Article

# Entrainment and control of bacterial populations; an in silico study over a spatially extended agent based model

Petros Mina, Krasimira Tsaneva-Atanasova, and Mario di Bernardo

*ACS Synth. Biol.*, **Just Accepted Manuscript** • DOI: 10.1021/acssynbio.5b00243 • Publication Date (Web): 25 Apr 2016

# Entrainment and control of bacterial populations; an *in silico* study over a spatially extended agent based model

Petros Mina[1,2*], Krasimira Tsaneva-Atanasova[2,3] Mario di Bernardo[2,4],

**1 Bristol Centre for Complexity Sciences, University of Bristol, Bristol, United Kingdom**

**2 Department of Engineering Mathematics, University of Bristol, Bristol, United Kingdom**

**3 Department of Mathematics, University of Exeter, Bristol, United Kingdom**

**4 Department of Information and Computer Engineering, University of Naples Federico II, Naples, Italy**

∗ **E-mail: Corresponding author petros.mina@bristol.ac.uk**

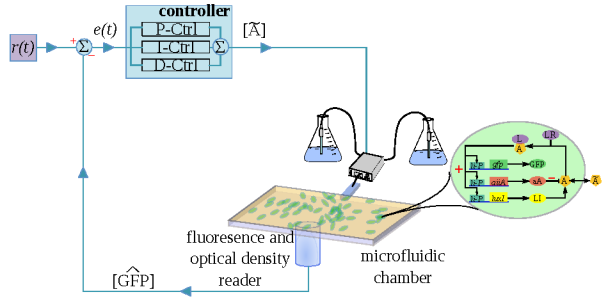
## Abstract

We extend a spatially explicit agent based model (ABM) developed previously to investigate entrainment and control of the emergent behaviour of a population of synchronized oscillating cells in a microfluidic chamber. Unlike most of the work in models of control of cellular systems which focus on temporal changes, we model individual cells with spatial dependencies which may contribute to certain behavioural responses. We use the model to investigate the response of both open-loop and closed-loop strategies, such as proportional control (P-control), proportional-integral control (PI-control) and proportional-integral-derivative control (PID-control), to heterogeneities and growth in the cell population, variations of the control parameters and spatial effects such as diffusion in the spatially explicit setting of a microfluidic chamber setup. We show that, as expected from the theory of phase locking in dynamical systems, open loop control can only entrain the cell population in a subset of forcing periods, with a wide variety of dynamical behaviours obtained outside these regions of entrainment. Closed-loop control is shown instead to guarantee entrainment in a much wider region of control parameter space although presenting limitations when the population size increases over a certain threshold. *In silico* tracking experiments are also performed to validate the ability of classical control approaches to achieve other reference behaviours such as a desired constant output or a linearly varying one. All simulations are carried out in BSim, an advanced agent-based simulator of microbial population which is here extended ad hoc to include the effects of control strategies acting onto the population.

## Key words

Control theory, microfluidics, quorum sensing, genetic regulatory network, agent based modelling

## Contents




# Introduction

Control theory models describe (simple) interconnected components, quantifying their input and taking action through some computation to modify their output [1]; much like living cells that process information from their surroundings (input) and act upon it (output) [2]. For example, pancreatic $\alpha$ and $\beta$-cells, sense and regulate blood glucose levels by increasing glucose uptake or promoting stored glycogen breakdown respectively [3].

Control theoretic approaches have been used alongside dynamical models in order to try and tightly regulate (biological) system output. These models usually describe the change of variables with respect to time, i.e. focus on temporal changes, and are used to calculate appropriate actions on the system in order to drive it to a particular state. In cell biology, dynamical models of control are more common in the field of neuroscience [4–11], where actuation can be in the form of electrical inputs, but with the development of microfluidics technologies over recent years [12] researchers begun to apply control in other cellular systems as well. For example Milias-Argeitis *et al.* [13] used a fourth-order linear model describing a light-responsive genetic network in order to control the gene expression of a microbial population, around a reference value, using optogenetics. Uhlendorf *et al.* [14] used a two-variable delay differential equation (DDE) model to capture the dynamics of the yeast hyperosmotic stress response and compute inputs to make a population of cells to follow a time-varying signal. Menolascina *et al.* modelled and implemented in yeast cells a complex synthetic genetic regulatory network (GRN), named IRMA [15], and utilised a five-variable DDE model to control the production of one of the network's genes. The control design was based on a switching strategy where yeast cells were alternatively fed with Galactose or Glucose [16], [17] to achieve the desired target behaviour. The feedback-based control strategy was validated *in vitro* firstly on a simple *Gal1* promoter expressing GFP and then on the more complex five-gene IRMA network [17].

However, the study of population-based phenomena under the tightly regulated experimental setups offered by microfluidics technologies [12] suggests that both time and space are of importance for observing certain phenomena. In the study of Cho *et al.* [18] the geometry of the cells' surroundings (i.e. the microfluidic chamber) was found to be important for the spatial self-organisation of bacterial colonies.

In Prindle *et al.* [19] *E.coli* cells were engineered to display global synchronised oscillations across physically disconnected, yet coupled via gaseous exchange, populations part of a microfluidics platform where the spatial arrangement of populations was important to the waveforms observed. Notable examples are also found outside the synthetic biology domain. For instance, upsetting the gradient of bicoid and nanos proteins during the growth of a drosophila embryo results in incorrect topological formation of the head and tail [2]. These examples illustrate that certain behavioural responses are dependent on the spatio-temporal organisation of populations, a result also confirmed through theoretical work in reaction-diffusion systems indicating that spatial structure affects qualitative behaviour [20].

Despite the importance of spatial dependence in certain systems, spatial aspects in models of control have not been studied as widely, with limited examples available in the literature. In order to gain insight into the mechanisms disrupting or promoting synchrony across coupled neuronal oscillators, Hauptmann *et al.* [21] developed a time-delayed feedback control model of neurons that are uniformly distributed over the area of a circle. Control was either applied locally, whereby each quadrant was affected using a dedicated controller, or globally where all controllers could have a (weighted) contribution on the entire cell population. Others have focused on continuum type spatial models and how control can be used to obtain specific spatio-temporal patterns. For example, in Alhborn and Parlitz [22] a 2D Ginzburg-Landau partial differential equation (PDE) model is presented [22] where the aim is to stabilize unstable oscillatory behavior (periodic orbits) or create spatio-temporal chaos using global and local time-delayed control input at different locations across the plane. Also, Ghosh [23] investigated global feedback control of an activator/inhibitor reaction-diffusion system spatially extended in a 2D-plane, modelled with two PDEs, in order to obtain stationary patterns in the plane.

The aim of this paper is to study the effect of control in a spatial setting by extending a previously developed, spatially resolved ABM [24] and its implementation in the software platform BSim [25, 26]. As such, we expand the model [24] to include open loop (non-feedback/feed-forward) and closed loop (feedback) control strategies. In the original model [24], which represented the experimental system of Danino *et al.* [27], cells were considered as agents in a microfluidic chamber, had explicit spatial positions and their intracellular dynamics were under the influence of an activation-inhibition type of GRN which can lead to oscillatory behaviour [28–33]. The GRN, based on a quorum sensing (QS) architecture [34], produced a small hormone molecule referred to as an autoinducer which was freely exchanged between cells and their environment leading to an all-to-all coupling across members of the population. In Mina *et al.* [24], we were able to convincingly illustrate that synchronised population-wide oscillations in the metabolic states of cells were likely to be an emergent population property and that synchronisation was dependent on the coupling between members of the population, which in turn depends on the cell density as well as the concentration and spatial diffusivity of the secreted hormone in the microfluidic chamber. By using as basis the model presented in [24] we take into account intracellular dynamics, as well as explicit spatial dependencies of coupled cells to investigate the application of classical proportional (P), proportional-integral (PI) and proportional-integral-derivative (PID) control strategies [1] to achieve tracking and regulation of a cell population. Unlike other computational studies where the focus is on the temporal aspects of a system [13, 14, 16], here we use BSim an agent based simulator of microbial populations able to simulate both the spatial and temporal dynamics of the cells and the embedded GRN [25, 26].

We study the effectiveness and performance of the control approaches and the effects of explicitly considering the spatial dynamics with *in silico* cell populations of increasing sizes, starting with a small number of 21 cells to make the computation easier and ending up with full scale simulations of several thousand of cells. (We also consider intermediate population sizes of 60 and 100 cells to reduce the computational cost and thus cover a greater range of values in control parameter space.) As an initial condition, each population is uniformly distributed over the chamber (see SI), each cell component is modelled with four non-linear ordinary differential equations (ODEs), and subject to control through the manipulation of the small hormone chemical field, which is modelled with a PDE. Specifically, we

investigate when entrainment of the population to a reference sinusoidal signal is possible in open and closed loop and whether the spatially resolved population can track a time-varying reference signal. Since we use an ABM approach, we also study the effect of cell-to-cell heterogeneity across the population, as in a physical setting cell dynamics will tend to vary slightly even between members of the same colony [2, 35, 36]. Finally, by extending BSim [25] to include cell death, we simulate a dynamic, motile population where cells not only grow and divide but also die, to evaluate how the performance of the control action is affected by an increasing and constantly varying population size.

Hence, we illustrate the model and extend the BSim simulation platform so as to provide an effective *in silico* test-bed where cellular populations are simulated in a spatially explicit 3D-environment and control strategies can be tested and validated via more realistic *in silico* experiments before their physical implementation. In so doing, we uncover the dynamic behaviour of the average population response and how it compares with the dynamics of individual cells, which are part of the model. We note that a copy of the BSim simulation platform, along with the developed control functions, can be obtained from the BSim repository [25, 37].

# Results and Discussion

## The spatiotemporal model with control

Our aim is to simulate and control the output of a population of cells, whose metabolic states undergo autonomous oscillations [27], by extending the spatially explicit agent-based model presented in Mina *et al.* [24]. In the autonomous system presented in Danino *et al.* [27] a GRN comprised of three genes was introduced into bacterial cells and allowed for oscillations to exist due to the presence of activation-inhibition feedback loops part of the GRN [24, 28–33]. As shown in **figure 1** these genes are, *luxI*, *aiiA* and *yemGFP* and all are under the influence of the same promoter, *li-P* [27]. The genes have a C-terminal degradation tag sequence that shortens the half-life of their protein products considerably [38] and are introduced into bacteria on separate plasmids [27]. The *luxI* gene encodes for the LuxI synthase (LI), a protein that produces acyl homoserine lactone (AHL). AHL, also known as an autoinducer, can interact with the constitutively expressed protein LuxR, the AHL autoinducer receptor (LR), to form the LuxR:AHL complex (L:A) and activate the promoter *li-P*, allowing for the transcription of all three genes [27]. The AHL molecule is removed from the system by interacting with the acyl homoserine lactonase (aA) enzyme that degrades AHL. AHL can also freely diffuse across the cell membrane allowing for communication, and hence coupling, between all cells in the population [27]. The ODEs used to model the internal dynamics of each cell are given in the supplementary information of this manuscript, as well as in reference [24].

To remain consistent with the experimental setup of Danino *et al.* [27] we define a three dimensional coordinate system, the $(x, y, z)$-plane, where $x \in [x_0, x_n]$, $y \in [y_0, y_n]$ and $z \in [z_0, z_n]$. The subscript $n$ in the axes coordinates represents the maximum distance of each coordinate from the origin, labelled with the subscript 0, in micrometers. As illustrated in **figure 1**, this defines the boundaries of the microfluidic chamber. We set $z_0 = 0$ and $z_n = 1$ to model a one cell thick microfluidic chamber as in Danino *et al.* [27]. Thus, we can restrict ourselves to spatial coordinates on a two-dimensional plane (here the $x$ and $y$ coordinates) and the concentration of external AHL, say $[\tilde{\mathrm{A}}]$, can be described using the reaction-diffusion PDE:

$$\frac{\partial\,[\tilde{\mathrm{A}}](x,y,t)}{\partial t} = \underbrace{D_{\tilde{\mathrm{A}}}\nabla^2[\tilde{\mathrm{A}}](x,y,t)}_{\text{brownian diffusion}} + \underbrace{\sum_{i=0}^{N}\eta_{\text{env}}([\mathrm{A}]_i(t) - [\tilde{\mathrm{A}}](x,y,t))}_{\text{AHL exchange}} - \underbrace{\tau_{\tilde{\mathrm{A}}}[\tilde{\mathrm{A}}](x,y,t)}_{\text{degradation}}, \tag{1}$$

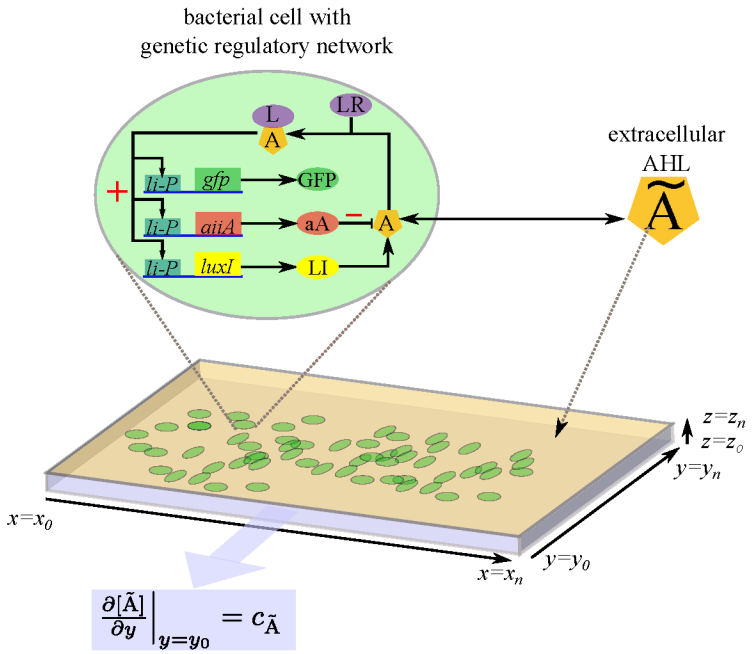



**Figure 1. The microfluidic chamber and the quorum sensing regulatory network of Danino *et al.* [27].** Cells (green ellipses) are found inside a one cell thick microfluidic chamber defined by the $(x, y, z)$ dimensions such that $x \in [x_0, x_n]$, $y \in [y_0, y_n]$ and $z \in [z_0, z_n]$ as illustrated in the figure. Each cell carries a GRN composed of three genes. The *luxI* gene encodes for the LuxI enzyme (LI-yellow), which produces the molecule acyl homoserine lactone (abbreviated as AHL in the main text and presented as variable A, with orange color, in the figure). AHL combines with the constitutively expressed LuxR (purple), to form a complex, L:A, that can activate the promoter *li-P* (blue). This drives the expression of all three genes, *gfp*, *aiiA* and *luxI*. AHL (A) is removed from the system by enzymatic degradation catalysed by the aA enzyme (aA-red). Positive and negative feedbacks steps within this network are indicated with a red plus/minus sign respectively. AHL can also freely diffuse through the cell membrane. The extracellular concentration of AHL is labelled as $\tilde{\text{A}}$ in the figure. Boundary conditions are such that extracellular AHL, $\tilde{\text{A}}$, can escape from the microfluidic chamber exit at a rate $c_{\tilde{\text{A}}}$, $\frac{\partial[\tilde{\text{A}}]}{\partial y}\Big|_{y=y_0} = c_{\tilde{\text{A}}}$.

where $[\text{A}]_i$ is the concentration of AHL produced by each cell, $D_{\tilde{A}}$ is the Brownian diffusion coefficient of extracellular AHL, $\eta_{\text{env}}$ describes the diffusion of AHL across each cell's membrane (each cell denoted with the index $i$ and $N$ cells are part of the model) while adjusting the concentration according to the microfluidic chamber's volume, and $\tau_{\tilde{\text{A}}}$ is a rate constant describing the half-life of AHL (see SI for additional details and derivation of (1)). The boundary conditions for this PDE are such that extracellular AHL can escape the microfluidic chamber through an exit, located at the $xz$-plane, such that $\frac{\partial[\tilde{\text{A}}]}{\partial y}\Big|_{y=y_0} = c_{\tilde{\text{A}}}$, as shown in **figure 1**. As with the experiments of Danino *et al.* [27] a $200 \times 50 \times 1$ $\mu\text{m}^3$ microfluidic chamber is modelled. We note that in the original work of Danino *et al.* [27] the optimal chamber height chosen to accommodate a one cell thick colony was 1.65 $\mu$m, instead of one micrometer, so as to allow optimal nutrient flow. The choice of a one micrometer tall chamber in the presented work is due to the need of still considering a one cell thick colony while fulfilling BSim [25] environment restrictions, as in the current version of BSim only integer values can be used to discretise space.

As was shown in Mina *et al.* [24], the state of the cells is affected by the availability of external AHL.

Thus, the action of an idealised controller setup, such as the one presented in **figure 2**, is to modify the concentration of external AHL in the microfluidic chamber, in an effort to make the cells output to follow a desired reference signal, $r(t)$. We investigate three types of reference signals, a constant reference value, $r_c(t)$, a sinusoidal function, $r_{\sin}(t)$, and a trapezoidal (ramp) function, $r_{\text{rmp}}(t)$ as shown in **figure 3**. The equations of the reference signals can be found in the SI accompanying this manuscript.

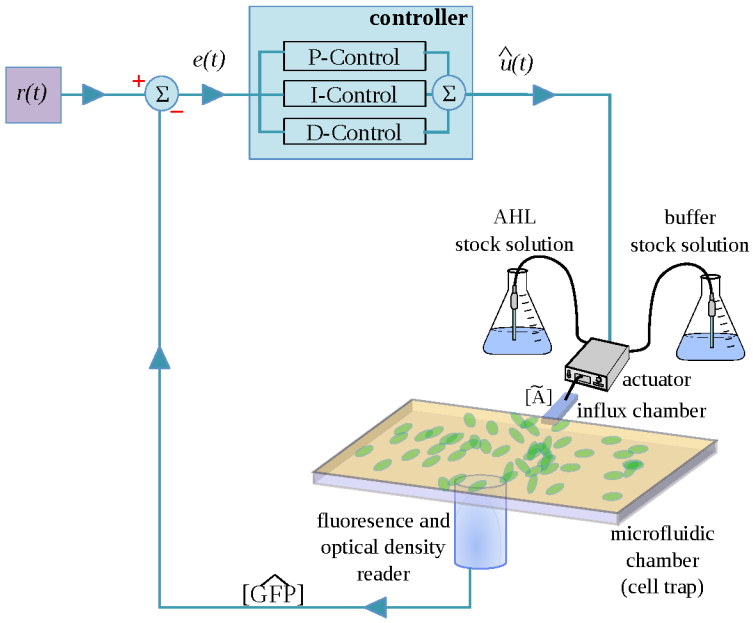



**Figure 2. Schematic of an idealised controller apparatus for implementing control over a population of cells in a microfluidic chamber.** The fluorescence and optical density meters are used to measure the population fluorescence and cell density whose values are used to estimate the average concentration of green fluorescent protein (GFP) ($\widehat{[\text{GFP}]}$) in the population [39]. The controller calculates the instantaneous error, $e(t)$, using $\widehat{[\text{GFP}]}$ and the current value of the reference signal, $r(t)$. For closed loop control, the instantaneous error is used to calculate the amount of external AHL, $\hat{u}(t)$, to be supplied to the chamber according to proportional (P), integral (I) and derivative (D) control. The modulation is accomplished using an actuator, which mixes AHL stock solution with a buffer to create the appropriate AHL concentration according to the controller's directives, and supplies the signal ($[\tilde{\text{A}}]$) to the cell trap via an influx chamber. Waste material (i.e. excess number of cells and/or chemical reaching the boundaries of the cell trap) exits the chamber from the open side, illustrated in the figure with a light blue color. In a physical implementation many cell traps can be connected in parallel and waste from all traps would be collected at a distal location [27]. For open loop control, the feedback connection between the chamber and the controller is overridden and the value $r(t)$ is used directly for controller actuation.

In the case of the periodic reference signal, $r_{\sin}(t)$, we consider the population to be entrained when the average period of the entire population is equal to the forcing period, $T_f$. The periodic behaviour of the population is assessed using Poincare sections, as described in the **Methods** section.

We investigate both open-loop and closed-loop control strategies. In the case of open-loop control, it is assumed that the reference signal, $r(t)$, modulates directly the chemical field of the entire microfluidics chamber. When closed-loop control is implemented, feedback from the cells would be required in order to calculate the correct amount of external AHL to be supplied to the chamber. As there is no direct measure of the intracellular concentrations (i.e. the system output), in an experimental

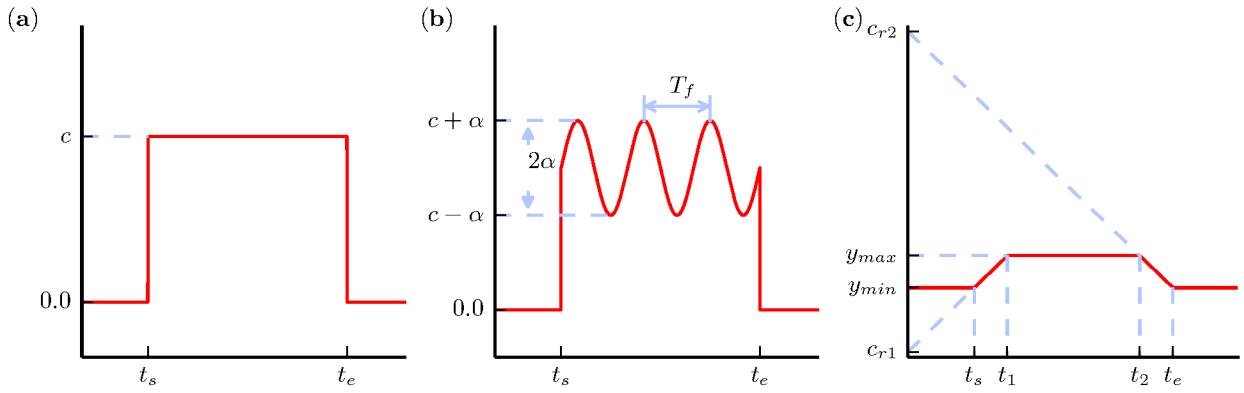



**Figure 3. The reference signals used for entrainment and tracking of the bacterial population.** The $y$-axis refers to the values of desired cell state (i.e. intracellular concentration of a protein variable) and the $x$-axis represents time. The reference signal is switched on at time $t_s$ and switched off at time $t_e$. The population is made to follow **(a)** a fixed reference value $c$, **(b)** a sinusoidal signal with period $T_f$ and **(c)** a ramp of height $h = y_{\max} - y_{\min}$. The equations producing the illustrated signals are given in the **SI** accompanying this manuscript.

setup the state of the system could be inferred from the output of GFP whose induction relies on the GRN dynamics [27]. As illustrated in **figure 2**, the population fluorescence and cell density can be obtained from measurements supplied by a fluorescence and optical density meter respectively. These measurements can be used to calculate the average concentration of GFP, $\widehat{[\mathrm{GFP}]}$, as shown in Canton *et al.* [39]. The difference between $\widehat{[\mathrm{GFP}]}$ and the reference value, $r(t)$, referred to as the instantaneous error, $e(t)$, is used by the controller to calculate the amount of external AHL to be supplied to the chamber according to proportional, integral and derivative control.

For simplicity, and in order to maintain the dimensionality of the model as low as possible so as to reduce the computational cost, and since GFP follows the dynamics of the regulatory elements of the GRN we use the internal variable LI as the reference read-out of the system (see **SI** for full set of ODEs). Thus we define the standard error between the signal's reference value and the system's current state as:

$$e(t) = r(t) - \widehat{[\mathrm{LI}]}(t) \tag{2}$$

where $\widehat{[\mathrm{LI}]}(t) = \frac{1}{N}\sum_i^N [\mathrm{LI}]_i(t)$ is the average concentration of LI of a population of size $N$ taken as the average output of the population to be controlled and $r(t)$ is the previously defined reference signal.

We note that the presented model is an idealised version of a physical implementation. Thus, with the exception of Brownian diffusion, any directional fluid flow that may result from any forcing is considered to be negligible and only modelled at the chamber exit by imposing the described boundary conditions which mimick a constant flux of autoinducer out of the chamber. Such a scenario might take place when the volume of autoinducer added by the controller during a given period of time is not forced through the chamber and is of comparable amount to that cleared at the chamber exit. Additional details with regards to the computational implementation of the model can be found in the **Methods** and **SI** sections.

## Open loop control

To investigate the feasibility of controlling an increasing population size *in silico* using a feedforward (open-loop) mechanism, we model a static (non-dividing) population and modulate the concentration of the external autoinducer chemical field, [Ã], using the time-varying input, $r_{\sin}(t)$ (equation (S18)). Specifically, the external chemical field equation becomes:

$$\frac{\partial\,[\tilde{\mathrm{A}}]}{\partial t}(x,y,t) = D_{\tilde{\mathrm{A}}}\nabla^2[\tilde{\mathrm{A}}](x,y,t) + \sum_{i=0}^{N}\eta_{env}\left([\mathrm{A}]_i(t) - [\tilde{\mathrm{A}}](x,y,t)\right) - \tau_{\tilde{\mathrm{A}}}[\tilde{\mathrm{A}}](x,y,t) + k_{\mathrm{o}}r_{\sin}(t), \qquad (3)$$

where $k_{\mathrm{o}}$, the open loop gain, is set one and we examine the effects on the population output when varying the parameters $c$, $\alpha$ and $T_f$ which determine the constant offset, the amplitude and the period of the input signal, as given in equation (S18). An example of a similar setup, where there is controlled chemical flow of an inducer molecule in a sinusoidal manner, can be found in Mondragon *et al.* [40].

We first investigate the effect of varying parameter $c$ of the sinusoidal signal, by fixing the external force frequency $T_f$ at a value away from $T_n$, the period that the cell population output illustrates when only constant forcing is considered and is in the range of approximately 360-370 minutes for the combination of $c$-values and population sizes examined in this manuscript. Results indicate that the population is completely entrained above a certain threshold for $c$. This is true even for small values of the amplitude, $\alpha$. Specifically, when setting $T_f$ to 250 minutes (this value is actually input in seconds in BSim) and $\alpha$ to 1, complete entrainment occurs when $c$ is greater than 1.5 for 21 cells, 3.4 for 60 cells and 8.4 for 100 cells. An example of the 60 cells response in the $(c,\alpha)$-plane, when $T_f$ is fixed at 250 minutes, is shown in **figure 4 (a)**. As such, for each investigated population we fix $c$ at a value lower than this threshold in order to investigate the effects of varying the forcing period, $T_f$, and amplitude, $\alpha$. The values of the control parameters used are given in the respective figures where results are illustrated and in the main text where results are described.

As shown in **figure 4 (b)**, a 21 cell population subject to variation of external AHL in the two parameters of amplitude and period $(\alpha, T_f)$ shows regions of entrainment in the $(T_f/T_n, \alpha)$-plane, close to multiples of the natural period, $T_n$. The value of the natural period, $T_n$, for a 21 cell population is approximately 368 min ($c = 1.0$). The increasing value of the amplitude $\alpha$ allows the system to be entrained to the forcing period even at values away from the natural frequency $T_n$, a behaviour typical of two distinct coupled oscillators [41]. Note that near $T_f/T_n \approx 1$, a minimal amplitude, $\alpha$, of the external input is sufficient to entrain the population (see also **supplementary figure S3**). Regions of entrainment are also found near $T_f/T_n \approx 2$, albeit at higher amplitude values. Phase diagrams of the system in the ([aA],[LI])-plane constructed from stroboscopic sections of the system's output illustrate the presence of quasi-periodic oscillations outside the entrained regions whilst a limit cycle exists within the entrained regions (see **supplementary figure S4**).

The same features observed in populations of 21 cells are also detected when larger populations are simulated. Qualitatively similar diagrams of the $(\alpha, T_f/T_n)$-plane for both 60 and 100 cells are shown in the **supplementary figure S5** where broad areas of entrainment near the natural period of the cell population can be seen. This natural period is approximately 360 minutes for both 60 and 100 cell populations ($c = 3.2$ and 5.0 respectively).

We also investigate the effect of heterogeneity by simulating a population of 60 cells and varying the production and degradation rates of the metabolic variables of each cell part of the population. To do so, we model the parameters ($\delta_1$, $\delta_2$, $\tau_A$, $\tau_{LA}$, $k_{p_{LI}}$, $k_{p_{aA}}$, $a_{0_L}$, $a_{0_A}$; see SI for further details) as random variables of a Gaussian distribution. The values are obtained using a random number generator which samples for each parameter a Gaussian distribution with mean value being the one given in **table S1** and standard deviation is a percentage of this value. Because of cell division in a controlled environment, we assume that cells will only vary slightly [35]. As such, starting from a homogeneous population

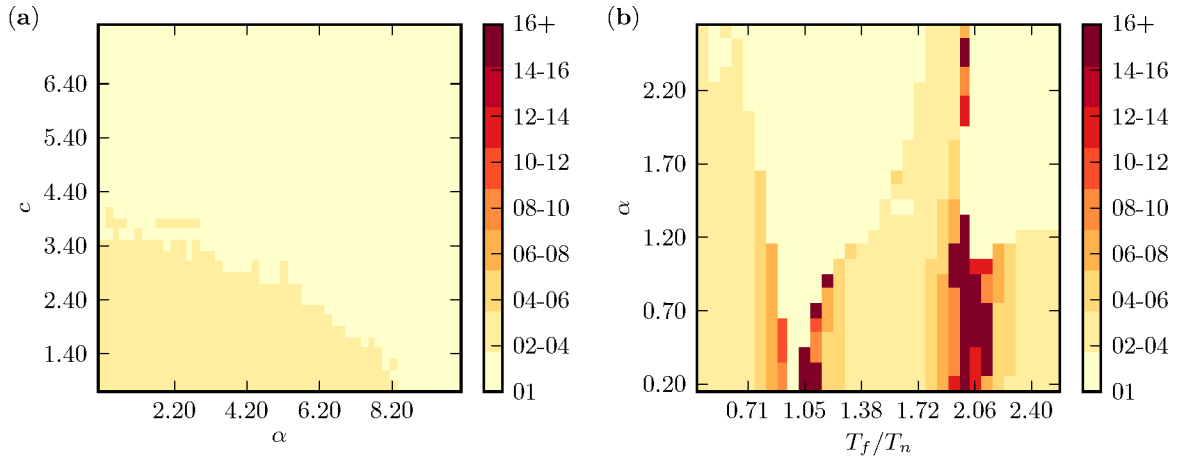



**Figure 4. Average behaviour of homogeneous cell populations under open loop control.** The ABM illustrates a wide range of dynamic behaviour when forcing the oscillating population during open loop control (equation (3)) according to the three parameters, $c$, $\alpha$, $T_f$ of $r_{\sin}(t)$ (equation (S18)). Each color pixel of the presented planes illustrates the periodic behaviour of the population average. The value of the periodic output is given as multiples of the forcing period and is illustrated in the colorbar. **(a)** The effect of increasing the constant offset, $c$, in open loop control (population of 60 cells, $T_f = 250$ min). Above a certain value for $c$ the population is completely entrained for all amplitude values, $\alpha$. This threshold increases with increasing population size. **(b)** The $(\alpha, T_f/T_n)$-plane presented above shows the periodic output a homogeneous population that is 21 cells strong ($c = 1.0$). Areas of entrainment appear at multiples of the natural period, $T_f/T_n \approx 1$ and $T_f/T_n \approx 2$. High periodic behaviour is seen before and after regions of entrainment. Similar behaviour is also seen in bigger sized populations as illustrated in the **supplementary figure S5**.

we introduce heterogeneity using a 2% step-size in the standard-deviation parameter of the Gaussian distribution, as defined above, and until synchronisation is greatly affected.

In the absence of heterogeneity all cells in the population oscillate with the same amplitude, constant offset and period, effectively responding to the forcing period in unison (see **supplementary figure S7**). However, as heterogeneity is increased the cells start to lose their synchronisation and each cell in the population starts to oscillate with slightly different amplitude and constant offset values. This becomes more pronounced at higher variability. For example, inspection of **figure 5** shows the population to be responding in unison in the 2-4% cases, even though there is no entrainment to the forcing signal. When heterogeneity is increased to 6%, the population follows the period of the forcing signal, but individual cell members illustrate great variability with respect to the average population response, in terms of amplitude and constant offset. In general, as the level of heterogeneity across the population increases the average amplitude of the population oscillatory behaviour decreases and a range of constant offset and amplitude values is seen across the oscillating population as illustrated by the histograms presented in **figure 5**.

The ABM also captures a variety of intracellular dynamical behaviours during external forcing. In **figure 6**, we present the average response of a population with 2% heterogeneity when subject to entrainment during open loop control. Note that similar regions of entrainment exist as with the homogeneous population presented in **figure 4**. The points labelled 1–6 in **figure 6** illustrate the average population response of the system dynamics and are expanded in the respective sub-panels showing the time-series and associated Poincare (stroboscopic) sections. At point one on the far left of the $(\alpha, T_f/Tn)$-plane presented in **figure 6** there is periodic behaviour on a torus governed by two

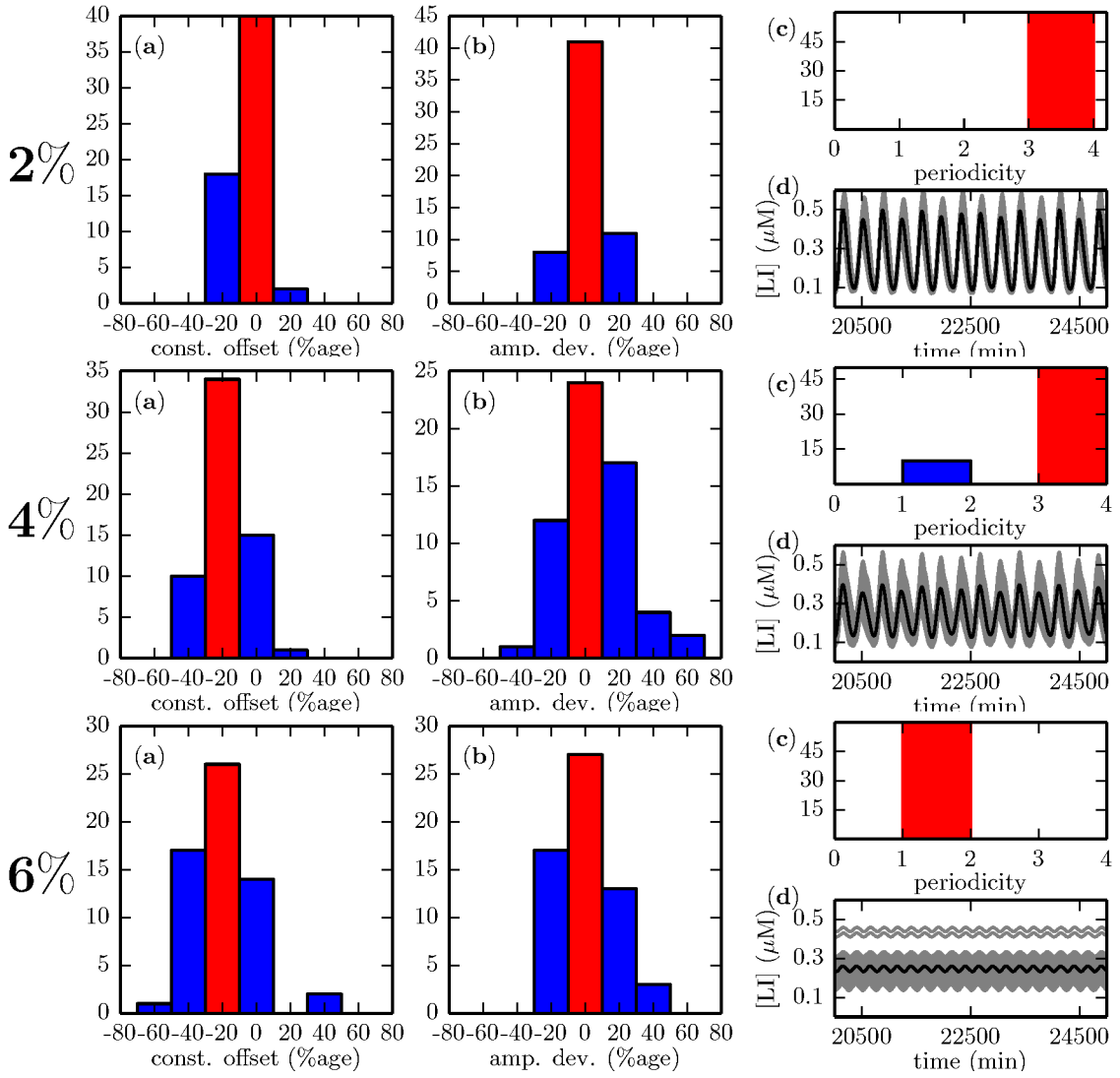



**Figure 5. The effect of heterogeneity on a 60-cell population during open loop control.** The introduction of heterogeneity in the population through parameter variation, as described in the main text, reduces the clustering effect of a population subject to forcing ($T_f = 250$ min, $c = 3.0$, $\alpha = 1.0$) as seen through variability of oscillatory output in the **(a)** constant offset and **(b)** amplitude of cells part of the population, given as a percentage deviation from the homogeneous case. **(c)** Periodicity of cells in the population. **(d)** Time-series of the entire population (grey) is compared with the population average response (black line). In the histograms (a)-(c), the $y$-axis represents the number of cells and the red column illustrates the population mode. The population response in the absence of heterogeneity, i.e. the 0% case, may be found in the **supplementary figure S7**.

periods (period-2 and period-30). As the forcing value approaches the region of entrainment the output becomes high-periodic and only dominated by a single-period as shown by sub-panel two. Entrainment is illustrated in sub-panel three. The behaviour illustrated in sub-panels four and five is similar to that of sub-panels two and three respectively. The torus behaviour governed by the two distinct periods (shown in sub-panels 1 and 4 of **figure 6**) is also sensitive to amplitude variation. For example, going from point 1 to point 6 on the $(\alpha, T_f/Tn)$-plane eliminates the high period that underlines the dynamics of point 1. At point 6 only the period-2 remains as illustrated in the power spectrums presented in **figure S6**.

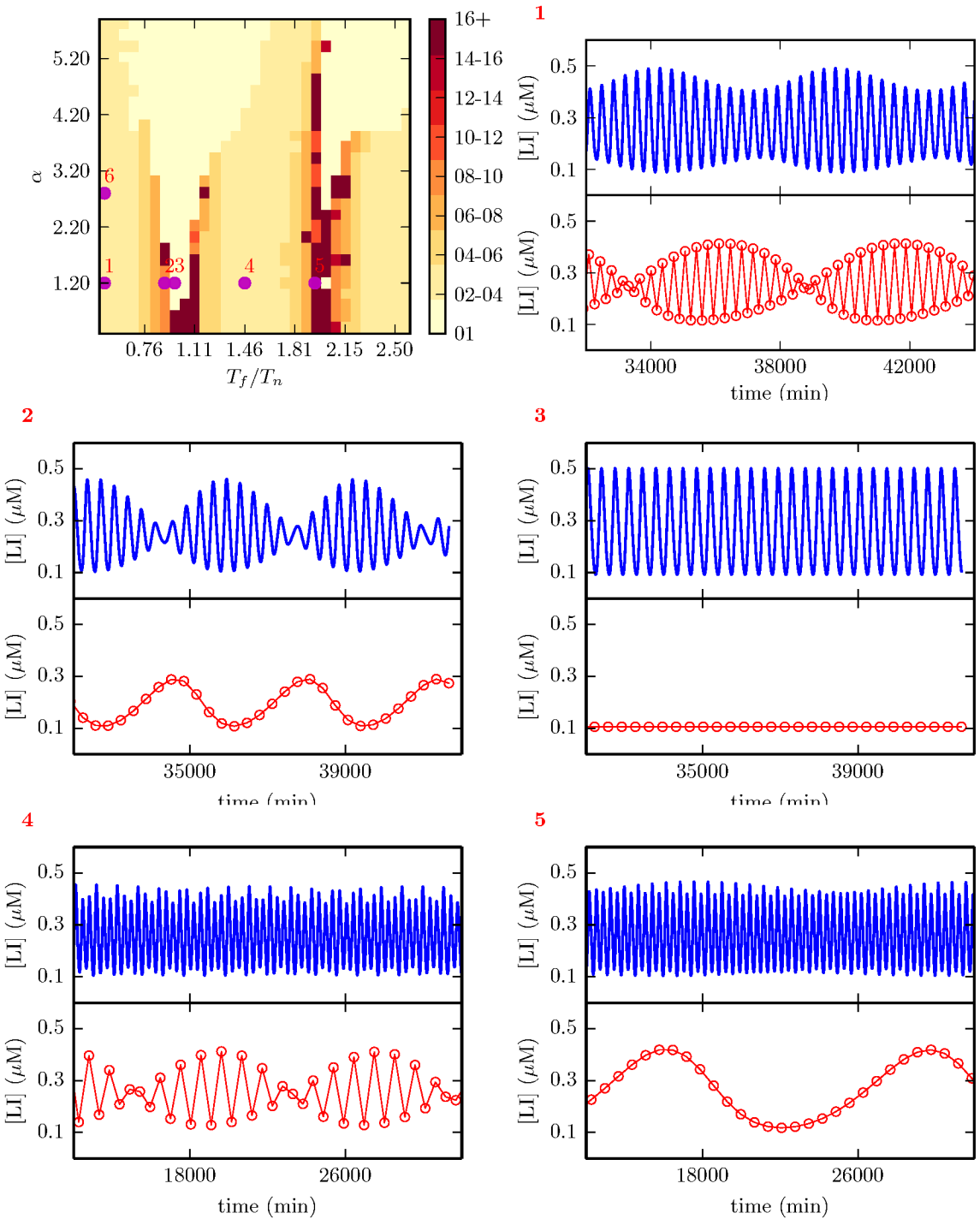



**Figure 6. Average behaviour of a 60-cell heterogeneous population during open loop control. Heterogeneity at 2%.** The agent based model illustrates a wide range of dynamic behaviour that can be seen when trying to entrain the oscillating population to the forcing period when we vary the amplitude, $\alpha$, and force period, $T_f$. The $(\alpha, T_f/T_n)$-plane presented illustrates the periodic output of the population average response. The average behaviour of the population in the annotated segments of the plane is presented in the panels that follow. These show the time-series of the output (top panels) and the Poincare (stroboscopic) sections (bottom panels) that result from strobing the output with the forcing period $T_f$. Details are given in the main text.

## Closed loop control

As seen in the previous section, and as expected from the literature [42], open loop control can only entrain a bacterial population of oscillating cells within a limited range of the control parameter values $\alpha$ and $T_f$. Specifically, entrainment is only possible near the natural frequency of the population, $T_f/T_n \approx 1$, for small values of $\alpha$, and as $\alpha$ increases so does the range of entrainment frequencies around this natural frequency (see **figure 4**). Also, there is entrainment near $T_f/T_n \approx 2$ subject however to the signal's amplitude value ($\alpha \geq 1.20$). Finally, increasing variability in the population causes cells to become less synchronised. Thus, we proceed to check whether closed loop control performs better in entraining the output of a population of oscillating cells to a desired forcing signal.

In contrast to open loop control, closed loop control continually adjusts the input to the system based on feedback from the system's current state. We implement P-control, PI-control and PID-control by adjusting the external chemical field equation as follows:

$$\begin{aligned}\frac{\partial\,[\tilde{\mathrm{A}}]}{\partial t}(x,y,t) =& D_{\tilde{\mathrm{A}}}\nabla^2[\tilde{\mathrm{A}}](x,y,t) + \sum_{i=0}^{N}\eta_{env}\left([\mathrm{A}]_i(t) - [\tilde{\mathrm{A}}](x,y,t)\right) - \tau_{\tilde{\mathrm{A}}}[\tilde{\mathrm{A}}](x,y,t) \\ & \underbrace{\underbrace{+k_p e(t)}_{\text{P-control}} + k_I\int_0^t e(\tau)d\tau}_{\text{PI-control}} + k_D\frac{de(t)}{dt} \\ & \underbrace{\qquad\qquad\qquad\qquad\qquad\qquad}_{\text{PID-control}}\end{aligned} \tag{4}$$

where $k_p$, $k_I$ and $k_D$ are the gains for the proportional, integral and derivative control action respectively and $e(t)$ is the instantaneous error defined in equation (2). Thus, the control input fluctuates the concentration of external AHL in the microfluidic chamber, based on the value of $e(t)$, which in turn affects the internal metabolic states of the cells such that the reference signal $r(t)$ is matched. Setting $k_D$ to zero allows us to investigate PI-control only and setting both $k_D$ and $k_I$ to zero allows us to investigate P-control only. Note that in classical control applications PID controllers are known to be effective to control the output of a system of interest to a constant reference signal. Here, because of the simplicity of their implementation, we also explore their ability to control the population onto a time-varying reference signal. Better performance could be certainly obtaining by considering more advanced control techniques such as model predictive control which has also been used in the literature on the control of biological systems [14]. In all simulations reported next, the control gains were selected empirically and chosen to give an acceptable performance. We firstly identified via simulation the value of control parameter $k_p$, giving an acceptable P-control performance. The process was repeated for control parameter $k_I$, where the previously identified value of $k_p$ was fixed and $k_I$ was varied until an acceptable PI-control performance was obtained. Finally, the procedure was repeated for $k_D$, where $k_p$ and $k_I$ were fixed at their identified values, and $k_D$ was varied until an acceptable PID-control performance was obtained. The closed loop controller values used are reported, where appropriate, along with the results.

We start by simulating a homogeneous population of 21 cells under the influence of P-control. We vary the amplitude and period of the reference signal $r_{\sin}(t)$ by changing the parameters $\alpha$ and $T_f$ (equation (S18)), to match the values of the $(\alpha, T_f/T_n)$-plane presented in **figure 4 (b)**. The constant offset, $c$, is fixed at the same value as for the open loop control of this population size ($c = 1.0$). As shown in **figure 7**, P-control can entrain the population over all amplitude and period values.

Furthermore, we assess the effect of the closed loop controller (P-control) with increasing levels of heterogeneity, as we did with open loop control. **Figure 8** shows the population response when 2%, 4% and 6% heterogeneity is introduced in the population (the homogeneous case, i.e. 0% heterogeneity, is presented in **supplementary figure S8**). As seen when comparing **figure 8** with **figure 5** the

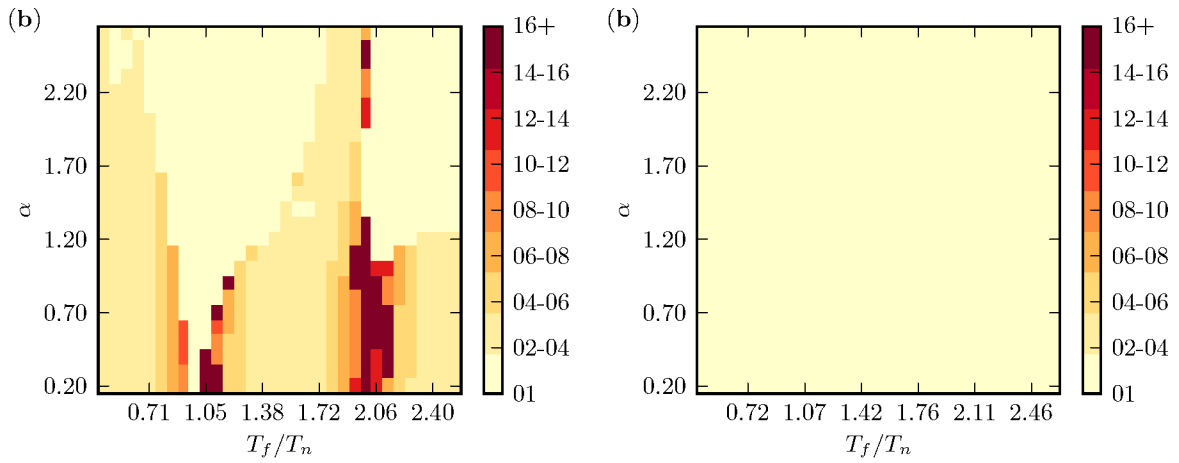



**Figure 7. Open vs closed loop entrainment for a homogeneous population of 21 cells.** P-control (right panel, $k_p = 500$) can entrain a population of 21 cells for all periodic signals whilst open loop control (left panel) can only entrain the population near multiples of natural frequencies.

closed loop controller performs better than the open loop control by minimising the variability of the population in terms of amplitude and constant offset. In addition to this, closed loop control does not reduce the average amplitude of the population, as opposed to the open loop controller. We also note that the closed loop controller manages to maintain the mode of the population in terms of constant offset and amplitude for all levels of heterogeneity. That is, the population mode for both the constant offset and amplitude is the same as with the homogeneous case (presented in **figure S8**), unlike the open loop results.

In addition to the sinusoidal reference signal, $r_{\text{sin}}$ (equation (S18)), we also test closed-loop control on heterogeneous populations of 21, 60 and 100 cells (heterogeneity at 2%) when using a constant or a ramp reference signal ($r_{\text{c}}$ and $r_{\text{rmp}}$ defined by equations (S17) and (S19) respectively). As seen in **figure 9**, where the response of a 60 cell population subjected to P-control with all three reference signals is shown as a representative example, a standard error difference between the population average and the reference value is maintained. This is true for all simulated population sizes.

To try and minimise the standard error we introduce integral and derivative actions to the control scheme. As a representantive example, the response of a 60 cell heterogeneous population undergoing PI-control and PID-control for the fixed reference signal, $r_{\text{c}}$, is presented in **figure 10**. PI-control minimises the standard error to 1.47%, as opposed to 28.2% when only P-control is used, but further improvement results to small amplitude oscillations around the reference value as illustrated in **figure 10 (a)**. PID-control, shown in **figure 10 (b)**, can reach 93% of the reference value (i.e. 1.4 $\mu$M when the reference value is 1.5 $\mu$M) in less than 48 hours ($\approx$ 46.5 hours) and also minimises the standard error to 0.33% of the reference value. In addition to this, the PID-controller has an improved response time over PI-control. Specifically for the 60-cell example illustrated in **figure 10**, PI-control has a settling time, the time needed to reach the peak value (the maximum value the population assumes with PI-control is 1.47 $\mu$M), of $\approx$ 1550 hours whilst PID-control can reach the same value in $\approx$ 245 hours and has a settling time of $\approx$ 737 hours (the maximum value the population assumes with PID-control is 1.495 $\mu$M).

As seen, PID-control guarantees the best performance in terms of the population's settling time and in minimising the standard error. However, in a physical setting the population will keep growing through cell division to fill-up the chamber, and not remain static as with the modelled population sizes. Also, some of the population members will undergo cell death. It is not unlikely that the constant

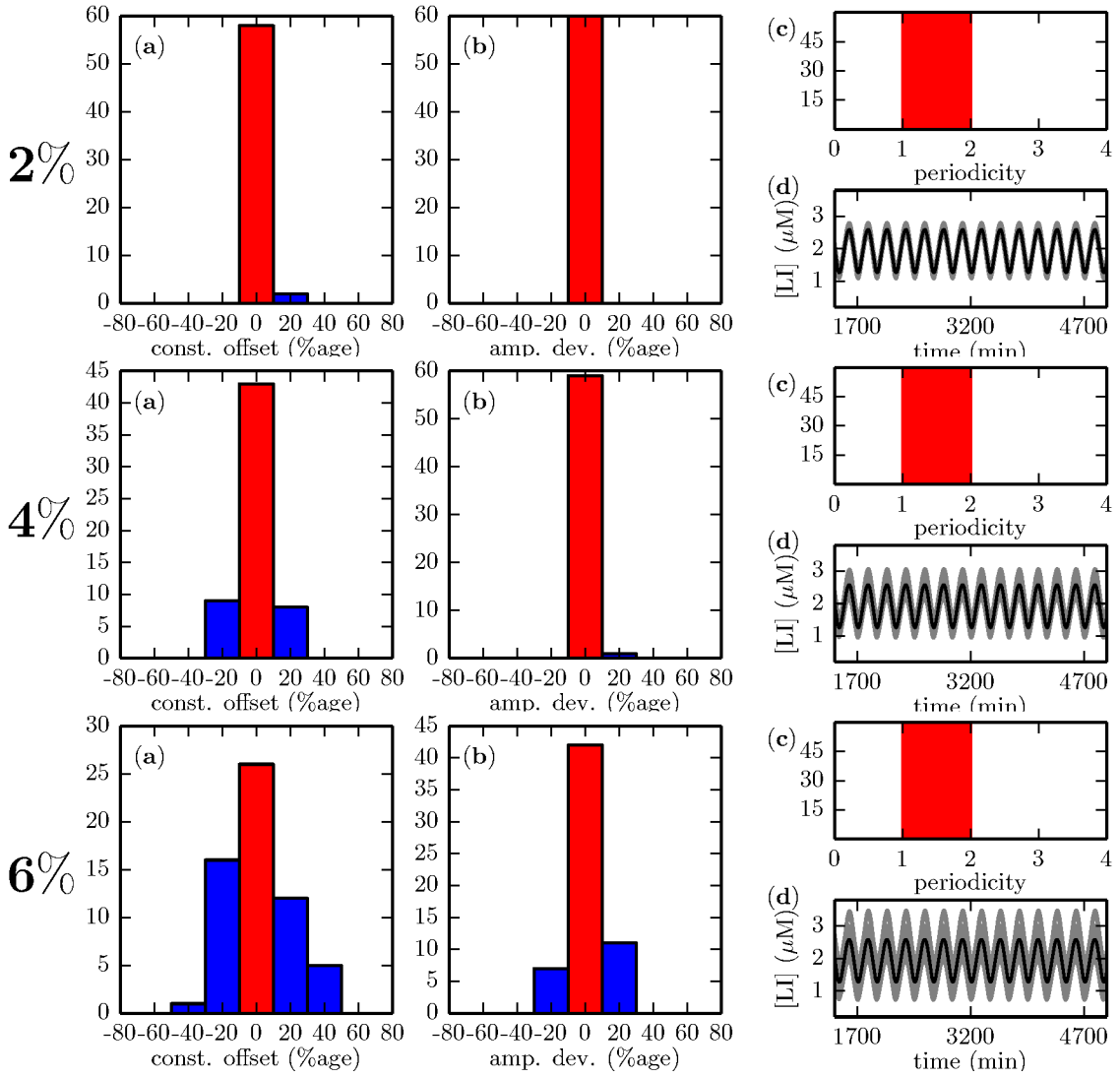



**Figure 8. The effect of heterogeneity on a 60-cell population under P-control.** The closed loop controller minimises the variability of oscillatory output in terms of the constant offset and amplitude in cells part of the population with respect to the open loop control (compare with **figure 5**, $T_f = 250$ min, $c = 3.0$, $\alpha = 1$). The histograms show the **(a)** constant offset and **(b)** amplitude of the cells oscillatory output, given as a percentage deviation from the homogeneous case. **(c)** Periodicity of cells in the population. **(d)** Time-series of the entire population (grey) compared with population average response (black line). In the histograms (a)-(c), the $y$-axis represents the number of cells and the red column illustrates the population mode. The homogeneous case, i.e. 0% heterogeneity, is given in the **supplementary figure S8**.

change in population numbers, brought about by cell division and cell death, may have an effect on the performance of the chosen control strategy. To investigate any such effects, we model a growing, motile, heterogeneous population that undergoes both cell division and cell death. In principle, and using the assumption that each *E.coli* cell has a volume of 1 $\mu$m$^3$ [43], a $200 \times 50 \times 1$ $\mu$m$^3$ chamber can hold up to 10000 cells. Hence, after confirming that a 10000 strong population, with average doubling time of 40 minutes, can oscillate autonomously (see **figure 11**) we proceed to test the PID-control strategy with fixed control gains ($k_p = 8.3$ min$^{-1}$, $k_I = 2.03E-4$ min$^{-1}$, $k_D = 8.3E4$ min$^{-1}$). According to the results presented earlier, the settling times of PID-control are much longer than the time required to

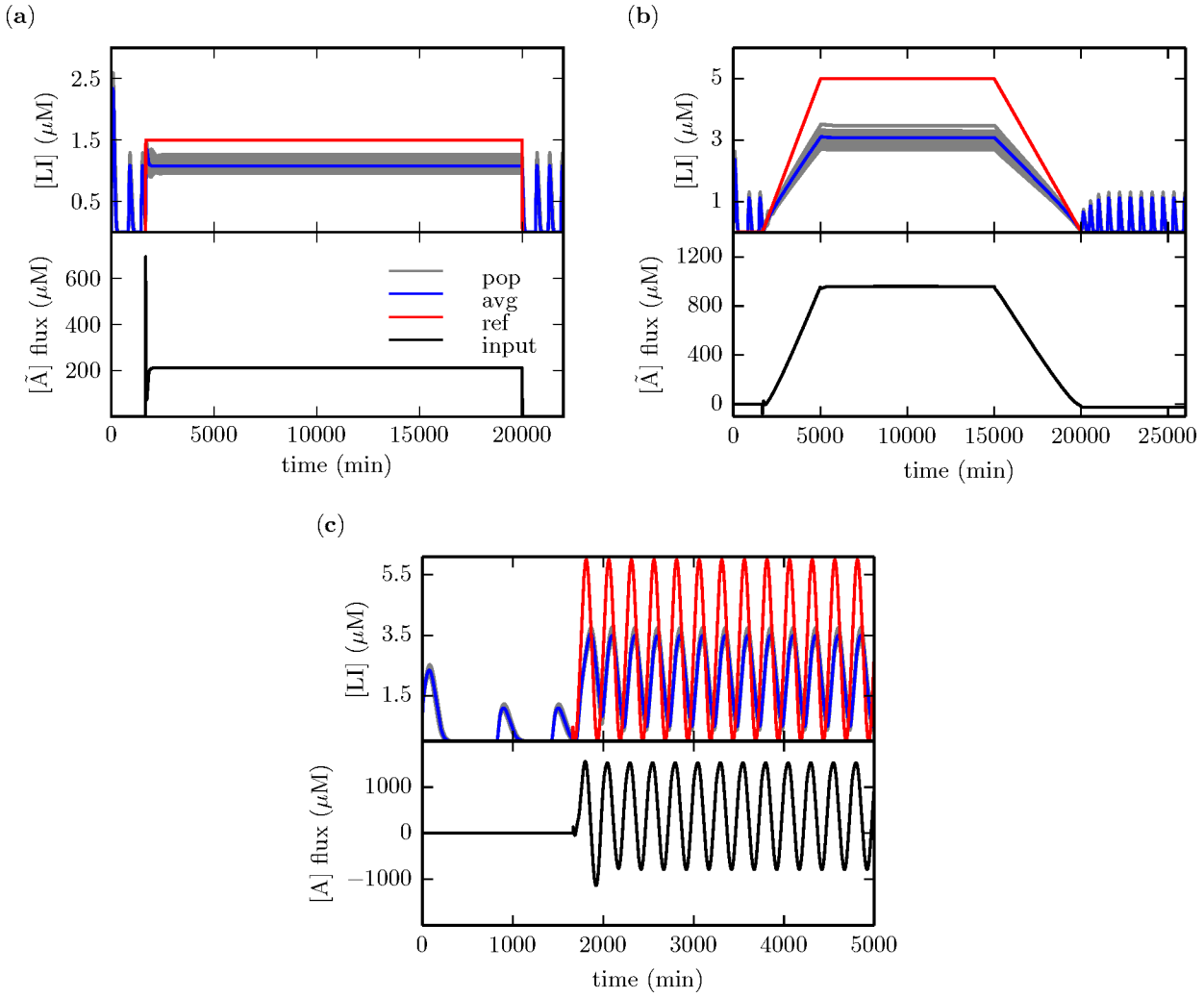



**Figure 9. A heterogeneous population of 60 cells under the action of P-control.** The figure illustrates time-series of a 60 cell population with 2% heterogeneity subject to P-control ($k_p = 8.3$ min$^{-1}, c = 1.5$ $\mu$M$, t_s \approx 1660$ min$, t_e = 20000$ min). The blue line is the population average,the grey lines are the plots of all the population members and the red line is the value of the reference signal. With P-control there always remains a standard error, a difference between the reference value and the population output (here the average response). P-control allows cells to follow non-oscillatory reference signals like **(a)** fixed values ($c = 1.5$ $\mu$M) and **(b)** ramps ($y_{\min} = 0$ $\mu$M$, y_{\max} = 5$ $\mu$M) in addition to **(c)** entraining the population to a periodic input ($c = 3$ $\mu$M$, \alpha = 3, T_f = 250$ min) away from the natural period ($T_n = 23100$ sec). The bottom time-series in each panel indicates the control effort, the concentration of AHL required as input from the controller for each respective action.

reach a 10000 strong population when cells have an average doubling time of approximately 40 minutes. As such, to identify if and when control fails we run individual simulations with different maximum population sizes, up until the 10000 cell limit.

Inspection of **figure 11** shows that our model predicts that PID-control with fixed gains can only harness a limiting population size. In all three reference-signal cases there is a notable decrease in the control performance with increasing population size. Finally, control, with the selected values of

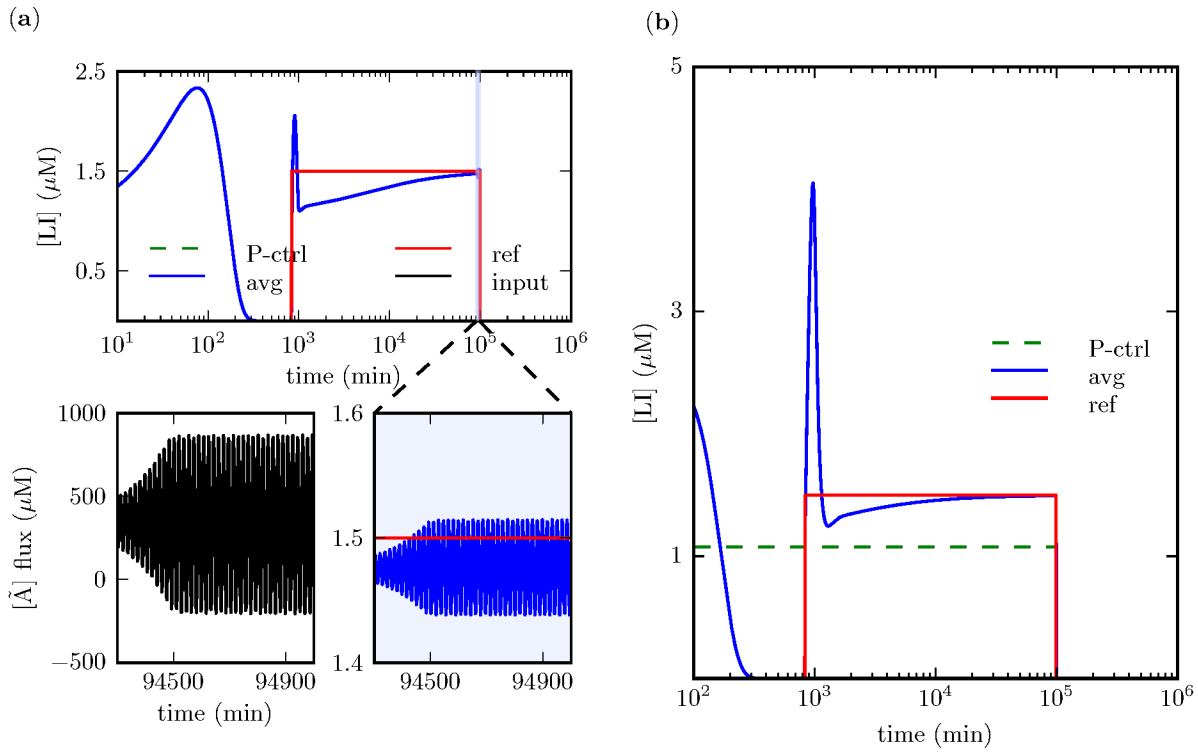



**Figure 10. A heterogeneous population of 60 cells under PI-control and PID-control.** The figure illustrates the time-series of a 60 cell population with 2% heterogeneity. The blue line is the population average response, the red line is the value of the reference signal, $r_c(t) = 1.5$, and the dashed green line the average response of the population under proportional only control (see **figure 9 (a)**). **(a)** PI-control reduces the standard error to 1.47% of the reference value, however it does so with a slow transient. The lower panels zoom into the highlighted region of the top panel and illustrate the high frequency oscillations that can occur near the reference value and the respective chemical fluctuations of the control action during that time period ($k_p = 8.3$ min$^{-1}$ ,$k_I = 3.97E-5$ min$^{-1}$, $c = 1.5$, $t_s = 833$ min, $t_e = 99166$ min). **(b)** PID-control reduces the standard error to 0.33% of the reference value. The lower panel illustrates the control effort, the amount of AHL required as input by the controller ($k_p = 8.3$ min$^{-1}$, $k_I = 2.03E-4$ min$^{-1}$, $k_D = 8.3E4$ min$^{-1}$).

the gains, completely fails for population sizes beyond 2500 cells (see supplementary **figure S9**). It is likely that the increase in population size affects the concentration of external AHL. The positive feedback loop present in the GRN also ensures (periodic) production of AHL from each cell which ultimately may contribute to an increase in the availability of external AHL in the microfluidic chamber at concentrations which may be higher than the control effort. This may render external contribution insufficient to control the system and cells may follow the stronger (of higher concentration) native signal when compared to the control effort. This speculative scenario seems to be suggested by **figure S9** as well, where in the presence of a constant reference signal the controller cannot harness a population of 3000 cells which seems to be in an unstable oscillatory state.

## Discussion

The investigation we presented show that spatially resolved agent based models such as the one presented in Mina *et al.* [24] can be useful to implement and test *in silico* strategies for exerting control over

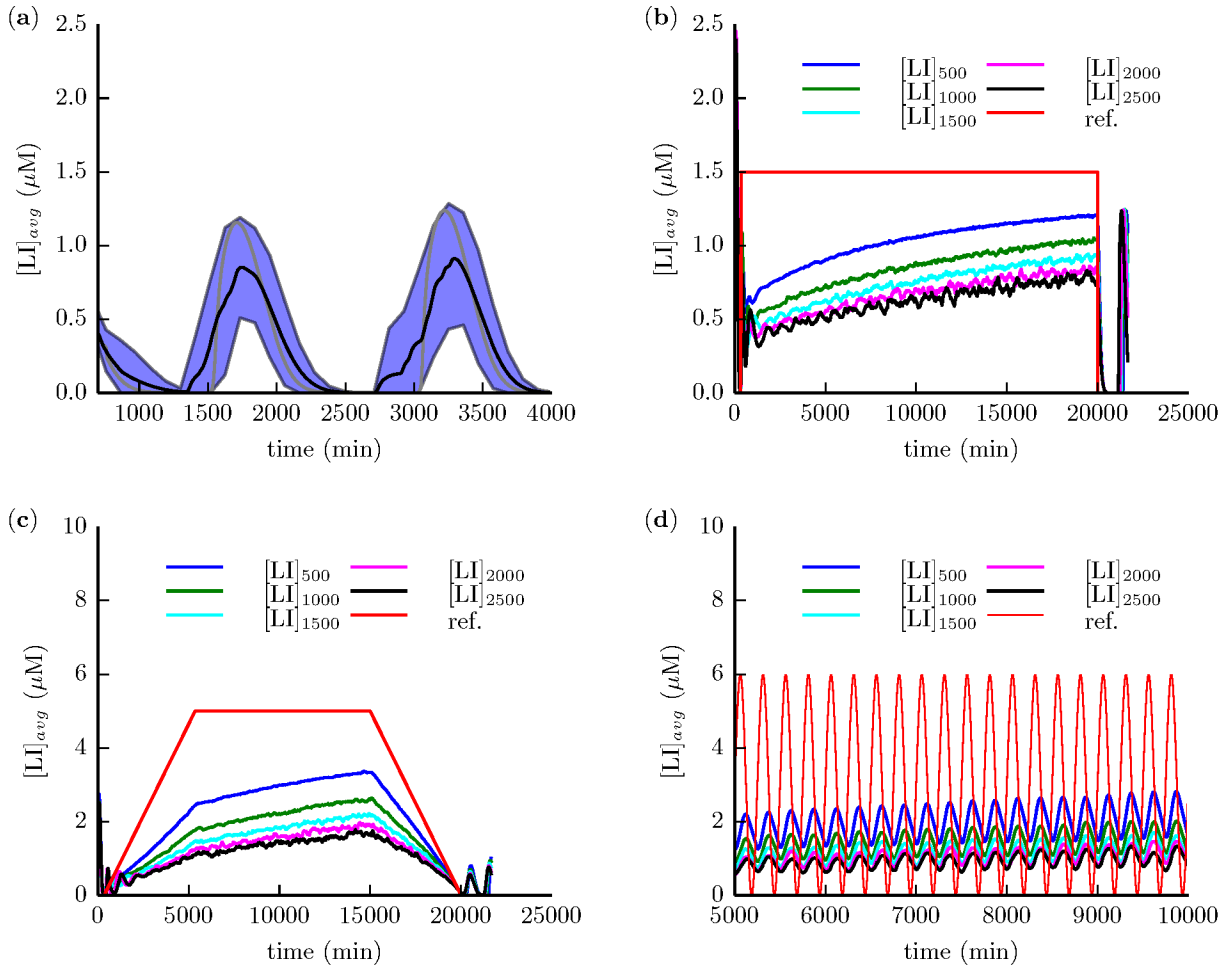



**Figure 11.** **(a)** The population average response for LI (black line illustrates the average, the blue envelope the standard deviation arising from 10 individual simulations and the grey line a typical time-series from one simulation) when there is cell division and cell death in a maximum population size of 10000 cells. Transient behaviour (first 500 min) not shown. **(b-d)** Time-series of a replicating population, as described in the main text, with maximum size of 500 (blue), 1000 (green), 1500 (cyan), 2000 (magenta) and 2500 (black) cells for a constant ($c$ =1.5 $\mu$M), ramp ($y_{min} = 0$ $\mu$M, $y_{max} = 5$ $\mu$M) and oscillatory ($c = 3$ $\mu$M, $\alpha = 3$ and $T_f = 250$ min) reference signals respectively. It is evident that there is a degrading performance of the PID-control controller with fixed gains as the population size increases. In the case of the constant signal **(a)**, there is an increasing non-smooth population output with an increasing population size whilst the signal is on. For the ramp signal **(b)**, the ramp's plateau is not reached during the specified time with the output's height diminishing whilst the population size grows. This is more pronounced in the 2500 population size scenario, where the ramp legs are indistinguishable from the ramp plateau and the population output looks parabolic instead of trapezoidal. For the oscillatory case **(d)**, even though the period of the reference signal is followed there is a diminishing amplitude in the population's output as the population grows (panel does not illustrate transient behaviour at the start of the simulation).

cellular populations. As application of control over cellular populations [13, 14, 16, 19, 40] is becoming more common, the use of ABMs and their numerical implementation via agent-based simulators such as BSim can indicate which control strategies would likely be more successful and also uncover the dynamic

behaviours to be expected.

More specifically, in this paper we adapted the spatiotemporal model presented in Mina *et al.* [24] to include open and closed loop strategies to assess if a population of oscillating cells can be entrained to a different input. We studied whether the cellular population can be forced to change dynamic behaviour by following various reference signals using linear methods of control, specifically open loop and classical feedback-based interventions (P-control, PI-control and PID-control), since in a physical setting these can be executed using real-time computation due to the simplicity of the underlying mathematics. In order to study the parameter dependence of the system's behaviour, we focused initially on studying small-sized populations.

Firstly, an open loop controller was implemented computationally by creating a sinusoidal flux of the extracellular coupling chemical, $\tilde{\mathrm{A}}$, through the microfluidic chamber. With increasing levels of heterogeneity, open loop control of the system resulted in loss of clustering between coupled cells as well as a diminished individual cell response in terms of amplitude (see **figure 5**). This is important as oscillatory biological mechanisms are known to operate within defined frequencies and signal strengths [44, 45] and any deviation away from these may be considered as biochemical noise and subsequently ignored by the system [36, 46]. As such, any signal obtained through control using open loop methods may be ineffective in initiating further response in the system of interest.

With open loop control, entrainment over populations of varying size was possible in a subset of the amplitude and forcing periods as shown in **figure 4 (b)**. As expected from the theory of phase locking in dynamical systems [41, 42], entrainment occurred near the natural periods of the oscillating population. Analysis of the model output showed that rich periodic behaviour can be found around entrainment regions with the system changing qualitative behaviour across the entrainment boundary as shown in **figure 6** and **S4**. Unexpected periodic behaviour could be of significance in biological systems as they may have a suboptimal response away from specific frequencies or amplitude thresholds [45, 47] or may result in different responses [44, 48]. For example, a system may have an additional GRN that responds to the same chemical input but only when this input is characterised by a specific oscillatory waveform which may appear during open loop control and subsequently lead to unexpected behaviour.

We note that the triangular regions of entrainment stemming from $T_f/T_n \approx 1$ in **figures 4 (b)** and **S5** are reminiscent of Arnold tongues, areas of synchronisation in the $(T_f, \alpha)$-plane bounded by two arcs that intersect the $\alpha = 0$ line [42]. In Arnold tongues, these arcs define a boundary of saddle-node type bifurcations [41], i.e. a change of the system from a stable to an unstable steady state. Such qualitative change of the system's behaviour across this boundary is shown in **figure S4** where phase diagrams are produced by varying the control parameter $T_f$. Phase diagrams of the system in the ([aA],[LI])-plane, constructed from stroboscopic sections of the system's output, illustrate the change from an unstable quasi-periodic oscillation to an attractor point in the phase plane, indicative of a stable limit cycle.

Also, according to linear oscillator theory such triangular regions of synchronisation (i.e. the Arnold tongues) appear at rational multiples of the system's natural frequency [42], here denoted as $T_n$, and when the forcing is not too strong. As can be seen from inspection of **figures 4 (b)** and **S5** there is also synchronisation at $T_f/T_n \approx 2$, although this does not touch the horizontal axis, i.e. when $\alpha = 0$. It is to be noted however that Arnold tongues, according to linear oscillator theory, get thinner (i.e. the arcs come closer together) as we move away from $T_f/T_n \approx 1$ [41, 42], so perhaps it is difficult to visualise in full such synchronised regions especially when using our brute-force method. Finally, we note that our system is far from linear and that we do not prove that the regions of synchronisation seen in **figures 4 (b)** and **S5** are indeed Arnold tongues. Such proof is beyond the scope of this paper but could be the subject of future work.

Physical implementations of similar controller apparatus have been shown possible. For example, in Mondragon *et al.* [40], a population of uncoupled oscillating cells is successfully entrained at regions near the system's natural frequency using an autoinducer which activates cells but is not responsible for cell-cell communication. As such, it would be very interesting to investigate experimentally whether

application of the open loop controller presented here, where cell-cell communication exists, creates areas of synchronisation confined to integer multiples of the system's natural frequency, as seen through our results. However, we note that in the context of our study, we consider a scenario where the controller action does not upset the system in terms of cell density or by the creation of chemical gradients. This is important, as both cell density and spatial dispersion of the coupling chemical can affect the synchrony of oscillations across the population [24]. For example, higher values of the Brownian diffusion coefficient which allow for rapid travel of AHL in the chamber promote synchrony across the population by disallowing the formation of local chemical gradients, whereas for smaller values of the diffusion constant the converse is true [24]. Thus, if the physical implementation of such a controller is inconsistent with the presented model, such as for example by forcing fluid through the chamber in a manner which upsets cell density and chemical gradients, we may end up with unexpected behaviour.

If the objective is to entrain a population of oscillators to a specific periodic behaviour then a closed loop strategy is preferable as shown in **figure 7**, since open loop control is only successful in entraining the system of interest near multiples of its natural frequency, $T_n$. Also, the model revealed that during open loop control cell-to-cell variability resulted in loss of clustering in the forced population as shown in **figure 5** but this level of variability was minimised for the closed loop controller as seen in **figure 8**. Furthermore, if we wish to stabilise the population around a specific value then PID-control action is preferred as it maintains a minimal standard error and allows to reach the reference value in the least amount of time (see **figures 9–10**). This is better understood when inspecting equation (4). Even though most of the control effort is contributed by P-control (see **figure 9**), a standard error is maintained between the reference value and the system's steady state [1]. A contribution to the control effort given by the integral control term generally eliminates the error in the presence of a constant reference signal but may cause overshooting as error correction is based on accumulation of past errors [1]. Finally, the settling time is minimised since the derivative control acts as a linear predictor for the system by estimating the rate of change of the error and contributes to minimising it based on this rate [1]. However, even with a simple closed loop controller one cannot control an ever growing population with fixed control gains as illustrated by the results shown in **figure 11**. This suggests the possible usage of more advanced control strategies such as adaptive controllers, where the control gain parameters are time-varying and adapted from the system's input/output response in conjunction with the reference signal [49].

In conclusion, testing control methods computationally may provide insight to which physical implementations of control would likely be more successful for harnessing a biological system. Furthermore, such testing may give insight to the dynamic behaviours to be expected in the presence of spatial dynamics that cannot be neglected, such as high periodic oscillatory behaviour seen near areas of entrainment, which may cause a biological system to have undesirable output [44,45,47,48]. As shown here, a breadth of responses may manifest depending on the strategy employed (e.g. open vs closed loop) and using agent based models one can quickly and inexpensively (when compared to the physical implementation of the controller) indicate the most promising strategies since both the average population response and the individual cell response may be retrieved from the model.

# Methods

## Numerical methods in BSim

As in Mina *et al.* [24], we implement the model in BSim [26] an open software platform [25] developed using the Java programming language [50] to study bacterial populations *in silico*. BSim is a 3D framework for simulating bacterial populations [25, 26] and has numerical solvers for both ODEs and partial differential equations (PDEs). We used the Runge-Kutta order four-to-five ODE solver when solving the system of ODEs presented in equations (S4)–(S7) in the supplementary information section.

A simple finite-differences scheme is used to solve for the reaction-diffusion PDE presented in (1). Depending on the population size modelled, a time-step of 0.01 or 0.05 seconds was used and time-series data of the variables were output every 50 or 100 seconds of the simulation. The bigger time-step is used when simulating the bigger population size to reduce the computational time. The spatial aspect of the model, i.e. the $200 \times 50 \times 1$ $\mu$m$^3$ microfluidic chamber, was discretised into $5 \times 5 \times 1$ $\mu$m$^3$ elements as this was shown to be the smallest element that did not cause the finite-difference scheme to be unstable during numerical solution of the model.

To ensure that results obtained from the model are numerically correct, solutions obtained from BSim were confirmed with the numerical solutions produced from other software. Specifically, the ODE–PDE coupling was momentarily uncoupled and the numerical solution of the separate components was compared with results obtained from XPPAUT [51] and MATLAB [52]. Results obtained and presented in the main text when solving the system of equations (S4)–(S7), presented in the supplementary information, in BSim using the Runge-Kutta order four-to-five ODE solver, were compared against the respective schemes of XPPAUT [51] (`Runge-Kutta`) and MATLAB [52] (`ode45`). The validity of the PDE solver was checked by introducing a fixed quantity of extracellular AHL in the environment, in the absence of cells, in order to obtain a degradation profile time-series (i.e. $\frac{\partial[\tilde{\mathrm{A}}]}{\partial t} = D_{\tilde{A}}\nabla^2[\tilde{\mathrm{A}}] - \tau_{\tilde{A}}[\tilde{\mathrm{A}}]$). This time-series was compared with results obtained when the equivalent ODE, $\frac{d[\tilde{\mathrm{A}}]}{dt} = -\tau_{\tilde{A}}[\tilde{\mathrm{A}}]$, was solved in XPPAUT [51] and MATLAB [52].

## Extending BSim to include open and closed loop control

In order to implement and test the open and closed loop controllers described in the main text we added the necessary functionality to the BSim source code by (i) extending the `BSimChemicalField()` class into the `BSimChemFieldExt()` class to include methods for open loop control and (ii) creating the `ControlMethods()` and `RefSignals()` classes to allow for the implementation of closed loop control methods and reference signals, respectively, in the BSim environment. The classes are summarised in **table 1**. For clarity, in the description that follows the parameters found within the brackets of each class or method correspond to the quantities defined in the main text.

The `BSimChemFieldExt()` class inherits all methods and constructors from the chemical field class already implemented in BSim, `BSimChemicalField()`. With the `BSimChemFieldExt()` class a chemical field is constructed in the BSim environment that can be manipulated with two extra methods not available in the `BSimChemicalField()` class. The `extConstantAdd(`$c$`)` method is used to add a fixed amount, $c$, of chemical in the field at each time-step as in $r_{\mathrm{c}}(t)$ (equation (S17)). The `extModSignal(`$T_f, \alpha, c, t_s$`)` class is used to modify the chemical field with the sinusoidal function $r_{\sin}(t)$ (equation (S18) in SI). The amount to be added at each time-step, calculated according to the respective equations, is divided into equal parts according to the number of discretised elements of the chemical field. Then each fraction is added to each discretised element even though in experimental setups there are limited entry and exit points providing access to the chamber [12, 53], as illustrated in **figure 2** of the main text.

We also note that directional fluid flow is only considered at the chamber exit, as per the boundary conditions of the PDE, where autoinducer is cleared from the chamber by a fixed amount, $c_{\tilde{A}}$. Variation of $c_{\tilde{A}}$ did not affect results (see **figure S10** in SI). Thus, if addition of extra autoinducer in the chamber produces negligible flux, such as for example the volume of added chemical is comparable to the volume of chemical cleared at the chamber exit and that the chemical is not forced through the chamber, then the physical setting and our modelling implementation should not produce conflicting results as the amount added to the chamber via the controller diffuses almost instantly throughout the whole chamber as shown in **supplementary figure S2**. It is also worth noting that the external AHL spatial profile remains uniform in space throughout the autonomous oscillations of the population (i.e. in the absence of control) as illustrated in the **supplementary figure S1**.

For closed loop control, relevant methods are implemented with the `RefSignals()` and

`ControlMethods()` classes. The `RefSignals()` class creates the reference signals $r_{\mathrm{c}}(t)$, $r_{\sin}(t)$ and $r_{\mathrm{rmp}}(t)$ presented in equations (S17)–(S19) of the SI. $r_{\mathrm{c}}(t)$ is created with the `refSignalConst(`$c$`)` method, $r_{\mathrm{rmp}}(t)$ with the `refSignalRamp(sim,`$y_{\mathtt{max}}, y_{\mathtt{min}}, t_s, t_1, t_2, t_e$`)` method and $r_{\sin}(t)$ with the `refSignalSin(sim,`$T_f, \alpha, c$`)` method. The `ControlMethods(`$k_p, \epsilon, I_i, D_i$`,chmfield,`$t_s, t_e$`)` class creates an object (where object refers to the Java technical term) that acts on the chemical field `chmfield`. The integral and derivative control gains are calculated as $k_I = k_p/I_i$ where $I_i$ is known as the integral time and $k_D = k_p D_i$ where $D_i$ is known as the derivative time and $k_p$ is the P-control gain. Control can be switched off when the standard error (the difference between the target value and the current value of the system) is equal to $\epsilon$. For these simulations we set $\epsilon = 0$. Three closed loop control methods are available, `PIDCtrl(),PICtrl(),PCtrl()` for PID-control, PI-control and P-control respectively.

Briefly, we describe the `PIDCtrl()` method. `PIDCtrl(sim,chmfield,`$r(t)$`,avg)` adds PID-control to the simulation environment `sim` that holds the chemical field `chmfield`. The controller action is calculated accordingly for proportional, integral and derivative action from the instantaneous error $e(t)$, using the `PCtrlRtnVal(avg,`$r(t)$`)`, `ICtrlRtnVal(sim,avg,`$r(t)$`)`, `DCtrlRtnVal(sim,avg,`$r(t)$`)` methods. These methods first calculate the instantaneous error by finding the difference between the average population response `avg` and the reference value $r(t)$ where $r(t)$ can be any of the defined reference signals given in equations (S17)–(S19) of the SI. The amount of chemical that needs to be added or removed from the field is carried out by the `distribchem(sim,chemicalinput)` method where `chemicalinput` is the total amount calculated by the controller and retrieved using the `getchemfluxinput()` method. The adjustment is carried out by adding or removing fractions from the discretised elements of the chemical chamber as explained earlier with the `extModSignal()` method. The PI-control and P-control methods work in similar fashion taking into account their respective controller actions.

## Assessing periodic behaviour with Poincare stroboscopic sections

Depending on the population size, a time-step of 0.01 or 0.05 sec was used and time-series data of the variables were output every 50 or 100 sec of the simulation. The average response of the population was calculated during the BSim simulations based on the population size and the variables presented in the ODE equations (S4)–(S7) and the PDE equation (3) presented in the main text. As already mentioned, BSim uses a Runge-Kutta 45 algorithm to solve the ODEs of the intracellular dynamics and a finite difference scheme is used to solve for the chemical field PDE.

Each BSim simulation was run for a prolonged time such that transient behaviour had died out prior to assessing the periodicity of the system's output when subject to control. Post-simulation the periodic behaviour was assessed by creating a Poincare section of the average population response using the final entries of the generated time-series. The procedure is outlined below.

Depending on the length of simulation time, the last 12000-20000 entries of the output data were used. Stroboscopic sections of this truncated series were generated by assessing the average response output at multiples of the forcing period, $T_f$. A linear interpolator between two time-segments was used when the stroboscopic section did not coincide with a 50 or 100 sec increment. The result of this processing was the generation of a Poincare time series of the average population response at multiples of the forcing period $T_f$. This post-simulation processing was carried out in MATLAB [52] using custom written software.

The Poincare time-series was then assessed for periodic behaviour using MATLAB's fast fourrier transform (FFT) algorithm [52] and generating a power spectrum. If more than one frequency was present, the most powerful was selected as the one to be used in the plotting of the $(\alpha, T_f)$-plane. If the strongest frequency present in the power spectrum was below a $5E-3$ threshold then the output was considered to be entrained to the forcing period.

| Method name | Description |
|---|---|
| | `BSimChemFieldExt()` |
| `extModSignal()` | Modulates chemical field with sin signal for open loop control. |
| `extConstantAdd()` | Adds constant amount of chemical into chemical field. |
| | `ControlMethods()` |
| `updateerrorlog()` | Updates vector with the standard error of the last two timesteps. |
| `distribchem()` | Adjusts the chemical field according to the control law action. |
| `PCtrlRtnVal()` | Calculates amount of chemical required according to P-control. |
| `ICtrlRtnVal()` | Calculates amount of chemical required according to I-control. |
| `DCtrlRtnVal()` | Calculates amount of chemical required according to D-control. |
| `PIDCtrl()` | Implements PID-control in the simulation environment. |
| `PICtrl()` | Implements PI-control in the simulation environment. |
| `PCtrl()` | Implements P-control in the simulation environment. |
| `getPctrlval()` | Returns the value calculated by `PCtrlRtnVal()`. |
| `getIctrlval()` | Returns the value calculated by `ICtrlRtnVal()`. |
| `getDctrlval()` | Returns the value calculated by `DCtrlRtnVal()`. |
| `getchemfluxinput()` | Returns the value calculated by the controller. |
| | `RefSignals()` |
| `refSignalConst()` | Creates a constant reference signal as in $r_{\mathrm{c}}(t)$. |
| `refSignalRamp()` | Creates a trapezoid reference signal as in $r_{\mathrm{rmp}}(t)$. |
| `refSignalSin()` | Creates a sinusoidal reference signal as in $r_{\sin}(t)$. |
| `getrefsignal()` | Returns the value of the reference signal at the current timestep. |

**Table 1. Java classes added to the BSim source code for implementing control methods.** The methods have been split into the three implemented classes. The `BSimChemFieldExt()` class is for open loop control and the `ControlMethods()` and `RefSignals()` classes for closed loop control.

## Supporting Information

The supporting information contains the model derivation, the supplementary figures/tables and the supplementary equations referred to in the main text.

## Author Information

Corresponding author e-mail address is petros.mina@bristol.ac.uk

## Acknowledgements

The authors declare no competing interests. We thank Dr. Nigel J. Savery at the University of Bristol for useful discussions around the subject of GRNs and for his help in developing the original ABM

model. We also wish to thank Dr Gianfranco Fiore at the University of Bristol and the anonymous reviewers for reading the revised manuscript carefully and providing insightful comments that led to a consistent revision of the original manuscript. PM was supported by EPSRC grant EP/E501214/1 and KT-A by EPSRC grant EP/I018638/1. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. This work was carried out using the computational facilities of the Advanced Computing Research Centre, University of Bristol - http://www.bris.ac.uk/acrc/.

## References

1. Aström KJ, Murray RM (2010) Feedback systems: an introduction for scientists and engineers. Princeton university press.

2. Alberts B, Johnson A, Lewis J, Raff M, Roberts K, et al. (2002) Molecular biology of the cell. Garland Science. New York. US.

3. Berg JM, Tymoczko JL, Stryer L (2002). Biochemistry. 5th.

4. Good LB, Sabesan S, Marsh ST, Tsakalis K, Treiman D, et al. (2009) Control of synchronization of brain dynamics leads to control of epileptic seizures in rodents. International journal of neural systems 19: 173–196.

5. Schnitzler A, Münks C, Butz M, Timmermann L, Gross J (2009) Synchronized brain network associated with essential tremor as revealed by magnetoencephalography. Movement Disorders 24: 1629–1635.

6. Hammond C, Bergman H, Brown P (2007) Pathological synchronization in parkinson's disease: networks, models and treatments. Trends in neurosciences 30: 357–364.

7. Titcombe MS, Glass L, Guehl D, Beuter A (2001) Dynamics of parkinsonian tremor during deep brain stimulation. Chaos: An Interdisciplinary Journal of Nonlinear Science 11: 766–773.

8. Xie Y, Chen L, Kang YM, Aihara K (2008) Controlling the onset of hopf bifurcation in the hodgkin-huxley model. Physical Review E 77: 061921.

9. Rusin CG, Johnson SE, Kapur J, Hudson JL (2011) Engineering the synchronization of neuron action potentials using global time-delayed feedback stimulation. Physical Review E 84: 066202.

10. Stefanatos D, Li JS (2012) Antiphase synchronization of phase-reduced oscillators using open-loop control. Physical Review E 85: 037201.

11. Batista C, Viana R, Ferrari F, Lopes S, Batista A, et al. (2013) Control of bursting synchronization in networks of hodgkin-huxley-type neurons with chemical synapses. Physical Review E 87: 042713.

12. Breslauer D, Lee P, Lee L (2006) Microfluidics-based systems biology. Mol BioSyst 2: 97–112.

13. Milias-Argeitis A, Summers S, Stewart-Ornstein J, Zuleta I, Pincus D, et al. (2011) In silico feedback for *in vivo* regulation of a gene expression circuit. Nature Biotechnology 29: 1114–1116.

14. Uhlendorf J, Miermont A, Delaveau T, Charvin G, Fages F, et al. (2012) Long-term model predictive control of gene expression at the population and single-cell levels. Proceedings of the National Academy of Sciences 109: 14271–14276.

15. Cantone I, Marucci L, Iorio F, Ricci MA, Belcastro V, et al. (2009) A yeast synthetic network for in vivo assessment of reverse-engineering and modeling approaches. Cell 137: 172–181.

16. Menolascina F, Di Bernardo M, Di Bernardo D (2011) Analysis, design and implementation of a novel scheme for in-vivo control of synthetic gene regulatory networks. Automatica 47: 1265–1270.

17. Menolascina F, Fiore G, Orabona E, De Stefano L, Ferry M, et al. (2014) In-vivo real-time control of protein expression from endogenous and synthetic gene networks. PLoS Comput Biol 10: e1003625.

18. Cho H, Jönsson H, Campbell K, Melke P, Williams JW, et al. (2007) Self-organization in high-density bacterial colonies: efficient crowd control. PLoS Biology 5: e302.

19. Prindle A, Samayoa P, Razinkov I, Danino T, Tsimring LS, et al. (2011) A sensing array of radically coupled genetic/biopixels/'. Nature 481: 39–44.

20. Naqib F, Quail T, Musa L, Vulpe H, Nadeau J, et al. (2012) Tunable oscillations and chaotic dynamics in systems with localized synthesis. Physical Review E 85: 046210.

21. Hauptmann C, OmelChenko O, Popovych O, Maistrenko Y, Tass P (2007) Control of spatially patterned synchrony with multisite delayed feedback. Physical Review E 76: 066209.

22. Ahlborn A, Parlitz U (2008) Control and synchronization of spatiotemporal chaos. Physical Review E 77: 016201.

23. Ghosh P (2011) Control of the hopf-turing transition by time-delayed global feedback in a reaction-diffusion system. Physical Review E 84: 016222.

24. Mina P, di Bernardo M, Savery NJ, Tsaneva-Atanasova K (2013) Modelling emergence of oscillations in communicating bacteria: a structured approach from one to many cells. Journal of The Royal Society Interface 10.

25. BSim source code on SourceForge [homepage on the Internet], [updated 2012; cited March 18, 2016]. Available from: `http://bsim-bccs.sourceforge.net/`.

26. Gorochowski T, Matyjaszkiewicz A, Todd T, Oak N, Kowalska K, et al. (2012) Bsim: An agent-based tool for modeling bacterial populations in systems and synthetic biology. PLoS One 7: e42790.

27. Danino T, Mondragón-Palomino O, Tsimring L, Hasty J (2010) A synchronized quorum of genetic clocks. Nature 463: 326–330.

28. Tsai TYC, Choi YS, Ma W, Pomerening JR, Tang C, et al. (2008) Robust, tunable biological oscillations from interlinked positive and negative feedback loops. Science Signalling 321: 126.

29. Novák B, Tyson JJ (2008) Design principles of biochemical oscillators. Nature Reviews Molecular Cell Biology 9: 981–991.

30. Smith H (1987) Oscillations and multiple steady states in a cyclic gene model with repression. Journal of Mathematical Biology 25: 169–190.

31. Bratsun D, Volfson D, Tsimring LS, Hasty J (2005) Delay-induced stochastic oscillations in gene regulation. Proceedings of the National Academy of Sciences of the United States of America 102: 14593–14598.

32. Müller S, Hofbauer J, Endler L, Flamm C, Widder S, et al. (2006) A generalized model of the repressilator. Journal of Mathematical Biology 53: 905–937.

33. Strelkowa N, Barahona M (2010) Switchable genetic oscillator operating in quasi-stable mode. Journal of The Royal Society Interface 7: 1071–1082.

34. Waters CM, Bassler BL (2005) Quorum sensing: cell-to-cell communication in bacteria. Annual Review of Cell and Developmental Biology 21: 319–46.

35. Huang S (2009) Non-genetic heterogeneity of cells in development: more than just noise. Development (Cambridge, England) 136: 3853–3862.

36. Maheshri N, O'Shea EK (2007) Living with noisy genes: how cells function reliably with inherent variability in gene expression. Annu Rev Biophys Biomol Struct 36: 413–434.

37. BSim source code on GitHub [homepage on the Internet], [updated 2012; cited March 18, 2016]. Available from: `https://github.com/BiocomputeLab/bsim/`.

38. Andersen JB, Sternberg C, Poulsen LK, Bjorn SP, Givskov M, et al. (1998) New unstable variants of green fluorescent protein for studies of transient gene expression in bacteria. Applied and Environmental Microbiology 64: 2240–6.

39. Canton B, Labno A, Endy D (2008) Refinement and standardization of synthetic biological parts and devices. Nature Biotechnology 26: 787–793.

40. Mondragón-Palomino O, Danino T, Selimkhanov J, Tsimring L, Hasty J (2011) Entrainment of a population of synthetic genetic oscillators. Science Signalling 333: 1315.

41. Stewart H, Thompson J (1986) Nonlinear dynamics and chaos. John Wiley & Sons.

42. Pikovsky A, Rosenblum M, Kurths J (2003) Synchronization: A universal concept in nonlinear sciences, volume 12. Cambridge university press.

43. Bionumbers database [homepage on the Internet], [updated: 2012; cited March 18, 2016]. Available from: `http://bionumbers.hms.harvard.edu/`.

44. Bootman M, Holmes A, Roderick H (2006) Calcium signalling and regulation of cell function. eLS .

45. Tsaneva-Atanasova K, Mina P, Caunt C, Armstrong S, McArdle C (2012) Decoding GnRH neurohormone pulse frequency by convergent signalling modules. Journal of The Royal Society Interface 9: 170–182.

46. Raser JM, O'Shea EK (2005) Noise in gene expression: origins, consequences, and control. Science 309: 2010–2013.

47. Tsaneva-Atanasova K, Caunt C, Armstrong S, Perrett R, McArdle C (2012) Decoding neurohormone pulse frequency by convergent signalling modules. Biochemical Society Transactions 40: 273.

48. Burger L, Haisenleder D, Aylor K, Marshall J (2008) Regulation of intracellular signaling cascades by GNRH pulse frequency in the rat pituitary: roles for CaMK II, ERK, and JNK activation. Biology of Reproduction 79: 947–953.

49. Åström KJ, Wittenmark B (2013) Adaptive control. Courier Dover Publications.

50. Java programming language [homepage on the Internet], [updated 2013; cited March 18, 2016]. Available from: `http://java.com/`.

51. Xppaut®–the differential equations tool, version 7.00; 2012, [homepage on the Internet], [updated 2012; cited March 18, 2016].

52. MATLAB (2010) version 7.10.0 (R2010a). Natick, Massachusetts: The MathWorks Inc.

53. Ferry M, Razinkov I, Hasty J (2011) Microfluidics for synthetic biology from design to execution. Methods Enzymol 497: 295.