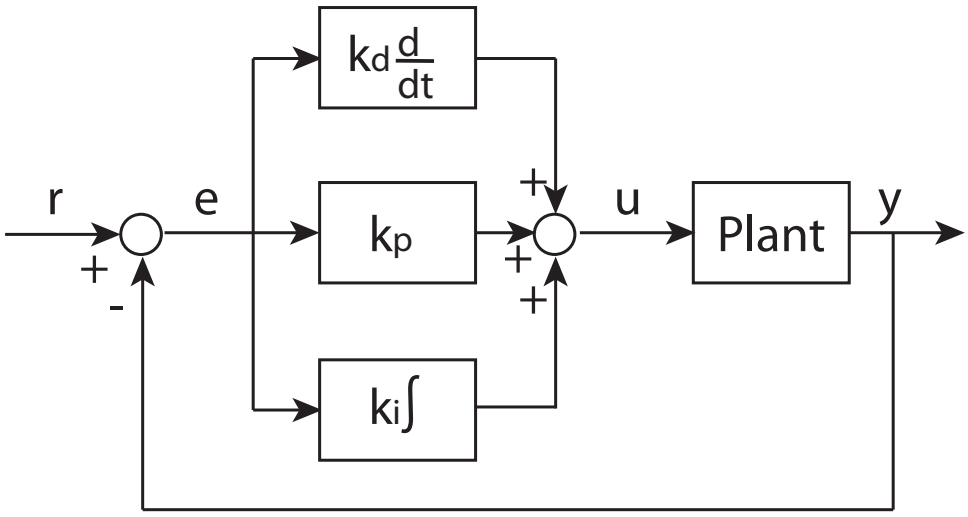


Figure **2-4**: **Block diagram of PID controller.** The controller has one output y, the control signal is u. For this controller, proportional, integral and derivative action acts on the control error $e = r - y$.

The action of the different terms can be illustrated by considering the response of the system output to a unit step as reference. As shown in Figure **2-6** A, by considering pure proportional control the output never reaches the reference and thus a non zero steady state error is present. The error decreases with increasing proportional gain k_p, but the system tendency towards oscillation increases. Figure **2-6** B illustrates how the zero steady state error is achieved when the integral term is added to the control action. With integral action the rise time is increased. The system is speeded up by increasing the integral gain k_i, but the closed loop system is likely to escape the stability region. Finally, Figure **2-6** C illustrates the effect of derivative action: the system is oscillatory when no derivative action is used, and it becomes more damped as the derivative gain is increased. Performance deteriorates if the derivative gain is too high. The derivative term not does

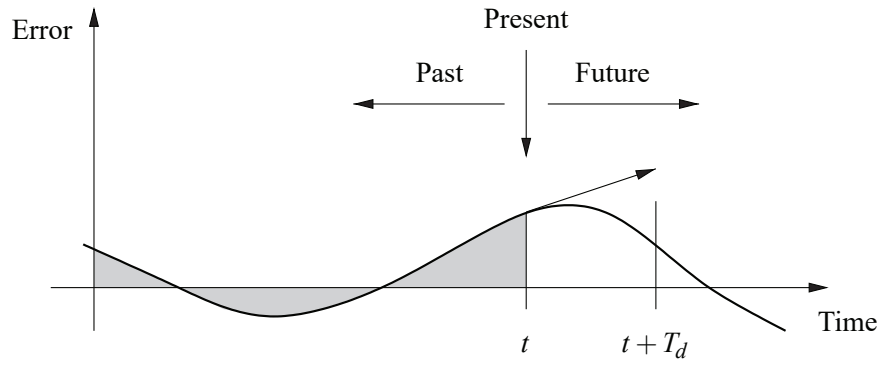


Figure **2-5**: **Action of a PID controller.** At time t, the proportional term $k_p e(t)$ depends on the instantaneous value of the error. The integral portion of the feedback $k_i \int_0^t e(\tau)d\tau$ is based on the integral of the error up to time t (shaded portion). The derivative term $k_d \frac{de(t)}{dt}$ provides an estimate of the growth or decay of the error over time by looking at the rate of change of the error. Adapted from [22].

not have any effect on the steady state error.

In the majority of practical applications of PID control, the derivative action is not used because of its variable impact on system stability in real-world applications and also because in the presence of noisy measurements as the control action would become noisy due to the high pass filter nature of the derivative term [25].

Integrator Windup In practical applications, the control action is fed to the plant by the actuator system that tipically has limitations (a motor has limited speed, a valve cannot be more than fully opened or fully closed, etc). For a control system with a wide range of operating conditions, it may happen that the control action reaches the actuator limits. When this happens the feedback loop is broken and the system runs as an open loop because the actuator will remain at its limit independently of the process output. If a controller with integral action is used, the error will continue to be integrated meaning that the integral term and the control output may become very large. The control signal will then remain saturated even when the error changes, and it may take a long time before the integrator and

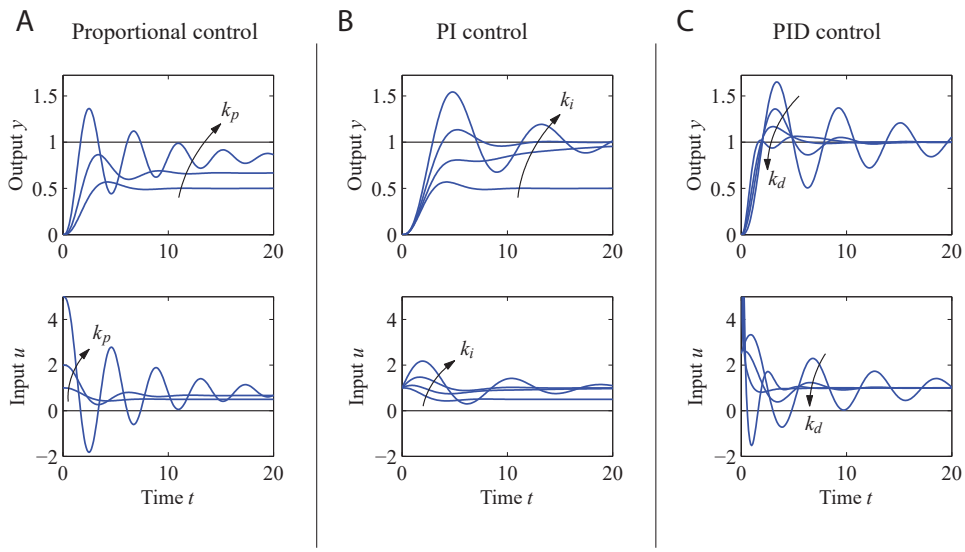


Figure **2-6**: **Effect of the gains of PID controller gain.** Responses to step changes in the reference value for a system with a proportional controller (A), PI controller (B) and PID controller (C). Adapted from [22].

the controller output come inside the saturation range. This phenomenon is called *integrator windup* [22].

To avoid the effects due of windup, a possible solution is to modify the PI scheme as shown in Figure **2-7** (see also [22]). The system has an extra feedback path that is generated by measuring the actual actuator output u and computing an error signal e_s as the difference between the the actuator output u and output of the controller $\hat{u}$. The signal e_s is added to the integrator input trough k_t. The signal e_s is zero when there is no saturation and thus it not has any effect on the normal operation when the actuator does not saturate. When the actuator saturates, the signal e_s is fed back to the integrator implying that that controller output $\hat{u}$ is kept close to the saturation limit. The controller output will then change as soon as the error changes sign and integral windup is avoided.

PI control with the anti-windup scheme shown in Figure **2-7** will be used in this Thesis to control gene expression of CHO *tetO7*-d2EYFP cells (see Chapter 4)

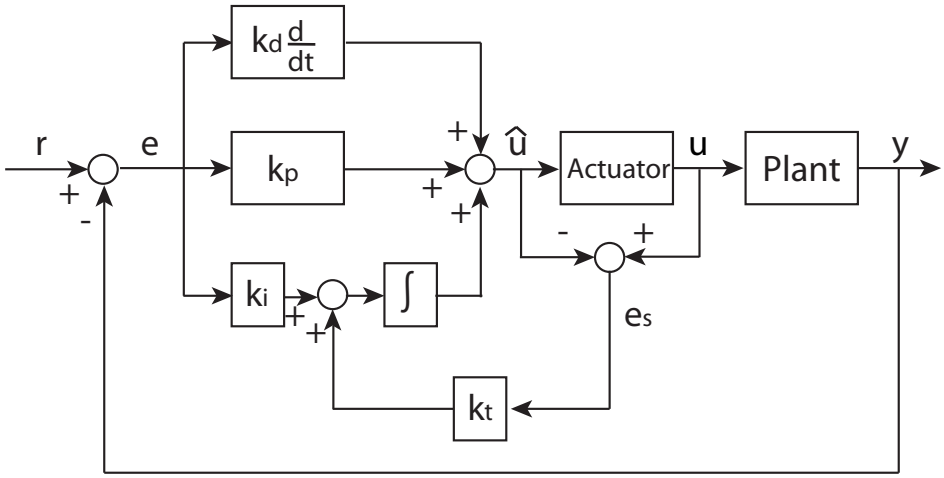


Figure **2-7**: **Block diagram of PID controller with anti-windup.** The input to the integrator consists of the error term e plus a "reset" e_s based on input saturation. If the actuator is not saturated, then $e_s = u\hat{u}$, otherwise e_s will decrease the integrator input to prevent windup.

The first advantage of PI controller is that, besides its very simple implementation, it guarantee zero steady-state error for constant reference and the rejection of constant disturbances at steady-state. Moreover, the PI controller does not require a model of the controlled system, although an idea of its dynamics is necessary for gains' tuning. On the other hand, a PI controller does not achieve a satisfactory performance for tracking time-varying references unless the reference dynamics are much slower than the closed-loop system dynamics.

These limitation can be addressed by an MPC controller even if the improved performance has a computational cost, since MPC is more complex than PI control [26].

2.2.3 Model Predictive Control

Model Predictive Control (MPC) is a well-established technique for controlling multivariable systems subject to constraints on manipulated variables and outputs in an optimized way. Following a long history of success in the process industries, in recent years MPC is rapidly expanding in several other domains, such as in the automotive and aerospace industries, smart energy grids, and financial engineering [27, 28, 29, 30].

In recent years, the MPC approach has been also used to control biological process [8, 12, 9, 11].

The conceptual structure of MPC is depicted in Figures **2-8** and **2-9** . The MPC arises from the idea of employing an explicit model of the plant to be controlled which is used to predict the future output behavior.

Specifically, based on the measurements obtained at time kT, the controller predicts the dynamic behaviour of the plant over a prediction horizon N in the future and solves optimal control problems on line (i.e. determines the optimal input sequence), where tracking error, namely the difference between the predicted output and the desired reference, is minimized, subject to constraints on the manipulated inputs and outputs.

To incorporate feedback and thus to make the MPC action robust to disturbance and model-plant mismatch, the result of the optimization is applied according to a *receding horizon* philosophy: at time kT only the first input of the optimal command sequence is actually applied to the plant until the next sampling instant. The remaining optimal inputs are discarded, and a new optimal control problem is solved at time $(k+1)T$. As new measurements are collected, using the new system state at time $(k+1)T$, the whole procedure (prediction and optimization) is repeated, moving the control and prediction horizon forward [31, 32].

The process model plays, in consequence, a decisive role in the controller. The chosen model must be able to capture the process dynamics to precisely predict the future outputs and be simple to implement and understand. Moreover the system model used to predict the future system behaviour at each sampling time, is initialized by the actual system state. Since all state

information is necessary for the prediction, the full state must be either measured or estimated.

2.2.4 Performance indices

To measure the performance of the control strategies adopted, I will consider the following metrics defined as integral expressions related to the control error $e = r - y$.

1. The Integral of Squared Error (ISE)

$$ISE = \int_0^t e(\tau)^2 \, d\tau \tag{2-4}$$

   ISE integrates the square of the error over time. ISE penalizes large errors more than smaller ones.

2. The Integral Absolute Error (IAE)

$$IAE = \int_0^t |e(\tau)| \, d\tau \tag{2-5}$$

   IAE integrates the absolute error over time. It doesn't add weight to any of the errors in a systems response.

3. The Integral Time-weighted Absolute Error (ITAE)

$$ITAE = \int_0^t \tau|e(\tau)| \, d\tau \tag{2-6}$$

   ITAE integrates the absolute error multiplied by the time over time. It is a weighted version of IAE. It weights errors which exist after a long time much more heavily than those at the start of the response.

2.3 System Identification

The design of a Model Predictive Control strategy requires a mathematical model of the dynamical system to be controlled.

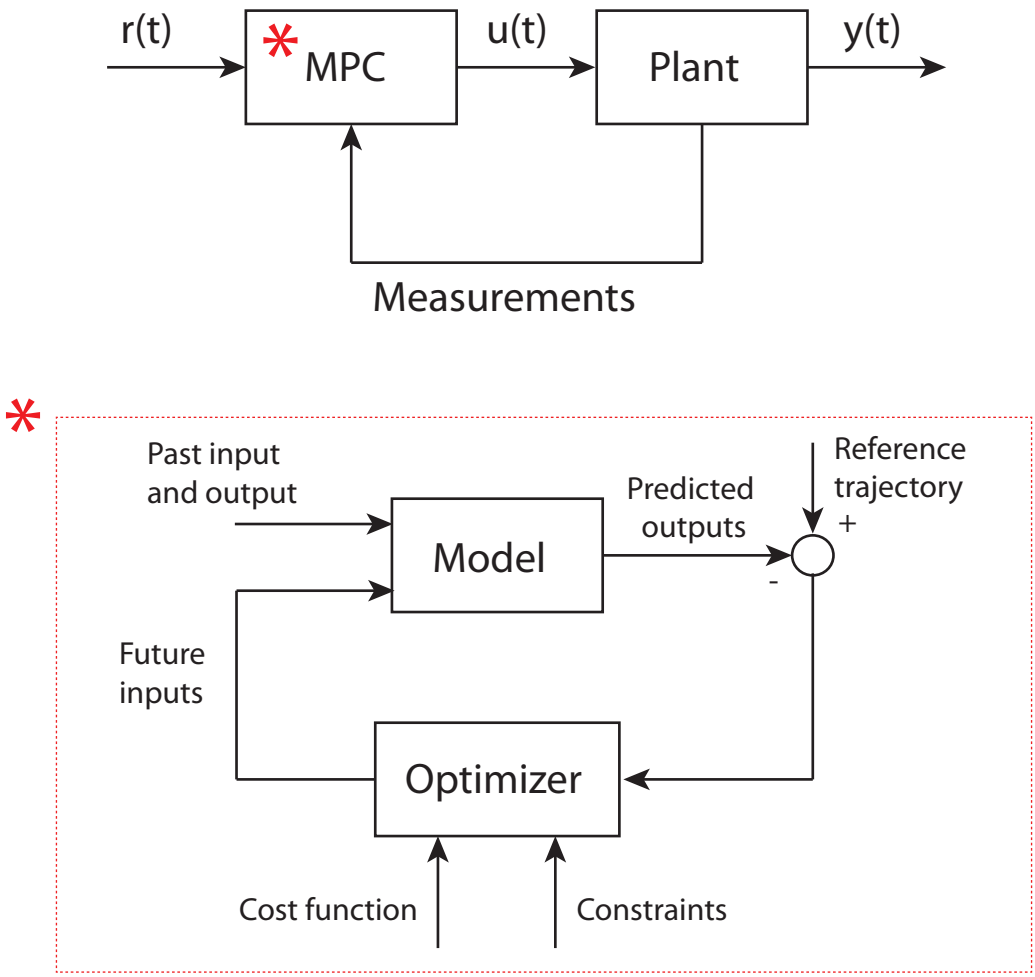


Figure **2-8**: **Block diagram of Model Predictive Control.** The MPC is based on the idea of employing an explicit model of the plant to be controlled which is used to predict the future output behavior. This prediction capability allows solving optimal control problems on line, where tracking error, namely the difference between the predicted output and the desired reference, is minimized over a future horizon, possibly subject to constraints on the manipulated inputs and outputs.

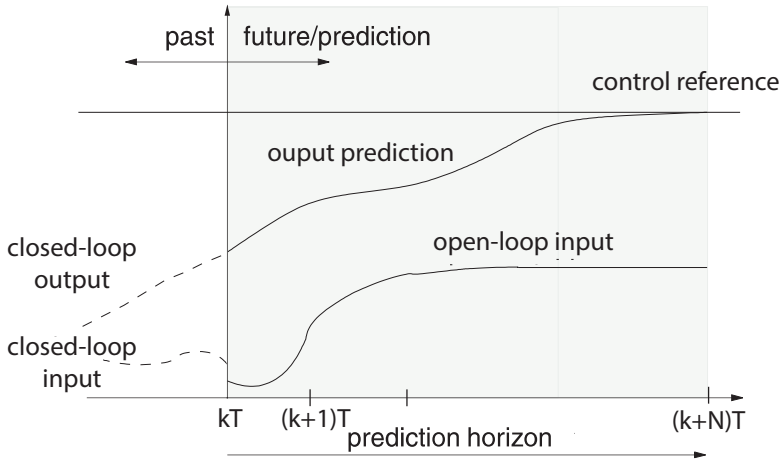


Figure **2-9**: **Principle of Model Predictive Control.** Starting from the measurements obtained at time kT, the controller predicts the dynamic behaviour of the plant over a prediction horizon N in the future and determines the input such that the difference between the predicted output and the desired reference is minimized.

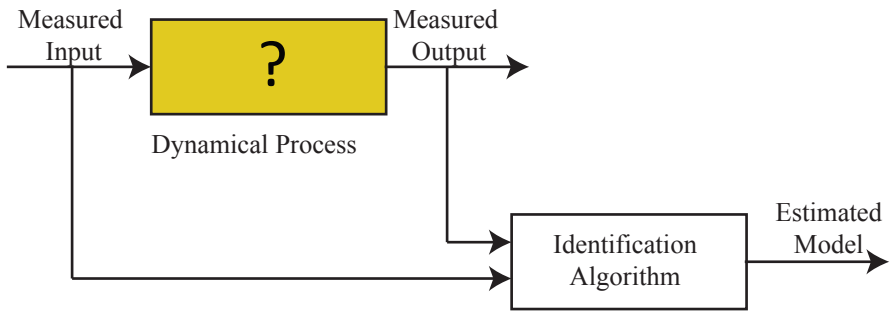


Figure **2-10**: **System identification.** System identification is a procedure to build a mathematical model of the dynamics of a system from measured input-output data.

System Identification deals with the problem of building mathematical models of dynamical systems based on the measured input-output data (Figure **2-10**.

The identification of a dynamical model from data can be partitioned into the following steps:

- **Collecting the data**. This is a very crucial (and most expensive) step. The input-output data are recorded during a specifically designed identification experiment. The purpose of the design of identification experiments is to collect maximally informative input-output data by choosing the signal to be measured, the sampling time at which carry out the measurement and the inputs to be apply to the system that has to be particularly suitable for identification.

- **Selecting a model set and choosing a selection criterion.** A specific model can be selected, on the basis of the observations, inside a family of models (e.g. linear or non linear, deterministic or stochastic, SISO or MIMO, continuous or discrete). These *a priori* assumptions on the set of the admissible models can be guided by the user's experience and improper past results. However any model can approximate the true system behaviour, but the selection criterion is determined by the planned use of the model in order to optimize its desired features. From the same data can be obtained several models that can be carried out following two major approaches: *black-box* and *gray-box*. In *black-box* mode the model is completely derived from input-output data and the model parameters are viewed as vehicles for adjusting the fit to the data and do not reflect physical considerations about the system. In the *gray-box* approach, first principles (e.g. Newton's law) and considerations on the physics of the system are used to partially derive the model structure, while parameters or terms in the model are determined by measurement data.

- **Determining the "best" model in the set, guided by data.** The choose of the "best" model in the considered class is typically

based on how the model perform when it attempt to reproduce the measured data. The basic approaches to this, like the Prediction-Error Minimization method and Least Square method [33], can be used independently of the model structure used.

- **Validating the model.** The model validation involves various procedures (like Final Prediction Error and Akaike Information Criterion [33]) to assess how the model relates to prior knowledge, to its intended used and to observed data, better by using collection data different from those used for identification. If the performance of the model is not good enough, the model has to be reject; while it is considered validated if it shows a good performance.

These choices are never univocal and can also involve some *a priori* knowledge of the system, for instance in the design of the experiments required to collect input-output data. When the model does not pass the final validation step , it is necessary to reconsider the choices and to restart the identification procedure from the beginning or from an intermediate step [34, 33, 35].

In my PhD work, I used System Identification approaches to identify dynamical models of gene expression in order to implement MPC strategies.

I considered the *grey-box* approach with a predefined model structure with unknown parameters to estimate linear and non linear SISO models in state space form describing the dynamics of a Tet-OFF system in mammalian cells. For the estimation of model parameters, I considered Least Squares criterion that minimizes the sum of squared residuals. The residual for the i-th data point r_i is defined as the difference between the observed response value y_i and the fitted response value $\hat{y}_i$ provided by the model ($r_i = y_i - \hat{y}_i$). The identification was implemented in the MATLAB System Idenfication toolbox (Mathworks Matlab R2016b) with the functions *greyest* for the linear model and *nlgreyest* for the non linear model on the collected input-output. The identification results will be shown in Sections 5.2 and 7.1 [33].

Moreover, I used the *black-box* identification approach to derive a mathematical model in the state space form for the accumulation dynamics of a small molecule in a microfluidic device (see Section 6.2). The model parameters were estimated by applying Prediction Error Minimization algorithm implemented in the MATLAB System Idenfication toolbox (Mathworks Matlab R2016b) with the function *pem* on the collected data [33, 36].

The input-output data were collected for dynamics identification of both Tet-OFF system and the accumulation of a small molecule by using the microfluidic-based platform that will be described in Chapter 3.

2.4 Kalman Filter

During *in-vitro* experiments it is only possible to measure directly system output and not the states that are instead needed by MPC algorithm to calculate the control input.

A standard approach for state estimation is to use a Kalman filter [37]. Using a LTI model, the past control inputs, and measurements, it gives an estimate of the state of a system together with its uncertainty.

In this work the systems under investigation will be described by single input-single output (SISO) discrete time model and the Kalman filter addresses the problem of estimating the process state by considering the following model:

$$\begin{aligned} \mathbf{x}(k+1) &= \mathbf{A}\mathbf{x}(k) + \mathbf{B}u(k) + \mathbf{v}(k) \\ y(k) &= \mathbf{C}\mathbf{x} + w(k) \end{aligned} \tag{2-7}$$

where $\mathbf{x}(k) \in \mathbb{R}^n$ is the system state, $u(k) \in \mathbb{R}$ is the system input and $y(k) \in \mathbb{R}$ is the output (measurement) of the system. $\mathbf{A} \in \mathbb{R}^{nxn}$ is the dynamic (or state) matrix, $\mathbf{B} \in \mathbb{R}^{nx1}$ and $\mathbf{C} \in \mathbb{R}^{1xn}$ are the input and output matrices. The $\mathbf{v}(k) \in \mathbb{R}^n$ and $w(k) \in \mathbb{R}$ are the process noise and the measurement noise having normal distributions $N(0, \mathbf{Q})$ nad $N(0, R)$, respectively.

At each instant k, by using a form of feedback control shown in the Figure **2-11**, the Kalman filter estimates the system state at some time and

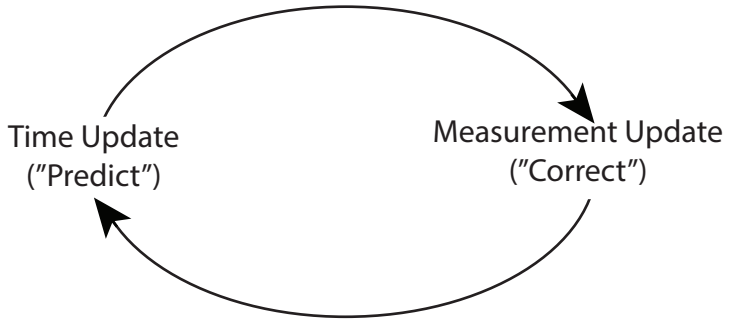


Figure **2-11**: **Discrete Kalman filter cycle.** The *time update* projects the current state estimate ahead in time. The *measurement update* adjusts the projected estimate by an actual measurement at that time.

then obtains feedback in the form of noisy measurements [38]. The equations for the Kalman filter are: *time update equations* and *measurement update equations*. The time update equations are responsible for projecting forward (in time) the current state and error covariance estimates to obtain the a priori estimates for the next time step, thus the time update equations are like predictor equations. The measurement update equations are responsible for incorporating a new measurement into the a priori estimate to obtain an improved a posteriori estimate (feedback), thus the measurement update equations can be considered corrector equations. Let $\hat{\mathbf{x}}(k)$ the state estimate at instant k, the obtained estimate error $\mathbf{e}(k) = \mathbf{x}(k) - \hat{\mathbf{x}}(k)$ and estimate error covariance $\mathbf{P}\left[\mathbf{e}(k), \mathbf{e}^T(k)\right]$, at each instant k, the specific equations for the time update are

$$\hat{\mathbf{x}}(k+1|k) = \mathbf{A}\hat{\mathbf{x}}(k|k) + \mathbf{B}u(k) \tag{2-8}$$

$$\mathbf{P}_p(k+1) = \mathbf{A}\mathbf{P}(k)\mathbf{A}^T + \mathbf{Q} \tag{2-9}$$

The Equations (2-8) and (2-9) evaluate the optimal prediction of the state $\hat{\mathbf{x}}(k+1|k)$ and the prediction error covariance $\mathbf{P}_p(k+1)$. The specific equations for the measurements update are

$$\hat{\mathbf{K}}(k+1) = \mathbf{P}_p(k+1)\mathbf{C}^T(\mathbf{C}\mathbf{P}_p(k+1)\mathbf{C}^T + R)^{-}1 \tag{2-10}$$

$$\hat{\mathbf{x}}(k+1|k+1) = \hat{\mathbf{x}}(k+1|k) + \mathbf{K}(k+1)(y(k+1) - \mathbf{C}\hat{\mathbf{x}}(k+1|k)) \tag{2-11}$$

$$\mathbf{P}(k+1) = [\mathbf{I} - \mathbf{K}(k+1)\mathbf{C}]\,\mathbf{P}_p(k+1) \tag{2-12}$$

The first task during the measurement update is to compute the Kalman gain $\mathbf{K}(k+1)$ (Equation (2-10)); the second step (Equation (2-11)) computes the optimal state estimate $\hat{\mathbf{x}}(k+1|k+1)$; the final step is to obtain the error covariance estimate $\mathbf{P}(k+1)$ via Equation 2-12.

In the implementation of the filter, the tuning of $\mathbf{Q}$ and R parameters are usually off-line. In this work this parameters are chosen empirically.

CHAPTER 3

Technological platform for controlling living cells

In this Chapter, I present the experimental platform for the external control of gene expression in a population of living cells. The requirements that the entire setup have to satisfy are: to guarantee the physiological conditions for the proper cells growth, to administer an external input and to monitor in real-time the desired output. To fit all these needs, I took advantage of a closed loop experimental platform based on a microfluidic device, a time-lapse microscopy apparatus and a set of automated syringes all controlled by a computer (Figure 4.1). Microfluidics allows to grow cells in living conditions and to precisely change their environmental conditions in real-time; the medium flowing in the microfludic device is modified by rising or lowering the height of the two motorized syringes attached to the inlet ports of the microfluidic device. Cells in the device can be imaged with the microscope at a given sampling time, in order to quantify protein expression in real-time. Thus from a Control Theory point of view, in the closed loop experimental platform the cells in the microfluidic device represent the plant to be controlled, the motorized syringes are the actuators and the sensor is represented by the microscopy unit.

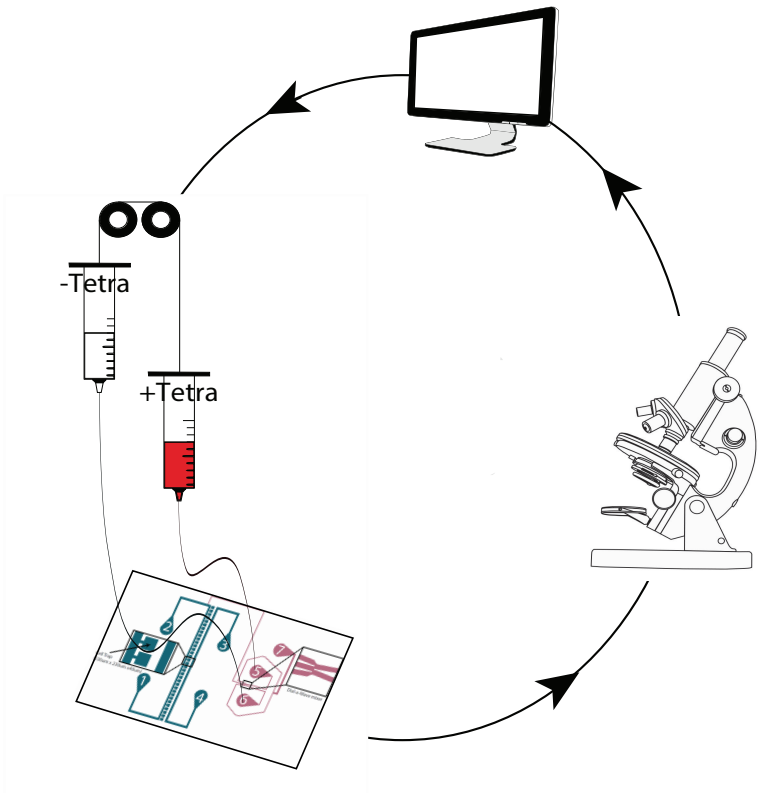


Figure **3-1**: **Technological platform to control mammalian cells.** Both the main components of the experimental platform, as well as the information flow, are highlighted in this scheme.

3.1 Microfluidics

Microfluidic cell culture technology has become a powerful tool in biological research by enabling the study of living cells in a precisely controlled microenvironment. Microfluidic devices have been used for everything from miniaturization of molecular biology reactions to platforms for cell growth and analysis [39, 40, 7, 41, 42, 43, 44, 45] .

A driving factor for increased use of microfluidics is the potential for more productive experiments, that is, accomplishing the same or more using

fewer resources (primarily less reagents, consumables, and time). Furthermore, microfluidic devices offer the unrivaled ability to precisely control and perturb the environment of single cells while capturing their behavior using high resolution microscopy. The microfluidic devices are, essentially, chips (mostly in PDMS), where fluid dynamics at the microliter scale are exploited. The principle is to have an area where the cells are forced to be in (cell trap), together with a series of channels to provide one or more compounds to the trap in order to regulate cells environment or to collect cells and fluids wastes.

While microbial microfluidic platforms have been successfully augmented and used in several applications, mammalian studies have been slower to adopt this approach due to the inherent complexities associated with mammalian microfluidic cell culture. The difficulties primarily arise because mammalian cells are much more sensitive to small changes in pH, osmolarity, shear stress and other external factors than most of the microbial model organisms. However, in [46] the authors developed a microfluidic device which can rapidly load a high density of cells into individual trapping regions that are extremely isolated from the shear stress effects of the main perfusion channel. They demonstrated the use of this platform in observing long-term growth of cell populations and in probing the response of a fluorescent reporter cell line to dynamic stimulation.

In this Thesys I thus decided to use the microfluidic device designed by Kolnik and colleagues [46] shown in Figure **3-2**. The device consists of 33 individual cuboid culture chambers (each has a 230 mm by 230 mm footprint, 40 mm height) adjoined to a main perfusion channel of identical height via a 50 mm wide opening on one side of each chamber. A separate channel for application of a temporary vacuum runs parallel to the column of culture chambers at a distance of 160 mm between the wall of this channel and the closest wall of each cuboid chamber. In this manner a cell suspension can flow into the main channel and upon the application of a temporary vacuum at gas-permeable PDMS interface, fluid containing cells is drawn into the culture chambers. Once captured in the traps, the cells are effectively

shielded from the fluid flow in the main channel and remain undisturbed even at high flow rates (Figure **3-2** B).

A fluidic mixer network named the Dial-a-Wave (DAW) [47], joins to the main perfusion channel and serves to deliver any desired waveform of biochemical inducer for dynamic stimulation of cells inside the culture chambers. The working principle of the DAW mixer is shown in Figure **3-3** and it is to precisely combine two incoming fluid streams (inducer and standard medium) in any desired ratio by adjusting the hydrostatic pressure at the corresponding inlets in order to increase the pressure at one inlet while decreasing it at the other by exactly the same amount so that the flow rate out of the junction remains constant. The DAW output signal is directed to the cells via a channel which contains staggered chaotic mixers [48] to enhance diffusive mixing of the combined laminar flow streams before the fluid reaches the main perfusion channel and the adjoining culture chambers.

For further details on the design, fabrication and cells loading refer to Appendix A and [46].

3.2 Actuator

As mentioned previously, the DAW is designed to combine the inputs from ports 6 and 7 of the microfluidic device in a precise ratio depending on the inlet hydrostatic pressures. To set the mixing ratio, the pressure of one inlet is increased and the other decreased by the same amount. By changing the inlet pressures in an opposing manner, the flow rate out of the junction remains constant and hence the flow to the cells is not altered (Figure **3-3** B).

By controlling the inlet pressures as a function of time, one can generate precise waves of inducer concentration reaching the cell trap. Therefore, the actuation aim is to establish this pressure difference.

To accomplish this, I used vertically two mounted linear actuators to physically move the filled syringes up and down, thereby altering their hydrostatic pressures. The automated syringes are shown in Figure **3-4**. Ob-

serve that the syringes in this design have no plunger, since they are used simply as tanks to store the media. Namely, each syringe will exert a hydrostatic force over the fluids in the channel that is simply proportional to its height.

The actuation system comprises two linear guides; every linear actuator is designed to move independently from the other; the motion is realised through a stepper motor, while the transmission by using a timing belt and two pulley. More details on the sizing and construction of the actuator can be found in [49, 50].

3.3 Microscopy and Image analysis

To monitor cellular processes dynamics, I have taken advantage of an inverted fluorescence microscope (Nikon Eclipse Ti), shown in Figure **3-4**, equipped with an automated and programmable stage, an incubator to guarantee fixed temperature and gasses to cell environment and a high sensitivity Electron Multiplying CCD (EMCCD) Camera (Andor iXON Ultra897). The microscope and the camera can be programmed to acquire, at regular time intervals, images from different points of the microfluidic device. At each time point the microscope is programmed to acquire two types of images: a bright field image (phase contrast image) and a fluorescence image.

Once cells have been imaged, image analysis algorithms can be applied to estimate their fluorescence. To this end, I used an image processing algorithm that is able to locate cells within each phase contrast image thus identifying all the pixels belonging to cells. This information is used to calculate the fluorescence expressed, for each fluorescent reporter, by the cells. Additional informations on the microscope set-up for real time image acquisition and on image analysis algorithm are available in Appendix A and [50].

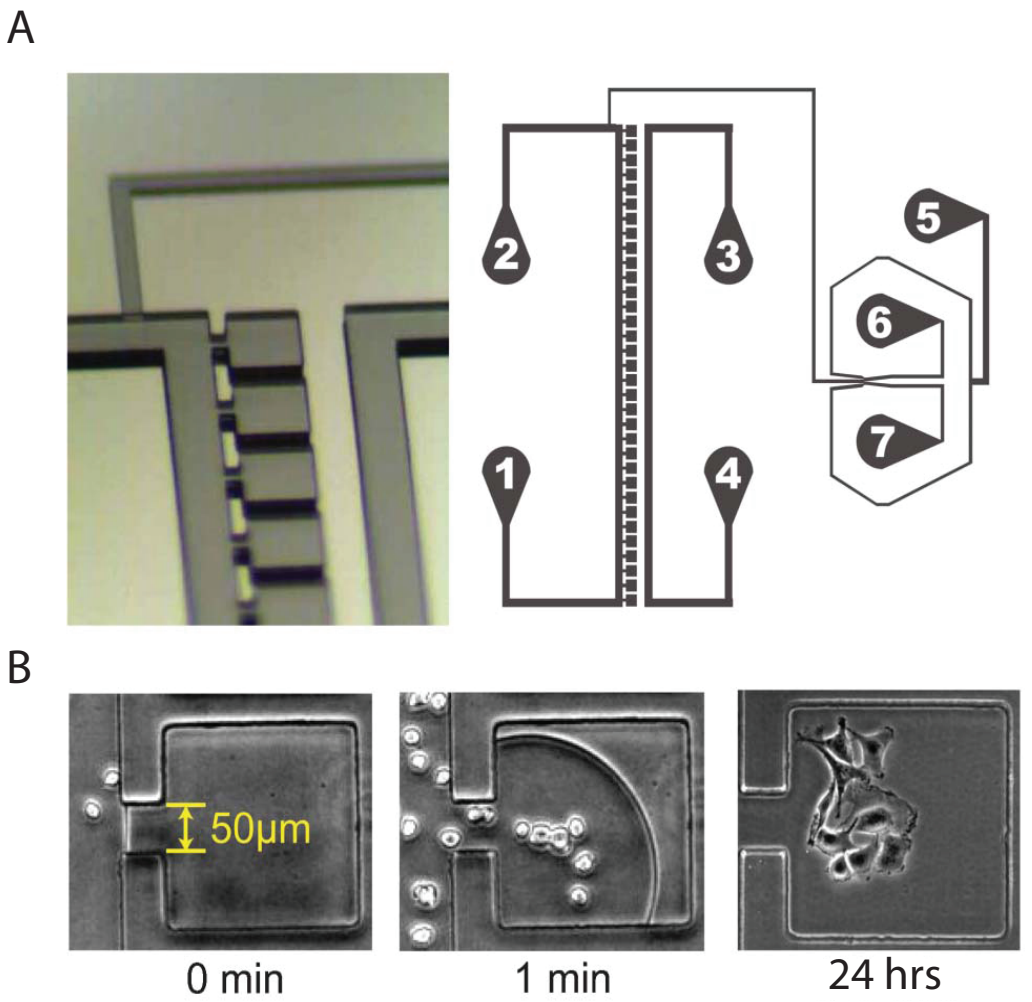


Figure **3-2**: **Microfluidic device for mammalian cells.**(A) Device design. Each of the 33 cuboid culture chambers is connected to a main perfusion channel that runs between ports 1 and 2. A separate air channel between ports 3 and 4 allows the application of a temporary vacuum at the PDMS interface to draw fluid from the main perfusion channel into the culture chambers. Ports 57 comprise the DAW dynamic stimulation generator. (B) Vacuum loading of cells. Upon application of a vacuum in the air channel at time 0, fluid containing cells is rapidly drawn into the culture chambers and fills the traps within 2 min, at which point the vacuum is turned off. Adapted from [46].

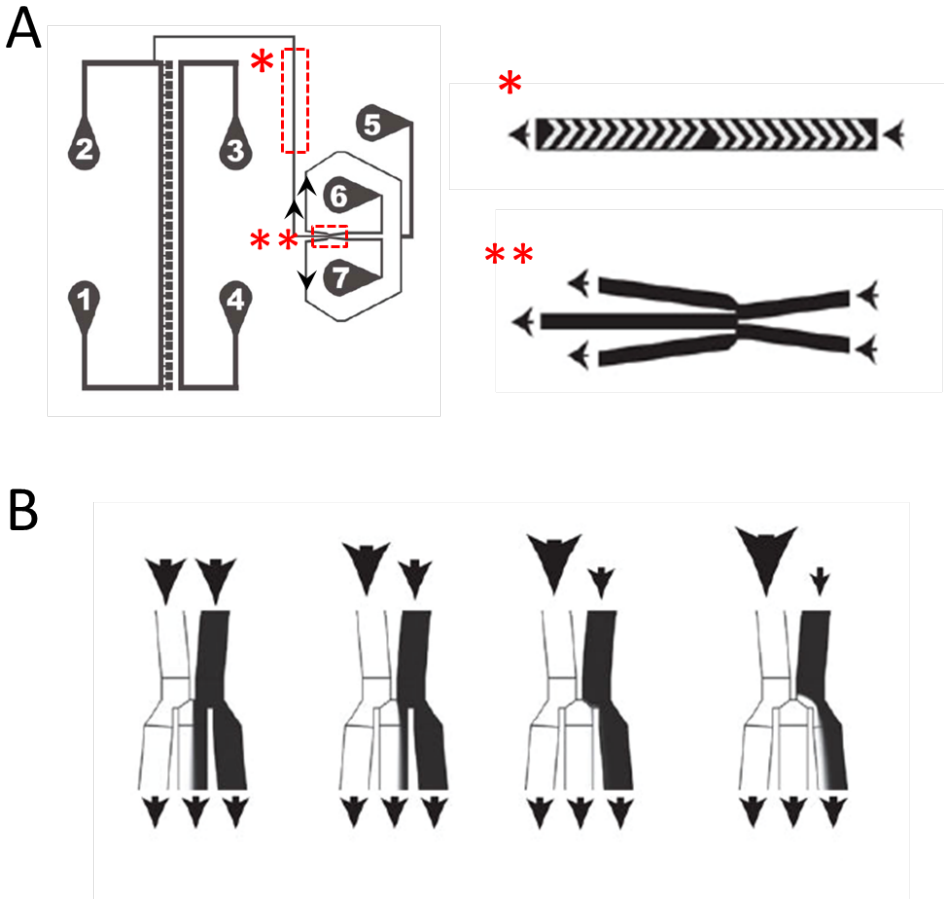


Figure **3-3**: **Overview of DAW junction and staggered herringbone mixers.**(A) The flow direction in DAW junction is indicated by the black arrows. The two inlets on the right come from ports 6 and 7. The flow from the inlets converges in a ratio dependent on the inlet pressures of each. The middle fork of the junction leads to the cell chambers while the two outer forks lead to port 5, the shunt waste port. The staggered herringbone mixers (SHM) immediately follow the DAW junction . It enhances diffusive mixing of the combined laminar flow streams before the fluid reaches the main perfusion channel and the adjoining culture chambers. (B) The panel B represents the DAW junction, the weight of the black arrows are an indication of the applied pressure. The ratio of the inputs from port 6 and 7 leaving the junction to the cell chamber is determined by each port's pressure By varying the pressures applied at the two inlets it is possible to obtain different mixing ratios.

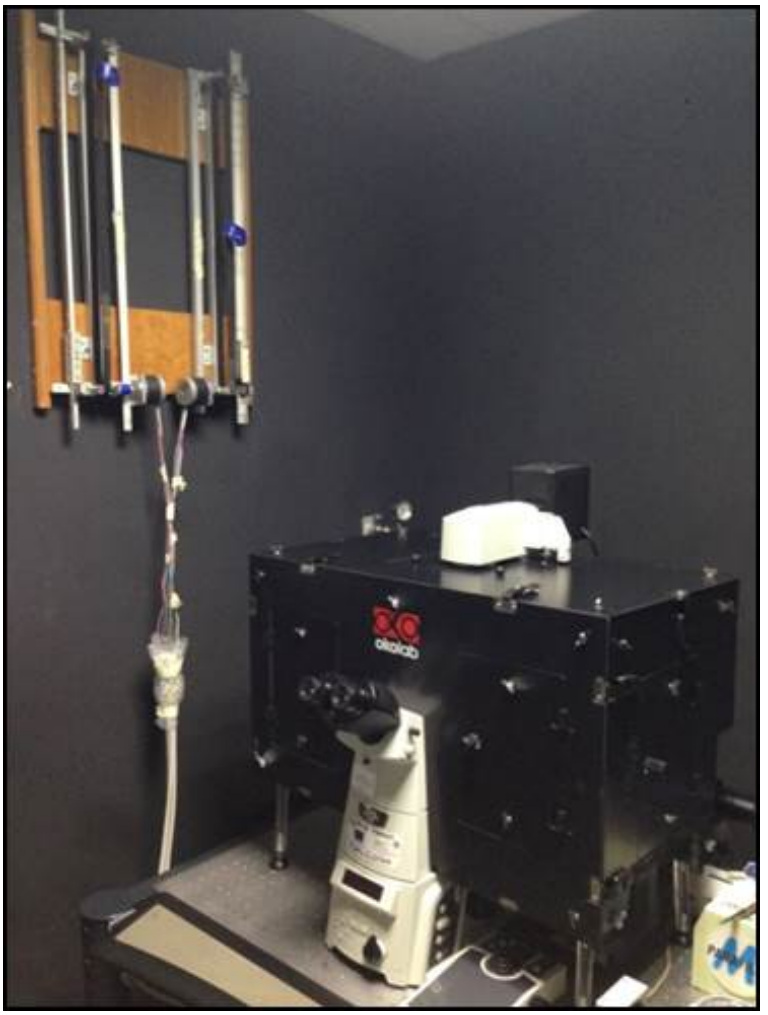


Figure **3-4**: **Linear actuators and Nikon Eclipse Ti fluorescence microscope.** The picture shows the fluorescence microscope (Nikon Eclipse TI) and the linear actuator used in this study.

CHAPTER 4

Relay and Proportional-Integral Control of Gene Expression from the *tetO7*-d2EYFP Synthetic Network.

In this Chapter, I first introduce a microfluidics-based method to automatically regulate gene expression from the tetracycline inducible promoter in mammalian cells in real time. The approach is based on the negative-feedback control engineering paradigm and it is validated in a monoclonal population of cells constitutively expressing a fluorescent reporter protein (d2EYFP) downstream of an inducible promoter consisting of a minimal *CMV* promoter with seven *tet*-responsive operator motifs (*CMVTET*). I then describe two different control algorithms (relay and PI control) to force a monoclonal population of cells to express an intermediate level of d2EYFP equal to 50% of its maximum expression level for up to 3500 min with a discrete control input that is either tetracycline-rich or standard growth medium. I show that both the relay and PI controllers can regulate gene expression at the desired level, despite oscillations (dampened in the case of the PI controller) around the chosen set-point. This work has been partially published in [51] demonstrating for the first time that it is possible to

regulate the gene expression from an inducible promoter in a population of mammalian cells using principles from control engineering.

4.1 The cellular model: Tet-OFF system

Control of gene expression in prokaryotes and lesser eukaryotes has been achieved on both endogenous and synthetic networks (refer to Chapter 2). However, transcriptional dynamics in mammalian cells is up to 1 order of magnitude slower than in bacteria or simple eukaryotes, and therefore, the time scale involved in the control of gene expression is much longer (days instead of hours). Moreover, mammalian cells have much more stringent requirements for culture conditions, such as temperature, pH and CO_2 concentration which have to be taken into account in order to successfully maintain a cell culture in a microfluidic device over several days. These needs are a distinctive feature of mammalian cells that have hindered the application of automatic control to mammalian systems and therefore, as described in Chapter 2, very few attempts have been made at applying Control Theory to control processes in mammalian cells.

In order to design and implement a controller for gene expression in mammalian cells, I decided to use as experimental testbed a simple inducible promoter expressing a fluorescent reporter protein and whose transcriptional activity can be modulated by varying the concentration of a small molecule in the cell growth medium. Specifically, I used the tetrycycline-responsive Tet-OFF system driving expression of a destabilised fluorescent protein, as shown in Figure **4-1**.

The Tet-OFF system, already described in the Section 2.1, has all the characteristics required for a successful control: it has a very tight regulation, hence low basal activity in the presence of the antibiotic tetracycline; once tetracycline is removed from the growth medium, gene expression is induced, with a high expression level compared to other inducible systems in mammalian cells; tetracycline (or doxycycline) is non-toxic and well characterised in the system in use. Moreover, the Tet-OFF system has been in use

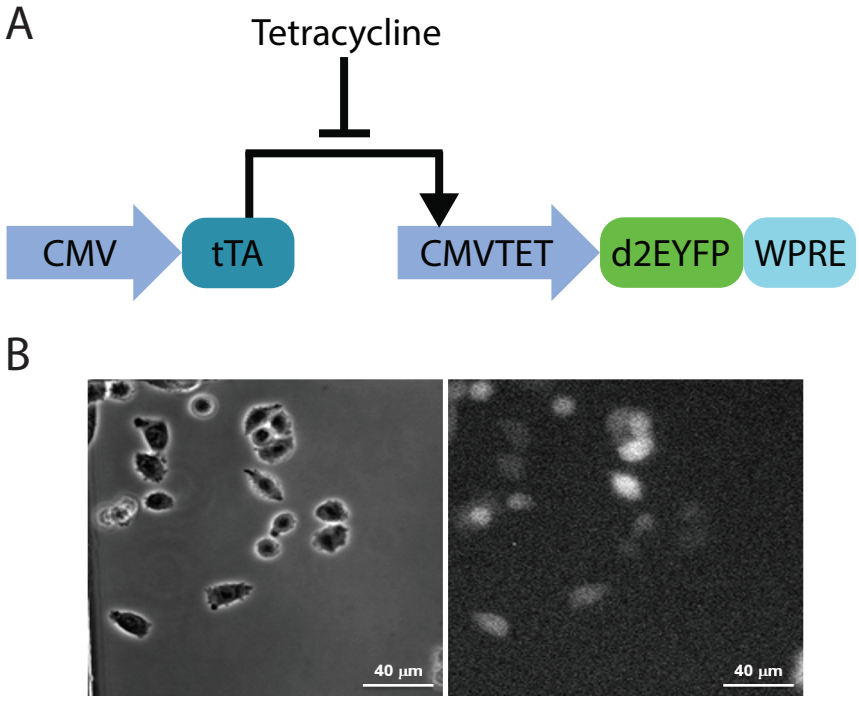


Figure **4-1**: **Experimental model: the *tetO7*-d2EYFP monoclonal cell line.** (A) Scheme of *tetO7*-d2EYFP circuit. The tetracycline transactivator protein (tTA) is constitutively expressed by a *CMV* promoter. tTA binds the *CMVTET* promoter, which harbours seven *tet*-responsive elements (*tetO7*) upstream of a minimal *CMV* promoter, thus driving the transcription of a destabilized enhanced yellow fluorescent reporter protein (d2EYFP). Tetracycline binds tTA and prevents it from binding to the *CMVTET* promoter, blocking the expression of the d2EYFP protein. (B) Representative phase-contrast (left) and fluorescence (right) images of *tetO7*-d2EYFP cells grown in the absence of tetracycline.

for decades, so that many cell lines and animal models using this promoter are available.

The experimental testbed for assessing the feasibility of automatic control of gene expression in mammalian cells consists of a monoclonal population of cells, named *tetO7*-d2EYFP cells, stably integrating the cassette shown in Figure **4-1**A where the Tet-OFF system drives the expression of a destabilized enhanced yellow fluorescent reporter protein d2EYFP [52]. In the Tet-OFF system, d2EYFP is expressed under the control of the *CMVTET* promoter [17, 53], which is composed of a minimal *CMV* promoter with seven *tet*-responsive operator motifs (*tetO7*), embedded in Chinese hamster ovary (CHO) cells, constitutively expressing the tetracycline Transactivator protein (tTA) (Figure **4-1**).

In cells grown in standard growth medium, tTA is able to bind to the *CMVTET* promoter, causing d2EYFP to be maximally expressed. Upon addition of tetracycline to the culture medium, tTA detaches from the *CMVTET* promoter, thus preventing the expression of d2EYFP.

4.1.1 Dynamical model of *tetO7*-d2EYFP system

The *tetO7*-d2EYFP cell line was first described in [54] and modeled using a three-dimensional a continuos-time linear time invariant dynamical system. For each of the species (mRNA and correspondent protein concentrations), an equation expresses the change in concentration as the result of a production term and a degradation term. The resulting model for the *tetO7*-d2EYFP system is:

$$\begin{aligned}
\frac{dx_1}{dt} &= v_1 \left(\alpha_1 + (1-\alpha_1) \frac{\left(\frac{\theta^{h_2}}{\theta^{h_2}+D^{h_2}} \bar{x}_2 \right)^{h_1}}{{K_1}^{h_1} + \left(\frac{\theta^{h_2}}{\theta^{h_2}+D^{h_2}} \bar{x}_2 \right)^{h_1}} \right) - d_1 x_1 \\
\frac{dx_3}{dt} &= v_2 x_1 - (d_3 + K_f) x_3 \\
\frac{dx_4}{dt} &= K_f x_3 - d_3 x4
\end{aligned} \tag{4-1}$$

where x_1 represents the d2EYFP mRNA concentration, x_3 is the unfolded d2EYFP protein concentration and x_4 is the folded d2EYFP protein concentration. The constant value $\bar{x}_2$ represents the concentration of tTA protein, constitutively expressed from the *CMVTET* promoter in *tetO7*-d2EYFP cell line. All parameters with the estimation procedure description are reported in [54].

From a control perspective, the model in Equation (4-1) is a single input-single output (SISO) dynamical system where the presence or absence of tetracycline (represented by parameter D) is the system input and the system output is the folded d2EYFP protein x_4. Thus it can be rewritten in state space form:

$$\begin{aligned} &\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{B}u + \mathbf{W}, \quad \mathbf{x}(0) = \mathbf{x}_0, \quad u \in \{0, 1\} \\ &y = \mathbf{C}\mathbf{x} \end{aligned} \tag{4-2}$$

where

$$\mathbf{A} = \begin{pmatrix} -0.0101 & 0 & 0 \\ 0.0271 & -0.0045 & 0 \\ 0 & 0.0012 & -0.0032 \end{pmatrix} \tag{4-3}$$

$$\mathbf{B} = \begin{pmatrix} 0.0727 \\ 0 \\ 0 \end{pmatrix} \tag{4-4}$$

$$\mathbf{C} = \begin{pmatrix} 0 & 0 & 1 \end{pmatrix} \tag{4-5}$$

$$\mathbf{W} = \begin{pmatrix} 8.52 \cdot 10^{-}6 \\ 0 \\ 0 \end{pmatrix} \tag{4-6}$$

$$\mathbf{x}(0) = \begin{pmatrix} 0.4319 \\ 2.6129 \\ 1.0212 \end{pmatrix} \tag{4-7}$$

In the model 4-2, $u \in \mathbb{R}$ is the system input and accounts for the presence ($u = 1$), or absence ($u = 0$), of tetracycline (or doxycyline) in the growth medium; $y \in \mathbb{R}$ is the output (measurement) of the system; $\mathbf{A} \in \mathbb{R}^{nxn}$ is the dynamic matrix, $\mathbf{B} \in \mathbb{R}^{nx1}$ and $\mathbf{C} \in \mathbb{R}^{1xn}$ are the input and output

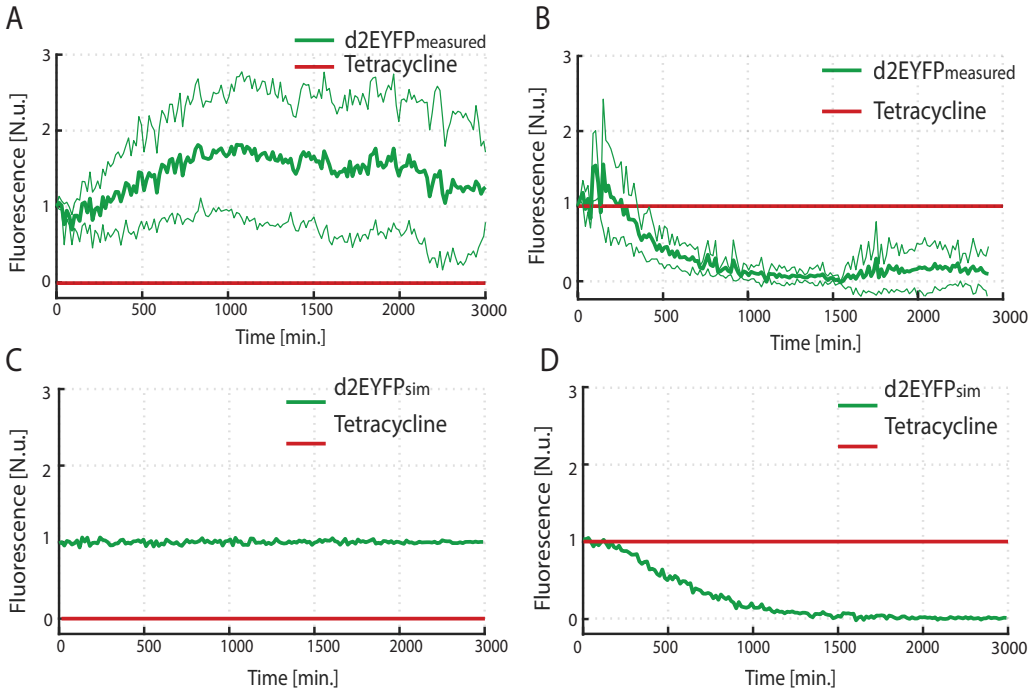


Figure **4-2**: **The *tetO7*-d2EYFP fluorescence in absence and in presence of tetracycline.** (A-B) The red line under each graph represents the respective tetracycline input provided to cells and the green line represents the measured mean fluorescence of the *tetO7*-d2EYFP cells (mean and standard deviation of technical replicates). (A) Fluorescence exhibited by the cells in the presence of growth medium with no tetracycline. (B) Fluorescence exhibited by the cells in the presence of 100 ng/mL tetracycline in the growth medium. (C-D) The red line under each graph represents the respective tetracycline input simulated, and the green line represents the simulated fluorescence output. To simulate biological variability Gaussian noise with zero mean and standard deviation equal to 1 is added to simulation output. (C) Simulation of the model when no tetracycline is provided to the cells. (D) Simulation of the model when a saturating concentration of tetracycline is applied to the cells.

matrices. $\mathbf{x}(0) \in \mathbb{R}^{nx1}$ is the initial state matrix and $\mathbf{W} \in \mathbb{R}^{nx1}$ matrix can be considered as constant disturbance on the input.

In the gene expression control experiments, the measured average fluorescence intensity of d2EYFP protein over the cell population is chosen as system output. Indeed, when dealing with living cells, one of the major issues is represented by the uncertainty affecting transcriptional and translational processes, introducing a remarkable cell-to-cell variability in mRNA and protein production. Thus individual cells may express different level of fluorescence. One of the way to account for this problem is to consider as the system output the average fluorescence intensity expressed by all cells, thus dampening the effects due to noisy measurements.

Figure **4-2** shows the simulated and measured fluorescence intensity expressed by the cells when they grow in standard medium or in tetracycline rich medium. The experimental data and the simulation results are consistent with the expected behaviour of the *tetO7*-d2EYFP cells under investigation. Indeed, when standard medium is provided to the cells, the tTA protein binds the *CMVTET* promoter and therefore d2EYFP fluorescence level is expected to increase or to stay at a high steady state value (Figure **4-2** A-C). Viceversa when tetracycline is given to the cells, tTA detaches from the *CMVTET*, thus d2EYFP fluorescence decreases, or stay at a low steady state (Figure **4-2** B-D).

4.2 Implementation of the experimental control platform to steer gene expression in mammalian cells

In order to carry out *tetO7*-d2EYFP gene expression, I used the integrated experimental platform described in the Chapter 3 and shown in Figure **4-3**, consisting in a disposable microfluidic device, a time-lapse microscope, and a set of automated syringes, all controlled by a computer. This platform enables both monitoring of fluorescent protein levels in real time, and control of the cellular environment. The microfluidic device presented in Chapter 3

was used to trap cells in a micrometric chamber. Either culture medium or an inducer molecule can be supplied to the chamber by changing the relative heights of two motorized syringes that are connected to it via capillary tubes, one containing just culture medium and the other culture medium plus the inducer molecule (tetracycline). The fluorescence produced by cells y is imaged by the microscope at regular sampling times and quantified via an image segmentation algorithm, which computes the average fluorescence intensity across the cells in the image y_m (see Appendix A for details on image segmentation). The measured fluorescence y_m is then compared to the desired fluorescence value r, and a control algorithm computes the necessary amount of inducer molecule u to be administered to the cells in order to minimize the control error e and then moves the syringes accordingly.

In principle, the hydrostatic pressure difference between the two syringes can be used to provide any desired concentration of the inducer molecule by mixing the two fluids through proper adjustment of the heights of the two syringes. However in this first pilot study a discrete control input was used, that is no inducer (i.e. just culture medium) or a single predetermined concentration of the inducer molecule is provided to the cells.

When designing a control platform for controlling a gene network in mammalian cells there are several constraints that have to be taken into account. Indeed during control experiments mammalian cells are subject to different types of stress that real-time imaging and microfluidics induce, such as phototoxicity, shear stress, overgrowth.

Phototoxicity is a deleterious effect of high-energy light irradiation that causes the production of ROS species, cell cycle arrest and ultimately cell death. It is an inevitable effect in time-lapse fluorescence microscopy over long periods of time, but some fine-tuning of acquisition parameters may limit the problem [55]. The mechanical stress that the flow induce in cells can be avoided in the microfluidic device by reducing the filling time of cell traps; however, its drawback is that full exposure to the desired concentration of drug occurs only after 10 minutes after the switching of the media. Hence, a time resolution of 15 minutes between experimental points is chosen as

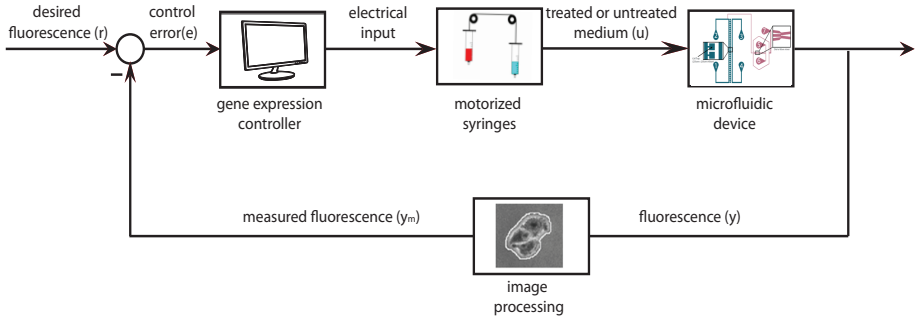


Figure **4-3**: **Control platform.** Cells are grown in a microfluidic device within a temperature- and CO_2-controlled environment under an inverted fluorescence microscope. Images are acquired at regular sampling time y, and quantification of the fluorescence is performed via an image-segmentation algorithm. The measured level of fluorescence y_m is compared to the desired reference value r obtaining the control error e. The control algorithm minimizes the control error and computes the input u to be provided to the cells and moves the syringes containing either untreated or tetracycline-treated medium.

a good trade-off among reduction of phototoxicity, approximation of drug concentration, and resolution of data for fitting and analysis. Finally, the dimension of the cell traps in the microfluidic device limits the length of the experiment to a maximum of $3-4$ days in order to avoid cell death because of overcrowding.

4.3 Control strategies and control objective

The task of controlling a gene network in mammalian cells has to face a number of challenges due to the biological nature of the system being controlled and to technical limitations. The choice of the type of controller is a trade-off between robustness and accuracy. In the case of the *tetO7*-d2EYFP system, the main constraints that the control law has to satisfy are:

1. It is possible to either feed standard medium or tetracycline rich medium

to cells, but not intermediate concentrations.

2. The output (d2EYFP fluorescence intensity) can be measured only at sampling time equal or greater than 15 minutes in order to avoid photoxicity.

Thus in this initial stage the control law should be as simple as possible, I thus started by implementing two simple controllers: Relay and Proportional-Integral (PI).

4.3.1 Relay control

The first control scheme I developed was the relay controller [22]. As already mentioned in the Chapter 2, the relay controller requires only the computation of the control error at each sampling time $e(kT)$, where $T = 15$ minutes, whose sign dictates which input must be provided to the cells. Specifically, cells are treated with standard growth medium for the next 15 min if $e(kT) > 0$ or tetracycline-containing medium otherwise.

The drawback of this controller is that it does not take into account the dynamics of the system and thus can be slow in converging to the desired set-point and show marked oscillations around it. In addition, when the measured value y_m is very close to the reference, the control error can rapidly change sign, thus causing the control input to continuously switch between the two growth media (chattering phenomenon [23]). To avoid chattering, a hysteresis interval ϵ equal to 5% of the reference value is added to the controller. Thus the relay control law used to regulate the *tetO7*-d2EYFP gene expression can be expressed as follows:

$$u(t) = \begin{cases} u_{ON} = 1 \quad if \quad e(t) \geq \epsilon \\ u_{OFF} = 0 \quad if \quad e(t) < -\epsilon \end{cases} \tag{4-8}$$

4.3.2 Proportional-Integral regulator

As detailed in Chapter 2, a PI controller is very useful and it is capable of solving a wide range of industrial control problems [22]. The output of PI

controller $\hat{u}(t)$ as a function of the control error $e(t)$ defined as

$$\hat{u}(t) = K_p e(t) + K_i \int_0^t e(\tau)\, d\tau \tag{4-9}$$

where the gains K_p and K_i have to be tuned to optimize controller yield.

Since in the *tetO7*-d2EYFP system, cells can be fed with or without Tetracycline in a mutually exclusive manner, the continuous signal $\hat{u}(t)$ has to be modulated in a discrete control input that is restricted to be either ON (standard medium) or OFF (tetracycline rich medium) so that:

$$u(t) = \begin{cases} u_{MAX} = ON & kT \leq t < (k + d_k)T \\ u_{MIN} = OFF & (k + d_k)T \leq t < (k + 1)T \end{cases} \tag{4-10}$$

Therefore at each sampling time kT the control input $\hat{u}(t)$ is encoded as the duty cycle d_k of a square wave with a period of 15 minutes according to the following law:

$$d_k = \frac{\hat{u} - u_{MIN}}{u_{MAX} - u_{MIN}} \tag{4-11}$$

where $\hat{u}$ is the output of the PI regulator saturated between $u_{MIN} = 0$ and $u_{MAX} = 1$.

This modulation introduces a saturation on the PI output $\hat{u}$. As already described in Chapter 2, the presence of a saturation downstream to the PI can lead to a break of the feedback loop affecting the control performance [22]. To overcome this issue I have implemented the anti-windup scheme described in Chapter 2 (Figure **2-7**) considering a feedback gain $K_t = .5$.

The proportional and integral gains, $K_p = 4.08$ and $K_i = 0.012$ were calculated with the Ziegler-Nichols' closed-loop tuning method [22] applied to the mathematical model of the *tetO7*-d2EYFP described by the Equations (4-2).

4.3.3 Control objective

In order to assess whether automatic control of gene expression in mammalian cells is feasible, I chose as control objective set-point regulation where

cells are required to express 50% of their maximum fluorescence for up to 3500 minutes.

The maximum fluorescence is estimated during the calibration phase at the beginning of each control experiment. During the calibration phase of 180 minutes, the cells in microfluidic device are fed with standard medium so that d2EYFP is maximally expressed. The images are acquired with a sampling time of 15 minutes and the fluorescence value is estimated as the time- and population-average fluorescence expressed by the cells during the first 180 minutes. The calibration phase is also needed to allow the cells to adapt to the microfluidic environment.

4.4 Numerical simulations of the Relay and PI controllers.

To validate the relay and PI control strategies, I simulated the control of *tetO7*-d2EYFP gene expression by using the dynamical model described by the Equations (4-2) as a proxy for the the behaviour of *tetO7*-d2EYFP cells. Thus the control objective was to regulate the output y of model in order to the reference signal r, calculated as the 50% of the steady state of the model's output for a control input equal to 1 (i.e. cells grown in standard medium). In order to account for measurement noise, a gaussian noise with zero mean and standard deviation equal to 1 is added to the model output before being fed back to the controller. For the sake of clarity and to make comparison with the experimental results easier, the simulated tetracycline concentration is reported in red in the Figure **4-4** as $(1 - u)$ where u is the input computed by the controllers. Thus, when the value of red signal is high, the cells are fed with tetracycline, viceversa standard growth medium is provided to the cells when the red signal is low (red line in Figure **4-4**).

Figures **4-4** A and B show the simulation results for the set-point control task using the relay and PI controllers, respectively. In both cases the controllers were able to control the output of the model y to the desired value

r but, as expected, oscillations around the set-point are present. These oscillations are due to the controllers' choice, as both relay and PI are not model-based and hence not able to predict the system's behaviour. Moreover, the controller performance are negatively affected also by the binary nature of the control input.

4.5 *In vitro* experimental implementation of the Relay and PI controllers.

In order to perform *in vitro* control experiments, the *tetO7*-d2EYFP cells were loaded in the microfluidic device and allowed to attach and grow for 24 hours under perfusion conditions in a cell culture incubator. The device was then moved to the inverted microscope, and image acquisition and control experiment were started.

The control task was to reach and maintain a set-point equivalent to 50% of the fluorescence value measured during the calibration phase. This value was obtained by imaging the cells in the absence of tetracycline for 180 minutes at 15 minutes intervals and then calculating the time-averaged value of the fluorescence across the images.

Figure **4-5** A and B show the results of the *in vitro* control experiments with the relay and PI controllers, respectively. In agreement with numerical simulations (Figure **4-4**), the system is unable to settle on the reference value but persistently oscillates around it for the whole length of the experiment (3500 minutes). The behaviour of the relay controller (Figure **4-5** A) is very close to that of the numerical simulations (Figure **4-4** A), exhibiting undamped oscillations around the set-point.

In the case of the PI controller (Figure **4-5** B), the oscillations have a decreasing amplitude with time, as predicted in the simulation . One difference between the experiment and the corresponding simulation in this case (Figure **4-4** B) is the longer initial delay following tetracycline administration before the fluorescence starts to decrease. Nevertheless, the controller is able to react accordingly, and the fluorescence value of the cells is slowly

brought toward the desired set-point. This slight difference in the initial dynamics is due to intrinsic variabilities caused by experimental conditions and cell batches.

Despite the oscillations around the set-point value, these experiments demonstrated that the developed control strategies are suitable to automatic control of gene expression in mammalian cells is feasible. This work was partially published in [51] and it was the first time automated control of gene expression from an inducible promoter has been achieved in mammalian cells.

The control performance can be greatly improved by moving to a model−based control strategy such as Model Predictive Control.

The implementation of an MPC strategy to regulate gene expression in mammalian cells will be the topic of the next Chapters.

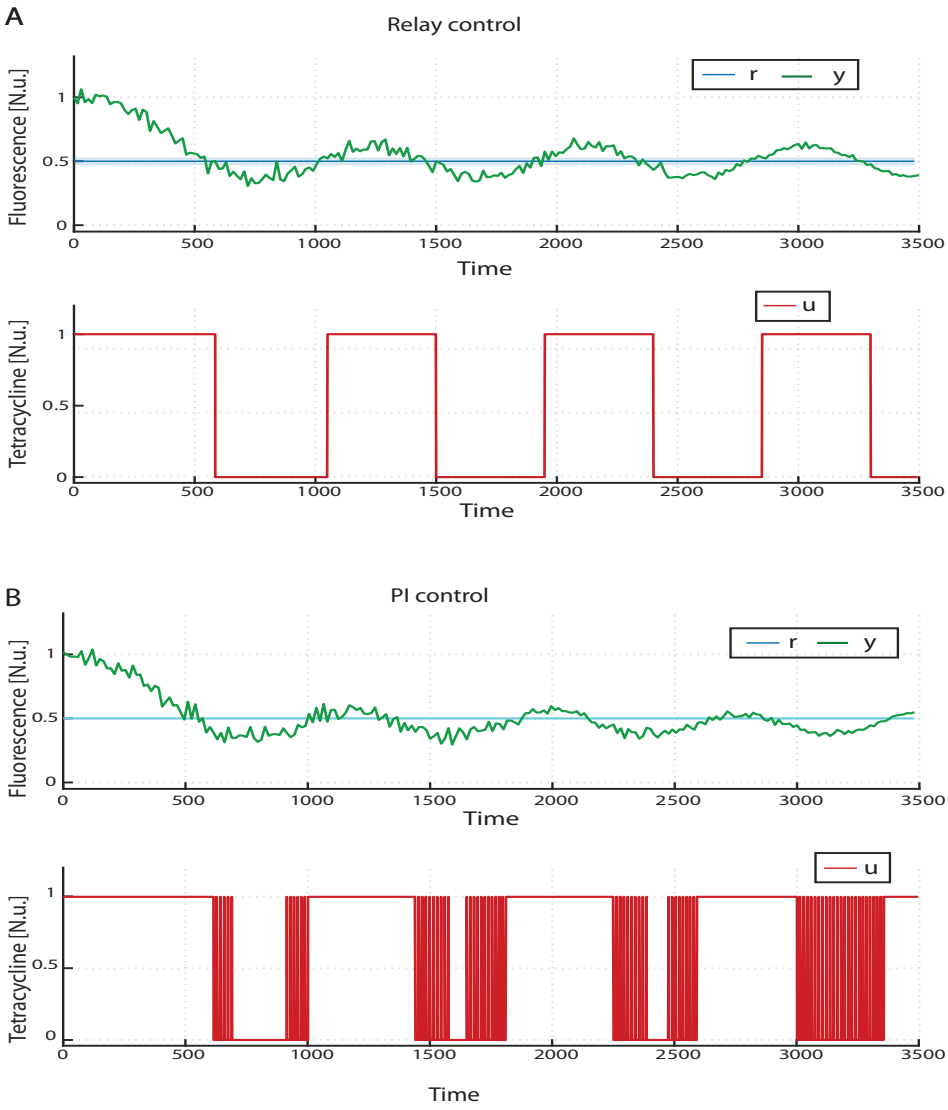


Figure **4-4**: **Numerical simulation of set-point control of *tetO7*-d2EYFP gene expression.** Simulations of control experiments performed on the *tetO7*-d2EYFP dynamical model by means of Relay with hysteresis with equal to the 5% of the set-point (A) and PI (B) controllers. The red line under each graph represents the respective tetracycline input simulated u, and the green line represents the simulated fluorescence output y. The reference value (r, blue line) is calculated as the 50% of the high steady state of the model's output, here normalized to 1.

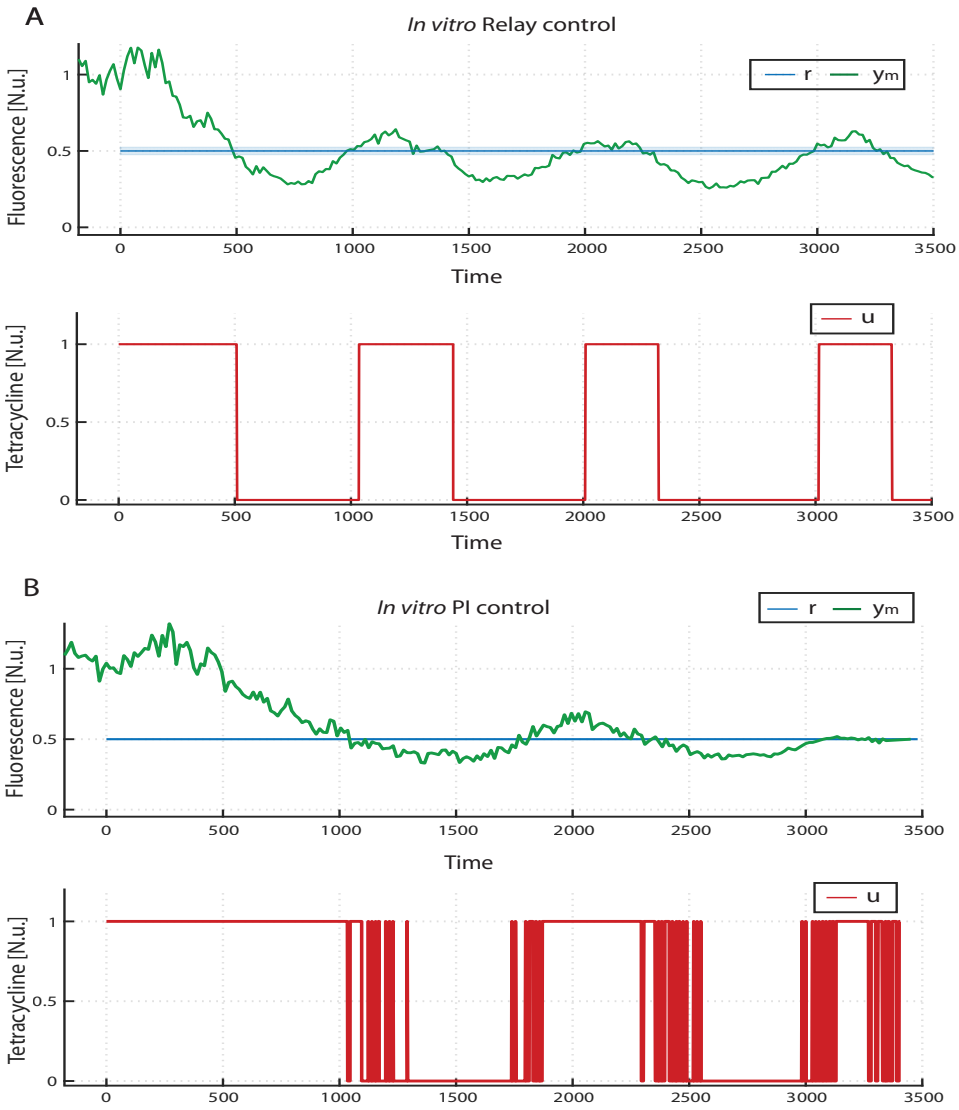


Figure **4-5**: ***In vitro* set-point control experiments.** Two *in vitro* set-point control of *tetO7*-d2EYFP gene expression by means of Relay with hysteresis with equal to the 5% of the set-point (A) and PI (B) controllers. The control action starts at time $t = 0$ min. The value of the desired level of fluorescence (r in blue) is calculated as the 50% of the d2EYFP fluorescence quantified before time $t = 0$ min. during the calibration phase of 180 min The red line under each graph represents the control input u (tetracycline provided to cells in the microfluidic device), and the green line represents the measured mean d2EYFP fluorescence (system output y, in green)

CHAPTER 5

Model Predictive Control of Gene Expression from the *tetO7*-Ub^{V76}GFP synthetic network with a Discrete Control Input

In this Chapter, I describe the construction of an inducible synthetic gene network in mammalian cells based on Tet-OFF system, which I named *tetO7*-Ub^{V}76GFP, exhibiting faster dynamics than *tetO7*-d2EYFP described in the Chapter 4. Then I discuss the use of the experimental platform presented in Chapter 3 to identify a dynamical model of the *tetO7*-Ub^{V76}GFP system . I show that this experimental set-up allows to infer quantitative dynamical models of transcriptional processes from measured input and output data. These models are useful to predict the behaviour of the system in response to various stimuli and to design feedback control strategies. Finally, I present a Model Predictive Control strategy to regulate the *tetO7*-Ub^{V76}GFP gene expression with a discrete control input. I will show how the application of a model based control strategy drastically improves the control performance.

5.1 Tetracycline inducible promoter driving the fast-degrading reporter protein Ub^{V76}GFP

The results presented in Chapter 4 actually demonstrate that microfluidics based automatic control of gene expression from an inducible promoter in mammalian cells is feasible. In Chapter 4, I presented the results obtained in controlling gene expression by using two controllers: a relay controller and Proportional-Integral (PI) controller. Despite the PI control scheme improves control performance compared to the relay, still oscillations around the reference value persist. Long timescales of transcriptional dynamics in mammalian cells are responsible for the low performance of the relay and PI control strategies. Therefore, I chose to deal with the inherent inertia of the biological system by applying model-based control strategies such as the Model Predictive Control (MPC) [8, 9, 11]. Indeed, the MPC allows to reduce or eliminate the oscillations around the reference value thanks to its predictive nature.

The technological platform for controlling living cells (discussed in Chapter 3) allows a control experiment to last at most 3 days, after which the microchannels clog because of cell proliferation. Therefore, I decided to modify the *tet07*-d2EYFP system used for testing the relay and PI controllers in Chapter 4 by replacing the fluorescent reporter protein (d2EYFP) with a reporter with faster mRNA and protein degradation rates, thus yielding a gene network with faster transcriptional and translational dynamics. This modification allows to run control experiments of shorter duration (less than 3 days).

Specifically, in this new *tetO7*-Ub^{V}76GFP network, the Tet-OFF system [17, 53] drives the expression of a fluorescent reporter gene with fast-degrading mRNA and protein. In order to a have a fast-degrading protein, I chose the Ub^{V76}GFP whose half-life is in the order of minutes rather than hours as in the case of $d2EYFP$, thanks to an ubiquitonanion signal tagging the protein for degradation by the proteasome [56].

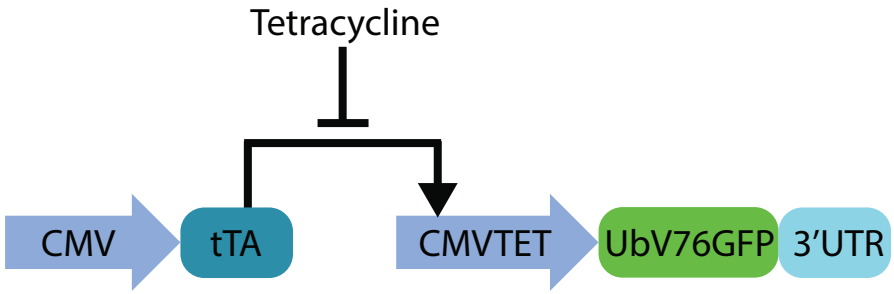


Figure **5-1**: **Scheme of the *tetO7*-Ub^{V76}GFP circuit.** A Tetracycline transactivator (tTA) is constitutively expressed from the *CMV* promoter, and it is able to activate the transcription of the Ub^{V76}GFP reporter from the *CMVTET* (or *tetO7*) promoter on a separate transcriptional unit. The 3'UTR sequence is cloned downstream of the Ub^{V76}GFP to increase its degradation rate. Tetracycline binds to the tTA, blocking its activity.

In order to enhance the degradation rate of the mRNA encoding for the reporter protein, the endogenous 3′ Untraslated Region (3′UTR) of the *Hes1* gene was cloned downstream the Ub^{V76}GFP sequence. Indeed, it has been hypothesized that this sequence contains destabilizing signals responsible for the rapid degradation of *Hes1* mRNA [57]. A scheme of the *tetO7*-Ub^{V76}GFP system under investigation is reported in the Figure **5-1**.

In these cells, the destabilized fluorescent reporter protein Ub^{V76}GFP is expressed under the control of the *CMVTET* promoter [17, 53], which harbours seven *tet*-responsive elements (*tetO7*) upstream of a minimal *CMV* promoter, embedded in CHO cells, constitutively expressing the tetracycline Transactivator (tTA) protein. In cells grown in standard growth medium, tTA protein is able to bind the *CMVTET* promoter causing Ub^{V76}GFP to be maximally expressed. Upon addition of tetracycline, or its homologous doxycycline, to the culture medium, tTA detaches from the *CMVTET* promoter thus preventing the expression of Ub^{V76}GFP.

Model identification and the control of *tetO7*-Ub^{V76}GFP system will be the subject of the next sections.

5.2 Modeling *tetO7*-Ub^{V76}GFP system dynamics in mammalian cells

The microfluidic-based experimental platform presented in the Chapter 3 can be used also in open-loop in order to infer quantitative models of transcriptional processes from measured input and output time-series data. Thus, I took advantage of this set-up in order to identify dynamical model for *tetO7*-Ub^{V76}GFP system.

As already described in the previous section, the biological system under investigation is CHO Tet-OFF cell line, expressing constitutively the tTA protein, in which the destabilised fluorescent reporter Ub^{V76}GFP is stably integrated under the control of *CMVTET* promoter. In untreated medium, the cells express Ub^{V76}GFP protein, but expression is inhibited upon addition of tetracycline to the culture medium (Figure **5-1**).

In order to perform identification experiments, *tetO7*-Ub^{V76}GFP cells were loaded in the microfluidic device and allowed to grow for 24 hours under perfusion conditions in a cell culture incubator. The device was then moved to the inverted microscope, and the images acquisitions and the identification experiment were started (for more details on experimental protocol see Appendix A).

In identification experiments, the strategy I followed was to give as input to the cells in the microfluidic device, a series of pulses of tetracycline concentration with variable duration but fixed amplitude by means of the automated syringes, and to measure Ub^{V76}GFP fluorescence over cell population, which I considered as the output of the system [58].

Experimental data. The green fluorescence of *tetO7*-Ub^{V76}GFP cells was captured for up to 20 hours each 15 minutes. During this interval, tetracycline rich medium at the concentration of 100 ng/ml and standard growth medium were alternatively provided to the cells chamber (Figure **5-2**). The concentration of tetracycline in the cells chamber was tracked with a red fluorescent dye (Sulforhodamine), so that it was possible to obtain a

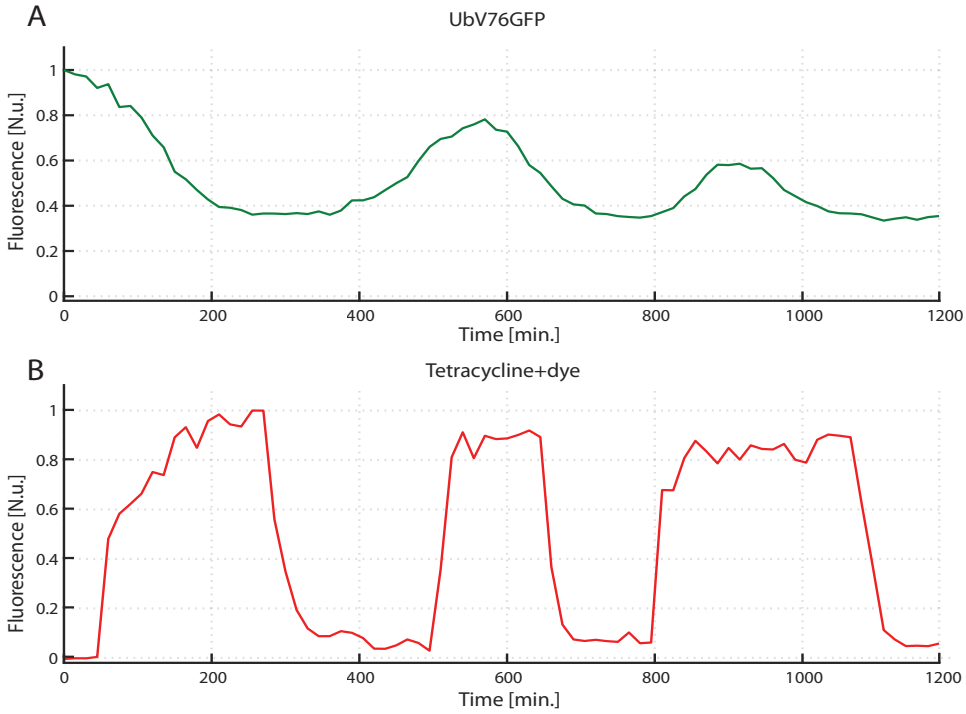


Figure **5-2**: ***TetO7*-Ub^{V76}GFP experimental data.** (A) The green signal is the average of the Ub^{V76}GFP fluorescence across the cell population measuring during the experiment. This signal is considered the output of the system. (B) The red signal is the fluorescence of the dye added to the tetracycline rich medium measured during the experiment. It is high when the cells are fed with tetracycline whereas it is low when the standard medium is provided to the cells.

time profile of the input provided to the cells by measuring the fluorescence of the medium in the red spectrum (Figure **5-2** B). The average fluorescence of the cells' population, in the green spectrum, which is proportational to the average expression of Ub^{V76}GFP protein, represents the system output (Figure **5-2** A).

The representative dataset shown in Figure **5-2** is consistent with the expected behaviour of the *tetO7*-Ub^{V76}GFP system under investigation. Indeed, when standard medium is provided to the cells, the tTA binds the *CMVTET* driving the expression of Ub^{V76}GFP fluorescent protein. Vicev-

ersa when tetracycline is given to the cells, tTA detaches from the *CMVTET* promoter, thus preventing the expression of Ub^{V76}GFP protein. Therefore, the fluorescence level is expected to reach a high steady state value when the cells are fed with standard medium. It should instead reach a lower steady state when tetracycline is provided. At the beginning of the experiment, the cells exhibit a high level fluorescence because the cells are grown in standard medium.

5.2.1 Grey box identification

In order to obtain a mathematical model for *tetO7*-Ub^{V76}GFP system, I followed the System Identification procedure described in Section 2.3 by starting from the input-output data depicted in Figure **5-2** where the system input is represented by pulses of tetracycline rich medium and standard medium and the system output is the average intensity of the cells fluorescence in the green spectrum.

In order to identify a dynamical model of the *tetO7*-Ub^{V76}GFP system with biologically meaningful parameters, I decided to use a gray-box identification approach with a predefined model structure with unknown parameters. I derived a set of three linear differential equations describing the production and degradation of the mRNA (Equation (5-1)), the unfolded reporter protein (Equation (5-2)) and the folded reporter protein (5-3)). Specifically, I assumed distinct dynamics for the unfolded (inactive) and folded (active) forms of the Ub^{V76}GFP reporter protein in order to take into account Ub^{V76}GFP protein maturation time needed for correct protein folding [56].

Letting x_1 be the Ub^{V76}GFP mRNA concentration, x_2 the unfolded Ub^{V76}GFP protein concentration and x_3 the folded Ub^{V76}GFP protein concentration, the *tetO7*-Ub^{V76}GFP system can be described as follows

$$\frac{dx_1}{dt} = -d_1 x_1 + \beta_1 u \tag{5-1}$$

$$\frac{dx_2}{dt} = \alpha_2 x_1 - d_2 x_2 \tag{5-2}$$

$$\frac{dx_3}{dt} = \alpha_3 x_2 - d_3 x_3 \tag{5-3}$$

with $u \in \{0, 1\}$.

In the Equation (5-1), u is the only external input to the model and it is assumed to be equal to 1 when cells are fed with standard medium, and 0 when tetracycline is provided to the cells (Figure **5-3** B). d_1 is a linear degradation coefficient for the mRNA and the input coefficient b is its production rate. In the Equations (5-2) and (5-3), d_2 and d_3 are the degradation rates of the unfolded Ub^{V76}GFP and folded Ub^{V76}GFP protein, and α_2 and α_3 represent the translation rate and the folding rate respectively.

The resulting SISO LTI model is:

$$\begin{aligned} &\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{B}u, \quad \mathbf{x}(0) = \mathbf{x}_0 \quad u \in \{0, 1\} \\ &y = \mathbf{C}\mathbf{x} \end{aligned} \tag{5-4}$$

where $\mathbf{A} = \begin{pmatrix} -d_1 & 0 & 0 \\ \alpha_2 & -d_2 & 0 \\ 0 & \alpha_3 & -d_3 \end{pmatrix}$, $\mathbf{B} = \begin{pmatrix} \beta_1 \\ 0 \\ 0 \end{pmatrix}$ and $\mathbf{C} = \begin{pmatrix} 0 & 0 & 1 \end{pmatrix}$

I have named this set of equations the *Ub^{V76}GFP.DI model*, where DI stands for Discrete Input.

The model parameters were obtained by using the grey-box identification technique described in [34, 59], and implemented in the MATLAB System Idenfication toolbox (Mathworks Matlab R2016b) with the function *greyest*, on the input-output data in Figure **5-3**, where the output is the average cells fluorescence of the reporter protein (x_3 in (5-3)) and the input represents the tetracycline rich or standard medium (u in (5-1)).

In order to validate the model, I used the following performance metrics: Akaike's Final Prediction Error (FPE), Information Criterion (AIC) and the fitting percentage (FIT%) [34, 36].

FPE and AIC metrics can be used to evaluate the ability of the identified model to capture the system response to a known input signal.

These performance indeces are computed as:

$$FPE = \frac{1 + \frac{d}{N}}{1 - \frac{d}{N}} V \tag{5-5}$$

$$AIC = log\left(V\left(1 + 2\frac{d}{N}\right)\right) \tag{5-6}$$

where d is the total number of estimated parameters and N is the length of the data record. V is the loss function defined as:

$$V = \frac{1}{N}\sum_{i=1}^{N} \epsilon^2\left(i, \theta_N\right) \tag{5-7}$$

where θ_N is the vector of estimated parameter and $\epsilon^2\left(i, \theta_N\right)$ is the prediction error.

The Fitting Percentage (FIT%) measures how well the response of the model fits the estimation data, expressed as a percentage. It is defined as:

$$FIT(\%) = 100\left(1 - \frac{\sqrt{\sum_{i=1}^{N}\left(\hat{y}_i - y_i\right)^2}}{\sqrt{\sum_{i=1}^{N}\left(\hat{y}_i - \bar{\hat{y}}\right)^2}}\right) \tag{5-8}$$

where for i-th datapoint, $\hat{y}_i$ is the measured output data, $\bar{\hat{y}}$ is its mean and y_i is the model output in response to the input that leads to the real system output $\hat{y}_i$.

The estimated model parameters are reported in Table **5-1**, while the response of the identified model to the input signal is shown in Figure **5-3** (A). The evaluated performance indices are reported in the Table **5-2**.

5.3 *In vitro* regulation of *tetO7*-Ub^{V76}GFP system with discrete control input

In order to implement a control strategy to steer Ub^{V76}GFP expression, I employed the experimental platform described in the Chapter 3 and used also for controlling the *tetO7*-d2EYFP (see Chapter 4 and [51]).

Briefly it consists of a closed loop control platform based on microfluidic device, featuring a computer implementation of the control algorithm and

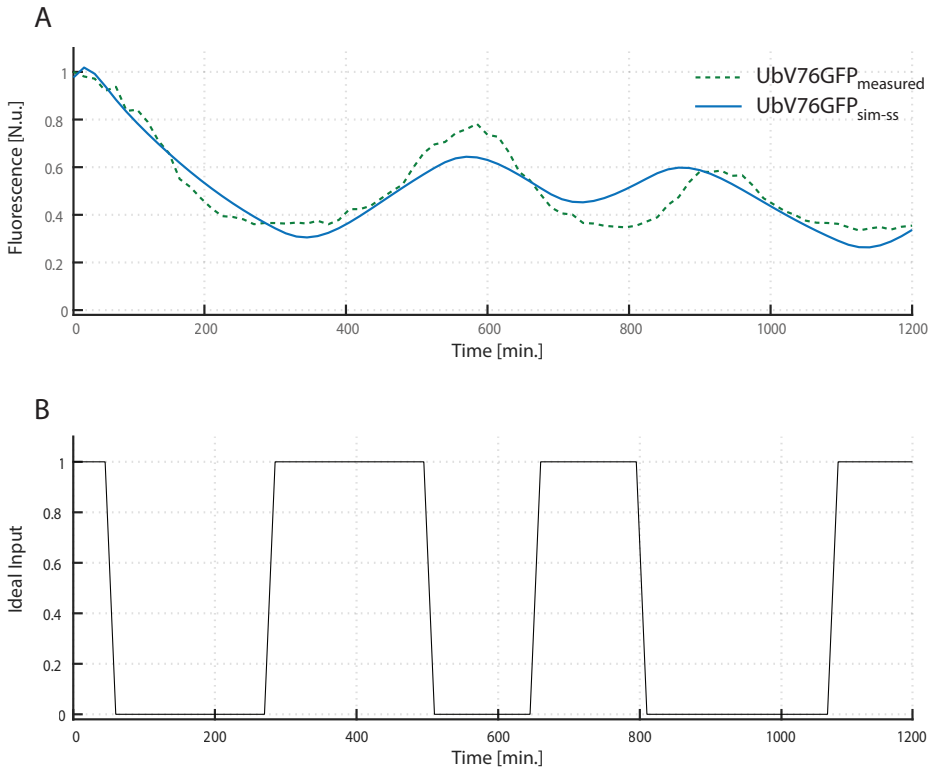


Figure **5-3**: **Grey box identification of the Ub^{V76}GFP model.** (A) The solid blue line is the output of the state-space model identified with *greyest* function from MATLAB System Idenfication toolbox (Mathworks Matlab R2016b). The green dashed line is the average cell fluorescence measured during the experiment. (B) The black signal is the ideal input and it can be either 0 (tetracycline-rich medium) or 1 (standard medium). The same input signal is used both to identify and validate the model.

an inverted fluorescence microscope. The *tetO7*-Ub^{V76}GFP cells are grown in the microfluidic device and are imaged by a time-lapse epifluorescence microscope. The computer uses the images to quantify cells fluorescence and compare it with the desired fluorescence amount at each sampling time. On the basis of the control error, the control algorithm computes the control action and varies the height of two motorized syringes filled with either tetracycline rich medium or standard medium. Hydrostatic pressure gener-

Table **5-1**: Parameters of *tetO7*-Ub^{V76}GFP model identified by with the grey box identification approach

Parameter	Description	Value
$d_1\,[min^{-1}]$	degradation rate of Ub^{V76}GFP mRNA	0.0256
$d_2\,[min^{-1}]$	degradation rate of Ub^{V76}GFP unfolded protein	0.0257
$d_3\,[min^{-1}]$	degradation rate of Ub^{V76}GFP folded protein	0.0045
$\beta_1\,[min^{-1}]$	production rate of Ub^{V76}GFP mRNA	0.0029
$\alpha_2\,[min^{-1}]$	production rate of Ub^{V76}GFP unfolded protein	0.0501
$\alpha_3\,[min^{-1}]$	production rate of Ub^{V76}GFP folded protein	0.0209

Table **5-2**: Values of the performance indices for the *tetO7*-Ub^{V76}GFP model

FPE	$6.50 * 10^{-3}$
AIC	-5.0362
FIT	64.3%

ated by the relative difference in the heights of the two syringes drives the flow in the microfluidic device and determines the type of growth medium cells will sense in the chamber.

As already mentioned in the previous sections, in the *tetO7*-Ub^{V76}GFP system, the Tta protein drives the expression of the UbV76GFP reporter protein in CHO cells (Figure **5-1**). It can be viewed as a single input-single output (SISO) dynamical system. The input $u(t)$ describes standard medium (1) or tetracycline-rich medium (0). The output $y(t)$ is the measured average level of fluorescence of the UbV76GFP protein over the cell population, which can be used as a proxy of the protein concentration [60].

In the control experiment, the cell fluorescence is imaged at 15 minutes intervals. This sampling time T is chosen as trade-off among reduction of phototoxicity and resolution of data for fitting and analysis.

5.3.1 Control Objective and control input

The control task consists in driving a growing population of CHO *tetO7*-Ub^{V76}GFP cells to reach and maintain over 2500 minutes a fixed amount of

the Ub^{V76}GFP protein equal to 50% of the maximum fluorescence exhibited by the cells when grown in standard medium. Specifically, the control objective is a set-point regulation where the reference value $r(t)$ is calculated as the 50% of the average of the fluorescence level expressed by the cells during the calibration phase that occurs at the beginning of each control experiment. In this phase, the cells, which were grown overnight in standard medium thus causing the Ub^{V76}GFP to be maximally expressed, are loaded in the microfluidic device and kept in standard medium for 180 minutes. This calibration phase allows to set the unit of measure of fluorescence that are subject to change in every experiment due to technical (fluorescent lamp usage) and biological (expression noise) issues. During the calibration phase the maximum fluorescence reached by the cells in quantified and it is then used to set the unit of measure of the reference signal.

The control input u is a discrete signal that is restricted to be either 1 (standard medium) or 0 (medium with tetracycline at a concentration of 100 ng/ml) so that:

$$u(t) = \begin{cases} u_{MAX} = ON & kT \leq t < (k + d_k)T \\ u_{MIN} = OFF & (k + d_k)T \leq t < (k + 1)T \end{cases} \tag{5-9}$$

5.3.2 Control strategy: Model Predictive Control

In order to regulate the Ub^{V76}GFP protein expression in CHO cells, I decided to implement a Model Preditive Control (MPC) feedback law, which has already been applied to control biological systems (refer to Chapter 2).

As described in the Section 2.2.3, given a dynamical model of the system being controlled and desired temporal profiles for the system's outputs, MPC aims at finding the control input to minimize the deviation between the outputs of the model and the desired outputs. As model of the *tetO7*-Ub^{V76}GFP system, I chose the *Ub^{V76}GFP.DI model* in the Equation (5-4). Moreover, to speed up the computation process I decided to use a discretised version of the *Ub^{V76}GFP.DI model*, assuming that the input is piece-wise

constant during the sampling period T (zero-order hold method described in [61]). I thus obtained

$$\begin{aligned} &\mathbf{x}_{k+1} = \mathbf{A}\mathbf{x}_k + \mathbf{B}u_k, \quad u_k \in \{0,1\} \quad \mathbf{x}(0) = \mathbf{x}_0 \\ &y_k = \mathbf{C}x_k \end{aligned} \tag{5-10}$$

where $\mathbf{x}_k = \begin{pmatrix} x_1(kT) \\ x_2(kT) \\ x_3(kT) \end{pmatrix}$ are the system states, $u_k = u(kT)$ is the control input and $y_k = y(kT)$ is the system output with k being a natural number ($k \in [1, 2...]$). The matrices $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{C}$ are:

$$\mathbf{A} = \begin{pmatrix} a_{11} & 0 & 0 \\ a_{2,1} & a_{2,2} & 0 \\ a_{3,1} & a_{3,2} & a_{3,3} \end{pmatrix} \tag{5-11}$$

$$\mathbf{B} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} \tag{5-12}$$

$$\mathbf{C} = \begin{pmatrix} 0 & 0 & 1 \end{pmatrix} \tag{5-13}$$

the discretised model preserves its dynamics, with eigenvalue of A being in modulus less than 1 (system asymptotically stable [22]).

At each sampling time kT, starting from the experimental data, the MPC controller uses the discrete model (5-10) to predict the dynamic behaviour of the *tetO7*-Ub^{V76}GFP system over a defined prediction horizon and to determine the input such that an open−loop objective function is minimized [28, 32].

As the open−loop objective function to be minimised, I chose the squared control error (SSE), defined as:

$$\begin{aligned} SSE_k &= \sum_{i=k+1}^{k+N} \left(N + 1 + k - i\right) \epsilon_i^2 \\ &= \sum_{i=k+1}^{k+N} \left(N + 1 + k - i\right) \left(r_i - y_i\right)^2 \end{aligned} \tag{5-14}$$

where $N = 15$ defines the length of the prediction horizon in terms of sampling intervals (corresponding to 225 minutes). $(N+1+k-i)$ is a weighting factor that weights more the control error samples at the beginning of the prediction horizon than those at the end in order to guarantee faster corrections of output deviations from the reference r [11].

Specifically at each sampling instants kT, starting from the current state information $\mathbf{x}(kT)$ the MPC solves the following open loop optimal control problem

$$\begin{aligned} &\min_{\{d_{k+1},\ldots,d_{k+N}\}\in[0,1]} SSE_k \\ &s.t. \\ &\quad \mathbf{x}_{i+1} = \mathbf{A}\mathbf{x}_i + \mathbf{B}u_i \\ &\quad y_i = \mathbf{C}\mathbf{x}_i \\ &\quad \mathbf{x}_k \quad is \quad given \\ &\quad u_i = d_i\,(u_{MAX} - u_{MIN}) \\ &\quad y_i \in [0,1] \\ &\quad i \in \{k+1,\ldots,k+N\} \end{aligned} \tag{5-15}$$

The state measurements enter the system via the initial conditions at the sampling time $\mathbf{x}(kT)$, i.e. the model is initialized by the actual system state. Since all state information is necessary to initialize the model, the full state must be either measured or estimated. In *in vitro tetO7*-Ub^{V76}GFP control experiments only the system output y_k (i.e. Ub^{V76}GFP fluorescence) can be measured, thus the system states $\mathbf{x}_k$ are reconstructed with a Kalman filter [37], implemented as described in section 2.4.

The solution to the optimal control problem in (7-6) is an array of N optimal duty cycles $\mathbf{d}_i$ and only the first element of this array (d_k in Figure **5-4**) is provided to the cells, in order to incorporate the feedback the optimal open loop input (d_k) is applying only until the next sampling instant $(k+1)T$ when the entire procedure (prediction and optimization) is repeated.

The optimization is carried out by adopting the Matlab implementation of the Genetic Algorithm described in [62].

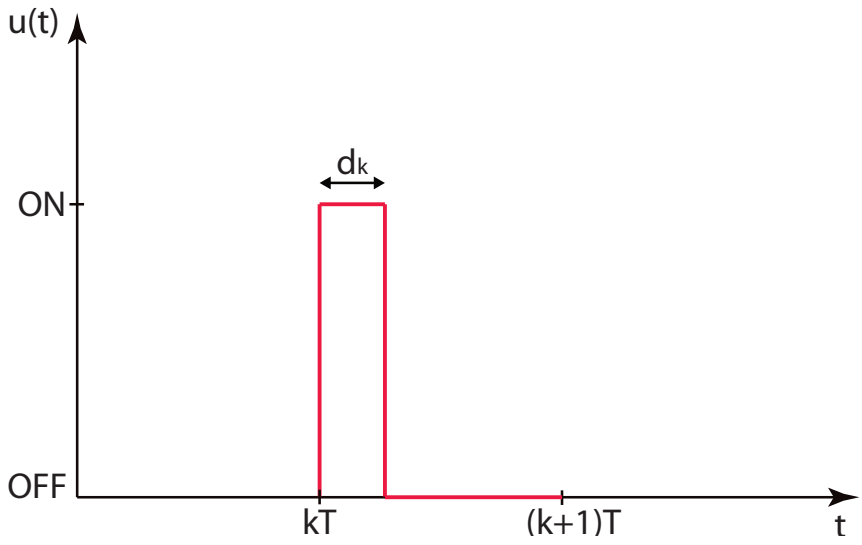


Figure **5-4**: **Discrete input signal.** During each sampling time T the input signal is a pulse whose duration t_{on} is defined by the duty-cycle $d_k = \frac{t_{on}}{T}$ with $d_k \in [0, 1]$.

5.3.3 Results

In order to test the MPC strategy described above, I simulated the control of *tetO7*-Ub^{V76}GFP gene expression by using the *Ub^{V76}GFP.DI model* in the Equation (5-4) as a proxy for the behaviour of *tetO7*-Ub^{V76}GFP cells. The control task was set−point regulation calculated as the 50% of the steady state of the model's output for a control input equal to 1 (i.e. cells grown in standard medium).

The result of simulation are reported in the Figure **5-5**. For the sake of clarity and to make comparison with the experimental results easier, the simulated tetracycline concentration is reported in the Figure **5-5** (B) as $(1 - u)$ where u is the input computed by the MPC controller inferred in Section 5.3.2. Thus, when the value of red signal is high, the cells are fed

with tetracycline, viceversa standard growth medium is provided to the cells when the red signal is low (Figure **5-5** (B)).

As shown in Figure **5-5** the controller is able to reach the control reference and to maintain its value without exhibiting oscillations at steady state, thus demonstrating that the developed control strategy is suitable to regulate the system under investigation.

The numerical results prompted the experimental implementation of the control strategy. To perform *in vitro* experiments, I integrated the devised control law into the experimental platform. The *tetO7*-Ub^{V76}GFP cells are loaded in the microfluidic device and allowed to grow in standard medium for 24 hours under perfusion conditions in cell culture incubator thus reaching their maximum level of fluorescence. Then the device is moved to the inverted microscope and the control experiment starts. At the beginning of each experiment, the cells are kept for 180 minutes in standard growth medium (calibration phase) in order to calculate their maximum fluorescence level, since the reference signal is expressed as a percentage of this value.

The control task is the same as for the numerical simulations, that is a set−point control task forcing the cell population to express an average fluorescence equal to 50% of their maximum fluorescence level calculated during the calibration phase.

As shown in Figure **5-6** the control action works effectively in keeping the output, namely the measured fluorescence, close to the desired set point for 2500 minutes.

Thus the results achieved *in vitro* (Figure **5-6**) are consistent with the corresponding numerical simulations (Figure **5-5**).

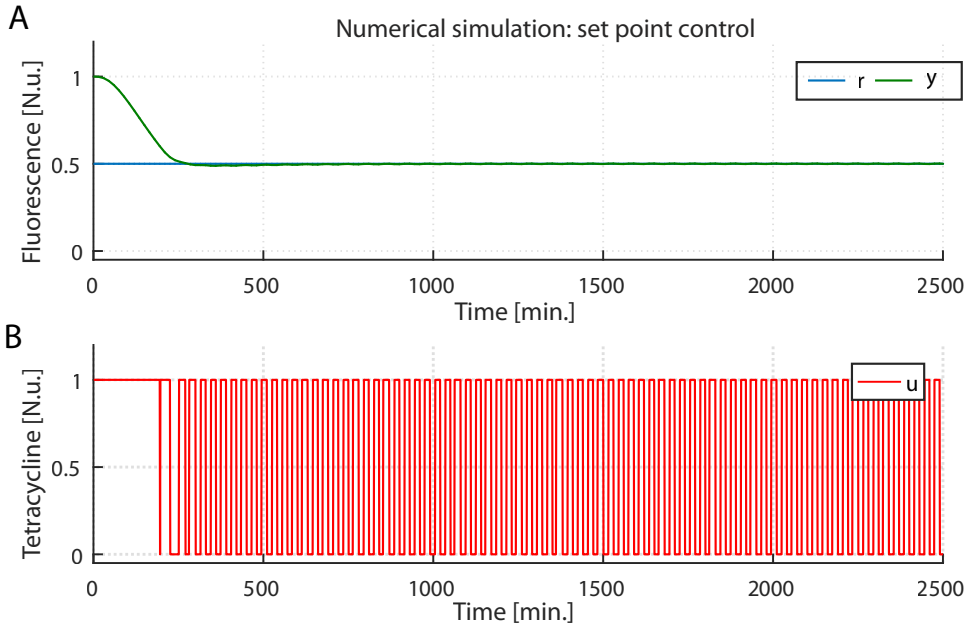


Figure **5-5**: **Numerical simulations of control *tetO7*-Ub^{V76}GFP model by means of MPC with discrete control input.** (A) The MPC algorithm is applied to control the dynamical model of the *tetO7*-Ub^{V76}GFP expression to a constant reference signal (r in blue). The set point is equal to 50% of the high steady state of the model's output, here normalized to 1. The green line is the system output y. (B) The control input is shown in red (u high level: tetracycline; low level: standard medium). The simulation was performed by controlling the dynamical *Ub^{V76}GFP.DI model* in the Equation (5-4).

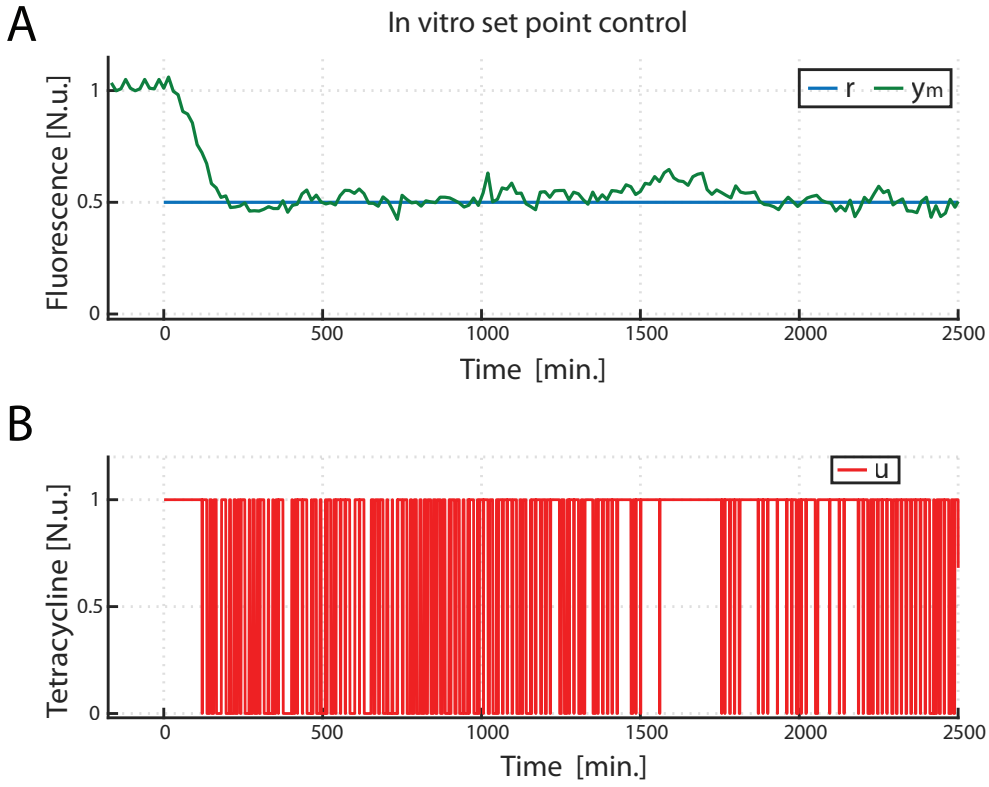


Figure **5-6**: ***In vitro* control of *tetO7*-Ub^{V76}GFP expression by means of MPC with discrete control input.** (A) *In vitro* set-point control experiment was performed on the *tetO7*-Ub^{V76}GFP cells. The control action starts at time $t = 0$. The set point (r in blue) is equal to 50 of the Ub^{V76}GFP fluorescence value quantified during the calibration phase, here normalized to 1. The green line is the average fluorescence measured over the population cells, y_m. (B) The input signal u, computed in real-time by the control algorithm, is shown in red: a high signal corresponds to tetracycline rich growth medium, a low signal to standard growth medium.

CHAPTER 6

Dynamic control of small molecule concentration in a microfludics device

In this Chapter, I describe the implementation of an active control strategy for regulating tetracycline concentration (or any other small molecule) within a microfluidic device. First, I identify a dynamical model describing the accumulation of tetracycline in the microfluidic device by using the experimental platform presented in Chapter 3. I then designed and implemented a Model Predictive Control strategy in order to regulate tetracycline concentration in the microfluidic device using as actuators two motorized syringes. Part of this work has been published in [63].

6.1 Motivation

The technological platform described in Chapter 3 allows to manipulate and observe mammalian cells in real-time thanks to a microfludics device with 33 cell chambers all connected to the same main channel that continuously delivers growth medium to the cells in the chambers (refer to Figure **3-2** in Chapter 3). Since the cell chambers are connected to the main channel by micro-channels perpendicular to it, cells are not subjected to a direct flow,

but rather the medium in the chambers is exchanged by diffusion. This prevents cell death due to mechanical shear stress on the cell membranes, at the cost of a slower "refresh rate" of the chamber content.

The flow with the main channel is constant but the on-chip Dial-A-Wave (DAW) hydrostatic pressure generator can be used to vary in real-time the concentration of a molecule of interest (e.g. tetracycline) within the main channel. Specifically the DAW consists of a set of microfludics channels connected to the main channel that allows to mix the content of two liquids in two syringes connected to two inlets by changing the relative height of the two syringes (refer to Figure **3-3** in Chapter 3). Specifically, if one syringe contains growth medium and the other the same growth medium but also a given concentration of a small molecule (e.g. tetracycline), it is then possible to modulate the concentration of compounds within the main channel (Figure **6-1** A,B).

In order to automate the platform, two vertically mounted linear actuators connected to the two syringes were designed and built to move independently. The motion is achieved through a stepper motor while the transmission by using a timing belt and two pulleys (for each of the rail). The actuator design is presented in Figure **6-2**. The syringes used in this design, have no plunger, since these are used simply as tanks to store the media. The details on the sizing of the stepper motors are presented in [21, 50].

From the experimental point of view, however, what is necessary is to be able to dynamically change the concentration of the small molecule in the cell chambers according to a desired reference, and not in the main channel. Since the main channel is connected to the chambers by microchannels perperdicular to it, diffusion phenomena from the main channel to the cell chamber must be taken into account in order to effectively control the concentration of the small molecule reaching the cells.

For these reasons, the technological platform presented in Chapter 3 was modified by including an active control of tetracycline concentration in the cell chambers.

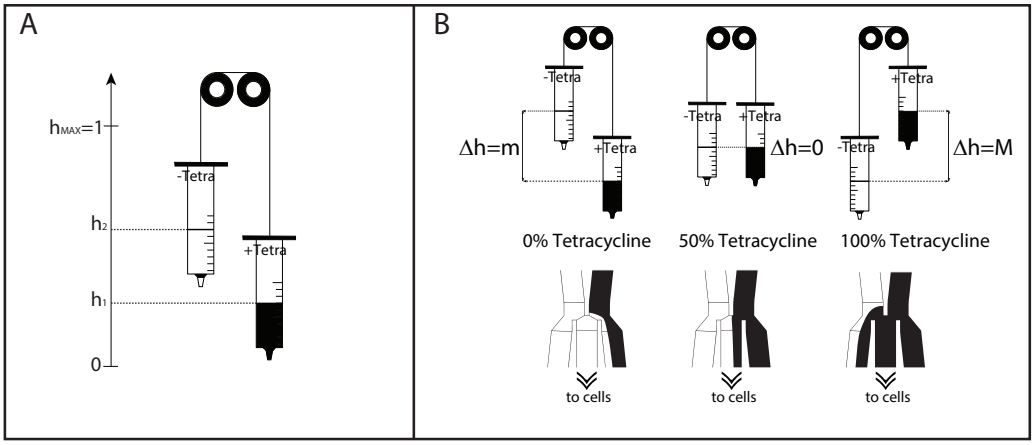


Figure **6-1**: **Actuation system and hydrostatic pressure effect in the DAW junction in the microfluidic device.** (A) Syringes are constrained to slide on pulleys. Regulating the stepper motors, each syringe is allowed to reach a desired height between the lowest 0 and the highest h_{max}. h_1 represents the height of the tetracycline filled syringe and h_2 is the height of the syringe with standard medium. The heights are taken at free surface of the medium. (B) For $\Delta h = h_1 - h_2 = m$ the tetracycline-medium/standard-medium mixing ratio is equal to 0%, meaning that only standard medium is given to the cells. For the dual situation ($\Delta h = M$ with $M = -m$) the mixing ratio is equal to 100% and tetracycline at its maximum concentration flows to the cells. When the two syringes are at the same level the mixing ratio is equal to 50%, thus the cells are fed with 50% of maximum tetracycline concentration.

The modeling and the control of the tetracycline concentration in the microfludics cell chambers will be the topics of the next sections.

6.2 Modeling tetracycline accumulation in cell chambers

The identification of the parameters of mathematical model able to capture the dynamics of the tetracycline accumulation in cells chambers has been carried out without assuming any *a priori* knowledge of the tetracycline flow in the microfluidic device. Input and output data have been used to fit the

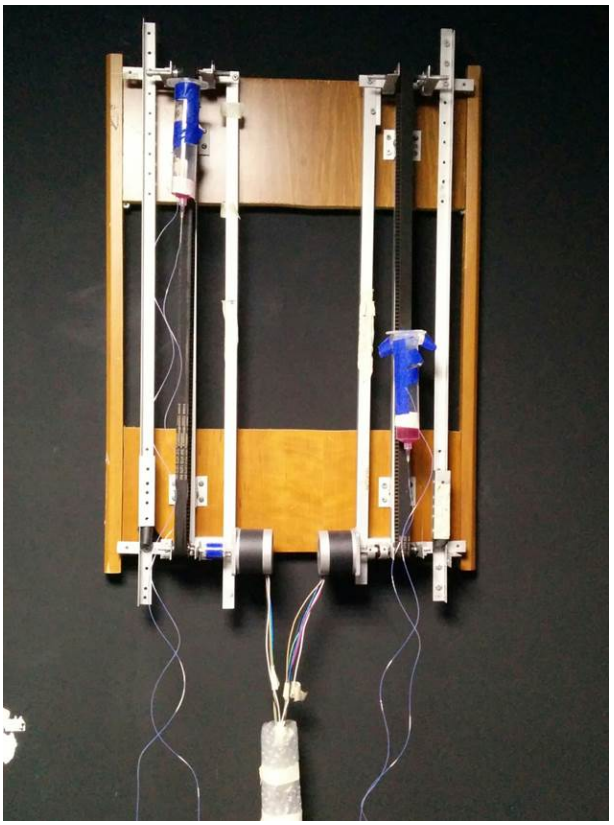


Figure **6-2**: **Motorised linear rails.** In this picture the two designed linear rails are presented. The external frame sustains the holder for the stepper motor and the two pulleys on which the timing belt is meant to slide. The sliding element is mechanically secured to the belt in order to allow coherent movements.

model structure, thus considering the system as a "black box" [34].

The strategy I followed is to dynamically modulate the differential heights between the two syringes, as input to the system, and to follow the dynamics of the tetracycline accumulation in the cells chambers in response to such an input (Figure **6-3**).

The concentration of tetracycline is tracked by adding a red fluorescent dye (Sulforhodamine) to the tetracycline-filled syringe and the medium fluorescence in the microfluidic chambers is imaged by a time lapse microscope at regular sampling time. Therefore it is possible to obtain a time profile

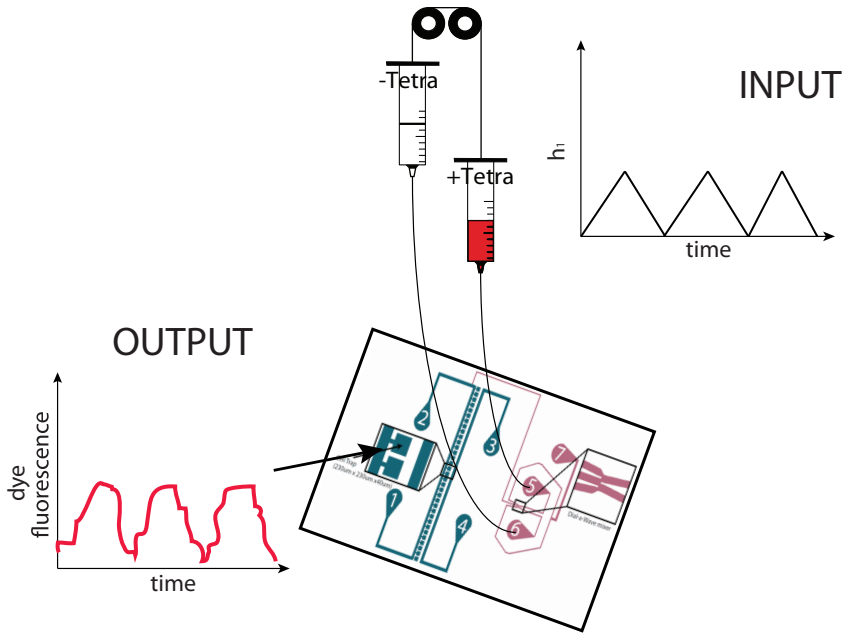


Figure **6-3**: **Technological platform.** The relative heights between the two syringes (filled with or without tetracycline) changes, thereby varying the tetracycline concentration in the microfluidic device. A fluorescent dye is added to the syringe with tetracycline. The dye fluorescence is imaged by a time lapse microscope and it is quantified to follow the dynamics of the tetracycline accumulation in the cells chambers in response to the variation of the differential heights of the automated syringes.

of the actual tetracycline concentration flowing in the microfluidic chambers by measuring the fluorescence of the medium in the red spectrum. The level of red fluorescence is considered as the output of the system.

For the sake of clarity, in this discussion the input signal is normalized between 0 and 1 as shown in Figure **6-1**. Thus, if we call h_1 and h_2 the current height of the syringe with tetracycline and with standard medium, respectively, we have

$$\begin{aligned} \bar{h}_1 &= \frac{h - h_{min}}{h_{max} - h_{min}}, \\ \bar{h}_2 &= 1 - \bar{h}_1 \end{aligned} \tag{6-1}$$

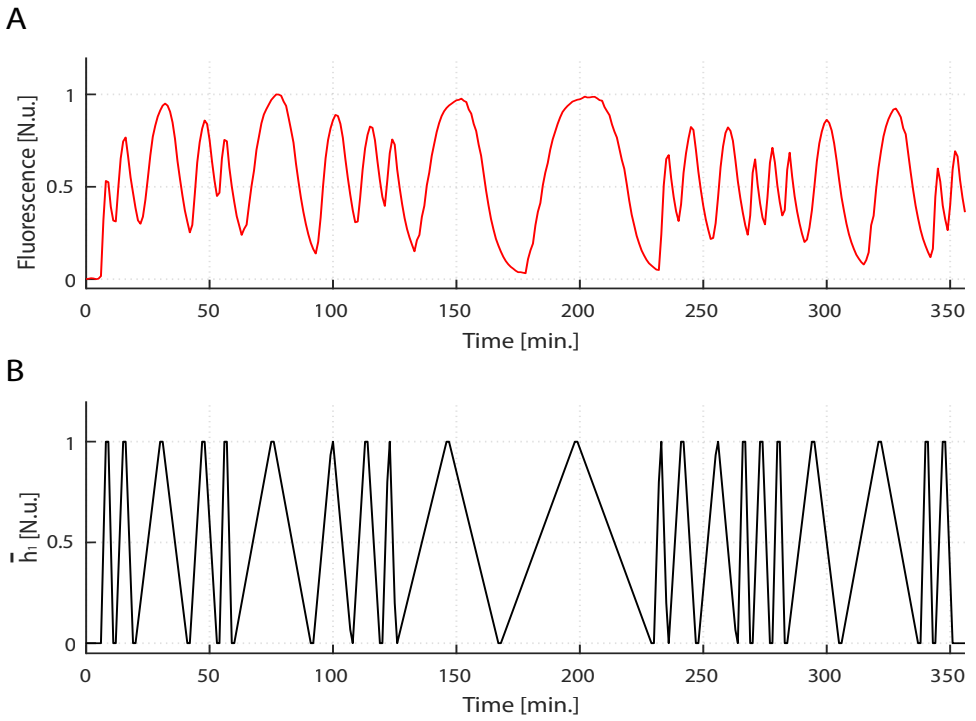


Figure **6-4**: **Experimental data.** (A) The red signal is the fluorescence of tetracycline rich medium measured during the identification experiment. This signal is considered the output of the system. (B) The input of the system (black line) is represented by the relative height between the syringes on the linear rail. $\bar{h}_1$ is the height of the tetracycline-filled syringe, here normalized between 0 and 1. The height of the syringe with standard medium is $\bar{h}_2 = 1 - \bar{h}_1$.

where h_{min} is the lowest height that a syringe can reach on the rail, h_{max} is the highest, $\bar{h}_1$ and $\bar{h}_2$ are the normalized heights.

Experimental data. The fluorescence of the tetracycline rich medium in the microfluidic chambers is captured every minute for up to 350 minutes (Figure **6-4** A). During this time, the relative height between the two syringes is varied according to the a triangular wave with several frequencies (Figure **6-4** B).

The representative data set shown in Figure **6-4** is consistent with the

expected behaviour of the DAW channel in the microfluidic device. Indeed when the syringe with tetracycline rich medium is in its high position on the rail ($\bar{h}_1 = 1$), the maximum level of tetracycline concentration flows in the microfluidic chambers. Viceversa when this syringe occupies the lowest position respect to the syringe with standard medium ($\bar{h}_1 = 0$), the tetracycline is absent in the cells chambers. Finally, by changing the respective position between the two syringes on the linear rail, any tetracycline concentration can flow in the microfluidic device.

Black box identification. State space model. I have used the technological platform depicted in Figure **6-3** and the experimental data shown in Figure **6-4** to identify the following one dimensional state space model describing the accumulation of tetracycline in the microfluidic device:

$$\begin{aligned} \dot{x} &= ax + bu, \quad x(0) = x_0, \quad u \in [0, 1] \\ y &= cx \end{aligned} \tag{6-2}$$

In the model 6-2, x is the system state, u is the system input (the relative height of the syringes on the linear rail), y_k is the system output (tetracycline accumulation in the microfluidic chambers) and x_0 is the initial condition.

The system order n, the model parameters and the initial condition were estimated from the data set in Figure **6-4** by applying the *prediction error minimization* algorithm (PEM, [34]).

The estimated model coefficients and the initial condition are:

$$a = -0.2234, \tag{6-3}$$

$$b = 0.0506, \tag{6-4}$$

$$c = 4.9643, \tag{6-5}$$

$$x_0 = 1.12 \tag{6-6}$$

The model was able to predict the experimental data across the identification scenario, as depicted in Figure **6-5**. Observe that according to parameter a, the settling time for tetracycline within the cell chamber is of the order of 20 minutes.

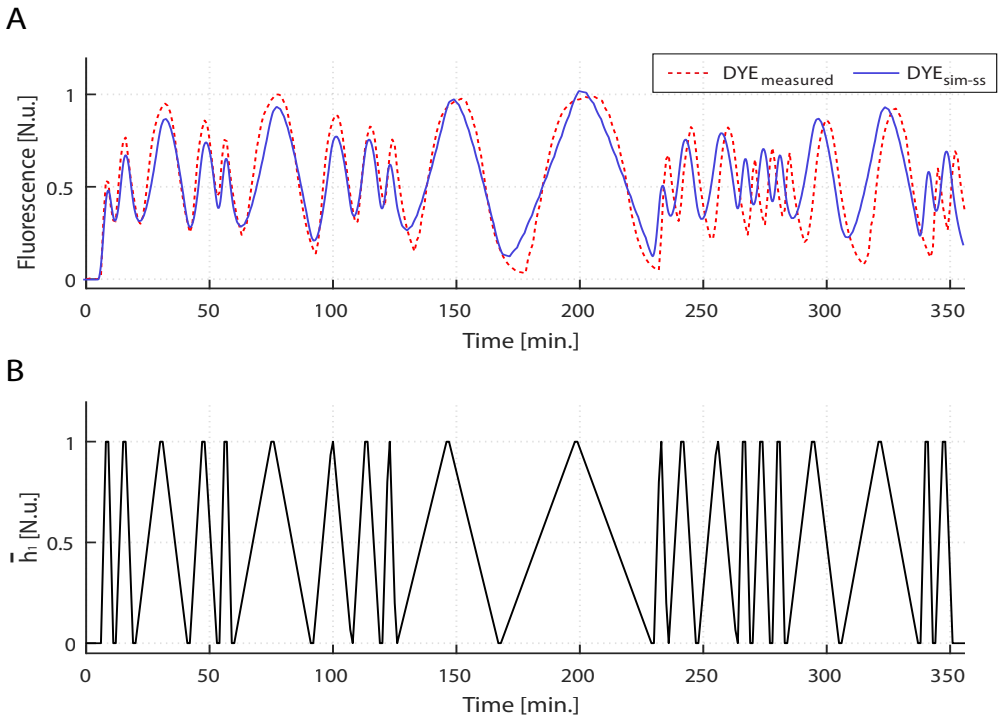


Figure **6-5**: **Fitting.** (A) The solid blue line represents the simulated output of the state space model identified with *prediction error minimization* algorithm and the red dashed line is the fluorescence of the tetracycline rich medium measured during the identification experiment (B) The input of the system (black line) is represented by the relative height of the syringes on the linear rail. This signal is used both to identify and to validate the model.

6.3 Dynamic control of small molecule concentration in cell chambers with an Model Predictive Control strategy

Once I inferred a dynamical model of tetracycline accumulation in the microfluidic cell chambers, I aimed at synthesizing a negative feedback control strategy to dynamically regulate the actual concentration of tetracycline in the microfluidic cell chambers using the experimental platfrom depicted in Figure **6-6**.

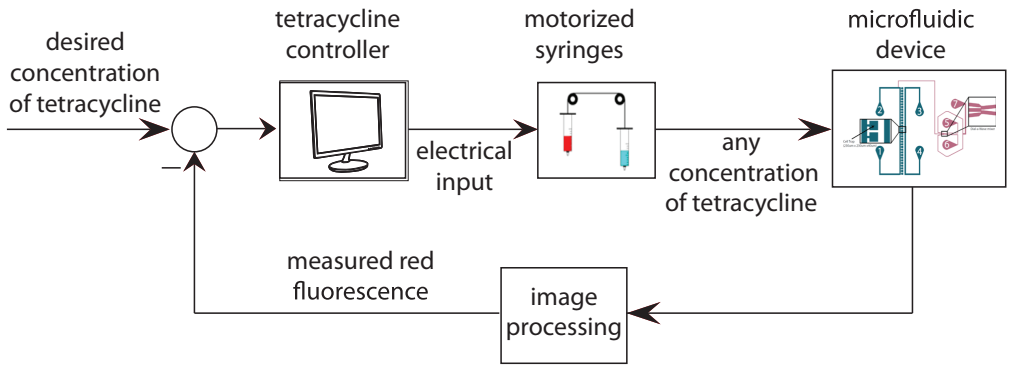


Figure **6-6**: **Negative feedback control of tetracycline concentration in microfluidics cell chambers.** The desired tetracycline concentration is the reference signal. The red fluorescence is measured at regular sampling time of 1 minute. The measured red fluorescence is compared to the reference and a MPC strategy modulates the differential positions of the automated syringes on the linear rail in order to have a the desired tetracycline concentration reaching the cells chambers in the microfluidic device.

The control input is a train of DC voltage pulses delivered to the two stepper motors guiding the two syringes on the linear rail in order to change their differential height and hence the concentration of tetracycline reaching the cells chambers via the Dial-a-Wave (DAW) mixing channel in the microfluidic device (see Figures **6-1** and **6-6**).

6.3.1 Model Predictive Control

In order to design the controller, I decided to implement a Model Predictive Control (MPC). The choice of the MPC was dictated by the constrain in the control input which is bounded, since the rails where the syringes are positioned has finite length. In the following, the relative height if the two syringes is normalized between 0 and 1 as shown in Figure **6-1**.

As described in Chapter 2, the MPC is an optimization based strategy

that predicts the future evolution of the process under investigation and optimizes the control input by minimizing at each sampling time a function of the control error over a finite future horizon.

In order to predict the future response of the controlled plant, I used the following discretised version of the tetracycline accumulation model in Equation (6-2) obtained by applying the zero-order hold (ZOH) method described in [61]:

$$\begin{aligned} &x_{k+1} = ax_k + bu_k, \quad x(0) = x_0 \quad u_k \in [0,1] \\ &y_k = cx_k \end{aligned} \tag{6-7}$$

where $a = 0.800$, $b = 0.045$, $c = 4.9643$ and $x_0 = 1.12$.

In the MPC implemented for the regulation of tetracycline, the cost function to be minimizes is the Sum of the Squared control Error (SSE), defined as

$$\begin{aligned} SSE_k &= \sum_{i=k+1}^{k+N} \left(N+1+k-i\right) \epsilon_i^2 \\ &= \sum_{i=k+1}^{k+N} \left(N+1+k-i\right) \left(r_i - y_i\right)^2 \end{aligned} \tag{6-8}$$

where the integer N defines the length of the prediction horizon in terms of sampling intervals (set as N = 15, corresponding to 15 minutes) and y_i is the output of the dynamical model describing the tetracycline accumulation in the cells chambers.

Therefore, at each sampling time kT the MPC algorithm solves the following optimal control problem over the prediction horizon of N steps starting from the current system state $x(KT)$:

$$
\begin{aligned}
&\min_{\{u_{k+1},\ldots,u_{k+N}\}\in[0,1]} SSE_k \\
&s.t. \\
&\quad x_{i+1} = ax_i + bu_i \\
&\quad y_i = Cx_i \\
&\quad x_k \quad is \quad given \\
&\quad y_i \in [0,1] \\
&\quad i \in \{k+1,\ldots,k+N\}
\end{aligned}
\tag{6-9}
$$

The optimization is carried out with *fmincon* routine of the MATLAB System Idenfication toolbox (Mathworks Matlab R2016b). This function solves linear least squares problem with non negative variables using the *active-set* algorithm described in [64]. The result of the optimization is an array of N optimal control inputs. The feedback loop is closed by applying only the first element of the calculated array of control inputs and at the next sampling time the entire procedure is repeated.

6.3.2 Experimental validation of the MPC strategy.

To perform the control experiments, I integrated the described control law into the experimental platform shown in Figure **6-6**. I then filled one syringe with standard growth medium and one syringe with standard growth medium plus tetracycline at a concentration of $100ng/\mu$l and sulforhodamine, a non toxic fluorescent dye emitting in the red spectrum, at concentration of 1μM in order to quantify tetracycline concentration via fluorescence microscopy.

At each sampling time ($T = 1$ min), an image is acquired by the microscope and the normalized red fluorescence in the microfluidic chambers is quantified. The actual concentration of tetracycline, which is proportional to the level of the red fluorescence in the microfluidic chambers, is compared to the reference. The control input is then computed by using the implementation of the MPC discussed in 6.3.1 thus determining the heights of the syringes on the linear rail through the stepper motors.

The control task consists in signal tracking of an descending staircase. The reference r is a descending staircase of 3 steps where each step has a duration of 50 min and an amplitude, respectively, of a 75%, 50% and 25% of the maximal red fluorescence (here normalized to 1) measured during a calibration phase of 10 min.

The results of the control experiments are reported in Figure **6-7**.

It can be appreciated in Figure **6-7** that MPC strategy described in 6.3.1 is not able to bring the tetracycline level (red line, y) to the desired value (reference, r) and a non zero control error is present over the entire control experiment. I found that this offset is due to the autofluorescence of the growth medium. Indeed, even in the absence of red fluorescent dye, the standard medium without tetracycline has an intrinsic fluorescence (named autofluorescence). However, this autofluorescence can be considered as a constant disturbance on the output and thus I moved to develop an offset-free Model Predictive Control as described in [65] in order to ensure a zero control error at steady state.

6.4 Dynamic control of small molecule concentration in cell chambers with an Offset-Free Model Predictive Control strategy

In order to remove offset, in [65] the authors design a control system that can remove asymptotically constant, nonzero disturbances on the output. To this end, the original dynamical model of the process being controlled is augmented with an additional state representing the unknown constant disturbance d_k.

Thus the discrete model (6-7) describing the accumulation of tetracycline in the microfluidic chambers, is augmented as follows:

$$\begin{aligned} &\begin{bmatrix} x_{k+1} \\ d_{k+1} \end{bmatrix} = \begin{bmatrix} a & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_k \\ d_k \end{bmatrix} + \begin{bmatrix} b \\ 0 \end{bmatrix} u_k, \quad \begin{bmatrix} x(0) \\ d(0) \end{bmatrix} = \begin{bmatrix} x_0 \\ d_0 \end{bmatrix}, \quad u_k \in [0,1] \\ &y_k = \begin{bmatrix} c & 1 \end{bmatrix} \begin{bmatrix} x_k \\ d_k \end{bmatrix} \end{aligned} \tag{6-10}$$

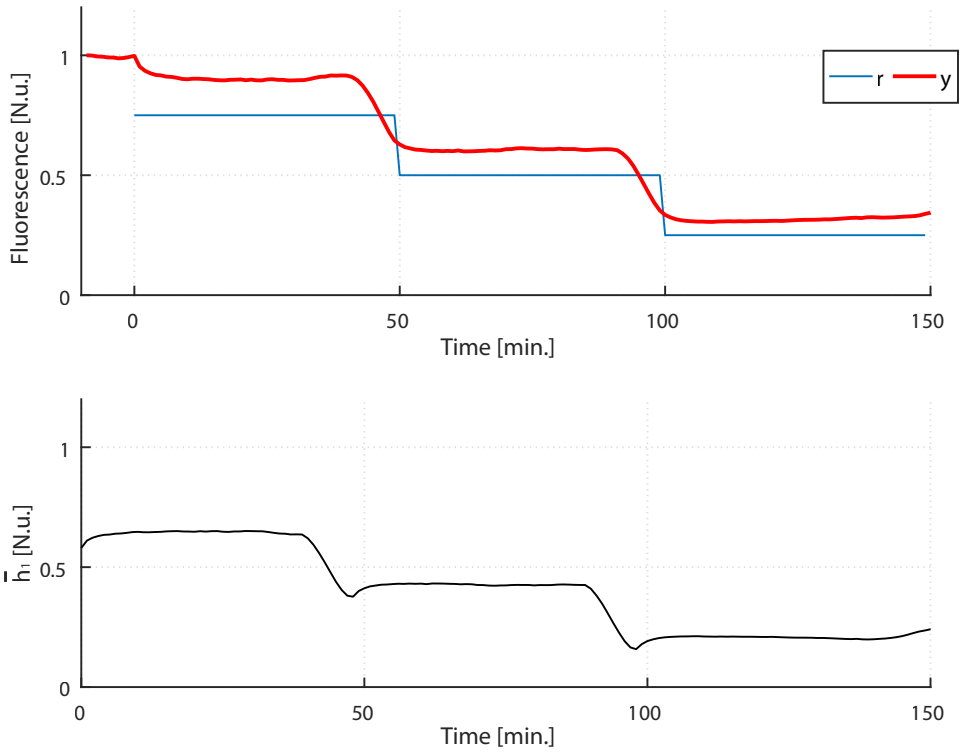


Figure **6-7**: **Experimental results following MPC of tetracycline concentration in the microfludics device.** (A) The control experiment starts at time $t = 0$ and lasts for 150 minutes. The reference r in blue is a step-like signal. Each step is calculated as a percentage (75%, 50% and 25%) of the fluorescence intensity quantified during the calibration phase of 10 minutes at the beginning of the experiment, here normalized to 1. The red line is the measured fluorescence signal (model output, y). (B) The control input is the black line and represents the relative height between the tetracycline-filled syringe and standard medium-filled syringe on the rail. $\bar{h}_1$ is the height of the tetracycline-filled syringe, here normalized between 0 and 1. The height of the syringe with standard medium is $\bar{h}_2 = 1 - \bar{h}_1$.

I then designed a Kalman filter as described in Section 2.4 in order to estimate the state x_k, and the disturbance d_k, given the output y_k.

In order to validate the Offset-Free Model Predictive Control strategy, I first validated the Offset-Free Model Predictive Control strategy described above by simulating the tetracycline accumulation in the microfluidic chamber using the dynamical model (6-2).

The control task was to follow two difference references r: (i) a step-like signal where each step lasts 100 minutes and is set equal to 75%, 50% and 25% of the initial red fluorescence value (Figure **6-8**) and (ii) a linear descending ramp of 150 minutes starting at the 100% of initial fluorescence and decreasing down to 25% (Figure **6-9**).

As I already described in section 6.3.2, tetracycline is tracked by adding a red fluorescent dye to the tetracycline-filled syringe, thus the value of the average fluorescence in the microfluidic chamber is the measured output of the system and it is proportional to the tetracycline concentration.

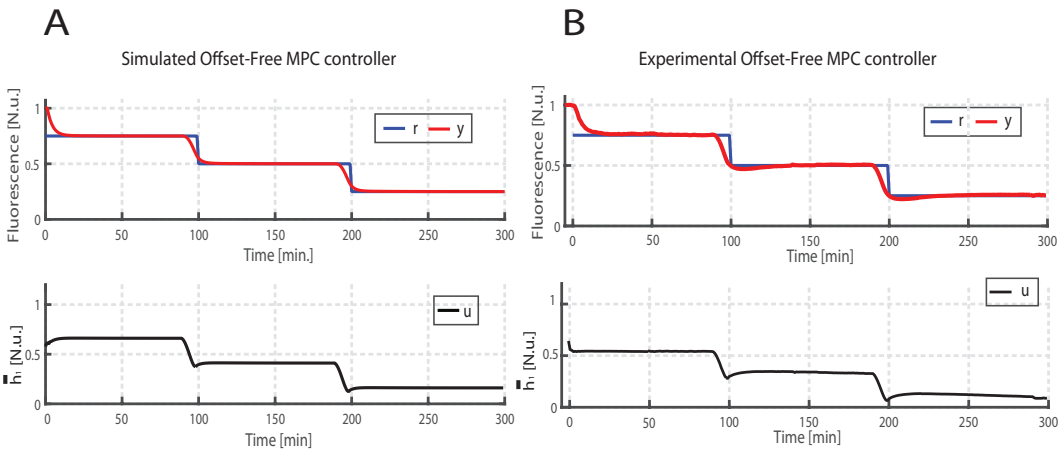
In the simulated experiments (Figures **6-8** A and **6-9** A), the initial value of the red fluorescence is calculated as the steady state of the output of model in the absence of tetracycline (6-2). The numerical simulations indicate that Offset-Free MPC strategy is able to steer tetracycline concentration to the desired value.

Therefore, I decided to perform the same control tasks experimentally. The results of experiments are shown in Figures **6-8** B and **6-9** B. At the beginning of each experiment, a calibration phase of 10 minutes is used to quantify the maximum level of red fluorescence, which is then used to set the unit of measure of the reference signal.

The experimental results confirm those obtained in simulations, as depicted also from the performance indices (Integral Square Error ISE, Integral Absolute Error IAE, Integral Time Absolute Error ITAE) in the Figures **6-8** C and **6-9** C.

In conclusion the control strategy to regulate tetracycline concentration in the microfluidic device presented in this Chapter allows to deliver any time-varying concentration of tetracycline (or any other small molecule) to

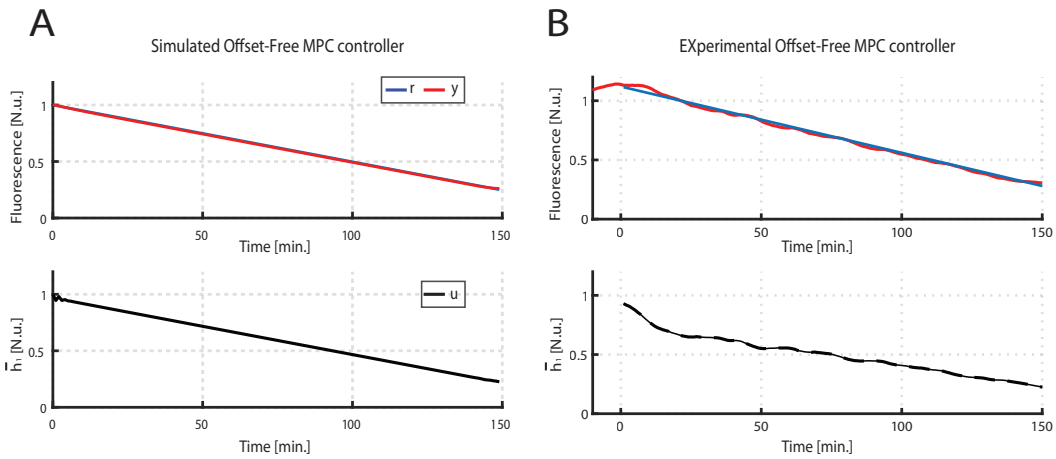
the cells growing in the cell chambers. This feature will be exploited in in Chapter 7 to control gene expression from a tetracycline inducible promoter.



C

	Simulated Offset-Free MPC	Experimental Offset-Free MPC
ISE	0.5462	0.9183
IAE	3.8670	7.1317
ITAE	350.5325	799.3528

Figure **6-8**: **Offset-free MPC staircase tracking control of tetracycline accumulation.** (A) *in silico* control experiment performed on the tetracycline accumulation dynamical model. The reference (r in blue) is a three steps signal where each step is calculating as a percentage (75%, 50% and 25%) of the high steady state of the model's output here normalized to 1. The green line is the model's output (y). (B) *in vitro* control experiment. The control action starts at time $t = 0$ min. The reference (r in blue) is a percentage of the initial value of medium red fluorescence calculated in the calibration phase. The red line is the measured red fluorescence (y). (C) Performance indices calculated over the control time interval: Integral Square Error (ISE), Integral Absolute Error (IAE), Integral Time Absolute Error (ITAE).



C

	Simulated Offset-Free MPC	Experimental Offset-Free MPC
ISE	0.0041	0.0408
IAE	0.7593	1.7972
ITAE	56.1999	97.2031

Figure **6-9**: **Offset-free MPC tracking control of tetracycline accumulation.** (A) *in silico* control experiment performed on the tetracycline accumulation dynamical model. The reference signal (r in blue) a linear descending ramp of 150 min. The green line is the model's output (y). (B) *in vitro* control experiment. The control action starts at time $t = 0$ min. The desired level of fluorescence (r, in blue) is a linear descending ramp starting at 100% of the initial value of medium fluorescence and going down to 25%. The red line is the measured red fluorescence (y). (C) Performance indices calculated over the control time interval: Integral Square Error (ISE), Integral Absolute Error (IAE), Integral Time Absolute Error (ITAE).

CHAPTER 7

Model Predictive Control of Gene Expression from the *tetO7*-Ub^{V76}GFP Synthetic Network with a Continuous Control Input

In this Chapter, I describe the implementation of a control strategy equipped with a continuous actuation by enabling any desired tetracycline concentration to be provided to the cells. To validate the implementation of this control strategy, I first carry out numerical simulations and I compare the performance achieved by the controller with either discrete or continuous actuation. Finally, I present preliminary experimental results of *in vitro* set point control experiment using continuous actuation. Part of this work has been published in [63].

7.1 Derivation of the mathematical model for *tetO7*-Ub^{V76}GFP network with continuous input

The experimental results described in Chapters 4 and 5 convincingly demonstrate that the expression of a protein can be controlled *in vitro* in real-time, using an inducer molecule, by applying a discrete actuation signal, i.e. either standard medium or medium with tetracycline at concentration of 100 ng/ml can be provided to the cells.

However, it is important to note that for the biological models in exam the concentration of tetracycline to be supplied to the cells could be graded, obtaining thus a continuous control input which may lead to improved control performances. This would require a highly refined calibration of the actuator that can be avoided by introducing in the experimental platform the tetracycline concentration controller implemented in the Chapter 6 .

As a testbed for the controller equipped with a continuous actuation, I chose to control the *tetO7*-Ub^{V76}GFP system described in Chapter 5. In this system, the tetracycline Transactivator (tTa) protein is expressed from a constitutive *CMV* promoter and it drives the expression of the destabilised reporter protein Ub^{V76}GFP from the inducible *CMVTET* promoter on a separate transcriptional unit. If tetracycline is added to the culture medium, it binds to tTa blocking its activity and thus the expression of Ub^{V76}GFP protein. The *tetO7*-Ub^{V76}GFP system can be viewed as a single input-single output (SISO) dynamical system. The input $u(t)$ is the concentration of tetracycline in the growth medium and the output $y(t)$ is the level of the Ub^{V76}GFP fluorescence across the cell population.

This system can be described by the *Ub^{V76}GFP.DI model* in Equation (5-4) inferred in the Chapter 5, but the effect of inducer molecule (tetracycline) on the *tetO7*-Ub^{V76}GFP system is now modelled by the Hill function

[15], thus obtaining the following dynamical model:

$$\begin{aligned}\frac{dx_1}{dt} &= -d_1 x_1 + \beta_1 \left(\frac{k^n}{k^n + u^n} + \gamma\right) \quad u \in [0,1] \\ \frac{dx_2}{dt} &= \alpha_2 x_1 - d_2 x_2 \\ \frac{dx_3}{dt} &= \alpha_3 x_2 - d_3 x_3\end{aligned} \tag{7-1}$$

where x_1 is the concentration of Ub^{V76}GFP mRNA, x_2 and x_3 are the concentrations of Ub^{V76}GFP unfolded protein and Ub^{V76}GFP folded protein, respectively. β_1, α_1 and α_2 represent the production rates and d_1, d_2 and d_3 are the degradation rates. u represents the tetracycline concentration while the system output (i.e. the measured fluorescence of the functional Ub^{V76}GFP protein) is x_3.

In order to obtain the model parameters, I performed two steps identification: (i) first I identified the coefficients of the Hill function $g(u) = \frac{k^n}{k^n+u^n} + \gamma$ and (ii) then the other model parameters were identified by using a grey-box identification approach from experimental data.

For identifying the Hill function coefficients, I took advantage of the tetracycline control platform discussed in the Chapter 6 to perform an automatic dose-response experiment (Figure **7-1**) by measuring cells fluorescence in response to tTa protein at several concentrations of tetracycline. In order to track tetracycline presence in the cell chambers, a fluorescent dye (Sulforhodamine) is added to the tetracycline rich medium. Cells' fluorescence is imaged at 15 minutes intervals by the time lapse microscope and it is quantified by image segmentation algorithm in Appendix A (Figure **7-1** A). I assumed that at each tetracycline concentration the value of fluorescence reaches the steady state level at the end of the step (Figure **7-1** B) and the coefficients of the Hill function were fitted to these experimental data, as shown in the Figure **7-2**.

In order to identify the other model parameters, I performed a grey-box identification by using *nlgreyest* function from MATLAB System Idenfication toolbox (Mathworks Matlab R2016b) from the input-output data in Figure **7-3** obtained as discussed in the section 5.2. I named the identified

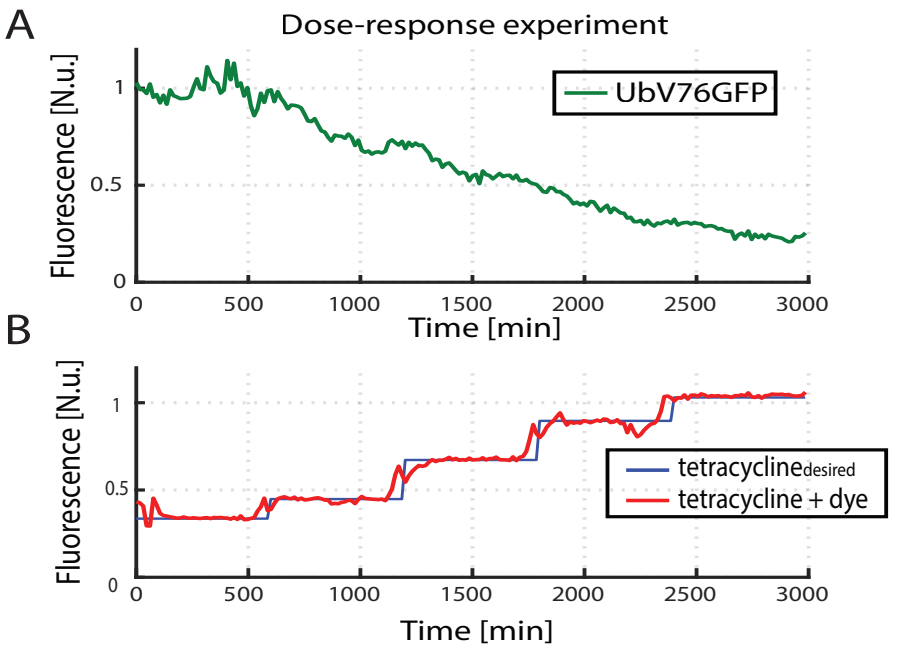


Figure **7-1**: **Automatic microfluidic-based dose-response experiment.** The variation of cells fluorescence in response to several tetracycline concentrations. (A) The average fluorescence across cell population is measured at 15 minutes intervals (green line).(B) Several concentration of tetracycline are given to the cells by means of the motorized syringes. The controller of tetracycline accumulation in microfluidic device ensures that the actual tetracycline concentration in the microfluidic device (red line) is equal to the desired value (blue signal).

model in Equations (7-1) $Ub^{V76}GFP.CI$ *model* where CI stands for Continuous Input.

A list of all model parameters and the performance indices AIC, FPE and FIT%, defined in Chapter 5, are reported in Tables **7-1** and **7-2**.

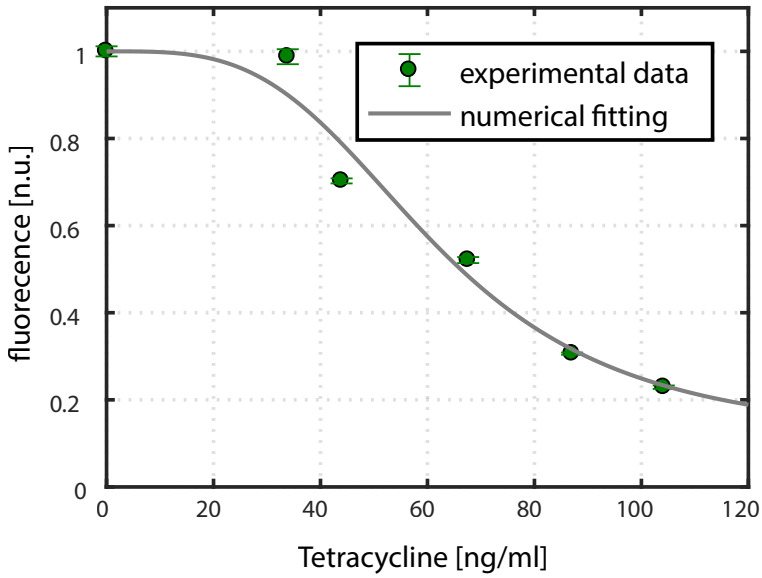


Figure **7-2**: **Hill function describing the effect of tetracycline on the *CMVTET* promoter.** Numerical fitting of the dose response curve of the *CMVTET* promoter to several concentration of tetracycline. The green dots are the experimental data and the black line is the numerical fitting of the HIll function to the data.

Table **7-1**: Model parameters

Parameter	Description	Value
$d_1\,[min^{-1}]$	degradation rate of Ub^{V76}GFP mRNA	0.0098
$d_2\,[min^{-1}]$	degradation rate of Ub^{V76}GFP unfolded protein	0.0199
$d_3\,[min^{-1}]$	degradation rate of Ub^{V76}GFP folded protein	0.0204
$\alpha_2\,[min^{-1}]$	production rate of Ub^{V76}GFP unfolded protein	0.0139
$\alpha_3\,[min^{-1}]$	folding rate of Ub^{V76}GFP protein	0.0172
$\beta_1\,[min^{-1}]$	CMVTET maximum transcriptional rate	0.0117
$\gamma\,[a.u.]$	CMVTET leakiness	0.011
$k\,[a.u.]$	Tetracycline concentration to achieve $\frac{\beta_1}{2}$	0.62
n	Hill coefficient	3.45

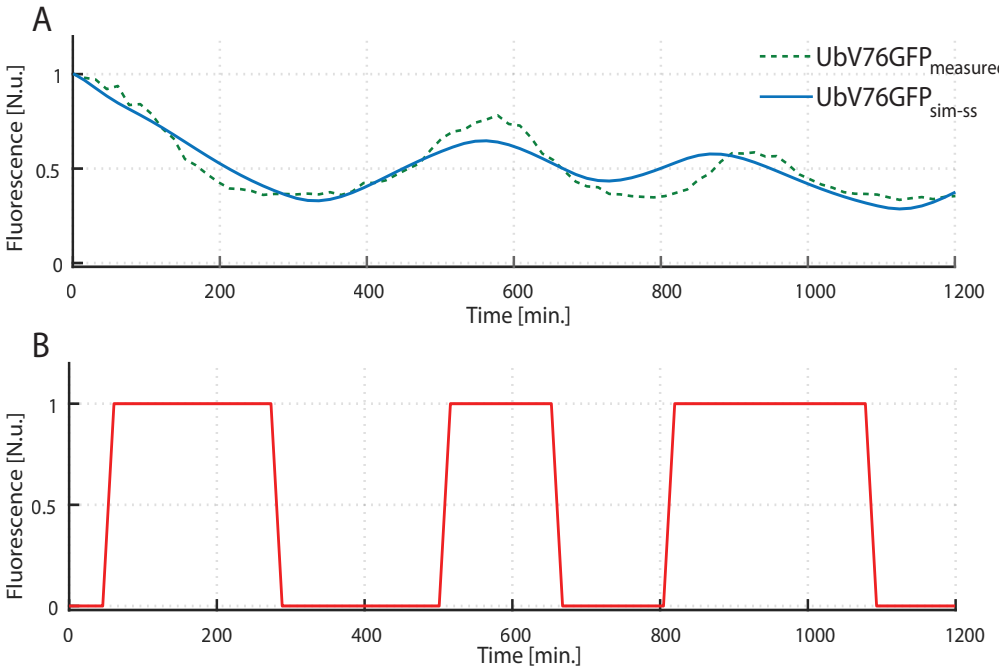


Figure **7-3**: **Identification experiment.** (A) The solid blue line is the output of the state space model identified with *nlgreyest* function from MATLAB System Idenfication toolbox (Mathworks Matlab R2016b). The green dashed line is the average cell fluorescence measured during the experiment. (B) The input (red line) is the tetracycline pulses. This signal is used both to identify and validate the obtained model. The red signal is the fluorescence of the dye added to the tetracycline rich medium measured during the experiment.

Table **7-2**: Values of the performance indices for $Ub^{V76}GFP.CI$ *model.*

FPE	$5.58 * 10^{-3}$
AIC	-5.1890
FIT	62.3%

7.2 Controlling *tetO7*-Ub^{V76}GFP gene expression with a continuous control input

In order to implement a control strategy to steer expression of the Ub^{V76}GFP with a continuous control input, I took advantage of the integrated exper-

imental platform described in Chapter 3. However, in the original control scheme, a discrete control input was used, that is only tetracycline rich medium or standard medium was provided to the cells. In order to move from a discrete (binary) control input to a continuous control input, I needed to finely control the action of the actuators (motorized syringes) and consequently the tetracycline concentration in the microfluidic device. I thus modified the technological platform as shown in Figure **7-4** to include an active control of tetracycline concentration in the microfluidic cell chambers.

As Figure **7-4** shows, the original platform is endowed with an inner loop (tetracycline concentration control loop) in which the measured level of tetracycline (u) is compared to the desired tetracycline concentration (u_d), i.e. the continuous control input computed by the gene expression controller. The tetracycline controller modulates the relative height of the syringes on the linear rails in order to provide cells with the desired amount of tetracycline. More details on the tetracycline concentration controller are reported in Chapter 6.

7.2.1 Control objective and controller design

The control tasks were either set point regulation or signal tracking. In the set point regulation, the reference value $r(t)$ is calculated as the 50% of the initial average fluorescence expressed by the cells in the **calibration phase** of 180 minutes at the beginning of the experiments. In the signal tracking contro,l I used two different references r: a sinusoidal wave of period $T = 3000$ min defined as $s = 0.5 + 0.25 * sin(2\frac{\pi}{T}(t-100) + \frac{\pi}{2})$ and a staircase function beginning at 50% of initial average fluorescence, then going to 25% and then 75%.

As the system output $y(t)$ I considered the average level of fluorescence in the cell population in order to reduce the noise measurements due to the cell-to-cell variability in the protein production.

The results achieved controlling gene expression by using a model predictive control with a discrete control input, as detailed in Chapter 5, convinced

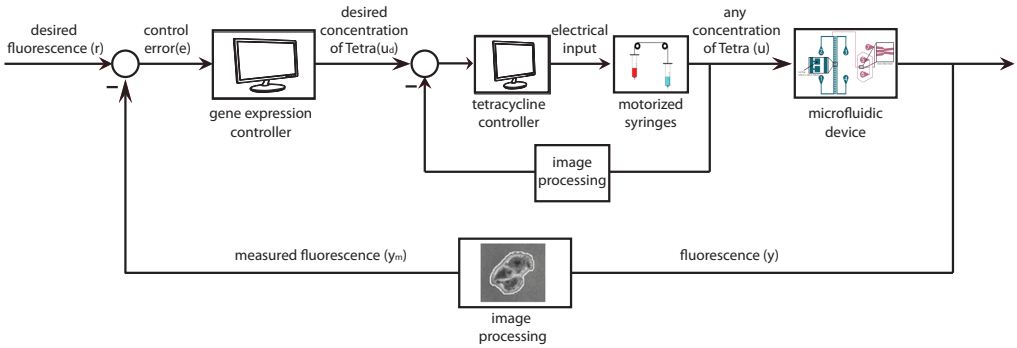


Figure **7-4**: **Technological control platform for continuous control input.** The experimental platform described in Figure **4-3** is endowed with a tetracycline controller that makes the actual concentration of input molecule (u) equal to the desired concentration (u_d) computed by the gene expression controller. The concentration of tetracycline in the growth medium is quantified by the addition of a fluorescent dye to the syringe containing tetracycline and it is imaged by the microscope at regular sampling time and converted to a quantitative value via an image processing algorithm.

me to implement an MPC based strategy also when regulating gene expression with a continuous control input.

As already discussed in the Chapter 2, the Model Predictive Control (MPC) algorithm is an iterative optimization-based strategy that, at each sampling time kT, uses mathematical model of the plant to predict the future value of a cost function over a finite time prediction horizon. The MPC strategy then calculates the optimal control input over the prediction horizon that minimizes the cost function.

I implemented the MPC strategy using a discretised version of the $Ub^{V76}GFP.CI$ *model* in Equation 7-1. Since the considered dynamical model is non linear in the input, I first performed an input transformation and then applied the zero-order hold method to convert the continuous model to a discrete one, as described in [61], thus obtaining the following discrete linear model:

$$\begin{aligned} \mathbf{x}_{k+1} &= \mathbf{A}\mathbf{x}_k + \mathbf{B}u_k, \quad u_k \in [0,1] \quad \mathbf{x}(0) = \mathbf{x}_0 \\ y_k &= \mathbf{C}\mathbf{x}_k \end{aligned} \tag{7-2}$$

where $\mathbf{x}_k = \begin{pmatrix} x_1(kT) \\ x_2(kT) \\ x_3(kT) \end{pmatrix}$ are the system states, $u_k = u(kT)$ is the input and $y_k = y(kT)$ is the system output. The matrices $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{C}$ are:

$$\mathbf{A} = \begin{pmatrix} a_{11} & 0 & 0 \\ a_{2,1} & a_{2,2} & 0 \\ a_{3,1} & a_{3,2} & a_{3,3} \end{pmatrix} \tag{7-3}$$

$$\mathbf{B} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} \tag{7-4}$$

$$\mathbf{C} = \begin{pmatrix} 0 & 0 & c_3 \end{pmatrix} \tag{7-5}$$

According to the model in Equation (7-2), at each sampling instants kT, starting from the current state information $\mathbf{x}(kT)$ the MPC solves the following open loop optimal control problem over the prediction horizon N

$$\begin{aligned} &\min_{\{u_{k+1},\ldots,u_{k+N}\}\in[0,1]} SSE_k \\ &s.t. \\ &\quad \mathbf{x}_{i+1} = \mathbf{A}\mathbf{x}_i + \mathbf{B}u_i \\ &\quad y_i = \mathbf{C}\mathbf{x}_i \\ &\quad \mathbf{x}_k \quad is \quad given \\ &\quad y_i \in [0,1] \\ &\quad i \in \{k+1,\ldots,k+N\} \end{aligned} \tag{7-6}$$

where the cost function SSE_k is the sum of the squared control error (SSE) defined in Equation (5-14). The state system are reconstructed using the Kalman filter described in Chapter 2. The optimization is carried out by using the Matlab implementation of the *fmincon* function that solves optimization problem with the *active-set*algorithm described in [64]. The result

of the optimization is an array of N optimal control inputs ($\mathbf{u}_i$). The optimal control input is used to decide the concentration of tetracycline to be provide to the cells. In principle in absence of external disturbances the optimal input could be applied over the entire prediction horizon. However in order to make the control action robust to any source of uncertainty and variability, the feedback loop is closed by applying the calculated control input only at the current sampling time and at the successive sampling time the entire procedure is repeated.

7.3 Numerical simulations and comparative analysis of the MPC strategy with either discrete or continuous control input for gene expression control regulation

7.3.1 Numerical validation

I numerically tested the control strategy with the continuous actuation by simulating the whole experimental setup consisting of both the tetracycline concentration controller (for more details see Chapter 6) and the model predictive controller of gene expression with continuous control input as described in the section 7.3.1. In the numerical simulations, in the place of the living cells integrating the *tetO7*-Ub^{V76}GFP system, I used the *Ub^{V76}GFP.CI model* in Equation (7-1) to simulate Ub^{V76}GFP protein expression. In order to simulate tetracycline accumulation in the cell chamber, I used the model described by the Equations 6-2, which I obtained as described in Chapter 6.

With continuous control input, cell can be provided with any intermediate concentration of tetracycline, thanks to the presence of the additional tetracycline controller ensuring that the desired tetracycline concentration calculated by the gene expression controller is indeed provided to the cells. With discrete control input, cells are provided with only two levels of tetra-

cycline, either the maximum concentration or no tetracycline is provided to the cells, as shown in detail in Chapter 5.

The controllers were simulated to perform both set-point regulation and signal-tracking of the model output.

As already described in Chapter 5, the MPC with discrete control input when applied to set point control (Figure **7-5** A), is able to force the model's output (y) to reach the control reference (r) and to maintain its value over time without exhibiting any oscillation. Comparing this result with the performance achieved by the MPC with continuous control input (Figure **7-5** B), the performance indices (ISE, IAE, ITAE in Figure **7-5** C) calculated considering the steady state error (over the last 3000 minutes of simulations) are of several orders of magnitude better in the continuous control input MPC strategy. Thus the gene expression control in mammalian cells is improved when using a continuous actuation based control strategy.

In the case of signal tracking control (Figure **7-6** and Figure **7-7**) the performance of the MPC with discrete control input is worst than that with the continuous actuation as confirmed by the performance indices.

Interestingly the continuous control input MPC strategy is able to achieve an optimal performance with a very small variations of tetracycline concentration. Furthermore from a biological point of view, slower is the variation in the tetracycline concentration (variation of control input), lower is the stress occurring to the cells.

7.3.2 *In vitro* experiments: preliminary results

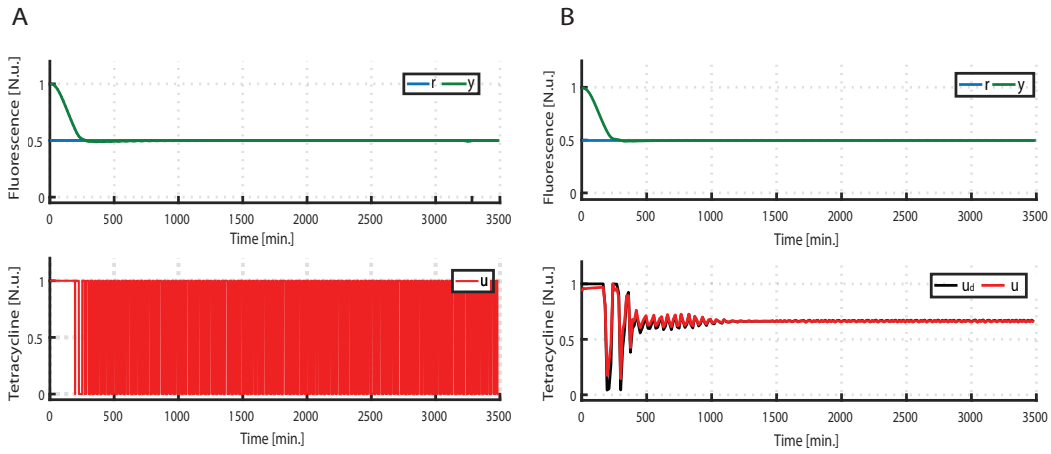
The results obtained by the numerical analysis show that the MPC based control strategy with continuous control input is feasible and advantageous over the MPC with discrete control input.

In order to verify whether the outcome of the numerical simulations can be confirmed experimentally controlling living cells using microfluidics, *tetO7*-Ub^{V76}GFP cells were loaded in the microfluidic device and allowed to grow for 24 hours under perfusion conditions in a cell culture incubator. Then the device was moved to the inverted microscope, and the images

acquisitions and the control experiment were started (the experimental protocol is reported in APPENDIX).

The control task consisted in reaching and maintaining 50% of the initial value of $Ub^{V76}GFP$ fluorescence calculated as the mean value of fluorescence over the **calibration phase** at the beginning of the control experiment. In this phase, lasting 180 minutes, the cells are kept in medium without tetracycline, thus the $Ub^{V76}GFP$ protein is switched on and its fluorescence is quantified in order to calculate the reference value. The control action starts at time $t = 0$ minutes and finishes at $t = 1000$ minutes. As shown in Figure **7-8**, the control action works effectively in keeping the output, namely the measured fluorescence, close to the desired set-point for 500 minutes. However in the both experiments, around the 650 minutes the cells fluorescence increases even if high concentration of tetracycline is given to the cells (the control input is high, red signal in the Figure **7-8**). This drawback is due to the fact that during these experiments some technical problems linked to cell death and the algorithm of segmentation failure have occurred. Indeed when cells die, they incorporate red dye that interferes in the quantification of the green fluorescence by means of the segmentation algorithm.

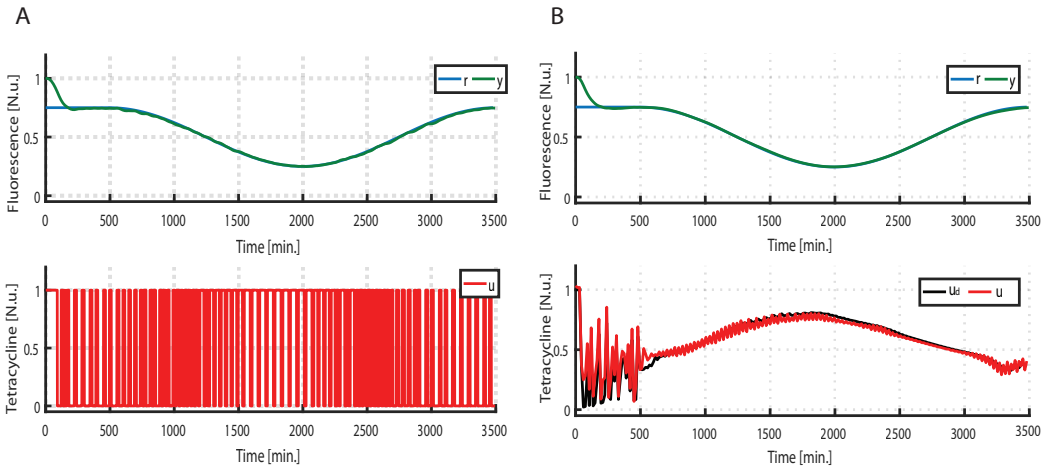
Nevertheless these preliminary results represents a good starting point to demonstrate that the implemented control strategy with a continuous actuation is suitable to regulate at will a protein of interest. Moreover the opportunity to slowly change the levels of the inducer molecule (control input) allows to reduce the stress filled by the cells.



C

	Binary actuation		Continuous actuation	
	control time	steady state	control time	steady state
ISE	0.9246	0.0054	1.1636	8.68E-05
IAE	3.7606	0.9156	3.8872	0.0395
ITAE	1.85E03	1.69E03	337.3563	38.6490

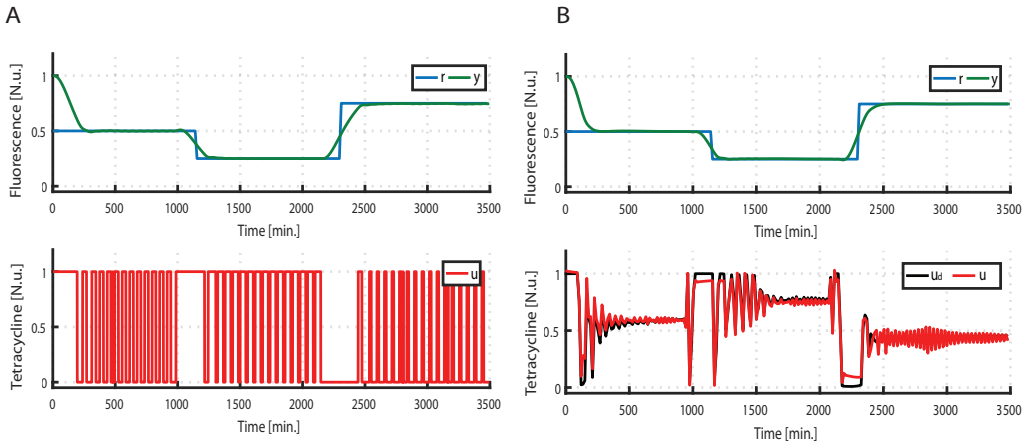
Figure **7-5**: **Numerical simulations of *tetO7*-Ub^{V76}GFP set point control.** (A-B) Two numerical simulations of set point control were performed on the *tetO7*-Ub^{V76}GFP dynamical model by means of either the MPC with discrete control input (A) or the MPC with continuous control input (B). The reference signal, r in blue, is equal 50% of the steady state of the model's output for a null concentration of tetracycline in input, here normalized to 1. The green line is the model's output (y) and the red line represents the control input (u). The black line in the MPC regulation with a continuous control input (B) is the desired control input (u_d) computed by the gene expression controller. (C) Performance indices: Integral Square Error (ISE), Integral Absolute Error (IAE), Integral Time Absolute Error (ITAE).



C

	Binary actuation	Continuous actuation
ISE	0.1941	0.0957
IAE	3.2783	1.4624
ITAE	4.0169E03	1.0023E03

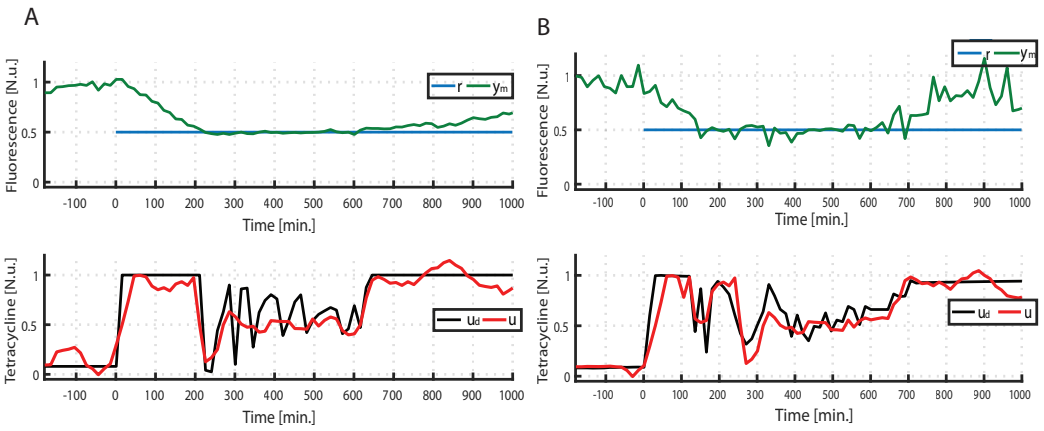
Figure **7-6**: **Numerical simulations of *tetO7*-Ub^{V76}GFP signal tracking control.** (A-B) Two numerical simulations of cosine wave tracking control experiments performed on the *tetO7*-Ub^{V76}GFP dynamical model by means of either the MPC with discrete control input (A) or the MPC with continuous control input (B). The reference signal (r in blue) is a steady state signal equal to 75% of the steady state of the model's output for a null concentration of tetracycline in input, here normalized to 1 with a duration of 500 min followed by a sinusoidal wave of period $T = 3000$ min defined as $s = 0.5 + 0.25 * sin(2\frac{\pi}{T}(t-100) + \frac{\pi}{2})$. The green line is the model's output (y) and the red line represents the control input (u). The black line in the MPC regulation with a continuous control input (B) is the desired control input (u_d) computed by the gene expression controller. (C) Performance indices: Integral Square Error (ISE), Integral Absolute Error (IAE), Integral Time Absolute Error (ITAE).



C

	Binary actuation	Continuous actuation
ISE	1.2890	0.5661
IAE	7.3568	3.7987
ITAE	1.2467E04	3.1211E03

Figure **7-7**: **Numerical simulations of *tetO7*-Ub^{V76}GFP staircase tracking control task.** (A-B) Two numerical simulations of staircase tracking control experiments performed on the *tetO7*-Ub^{V76}GFP dynamical model by means of either the MPC with discrete control input (A) or the MPC with continuous control input (B). The reference signal (r in blue) is a three steps signal where each step is calculating as a percentage (50%, 25% and 75%) of the high steady state of the model's output, here normalized to 1. The green line is the model's output (y) during the control time interval and the red line represents the simulated control input (u). The black line is the desired control input (u_d) computed by the gene expression controller. (C) Performance indices calculated over the control time interval: Integral Square Error (ISE), Integral Absolute Error (IAE), Integral Time Absolute Error (ITAE).



C

	Experiment #1	Experiment #2
ISE	1.5841	1.3825
IAE	5.0396	5.7051
ITAE	866.6737	2.1584E03

Figure **7-8**: ***In vitro tetO7*-Ub^{V76}GFP set point control.** (A-B) Two *in vitro* set point control experiments performed on the *TetO7*-Ub^{V76}GFP system by means of the MPC with continuous control input. The control action starts at time $t = 0$ min. The reference signal (r in blue) is equal to 50% of the initial value of Ub^{V76}GFP fluorescence calculated as the mean value over 180 minutes prior to starting of control action (calibration phase). The green line is the measured Ub^{V76}GFP fluorescence (y_m). The tetracycline concentration is quantified by adding a fluorescent red dye in the tetracycline-filled syringe, therefore the control input is the measured red fluorescence (red signal, u_m) while the black line is the desired control input u_d computed by the gene expression controller. (C) Performance indices: Integral Square Error (ISE), Integral Absolute Error (IAE), Integral Time Absolute Error (ITAE).

CHAPTER 8

Conclusions

In this Thesis, I proposed a microfluidics-based feedback control strategy to quantitatively regulate gene expression from a tetracycline-inducible promoter in mammalian cells. The strategy proposed is based on the regulation of gene expression by means of an inducer molecule (i.e. tetracycline), which is able to block the transactivational activity of a constitutively expressed transcription factor.

In Chapter 4, I presented a proof-of-principle study to demonstrate that microfluidics-based control of gene expression in mammalian cells is feasible. In order to do that, I chose two simple control strategies that do not require a detailed quantitative model of gene expression (relay and PI). I showed that it is indeed possible to force a population of cells harboring the inducible promoter to express a predetermined level of a protein of interest by automatically administering to the cells pulses of tetracycline whose duration is computed in real time by the Relay and PI control algorithms. Oscillations around the set point are to be expected when using these controllers. For example, the relay control algorithm checks at each sampling time whether the fluorescence has reached the desired value, and if it has not, then the algorithm treats the cells with either tetracycline (to reduce the fluorescence) or standard medium (to increase the fluorescence) for the whole duration of the next sampling interval. This may cause fluorescence to decrease/increase too much, so that at the next sampling interval its value will be below/above

the desired level, giving rise to oscillations around the set point. Oscillations around the set point are well-documented in the control literature, for example in the application of "bang-bang" control [66, 67]. This is an optimal control strategy to minimize the time required for a physical system to achieve the desired set point.

Despite the oscillations around the set-point value, the *in vitro* experiments showed in Chapter 4 demonstrated that it is possible to regulate the expression of a reporter protein from the tetracycline-inducible promoter in a population of mammalian cells using principles from automatic control engineering. This study, published in [51], was the first in which automated control of gene expression from an inducible promoter has been achieved in mammalian cells.

However, the control performance can be greatly improved by moving to a model-based control strategy such as Model Predictive Control.

In Chapter 5 I presented a Model Predictive Control strategy to regulate the gene expression with a discrete control input. Specifically the controller automatically computed the length of pulses of tetracycline to give to cells. I experimentally demonstrated how it is possible to eliminate oscillations around the set-point by applying a model based control strategy thanks to its predictive nature.

The experimental results described in Chapters 4 and 5 convincingly demonstrated that the expression of a protein can be controlled *in vitro* in real-time, using an inducer molecule, by applying a discrete actuation signal, i.e. either standard medium or medium with tetracycline at predetermined concentration be provided to the cells. However, for the biological models in exam, the concentration of tetracycline to be supplied to the cells could be graded, obtaining thus a continuous control input which may lead to improved control performances. For this reason, in Chapter 6, I described the implementation of an active control strategy to precisely regulate tetracycline concentration in the microfluidic device. This strategy allows to deliver any time-varying concentration of tetracycline to the cells growing in the cell

chambers. This feature is then exploited in Chapter 7 to control gene expression from a tetracycline inducible promoter using continuous actuation.

As discussed in Chapter 2, the exploitation of Control Theory in biological systems began only recently, and with limited applications only in lesser eukaryotes. In this work, I have generated preliminary results that demonstrate that control gene expression in mammalian cells is possible. The ability to set a desired level of protein inside cells, even at a constant rate, will allow biologists to perform dose-response analyses that will provide a plethora of new quantitative data: activation thresholds, enzymatic kinetics, protein aggregation studies and so on. This will be of paramount interest to enrich our knowledge of basic biological processes, which is still qualitative.

APPENDIX A

Material and Methods

A.1 Cell culture

Chinese Hamster Ovary (CHO) cells (both *tetO7*-d2EYFP and *tetO7*-Ub^{V76}GFP cells) are cultured in αMEM (EuroClone) supplemented with 10% tetracycline-free FBS (EuroClone), 1% Penicillin/Streptomycin (EuroClone), 2 mM Glutamine (EuroClone) and 3μg/ml blasticidine (Sigma-Aldrich) and are kept in a standard tissue culture incubator at 37$\circ$C, 5% CO_2, 98% of humidity.

Cells are splitted twice per week using PBS (Gibco) and trypsin (Gibco) and seeded at 2000 cells/cm^2.

The day before the loading of the microfluidic device cells were splitted 1:3 and placed back into the incubator.

A.2 Microfluidics

All the experiments presented in this study were performed by using the microfluidic device for mammalian cells designed by Kolnik and colleagues in the Jeff Hasty's Bioengineering Laboratory at the University Of California, San Diego [46]. This device allows to grow cells in temperature- and CO_2-controlled environment isolated from potentially damaging flow effects, to administer in a precisely manner external input via a complex topology of

channels connecting inlets between each other and the cells chambers (a detailed description of its topology is in Chapter 3 and [46].

A.2.1 Device fabrication protocol

I have used replica molding technique to obtain polydimethylsiloxane (PDMS) replicas of the device presented in [46] by using a master mold of the features as blueprint. The master mold has been produced using a 4" silicon wafer as substrate (Silicon Valley Microelectronics, US) by a way of layer-by-layer photolithography [68].

Before the fabrication of the microfluidic devices the master is exposed to chlorotrimethylsilane (Sigma-Aldrich Co.) vapours for 10 minutes in order to create an anti-sticking silane layer for PDMS. PDMS is prepared by mixing Sylgard 184 Elastomer curing agent and base (DOW corning) in a 1 : 10 ratio. PDMS is poured onto the master mold, degassed for 30 minutes, cured for 1 hour at 80°C, allowed to cool to room temperature and then peeled from the wafer . PDMS was then autoclaved for 30 minutes at 121°C to ensure long-term viability of cells in the devices [69]. Holes for the 7 ports were punched using a 24-gauge blunt needle in order to create fluidic ports for the access of cells and liquid substances. The PDMS devices obtained are rinsed in isopropyl alcohol and distilled water to remove debris. For each PDMS piece containing microchannels a thin glass slide (150 μm) is cleaned in methanol and 70% ethyl alcohol. Finally the PDMS layers and glass slides are exposed to oxygen plasma in Plasma Cleaner machine (ZEPTO version B, Diener electronic GmbH) for 2 minutes and brought into contact forms a strong irreversible bond between two surfaces. As last step all devices were checked for faults inside and outside the channels.

A.2.2 Device cells loading protocol

For device loading, cells are washed with sterile Phosphate-Buffered Saline (PBS, Gibco), detached from the culture dishes by exposing to 0.25% Trypsin EDTA for 1 minute and centrifuged to form a pellet. Then in order to obtain

a seeding density of $3-5$ cells per chamber, the cells pellet is resuspended in complete media at a density of 0.064cells per mm^2.

The channel of the device are then completely filled with fluid (except for the culture chambers) by applying complete media first through port 5 and then through port 2 once it is filled with fluid. The cell suspension is loaded into the main channel of the device from port 2 and a vacuum is applied in the channel adjacent to the culture chambers (ports 3 and 4) to evacuate air and replace the chamber volume with fluid containing cells. Remaining untrapped cells in the main channel are washed away at a high flow rate without disturbing cells inside the traps. Fluidic connections from the ports of the device to syringes containing growth medium are then established using 24 gauge PTFE tubing (Cole-Parmer Inc.) interfaced via 22 gauge stainless steel luer stub pins.

Once cells are loaded in the microfluidc device, they are allowed in a cell culture incubator for 24 hours in perfusion conditions providing that the cells in chambers receive fresh medium from syringe connected to port 5 and the waste medium is washed trough port 1, while all the other ports (2,6 and 7) are plugged.

For both identification and control experiments, the device is secured on the microscope stage within an environmental chamber maintained at 37°C with humidified 5% CO_2 and 60ml syringes are connected to its ports and hanged at different heights, in order to regulate the flow from the inlets (ports 6 and 7) to the outlets (ports 5, 1 and 2) according to hydrostatic pressure. The syringes connected to the outlets port contain 10ml of standard complete culture medium and they serve as a waste tanks. The syringes connected to the inlets port are filled with standard medium and medium with tetracycline at a concentration of 100ng/ml (Clontech) and 1μM Sulforhodamine 101 (Sigma) and secured on the linear actuator. The correct functioning is checked by inspecting the red fluorescence emitted by Sulforhodamine as a result of the automatic height control of syringes.

A.2.3 Device loading protocol for the identification and control of the dynamic of small molecule concentration in cell chambers.

First the device is completely filled with small molecule filled fluid (except for the culture chambers) by applying complete media first through port 5 and then through port 2 once it is filled with fluid. Then a vacuum is applied in the channel adjacent to the culture chambers (ports 3 and 4) to evacuate air and replace the chamber volume with fluid containing small molecule.

For both identification and control experiments, the device is the time lapse and 60ml syringes are connected to its ports and hanged at different heights, in order to regulate the flow from the inlets (ports 6 and 7) to the outlets (ports 5, 1 and 2) according to hydrostatic pressure. The syringes connected to the outlets port contain 10ml of standard complete culture medium and are used as a waste tanks. The syringes connected to the inlets port are filled with standard medium and medium with small molecule and 1μM Sulforhodamine 101 (Sigma) and secured on the linear actuator. The correct functioning is checked by inspecting the red fluorescence emitted by Sulforhodamine as a result of the automatic height control of syringes. the fluorescence emitted by Sulforhodamine in cell chambers is used as a read out in the identification and control experiments (see Chapter 6).

A.3 System identification experiment

A.3.1 *TetO7*-Ub^{V76}GFP system identification experiment

Once *TetO7*-Ub^{V76}GFP cells loaded in the microfluidic device are put under the microscope and the syringes are connected, they are kept in a standard medium to allow cells adapt to the microchemostat environment simply by holding the syringe filled with standard medium in a higher position respect to the one carrying tetracycline rich medium. After this calibration phase of the experiment, a sequential MATLAB script controls syringes' positions

over the time to obtain desired time profile for the input fed to cells. The correct functioning can be check by inspecting the red fluorescence emitted by Sulforhodamine in cells chamber.

The image processing algorithm, running in real time, calculates the absolute fluorescence emitted by the entire cell population as well as its normalised value by dividing the time course of fluoresce by the average fluorescence intensity measured during the initial calibration phase, obtaining the output data.

Input and output time series thus generated are used, as discussed in Section 5.2, to apply System Identification techniques.

A.3.2 Identification experiment for modelling tetracycline accumulation in cell chambers

Once the microfluidic device filled with small molecule and Sulforhodamine is under the microscope the small molecule filled syringe is keeping in higher position respect to the one carrying medium without the small molecule. After this calibration phase of the experiment, a sequential MATLAB script controls syringes' positions and dynamically modulates the differential heights between the two syringes over the time as system input.

The dynamics of the small molecule accumulation in the cells chambers in response to such an input is followed by measuring red dye fluorescence. The absolute fluorescence emitted by the Sulforhodamine in cells chambers is calculated and it is normalised by dividing the time course of fluoresce by the average fluorescence intensity measured during the initial calibration phase, obtaining the output data.

Input and output time series thus generated are used, as discussed in Section 6.2, to apply System Identification techniques.

A.4 Control experiment

A.4.1 Gene expression control experiment

Both *tetO7*-d2EYFP and *tetO7*-Ub^{V76}GFP cells control experiments are carried out by using the same procedures.

Once microfluidic device with cells is secured on the microscope stage and the syringes are connected, the user has to start a custom MATLAB script, that manages the entire experimental platform (controller implementation, actuation, image analysis), and has to set the duration (in minutes) of the control, the crop area and the threshold needed for image analysis.

The script calculates the set point for the control as a percentage (indicated by the user at the beginning of the experiment) of the average of the fluorescence measured by the image processing algorithm during the calibration phase previously described. After this the implemented script proceeds in executing all the code blocks necessary to reach and maintain the fluorescence reference.

A.4.2 Small molecule concentration control experiment

Also for the control of molecule concentration in the microfluidic device, once the microfluidic is under the microscope a MATLAB script manages the entire experimental platform. The duration of the control and the crop area are setted at the beginning of the experiment. The script calculates the length and the values of the steps in the step - like time varying reference control and the value of reference in the tracking control as percentages (indicated by the user at the beginning of the experiment) of the average of the red fluorescence emitted by Sulforhodamine measured during the calibration phase. In this phase the syringe with small molecule and Sulforodhamine is higher than the other syringe. At the end of the calibration, the implemented script proceeds in executing all the code blocks necessary to reach and maintain the fluorescence reference.

A.5 Fluorescence microscopy

The closed loop control platforms for gene expression control described in the Chapter 3 and 7, employs an inverted fluorescence Nikon-TI Eclipse microscope equipped with a digital camera (Andor iXon897, Andor) and an incubation chamber (H201-OP R2, Okolab) to acquire images from cells in the microfluidic device. To overcome the problem of the focus drift due to cells growing and replicating and to the length of the experiments performed, the microscope has been equipped with the Nikon Perfect Focus System (PFS) that is able to compensate for axial focus fluctuations in real time during long-term imaging experiments.

The microscope is programmed to acquire two types of images: a phase contrast image (PhC) and a fluorescence image in green spectrum for the fluorescent reporter used to track cell state and a fluorescence images in the red spectrum for Sulforhodamine. Both PhC and fluorescence images are acquired with the same objective (Obj. PlanFluor DLL 40X Ph2 Nikon, NA 0.63) at intervals of 15 minutes. An automated shutter is used to finely control the exposure times for each type of image acquired that are set to 286 ms for PhC image, to 300 ms for green spectrum with Nikon GFP BP HYQ filter (Ex $490-510$ nm, Em $520-550$ nm) for Ub^{V76}GFP and Nikon FITC Filter (Ex $465-495$ nm, Em $515-555$ nm) for the d2EYFP , to 300 ms for red spectrum, with a Nikon TRITC HYQ filter (Ex $530-560$ nm, Em $590-650$ nm). The exposure times and the acquisition interval of 15 minutes have been chose to avoid phototoxicity damages to cells and photobleaching of the fluorescent proteins/dyes [47]. Temperature is maintained constantly at 37°C, and CO_2 concentration is set to 5% of the total air volume injected in the incubation chamber.

For the small molecule concentration control the same microscope is used. Since the Sulforhodamine is added to follow the dynamics of the small molecule accumulation, the microscope is programmed to acquire a fluorescence images in the red spectrum with Nikon TRITC HYQ filter (Ex $530-560$ nm, Em $590-650$ nm) at exposure times setted to 300ms. The

images are acquired with the Nikon 40X dry objective, NA 0.63 at intervals of 1 minutes.

A.6 Image analysis

Image Analysis, together with the microscope, composes the sensing apparatus of the experimental platform. The outcome of the experiments shown in this study is strictly dependent on the accuracy of the real time image analysis performed. For this reason I used an image processing algorithm, developed in the laboratory of Diego di Bernardo at TIGEM by Dr. Gianfranco Fiore, that is meant to locate cells within each PhC frame, and to use this information to calculate the fluorescence (corresponding to each reporter or fluorophore) [50].

In this study I have considered Chinese Hamster Ovary (CHO) cells that as other mammalian cells, do not have a particular geometrical shape [70], the used image processing algorithm is not based on cells' morphological features but it exploits the property of cells to exhibit a white halo in Phase contrast image, shown in Figure to locate the cells within image. Briefly, the image analysis algorithm works as follow: (a) it defines a threshold to generate a first binary image selecting only pixels belonging to cells' edges (Figure); (b) it obtains a second binary image (mask) in which the cell area is overestimates by using dilation and filling operators; (c) by subtracting from the mask obtained at point (b) that obtained at point (a), it derives a binary image (Figure 8.2 C) that select the portion of the original image covered by cells (Figure); (d) the binary filter obtained, is applied to green field image to calculate the average intensity fluorescence of pixels belonging to cells, subtracting the background signal.

Bibliography

[1] Pablo A Iglesias and Brian P Ingalls. *Control theory and systems biology*. MIT Press, 2010.

[2] Giovanni Russo, Mario Di Bernardo, and Eduardo D Sontag. Global entrainment of transcriptional systems to periodic inputs. *PLoS Comput Biol*, 6(4):e1000739, 2010.

[3] Fred S Grodins, John S Gray, Karl R Schroeder, Arthur L Norins, and Richard W Jones. Respiratory responses to co2 inhalation. a theoretical study of a nonlinear biological regulator. *Journal of applied physiology*, 7(3):283–308, 1954.

[4] Fred S Grodins. *Control theory and biological systems*. Columbia University Press, 1963.

[5] Howard T Milhorn. Application of control theory to physiological systems. 1966.

[6] Hans Kalmus et al. Regulation and control in living systems. 1966.

[7] Tal Danino, Octavio Mondragón-Palomino, Lev Tsimring, and Jeff Hasty. A synchronized quorum of genetic clocks. *Nature*, 463(7279):326–330, 2010.

[8] Andreas Milias-Argeitis, Sean Summers, Jacob Stewart-Ornstein, Ignacio Zuleta, David Pincus, Hana El-Samad, Mustafa Khammash, and

John Lygeros. In silico feedback for in vivo regulation of a gene expression circuit. *Nature biotechnology*, 29(12):1114–1116, 2011.

[9] Jannis Uhlendorf, Agnès Miermont, Thierry Delaveau, Gilles Charvin, François Fages, Samuel Bottani, Gregory Batt, and Pascal Hersen. Long-term model predictive control of gene expression at the population and single-cell levels. *Proceedings of the National Academy of Sciences*, 109(35):14271–14276, 2012.

[10] Filippo Menolascina, Gianfranco Fiore, Emanuele Orabona, Luca De Stefano, Mike Ferry, Jeff Hasty, Mario di Bernardo, and Diego di Bernardo. In-vivo real-time control of protein expression from endogenous and synthetic gene networks. *PLoS Comput Biol*, 10(5):e1003625, 2014.

[11] Gianfranco Fiore, Giansimone Perrino, Mario di Bernardo, and Diego di Bernardo. In vivo real-time control of gene expression: A comparative analysis of feedback control strategies in yeast. *ACS synthetic biology*, 5(2):154–162, 2015.

[12] Andreas Milias-Argeitis, Marc Rullan, Stephanie K Aoki, Peter Buchmann, and Mustafa Khammash. Automated optogenetic feedback control for precise and robust regulation of gene expression and cell growth. *Nature Communications*, 7, 2016.

[13] Xavier Darzacq, Yaron Shav-Tal, Valeria De Turris, Yehuda Brody, Shailesh M Shenoy, Robert D Phair, and Robert H Singer. In vivo dynamics of rna polymerase ii transcription. *Nature structural & molecular biology*, 14(9):796–806, 2007.

[14] Eduardo D Sontag. Molecular systems biology and dynamics: An introduction for non-biologists.

[15] Uri Alon. *An introduction to systems biology: design principles of biological circuits*. CRC press, 2006.

[16] Jared E Toettcher, Delquin Gong, Wendell A Lim, and Orion D Weiner. Light-based feedback for controlling intracellular signaling dynamics. *Nature methods*, 8(10):837–839, 2011.

[17] Manfred Gossen and Hermann Bujard. Tight control of gene expression in mammalian cells by tetracycline-responsive promoters. *Proceedings of the National Academy of Sciences*, 89(12):5547–5551, 1992.

[18] Manfred Gossen, Gabriele Bender, Gerhard Muller, Sabine Freundlieb, et al. Transcriptional activation by tetracyclines in mammalian cells. *Science*, 268(5218):1766, 1995.

[19] Stefanie Urlinger, Udo Baron, Marion Thellmann, Mazahir T Hasan, Hermann Bujard, and Wolfgang Hillen. Exploring the sequence space for tetracycline-dependent transcriptional activators: novel mutations yield expanded range and sensitivity. *Proceedings of the National Academy of Sciences*, 97(14):7963–7968, 2000.

[20] Evan J Olson, Lucas A Hartsough, Brian P Landry, Raghav Shroff, and Jeffrey J Tabor. Characterizing bacterial gene circuit dynamics with optically programmed gene expression signals. *Nature methods*, 11(4):449–455, 2014.

[21] Filippo Menolascina, Mario Di Bernardo, and Diego Di Bernardo. Analysis, design and implementation of a novel scheme for in-vivo control of synthetic gene regulatory networks. *Automatica*, 47(6):1265–1270, 2011.

[22] Karl Johan Aström and Richard M Murray. *Feedback systems: an introduction for scientists and engineers*. Princeton university press, 2010.

[23] Vadim Utkin and Hoon Lee. Chattering problem in sliding mode control systems. In *Variable Structure Systems, 2006. VSS'06. International Workshop on*, pages 346–350. IEEE, 2006.

[24] Tau-Mu Yi, Yun Huang, Melvin I Simon, and John Doyle. Robust perfect adaptation in bacterial chemotaxis through integral feedback control. *Proceedings of the National Academy of Sciences*, 97(9):4649–4653, 2000.

[25] Kiam Heong Ang, Gregory Chong, and Yun Li. Pid control system analysis, design, and technology. *IEEE transactions on control systems technology*, 13(4):559–576, 2005.

[26] Eduardo F Camacho and Carlos Bordons Alba. *Model predictive control*. Springer Science & Business Media, 2013.

[27] S Joe Qin and Thomas A Badgwell. An overview of industrial model predictive control technology. In *AIChE Symposium Series*, volume 93, pages 232–256. New York, NY: American Institute of Chemical Engineers, 1971-c2002., 1997.

[28] Manfred Morari and Jay H Lee. Model predictive control: past, present and future. *Computers & Chemical Engineering*, 23(4):667–682, 1999.

[29] S Joe Qin and Thomas A Badgwell. An overview of nonlinear model predictive control applications. In *Nonlinear model predictive control*, pages 369–392. Springer, 2000.

[30] S Joe Qin and Thomas A Badgwell. A survey of industrial model predictive control technology. *Control engineering practice*, 11(7):733–764, 2003.

[31] Alberto Bemporad and Manfred Morari. Robust model predictive control: A survey. In *Robustness in identification and control*, pages 207–226. Springer, 1999.

[32] Rolf Findeisen, Lars Imsland, Frank Allgower, and Bjarne A Foss. State and output feedback nonlinear model predictive control: An overview. *European journal of control*, 9(2-3):190–206, 2003.

[33] Lennart Ljung. System identification: Theory for the user, ptr prentice hall information and system sciences series, 1999.

[34] Lennart Ljung. *System identification*. Springer, 1998.

[35] Roberto Guidorzi. *Multivariable system identification: from observations to models*. Bononia University Press Bologna, 2003.

[36] Lennart Ljung. System identification toolbox for use with {MATLAB}. 2007.

[37] Rudolph Emil Kalman. A new approach to linear filtering and prediction problems. *Journal of basic Engineering*, 82(1):35–45, 1960.

[38] Greg Welch and Gary Bishop. An introduction to the kalman filter. 1995.

[39] Matthew R Bennett, Wyming Lee Pang, Natalie A Ostroff, Bridget L Baumgartner, Sujata Nayak, Lev S Tsimring, and Jeff Hasty. Metabolic gene regulation in a dynamically changing environment. *Nature*, 454(7208):1119–1122, 2008.

[40] Scott Cookson, Natalie Ostroff, Wyming Lee Pang, Dmitri Volfson, and Jeff Hasty. Monitoring dynamics of single-cell gene expression over multiple cell cycles. *Molecular Systems Biology*, 1(1), 2005.

[41] Pascal Hersen, Megan N McClean, L Mahadevan, and Sharad Ramanathan. Signal processing by the hog map kinase pathway. *Proceedings of the National Academy of Sciences*, 105(20):7165–7170, 2008.

[42] F Kurth, CA Schumann, LM Blank, A Schmid, A Manz, and PS Dittrich. Bilayer microfluidic chip for diffusion-controlled activation of yeast species. *Journal of Chromatography A*, 1206(1):77–82, 2008.

[43] Philip J Lee, Noah C Helman, Wendell A Lim, and Paul J Hung. A microfluidic system for dynamic yeast cell imaging. *Biotechniques*, 44(1):91–95, 2008.

[44] Amy C Rowat, James C Bird, Jeremy J Agresti, Oliver J Rando, and David A Weitz. Tracking lineages of single cells in lines using a microfluidic device. *Proceedings of the National Academy of Sciences*, 106(43):18149–18154, 2009.

[45] RJ Taylor, D Falconnet, A Niemistö, SA Ramsey, S Prinz, I Shmulevich, T Galitski, and CL Hansen. Dynamic analysis of mapk signaling using a high-throughput microfluidic single-cell imaging platform. *Proceedings of the National Academy of Sciences*, 106(10):3758–3763, 2009.

[46] Martin Kolnik, Lev S Tsimring, and Jeff Hasty. Vacuum-assisted cell loading enables shear-free mammalian microfluidic culture. *Lab on a chip*, 12(22):4732–4737, 2012.

[47] MS Ferry, IA Razinkov, and J Hasty. Microfluidics for synthetic biology: from design to execution. *Methods in enzymology*, 497:295, 2011.

[48] Abraham D Stroock, Stephan KW Dertinger, Armand Ajdari, Igor Mezić, Howard A Stone, and George M Whitesides. Chaotic mixer for microchannels. *Science*, 295(5555):647–651, 2002.

[49] Filippo Menolascina. Synthetic gene networks identification and control by means of microfluidic devices. 2011.

[50] Gianfranco Fiore. Identification and control of gene networks in living cells. 2015.

[51] Chiara Fracassi, Lorena Postiglione, Gianfranco Fiore, and Diego di Bernardo. Automatic control of gene expression in mammalian cells. *ACS synthetic biology*, 5(4):296–302, 2015.

[52] Xianqiang Li, Xiaoning Zhao, Yu Fang, Xin Jiang, Tommy Duong, Connie Fan, Chiao-Chain Huang, and Steven R Kain. Generation of destabilized green fluorescent protein as a transcription reporter. *Journal of Biological Chemistry*, 273(52):34970–34975, 1998.

[53] Manfred Gossen, Angelika L Bonin, Sabine Freundlieb, and Hermann Bujard. Inducible gene expression systems for higher eukaryotic cells. *Current opinion in biotechnology*, 5(5):516–520, 1994.

[54] Velia Siciliano, Filippo Menolascina, Lucia Marucci, Chiara Fracassi, Immacolata Garzilli, Maria Nicoletta Moretti, and Diego di Bernardo. Construction and modelling of an inducible positive feedback loop stably integrated in a mammalian cell-line. *PLoS Comput Biol*, 7(6):e1002074, 2011.

[55] Valentin Magidson and Alexey Khodjakov. Circumventing photodamage in live-cell microscopy. *Methods in cell biology*, 114, 2013.

[56] Marcel Tigges, Tatiana T Marquez-Lago, Jörg Stelling, and Martin Fussenegger. A tunable synthetic mammalian oscillator. *Nature*, 457(7227):309–312, 2009.

[57] Hiromi Hirata, Shigeki Yoshiura, Toshiyuki Ohtsuka, Yasumasa Bessho, Takahiro Harada, Kenichi Yoshikawa, and Ryoichiro Kageyama. Oscillatory expression of the bhlh factor hes1 regulated by a negative feedback loop. *Science*, 298(5594):840–843, 2002.

[58] G Fiore, Filippo Menolascina, M di Bernardo, and D di Bernardo. An experimental approach to identify dynamical models of transcriptional regulation in living cells. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 23(2):025106, 2013.

[59] Adrian Wills and Brett Ninness. On gradient-based search for multivariable system estimates. *IEEE Transactions on Automatic Control*, 53(1):298–306, 2008.

[60] Andrew Gordon, Alejandro Colman-Lerner, Tina E Chin, Kirsten R Benjamin, C Yu Richard, and Roger Brent. Single-cell quantification of molecules and rates using open-source microscope-based cytometry. *Nature methods*, 4(2):175–181, 2007.

[61] Gene F Franklin, J David Powell, and Michael L Workman. *Digital control of dynamic systems*, volume 3. Addison-wesley Menlo Park, 1998.

[62] Goldberg E David. Genetic algorithms in search optimization and machine learning. *Addison-Wesley*, 1989.

[63] Lorena Postiglione, Marco Santorelli, Barbara Tumaini, and Diego di Bernardo. From a discrete to continuous actuation for improved real-time control of gene expression in mammalian cells. *IFAC-PapersOnLine*, 49(26):14 – 19, 2016. Foundations of Systems Biology in Engineering - {FOSBE} 2016Magdeburg, Germany, 9âĂŤ12 October 2016.

[64] Luís F Portugal, Joaquim J Judice, and Luís N Vicente. A comparison of block pivoting and interior-point algorithms for linear least squares problems with nonnegative variables. *Mathematics of Computation*, 63(208):625–643, 1994.

[65] Gabriele Pannocchia and James B Rawlings. Disturbance models for offset-free model-predictive control. *AIChE journal*, 49(2):426–437, 2003.

[66] IÍąAkov Zalmanovich TÍąSypkin. *Relay control systems*. CUP Archive, 1984.

[67] Yoshito Hirata, Mario di Bernardo, Nicholas Bruchovsky, and Kazuyuki Aihara. Hybrid optimal scheduling for intermittent androgen suppression of prostate cancer. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 20(4):045125, 2010.

[68] George M Whitesides. The origins and the future of microfluidics. *Nature*, 442(7101):368–373, 2006.

[69] Daniel T Chiu, Noo Li Jeon, Sui Huang, Ravi S Kane, Christopher J Wargo, Insung S Choi, Donald E Ingber, and George M Whitesides.

Patterned deposition of cells and proteins onto surfaces by using three-dimensional microfluidic systems. *Proceedings of the National Academy of Sciences*, 97(6):2408–2413, 2000.

[70] Yu P Petrov and NV Tsupkina. Growth characteristics of cho cells in culture. *Cell and Tissue Biology*, 7(1):72–78, 2013.