bioRxiv preprint doi: https://doi.org/10.1101/2020.03.13.987032; this version posted March 13, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

# Synchronised oscillations in growing cell populations are explained by demographic noise

Enrico Gavagnin*[1], Sean T. Vittadello[2], Gency Guanasingh[3], Nikolas K. Haass[3], Matthew J. Simpson[2], Tim Rogers[1], and Christian A. Yates[1]

[1] *Department of Mathematical Sciences, University of Bath, Claverton Down, Bath, UK*

[2] *School of Mathematical Sciences, Queensland University of Technology, Brisbane, Australia*

[3] *The University of Queensland, The University of Queensland Diamantina Institute, Brisbane, Queensland, Australia*

**Abstract**

Understanding synchrony in growing populations is important for applications as diverse as epidemiology and cancer treatment. Recent experiments employing fluorescent reporters in melanoma cell lines have uncovered growing subpopulations exhibiting sustained oscillations, with nearby cells appearing to synchronise their cycles. In this study we demonstrate that the behaviour observed is consistent with long-lasting transient phenomenon initiated, and amplified by the finite-sample effects and demographic noise. We present a novel mathematical analysis of a multi-stage model of cell growth which accurately reproduces the synchronised oscillations. As part of the analysis, we elucidate the transient and asymptotic phases of the dynamics and derive an analytical formula to quantify the effect of demographic noise in the appearance of the oscillations. The implications of these findings are broad, such as providing insight into experimental protocols that are used to study the growth of asynchronous populations and, in particular, those investigations relating to anti-cancer drug discovery.

*Keywords: cellular population dynamics, cell-cycle time distribution, demographic noise, finite-sample effects, multi-stage model*

*Corresponding author: enrico.gavagnin@bristol.ac.uk

bioRxiv preprint doi: https://doi.org/10.1101/2020.03.13.987032; this version posted March 13, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

# 1 Introduction

Growing populations are a crucial feature of many biological phenomena, from the clonal expansion of cancer cell lines to the increase in numbers of infected individuals during a disease outbreak. A deeper understanding of cell proliferation sheds light on a vast range of biological processes, from morphogenesis to tumour growth [Gilbert, 2000, Evan and Vousden, 2001], understanding and predicting the time evolution of these growing population is, therefore, of fundamental biological relevance [Mort et al., 2016, Haass and Gabrielli, 2017]. The initial stages of growth in both these scenarios are typically considered to be exponential as cells replicate without restriction or disease spreads into an otherwise susceptible population.

Standard mathematical modelling approaches assume that cell divisions are independent events with exponentially distributed waiting times. This gives rise to exponential growth in unstructured populations [Murray, 2007]. In cell biology, this approach has been supported by classic experimental studies for large populations under favourable growth conditions [Monod, 1949, Laird, 1965]. However, when smaller populations are considered - for example clones of a single progenitor cell - the classical model of exponential growth fails to capture the variable per capita growth rates caused by non-exponentially distributed cell cycle times and more sophisticated models are necessary [Baker and Simpson, 2010, Yates et al., 2017, Jafarpour, 2019, Pirjol et al., 2017, Lang et al., 2009, Kuritz et al., 2018].

Due to recent technological advances, we are now able to access accurate data revealing the structure of dynamic cell populations [Chao et al., 2018, Vittadello et al., 2018, Simpson et al., 2018] using, amongst other tools, proliferation assays: an *in vitro* experimental protocol in which the growth of cell populations is monitored over time [Riss et al., 2016]. In particular, in recent work [Vittadello et al., 2019], we assayed the proliferation of melanoma cells labelled with FUCCI (Fluorescent Ubiquitous Cell Cycle Indicator [Sakaue-Sawano et al., 2008, Haass et al., 2014] - see Figure 1 and Section 4.2 of Materials and Methods) which allowed us to track the number of cells

bioRxiv preprint doi: https://doi.org/10.1101/2020.03.13.987032; this version posted March 13, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

in particular phases of the cell cycle over a timespan of 48 hours (h). Strikingly, the proportion of cells in the first phase of the cell-cycle, gap 1 (G1), demonstrates clear and unexpected fluctuations during the entire duration of the experiment. This synchrony, reproducible over multiple cell lines and experimental replicates, has potentially serious implications for studies of rapidly dividing cell populations and demonstrates that classical, and widely adopted exponential models of population growth are insufficient to capture either individual-level or population-scale growth dynamics.

Inspired by our experimental findings, we build a multi-stage mathematical model for cell proliferation which represents the cell-cycle as a series of discrete stages [Vittadello et al., 2019]. The waiting time distribution between consecutive stages is exponential, meaning that the cell cycle time (CCT) follows a more general class of distributions, known as hypo-exponential. This family of distributions has been shown to provide good agreement with the experimental cell-cycle time distribution data [Yates et al., 2017, Gavagnin et al., 2019, Simpson et al., 2018, Vittadello et al., 2018, Golubev, 2016]. By deriving a deterministic representation of the population dynamics under the multi-stage approach we reproduce the cell-cycle fluctuations observed in the experiments. This suggests that multi-stage models are a suitable framework for investigating the phenomenon of cell-cycle synchronisation. However, since in [Vittadello et al., 2019] the parametrisation of the model is carried out individually for each experimental trajectory, our previous study did not explain the origin of such oscillatory phenomena nor their asymptotic behaviour.

The fluctuations which appear in the multi-stage model are a typical example of damped oscillations about an equilibrium [Harrison and Topiwala, 1974]. In other words, if the initial phase-distribution - the distribution of cells in each phase of the cell cycle - is sufficiently far from its invariant distribution, the system enters a *transient* phase characterised by oscillations of decaying amplitude, followed by an *asymptotic* phase in which the invariant distribution is reached and the total population grows exponentially. The presence of these two regimes (the

bioRxiv preprint doi: https://doi.org/10.1101/2020.03.13.987032; this version posted March 13, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

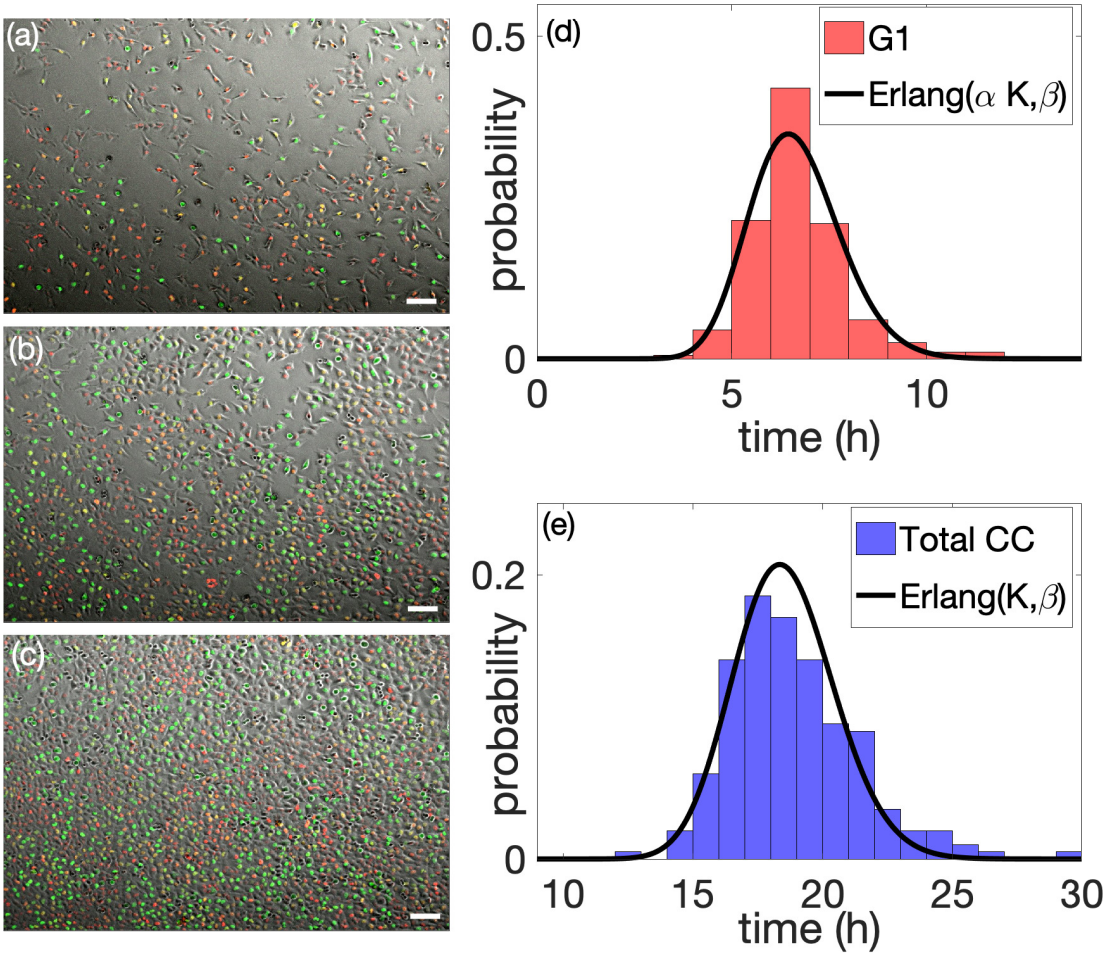



Figure 1: Data from a proliferation assay. Panels (a) to (c) show three snapshots of an epifluorescence microscopy image time series - merged images of bright-field and the red and green fluorescence channels. Panel (a) is taken at the beginning of the recording, $t = 24h$, panel (b) is taken half-way through the recording, $t = 48h$, and at panel (c) at the end of the recording, $t = 72h$ (scale bar $100\mu m$). The cells with red nuclei are in the G1 (gap 1) phase of the cell cycle, the ones with yellow nuclei are in the eS (early-synthesis) phase and those with green nuclei are in one of the remaining consecutive phases: S (synthesis), G2 (gap 2) or M (mitosis). Panel (d) shows a comparison between the distribution of the duration of the G1 phase obtained by tracking 200 randomly chosen cells (red histogram) and the Erlang distribution with the same mean and variance (black curve). Panel (e) shows the comparison between the distribution of the full cell cycle time of the same tracked cells (blue histogram) together with the corresponding Erlang distribution. See Section section S.4 of Supplementary Materials for full details.

bioRxiv preprint doi: https://doi.org/10.1101/2020.03.13.987032; this version posted March 13, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

transient-oscillatory regime and asymptotic-exponential regime) is a common feature of many structured growing population [Jafarpour, 2019, Jafarpour et al., 2018, Pirjol et al., 2017, Baker and Röst, 2019]. It is not surprising, therefore, that these two phases play distinct but critically important roles in the dynamics of a growing cell population.

In this study, we establish that strong oscillations observed in growing populations of cells can be the result of finite-sample stochasticity alone. This finding demonstrates that it may not be necessary to appeal to an external synchronisation mechanism requiring cell-cell communication to explain synchronisation observed in experiments. To do this we first analyse the multi-stage model with a particular focus on characterising the transient and asymptotic phases. By deriving a stochastic mesoscopic model, we study the effect of stochasticity in the system and obtain an analytical formula that can be used to quantify the amplitude of the fluctuations due to finite-sample effects. Finally, we parametrise the multi-stage model by fitting the G1 and total cell-cycle time distributions, obtained from single-cell tracking data, and compare our predictions with the time series obtained in the experiments.

Our central finding is that the fluctuations in the subpopulation of G1 cells in the proliferation assay are of the same magnitude as those induced by demographic noise alone, which suggests finite-sample effects as the main origin of the synchronisation. Our study also examines the specific impact of demographic noise on the dynamics, predicting that the observed oscillations are a transient phenomenon for which we can predict the corresponding characteristic decay time.

The fact that the observed synchrony is a result of demographic noise, and not a feature peculiar to the cell line we studied, means that we expect the same phenomenon to be observed in a wide range of other populations undergoing stochastic growth, as exemplified in a number of studies [Riss et al., 2016, Beaumont et al., 2016, Haass and Gabrielli, 2017, Welsh et al., 2016, Hill et al., 2009]. The possible implications of our findings therefore range from revised experimental protocols, to altered cancer treatment schedules and from new ways of understanding early

bioRxiv preprint doi: https://doi.org/10.1101/2020.03.13.987032; this version posted March 13, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

infection progression within the body to strategies for prevention of the spread dissemination of disease in the early stages of an outbreak. The characterisation of the transient nature of synchronised behaviour may also impact on the broad range of experimental protocols in which cell cycle synchronisation is of vital importance [Banfalvi, 2017].

# 2 Results

## 2.1 Multistage model recapitulates experimental observations

We adopt an agent-based model (ABM) for the growth and division of cells, following [Yates et al., 2017, Gavagnin et al., 2019, Vittadello et al., 2018]. In this formulation the cell-cycle is represented as a series of $K$ stages through which a cell progresses before it divides. We choose the waiting time to progress from one stage to the next to be exponentially distributed with rate $\beta$, independent from all other events. When a cell passes through the final stage, it divides into two new daughter cells, both initialised at stage one. This is a simplified model of the cell cycle, however, it is sufficient for the purposes of this study and (as we will show later) it gives a good fit to experimentally observed distributions of cell cycle time.

The $K$ stages of our model are grouped into sections corresponding to the known phases of the cell cycle. In particular, we say that a cell is in the G1 phase if it is in one of the first $\alpha K$ stages, where $\alpha \in [1/K, 2/K, \ldots, 1]$ is a constant to be determined by comparison with data. Expressed as a sum of exponential random variables, the duration of both the G1 and the entire cell cycle are Erlang distributed with parameters $(K, \beta)$ and $(\alpha K, \beta)$, respectively. Figure 1 (d) and (e) show the maximum likelihood fit of the model simultaneously to both the duration of the G1 phase and the total cell-cycle time for the melanoma cell line C8161. In this example we find parameters $K = 92, \alpha K = 33, \beta = 4.96h^{-1}$. The measured cell cycle time has an average of 18.5 hours with standard deviation around 2 hours (see Section S.4 of Supplementary Materials).

To keep track of the growth of a population of cells, we define the state vector $\boldsymbol{X}(t) =$

bioRxiv preprint doi: https://doi.org/10.1101/2020.03.13.987032; this version posted March 13, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

$(X_1(t), X_2(t), \ldots, X_K(t))$, where $X_k(t)$ denotes the number of cells in stage $k$ at time $t$. Our model is then represented as a series of chemical reactions, namely

$$\underbrace{\overbrace{X_1 \xrightarrow{\beta} X_2 \xrightarrow{\beta} \cdots \xrightarrow{\beta} X_{\alpha K}}^{\text{G1 phase}} \xrightarrow{\beta} X_{\alpha K+1} \xrightarrow{\beta} \cdots \xrightarrow{\beta} X_K}_{\text{Cell cycle}} \xrightarrow{\beta} 2X_1\,. \tag{1}$$

We write $N = \sum_{k=1}^{K} X_k$ for the total number of cells, $G = \sum_{k=1}^{\alpha K} X_k$ for the number of cells in G1 phase. As the population grows, the proportion of cells in each stage will eventually converge to a fixed value $X_i/N \xrightarrow{t\to\infty} u_i$, for $i = 1, \ldots, N$. Together, these proportions are known as the *invariant stage distribution*, $\boldsymbol{u}$. In Section 2.2 we prove this fact and derive an exact expression for the limit $\boldsymbol{u}$. On shorter time horizons, the system exhibits previously transitory oscillations about the invariant distribution [Strässle et al., 1988, 1989].

To assess the strength of oscillations, in what follows we develop a mathematical theory for the behaviour of the proportion, $Q = G/N$, of G1-phase cells. The first part of our analysis reveals long-lived damped oscillations in the expected value of $Q$ in a growing population, while the second shows how this effect is initiated and sustained by demographic noise.

Our experimental data are 30 time series of images taken from proliferation assays, as previously reported in Vittadello et al. [2019] - see Figure 1 (a), (b) and (c) for three snapshots of the microscopy images and Section 4.1 of Materials and Methods. Each time series captures a 48h time window following an incubation period of 24h. In Figure 2 we present a comparison between an experimental time series of $Q$ (the proportion of G1-stage cells) (blue line) and the envelope of two standard deviations about the mean, $\Omega$ (light grey region) obtained from the multi-stage model (see Section S.3 of Supplementary Materials). Although the trajectory shows clear oscillations about the mean (about three complete cycles from time 24h to time 72h), 97% of the data points lie inside $\Omega$. We repeat the comparison for all 30 time series from the experiments (reported in the Supplementary Materials). The envelope, $\Omega$, provides a good approximation for the amplitude of the fluctuations for most of the experimental trajectories.

bioRxiv preprint doi: https://doi.org/10.1101/2020.03.13.987032; this version posted March 13, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

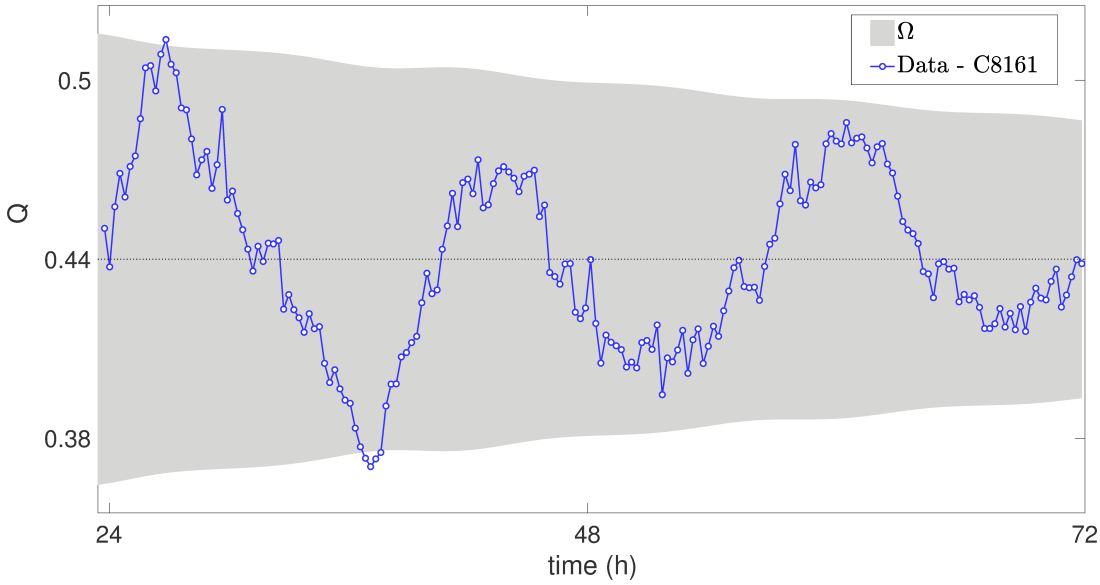



Figure 2: Comparison between experimental data and model prediction for the time evolution of $Q(t)$. One time series trajectory obtained from the experiments is plotted (blue line), together with the envelope of two standard deviations from the mean (light grey region) predicted using the multi-stage model. The parameters of the multi-stage models are obtain by fitting the distribution of the total cell-cycle time and G1 duration (see Section S.4 of Supplementary Materials): $K = 92, \alpha K = 33, \beta = 4.96h^{-1}$ and $N_0 = 155$.

## 2.2 Understanding the transient and asymptotic dynamics

In order to understand the interplay between the transient oscillatory dynamics and asymptotic exponential growth, we begin by writing down the equations governing the dynamics of the expected number of cells in each stage $\bar{\boldsymbol{x}} = \mathbb{E}[\boldsymbol{X}]$. Here *expected* should be interpreted to mean the average over many experiments with precisely the same initial condition — we will later see that the variability of the initial condition is a different feature that can also lead to the emergence of oscillations. From the model formulation we directly obtain

$$\frac{\mathrm{d}\bar{\boldsymbol{x}}}{\mathrm{d}t} = \beta \mathcal{S} \bar{\boldsymbol{x}}\,, \tag{2}$$

bioRxiv preprint doi: https://doi.org/10.1101/2020.03.13.987032; this version posted March 13, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

where $\beta$ is the rate of progression through the model stages, and $\mathcal{S}$ is the corresponding stoichiometry matrix. This matrix has non-zero entries $\mathcal{S}_{k,k} = -1, \mathcal{S}_{k,k+1} = 1$ for $k = 1, \ldots, K-1$ describing progression between stages, and $\mathcal{S}_{K,K} = -1, \mathcal{S}_{1,K} = 2$ describing cell division.

For the purpose of the analysis, we assume $\beta = K$ throughout, so that the average cell-cycle time is normalised to unity. The characteristic polynomial of the matrix $\mathcal{S}$ is given by $\mathcal{P}(y) = (y+1)^K - 2$, from which the eigenvalues of $\mathcal{S}$ are $\lambda_k = \xi_k\, 2^{\frac{1}{K}} - 1$ for $k = 1, \ldots, K$, where $\xi_k = e^{2\pi i k/K}$ is a $K$-th root of unity. By solving a series of recursive equations, one can write down the left- and right-eigenvectors associated with the $k$-th eigenvalue of $\mathcal{S}$, which we denote $\boldsymbol{u}^k$ and $\boldsymbol{v}^k$, respectively. Specifically, we have

$$u_j^k = \frac{2\lambda_K}{(1+\lambda_k)^j}\,, \qquad v_j^k = \frac{1}{K}\frac{(1+\lambda_k)^j}{2\lambda_K}\,. \tag{3}$$

We drop the index $k$ whenever we refer to the eigenvalue with maximum real part and the corresponding eigenvectors, i.e. $\lambda = \lambda_K = 2^{\frac{1}{K}} - 1$, $\boldsymbol{u} = \boldsymbol{u}^K$ and $\boldsymbol{v} = \boldsymbol{v}^K$.

Notice that from system (2) we can write $\bar{\boldsymbol{x}}(t) = e^{Kt\mathcal{S}}\bar{\boldsymbol{x}}_0$, where $\bar{\boldsymbol{x}}_0$ denotes the initial number of cells per stage. In order to study the matrix exponential $e^{Kt\mathcal{S}}$, we first notice that we can write down the $(i,j)$ element in terms of the eigenvalues and eigenvectors of $\mathcal{S}$ as

$$\left[e^{Kt\mathcal{S}}\right]_{i,j} = \sum_{k=1}^{K} u_i^k v_j^k e^{K\lambda_k t}\,. \tag{4}$$

Notice that as $t \to \infty$ the leading term of equation (4) is $u_i v_j e^{K\lambda t}$ and, hence $\bar{\boldsymbol{x}}(t) \sim \boldsymbol{u}e^{K\lambda t}$ determines the long-time behaviour of system (2). We can use this fact to study the limiting behaviour of $Q$: we write $\bar{Q}(t) = \sum_{i=1}^{\alpha K}[e^{K\mathcal{S}t}\bar{\boldsymbol{x}}_0]_i / \sum_{i=1}^{K}[e^{K\mathcal{S}t}\bar{\boldsymbol{x}}_0]_i$ and by looking at the first two terms of (4), we obtain $\lim_{t\to\infty} Q(t) = \sum_{i=1}^{\alpha K} \boldsymbol{u}_i = 2(1-2^{-\alpha}) = Q_*$. Notice that convergence to $Q_*$ occurs with an exponential decay rate given by the spectral gap of the stoichiometry matrix, $\Re[\lambda_{K-1}] - \lambda_K$ (see Figure 3).

bioRxiv preprint doi: https://doi.org/10.1101/2020.03.13.987032; this version posted March 13, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

We now focus on the transient behaviour of the system (2). We substitute the expressions (3) into the formula (4) and, by exploiting a remarkable identity of the Mittag–Leffler function [Paris, 2002], we transform the finite sum over eigenvalues on the right-hand side of (4) into an infinite sum over the cycles of the oscillatory solutions. Precisely, we write

$$\left[e^{Kt\mathcal{S}}\right]_{i,j} = \frac{1}{K}\sum_{k=1}^{K}(1+\lambda_k)^{j-i}e^{K\lambda_k t} \tag{5}$$

$$= \sum_{n=0}^{+\infty}\frac{1}{2\pi i}\oint 2^n(1+z)^{-1-Kn-(i-j)}e^{Kzt}dz \tag{6}$$

$$= \sum_{n=0}^{+\infty}\varphi_n(t,i,j)\,, \tag{7}$$

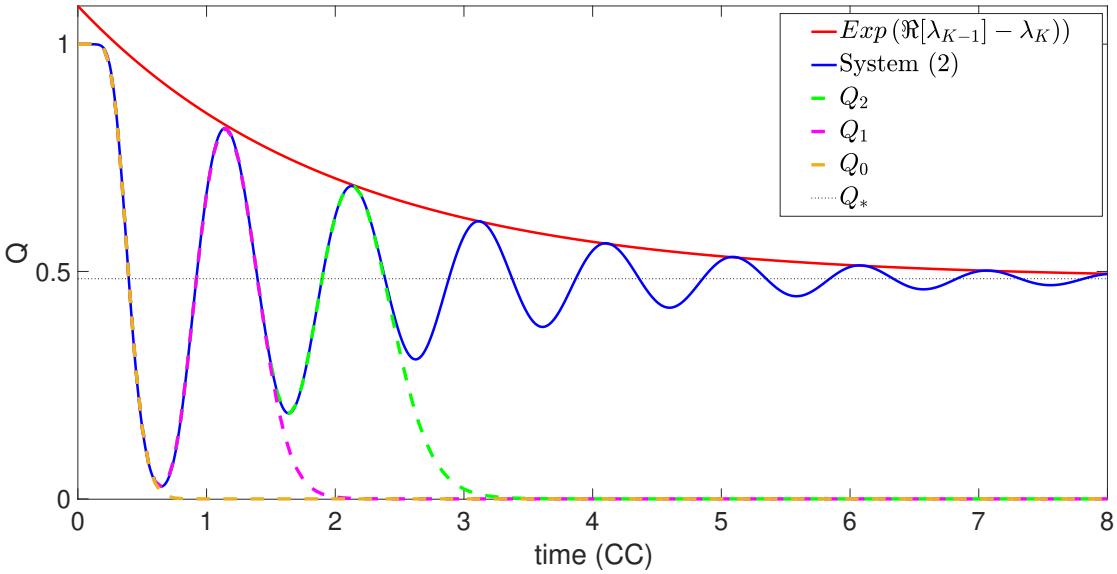
where $\varphi_n(t,i,j) = 2^n e^{-Kt}(Kt)^{Kn+i-j}/(Kn+i-j)!$.

We can now use the expression (7) to approximate $e^{Kt\mathcal{S}}$ for short times, by truncating the sum over $n$ to a finite index, $\bar{n}$. For example, let us consider an initial population of $N_0$ cells perfectly synchronised at the beginning of the cell cycle, *i.e.* $\bar{\boldsymbol{x}}_0 = N_0\mathbf{e}_1$. Then we define $G_{\bar{n}} = N_0\sum_{k=1}^{\alpha K}\sum_{n=0}^{\bar{n}}\varphi(n,i,1)$ and $N_{\bar{n}} = N_0\sum_{k=1}^{K}\sum_{n=0}^{\bar{n}}\varphi(n,i,1)$. In Figure 3 we plot $Q_{\bar{n}} = G_{\bar{n}}/N_{\bar{n}}$, for $\bar{n} = 0, 1$ and 3, together with $Q$ obtained by solving system (2) numerically. The plot illustrates how each term of the sum (7) contributes one additional oscillation to the transient dynamics of the proportion of G1-phase cells. We now have a complete picture of how oscillations propagate on average in the growing population. It remains for us to show how those oscillations are created and sustained.

### 2.3 Finite-sample effects trigger and amplify oscillations

There are two sources of randomness that are relevant to our model cell population growth: the choice of the initial state, and the timing of cell division events. In order to take into account the stochasticity in the initial population of cells we mimic the sampling procedure of the experiments. We assume that a sample of average size $N_0$ cells is drawn from the flask at

bioRxiv preprint doi: https://doi.org/10.1101/2020.03.13.987032; this version posted March 13, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.




(a)

Figure 3: The transient oscillatory dynamics. The figure shows the plot of the ratio $Q(t)$ obtained by solving the deterministic system (2) numerically (blue solid) initialised with $\boldsymbol{X}(0) = N_0\mathbf{e}_1$ and parameters $K = 40$, $N_0 = 100$ and $\alpha = 0.4$. The dashed lines represent the short-time approximation obtained by truncating expression (7) up to $\bar{n} = 0$ (dashed yellow), $\bar{n} = 1$ (dashed pink) and $\bar{n} = 2$ (dashed green). The red solid line shows the exponential decay of the oscillations.

time $t = 0$ (we ignore cell death due to detachment - see Section 4.1 of Material and Methods). As the cell culture in the flask has been maintained at a subconfluent cell density in fresh growth medium to prevent any synchronisation due to G1-arrest [Beaumont et al., 2016], we consider the population of the flask to have reached the invariant stage distribution $\boldsymbol{u}$. However, since only a small proportion of the total population is recorded in each experiment, there is a natural form of finite-sample departure from the theoretical invariant distribution due to initial sampling procedure. The random sampling of the number of cells in each stage is modelled by using $K$ independent Poisson random variables, $X_i(0) \sim \text{Po}(u_i N_0)$, for $i = 1, \ldots, K$, describing the number of cells initialised in each stage.

Next, we aim to quantify the effects of inherent stochasticity in the agent-based model.

bioRxiv preprint doi: https://doi.org/10.1101/2020.03.13.987032; this version posted March 13, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Performing a finite-size expansion of the master equation associated with the model [Morris and Rogers, 2014, Gardiner, 2009], we derive a system of stochastic differential equations for the density of cells relative to the initial population size. Let $\boldsymbol{x} = \boldsymbol{X}/N_0$, then for large but finite $N_0$ we obtain the Langevin equation

$$\frac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}t} = K\mathcal{S}\,\boldsymbol{x} + \sqrt{\frac{K}{N_0}}\mathcal{S}\,\boldsymbol{\eta}(t)\,, \tag{8}$$

where $\boldsymbol{\eta}(t)$ is a $K$-dimensional white noise vector with correlator $\mathbb{E}[\eta_i(t)\eta_j(t')] = x_i\delta_{ij}\delta(t-t')$. The first term on the right describes the average behaviour of the model, and is the same as in equation (2). The second term captures the stochastic contributions arising from the finiteness of the population.

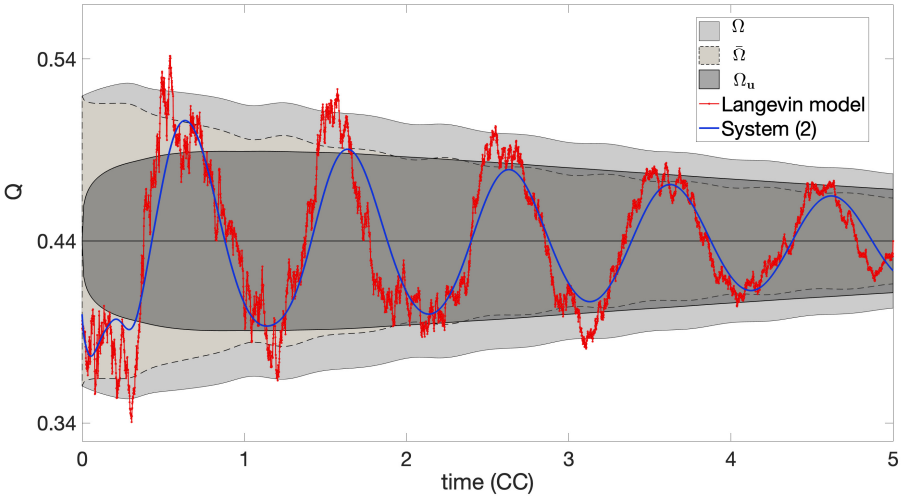
To gain more insight into the behaviour of this model, we first write down an Ornstein–Uhlenbeck (OU) model which approximates the behaviour of the Langevin equation (see Section S.1 of Supplementary Materials). By studying the OU process, and in particular its correlation matrix (see Section S.2 of Supplementary Materials), we calculate the envelope of two standard deviations about the mean of $Q$, defined as

$$\Omega(t) = [Q_* - 2\sigma_Q(t), Q_* + 2\sigma_Q(t)]\,, \tag{9}$$

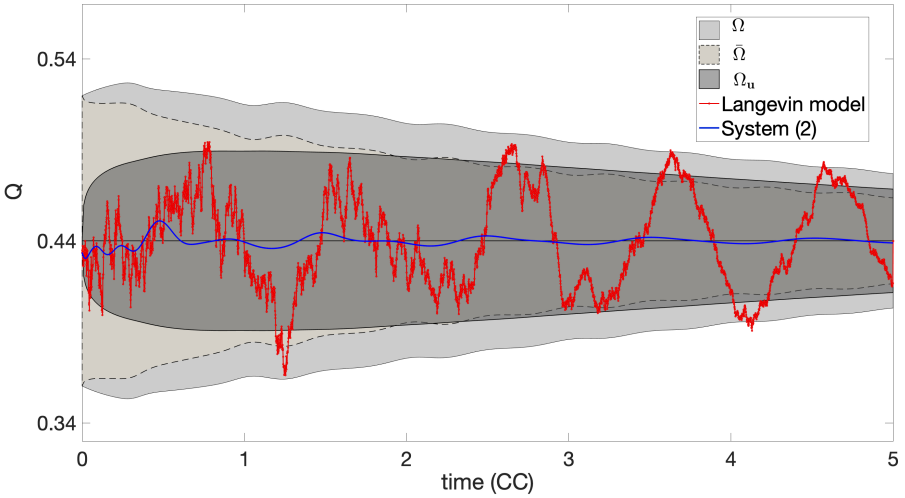
where $\sigma_Q(t)$ denotes the standard deviation of $Q(t)$. All the details of the derivation are discussed in the Supplementary Materials.

We denote with $\Omega$ the envelope obtained from the Langevin model with initial random sampling, with $\bar{\Omega}$ the envelope obtained from the system (2) with initial random sampling and with $\Omega_{\boldsymbol{u}}$ the envelope obtained from the of the Langevin model with the deterministic initial condition $\boldsymbol{x}_0 = \boldsymbol{u}$. In the two panels of Figure 4 we overlay $\Omega, \bar{\Omega}$ and $\Omega_{\boldsymbol{u}}$, together with two numerical trajectories of $Q$: one (red) obtained by solving the Langevin equation (8) and one (blue) by

bioRxiv preprint doi: https://doi.org/10.1101/2020.03.13.987032; this version posted March 13, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.




(a)




(b)

Figure 4: Finite-sample effects amplify the oscillations of $Q$. The two panels show the overlay of the three envelopes $\Omega$ (light grey region with solid line), $\bar{\Omega}$ (medium grey region with dashed line) and $\Omega_{\boldsymbol{u}}$ (dark grey region with solid line) together with two trajectories of $Q(t)$ obtained by solving numerically (by using the Euler-Maruyama method with time step $\Delta t = 10^{-3}$) the Langevin model (red line) and the deterministic system (2) (blue line) with the same, random initial condition. The two panels show two independent realisations of the models with different stochastic, initial conditions. The plots provide an example of two possible scenarios in which the intrinsic stochasticity of the Langevin model amplifies the oscillations of the deterministic system (panel b), or it triggers the emergence of new oscillations (panel b).

bioRxiv preprint doi: https://doi.org/10.1101/2020.03.13.987032; this version posted March 13, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

solving system (2), both initialised with the same random initial condition.

The results of Figure 4 show that in all three cases considered, accounting for the finite-sample stochasticity can lead to a persistent departure of $Q$ from the equilibrium value. In the two cases which account for the initial random sampling, $\Omega$ and $\bar{\Omega}$, the envelopes present an evident initial departure from equilibrium, which is sustained for several cell-cycle times, halving after approximately four periods. The inherent dynamical stochasticity of the Langevin model tends to amplify the departure from the equilibrium as evident in $\Omega$. Interestingly, both these envelopes have slightly fluctuating edges. In contrast, the envelope initialised at the invariant distribution, $\Omega_{\boldsymbol{u}}$ shows an initial, fast expansion, followed by a phase of slower decay. Notice that $\Omega_{\boldsymbol{u}}$ lies well inside $\Omega$ for all time. This suggests that the initial random sampling plays a role for the entire duration of the experiment. The numerical trajectories overlaid show good agreement with the these findings. In particular, the solution of system (2) (blue line) lies well inside $\bar{\Omega}$ while the simulation of the Langevin equation (red line) shows a larger departure but it remains almost entirely inside the envelope $\Omega$.

Notice that both trajectories in Figure 4 (a) show clear oscillations about the origin with similar phases. The Langevin solution has increased oscillation amplitude in comparison to the solution of the deterministic system (2). Although this amplification phenomenon, due to the stochasticity of the Langevin model, is common, we also find that, for some initialisations, the oscillations appear only in the Langevin model and not in the deterministic model as shown in Figure 4 (b).

In order to quantify the appearance of the oscillations, we look at the time-autocorrelation function of $G(t)$ which we define as

$$A(t,t') = \rho\left[G(t), G(t')\right] \ , \tag{10}$$

bioRxiv preprint doi: https://doi.org/10.1101/2020.03.13.987032; this version posted March 13, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

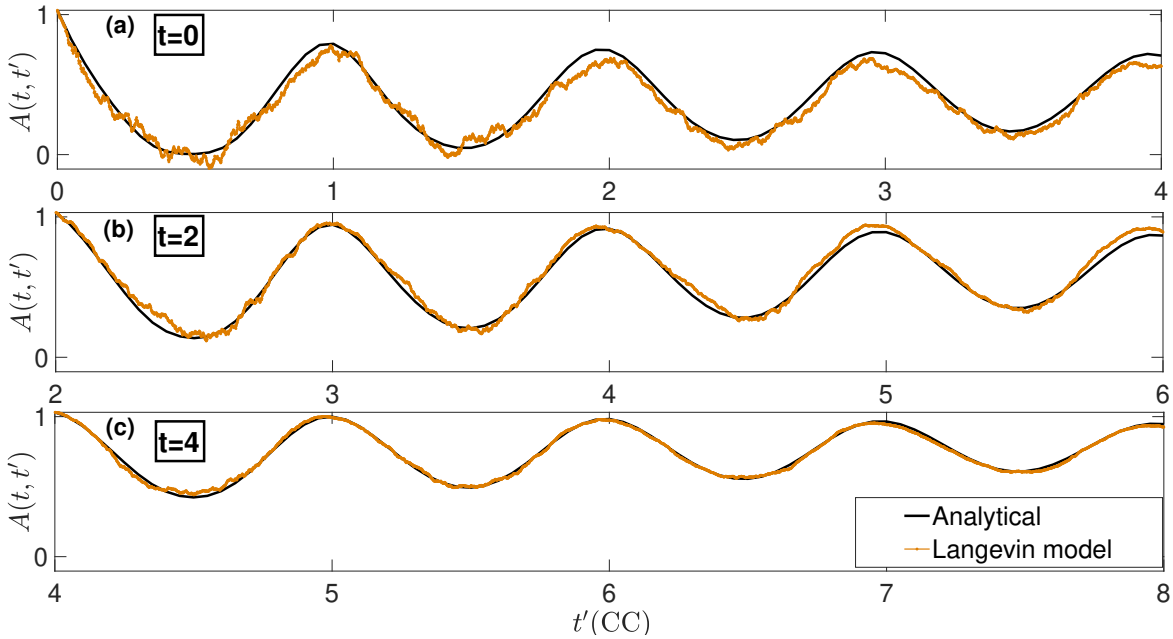



Figure 5: The time-autocorrelation function. The three panels (a), (b) and (c) show the time-autocorrelation function, $A(t,t')$ at time $t = 0, 2$ and 4, respectively, obtained analytically from the full stochastic model (back line) and by averaging over 50 independent simulations of the Langevin model (orange line). The parameters of the model are the same as in Figure 2. Time is normalised with respect to the average cell-cycle time.

where $\rho$ denotes the correlation coefficient (defined in equation (S.9) of Supplementary Materials) and can be computed using the formula for the correlation matrix (see Section S.2 of Supplementary Materials). Figure 5 shows the evolution of the autocorrelation function, $A(t,t')$ as function of $t'$, for $t = 0, 2$ and 4, respectively. In each panel we plot $A(t,t')$ calculated analytically, using the correlation matrix, (black solid line) and the simulated value obtained by averaging 50 independent trajectories of the Langevin model (orange line). All three panels show a good agreement between the analytical formula and the simulated counterpart. Moreover, the results confirm the presence of strong fluctuations on the time autocorrelation of $G$, *i.e.* the number of cells in the G1-phase, with a period of exactly one cell cycle. As $Q$ converges to the equilibrium, the amplitude of the oscillations decreases and the autocorrelation function $A(t,t')$ tends to unity.

bioRxiv preprint doi: https://doi.org/10.1101/2020.03.13.987032; this version posted March 13, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

# 3 Discussion

In this work we highlight the importance of demographic noise to the early dynamics of growing populations with non-exponentially distributed generation times. We demonstrate that finite-sample effects can recapitulate the synchronisation in the cell-cycle phase that we previously observed [Vittadello et al., 2019]. To provide insight in these observations we adopt a multi-stage approach to model both the total cell-cycle time distribution and the distribution of the G1 duration and we derive both a deterministic and a stochastic representation for the time evolution of the ratio $Q$. We find that the stochasticity in the initial sampling of cells leads to a departure from the invariant distribution which triggers a transient oscillatory phase. The presence of intrinsic stochasticity in the dynamics tends to amplify these oscillations and delay their exponential decay.

We characterise the transient and asymptotic phases of the multi-stage model by deriving an analytical formula for the variance of the amplitude of the oscillations. By comparing our results with the experimental data from a proliferation assay of C8161 melanoma cells, we find that the amplitude of the experimentally observed fluctuation lies inside the envelope predicted by the model. Our findings suggest that finite-sample stochasticity plays a crucial role in the early stage dynamics of growing populations and that it can provide an explanation for observed synchronisation in the subpopulation of the cell cycle phases.

From a theoretical point of view, our study provides a further understanding of the relation between cell-cycle distribution and global population dynamics. Whilst our analysis employed a multi-stage model, the results of our analysis are amenable to extension to more general type of cell-cycle distributions. In fact, for certain choices of the model parameters, the Erlang distribution adopted in this paper is an excellent approximation of a Gaussian distribution. In Section S.5 of the Supplementary Materials we compute the relative entropy (Kullback-Leibler divergence) between Erlang and Gaussian distributions to show that the Erlang distribution tend to a

bioRxiv preprint doi: https://doi.org/10.1101/2020.03.13.987032; this version posted March 13, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Gaussian distribution as $K \to \infty$. In principle, one could use this fact to study the applicability of our findings to Gaussian cell-cycle time and, hence, compare our results with other relevant studies which rely on a Gaussian approach [Jafarpour, 2019, Pirjol et al., 2017]. For example, Jafarpour [2019] used a Gaussian model to study the connection between mother-daughter size regulation in bacteria and the decay of transient fluctuations. Our study focuses on synchrony emerging even in the absence of correlation of CCTs, we expect that accounting for such mechanisms will tend to amplify the amplitude of the oscillations predicted by the model, however, the analysis of this phenomenon is left for future study.

From an experimental point of view, our results highlight the routinely-overlooked importance of the sample size when performing experiments which involve small populations. In particular, any data interpretation should be carried out with the role played by finite-sample stochasticity in mind. Employing larger initial populations, for example by increasing the size of microscopy images, would diminish the amplitude of the synchronisation. However, since the oscillations described in this paper are an intrinsic phenomenon due to the finiteness of the population, the aim of any intervention would be to mitigate their effects, since they cannot be completely eradicated. Our analysis provides a novel platform to quantify the extent of finite-sample effects which can then be used to assess their relevance in experimental contexts. While we primarily focus on explaining the synchronised oscillatory behaviour observed in subpopulations of melanoma cells in a proliferation assay, we stress that the general protocol developed here will be useful in order to determine the extent of effects due to finiteness of the cell population in a broad class of applications. Our proliferation assays are typical experimental protocols used to investigate the efficiency of cell-cycle-inhibiting drugs [Beaumont et al., 2016, Haass and Gabrielli, 2017], hence our findings may impact upon the reproducibility of such experiments, the efficacy of treatment protocols [Welsh et al., 2016, Hill et al., 2009] and the findings of mathematical models of these experiments [Altinok et al., 2009, Clairambault, 2011, Lévi, 2006]. Our work suggests that in-

bioRxiv preprint doi: https://doi.org/10.1101/2020.03.13.987032; this version posted March 13, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

herent synchronisation will also occur in bacterial populations and consequently that studies of bacterial pathogen growth [Jafarpour et al., 2018] may be impacted. Many experimental protocols rely on the synchronisation of cell populations in order to study the structural and molecular events that occur throughout the cell cycle, providing information about gene expression patterns, post transcriptional modification and contributing to drug discovery [Banfalvi, 2017]. The improved understanding of the impacts of demographic noise on the evolution of synchronous populations provided here will shed light on the potential impact that desynchronisation has on the results of these studies.

# 4 Materials and Methods

## 4.1 Cells and cell culture

The human melanoma cell lines C8161 (kindly provided by Mary Hendrix, Chicago, IL, USA) were genotypically characterised [Davies et al., 2009, Hoek et al., 2006, Smalley et al., 2007a,b], grown as described by Spoerri et al. [2017] and authenticated by STR fingerprinting (QIMR Berghofer Medical Research Institute, Herston, Australia)

We maintain the cell cultures to prevent any induced synchronisation from cell cycle arrest in G1 phase. In general, such induced synchronisation can occur through various experimental conditions, namely contact inhibition of proliferation at relatively high population densities [Davis et al., 2001], decreased pH of the growth medium due to the concentration of acidic cell-metabolites such as lactic acid [Taylor and Hodson, 1984], and reduced availability of nutrients such as serum [Beaumont et al., 2016]. We prevent induced synchronisation by passaging the cells every three days, and on the day prior to setting up an experiment, to maintain a subconfluent cell density and a fresh growth medium, so that the cell culture conditions are never such that they cause G1 arrest.

We note that there are other factors that can induce cell synchronisation. For example, during

bioRxiv preprint doi: https://doi.org/10.1101/2020.03.13.987032; this version posted March 13, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

the suspension, prior to seeding, some cells may die due to detachment - in particular those close to mitosis. Clearly, this phenomenon might lead to further deviation from the invariant distribution and hence, to amplification of the appearance of the oscillations. Since the aim of our study is to quantify the oscillations arising only as a result of finite-sample effects, we do not account for this phenomenon in our model.

### 4.2 Fluorescent ubiquitination-based cell cycle indicator

To generate stable melanoma cell lines expressing the FUCCI constructs, mKO2-hCdt1 (30–120) and mAG-hGem (1–110) [Smalley et al., 2005] were subcloned into a replication-defective, self-inactivating lentiviral expression vector system as previously described [Smalley et al., 2005]. The lentivirus was produced by cotransfection of human embryonic kidney 293T cells with four plasmids, including a packaging defective helper construct (pMDLg/ pRRE), a Rev plasmid (pRSV-Rev), a plasmid coding for a heterologous (pCMV-VSV-G) envelope protein, and the vector construct harboring the FUCCI constructs, mKO2-hCdt1 (30–120) and mAG- hGem (1–110). High-titer viral solutions for mKO2-hCdt1 (30/120) and mAG-hGem (1/110) were prepared and used for co-transduction into eight biologically and genetically well-characterized melanoma cell lines (see above), and subclones were generated by single-cell sorting.

### 4.3 Image processing and analysis

The microscopy data consist of multi-channel time-series stacks which are processed and analysed automatically with Fiji/ImageJ and MATLAB as described in Vittadello et al. [2019].

To obtain the time distribution of the G1 phase cells and of the cell cycle we selected 200 cells towards the beginning of the experiment. To do this, we first labelled all the automatically detected cells on the first frame of the processed merged image of the red and green channels (using the routine `Analyse particle` of Fiji/ImageJ), we then selected 100 labelled mother cells uniformly, without replacement. For each selected mother cell we manually recorded the time

bioRxiv preprint doi: https://doi.org/10.1101/2020.03.13.987032; this version posted March 13, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

intervals corresponding to the G1 phase (*i.e.* between the mitosis event of the mother cell and the first appearance of the cell in the green channel) and to the cell cycle (*i.e.* between the mitosis event of the mother cell and the last appearance of the cell in the green channel) of its two daughter cells. We ignored cells which did not reach mitosis before the end of the experiment or move out of the microscopy window (1% of the selected cells).

## References

A. Altinok, F. Lévi, and A. Goldbeter. Identifying mechanisms of chronotolerance and chronoefficacy for the anticancer drugs 5-fluorouracil and oxaliplatin by computational modeling. *Eur. J. Pharm. Sci*, 36(1):20–38, 2009.

R.E. Baker and G. Röst. Global dynamics of a novel delayed logistic equation arising from cell biology. *arXiv:1901.07817*, 2019.

R.E. Baker and M.J. Simpson. Correcting mean-field approximations for birth-death-movement processes. *Phys. Rev. E*, 82(4):041905, 2010.

G. Banfalvi. *Cell cycle synchronization*. Springer, 2017.

K. A Beaumont, D.S. Hill, S.M. Daignault, Goldie Y.L.L., D.M. Sharp, B. Gabrielli, W. Weninger, and N.K. Haass. Cell cycle phase-specific drug resistance as an escape mechanism of melanoma cells. *J. Investig. Dermatol.*, 136(7):1479–1489, 2016.

H.X. Chao, R.I. Fakhreddin, H.K. Shimerov, R.J. Kumar, G.P. Gupta, and J.E. Purvis. Evidence that the cell cycle is a series of uncoupled, memoryless phases. *bioRxiv*, page 283614, 2018.

J. Clairambault. Optimizing cancer pharmacotherapeutics using mathematical modeling and a systems biology approach. *Pers. Med.*, 8(3):271–286, 2011.

bioRxiv preprint doi: https://doi.org/10.1101/2020.03.13.987032; this version posted March 13, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

M.A. Davies, K. Stemke-Hale, E. Lin, C. Tellez, W. Deng, Y.N. Gopal, S.E. Woodman, T.C. Calderone, Z. Ju, A.J. Lazar, et al. Integrated molecular and clinical analysis of akt activation in metastatic melanoma. *Clin. Cancer Res.*, 15(24):7538–7546, 2009.

P.K. Davis, A. Ho, and S.F. Dowdy. Biological methods for cell-cycle synchronization of mammalian cells. *BioTechniques*, 30(6):1322–1331, 2001.

G.I. Evan and K.H. Vousden. Proliferation, cell cycle and apoptosis in cancer. *nature*, 411(6835): 342, 2001.

C. Gardiner. *Stochastic methods*, volume 4. Springer Berlin, 2009.

E. Gavagnin, M.J. Ford, R.L. Mort, T. Rogers, and C.A. Yates. The invasion speed of cell migration models with realistic cell cycle time distributions. *J. Theor. Biol.*, 481:91–99, 2019.

S.F. Gilbert. *Developmental Biology. Sunderland, MA*. Sinauer Associates, Inc, 2000.

A. Golubev. Applications and implications of the exponentially modified gamma distribution as a model for time variabilities related to cell proliferation and gene expression. *J. Theor. Biol.*, 393:203–217, 2016.

N.K. Haass and B. Gabrielli. Cell cycle-tailored targeting of metastatic melanoma: Challenges and opportunities. *Exp. Dermatol.*, 26(7):649–655, 2017.

N.K. Haass, K.A. Beaumont, D.S. Hill, A. Anfosso, P. Mrass, M.A. Munoz, I. Kinjyo, and W. Weninger. Real-time cell cycle imaging during melanoma growth, invasion, and drug response. *Pigment Cell Melanoma Res.*, 27(5):764–776, 2014.

D.E.F. Harrison and H.H. Topiwala. Transient and oscillatory states of continuous culture. In *Adv. Bio. Eng. Vol. 3*, pages 167–219. Springer, 1974.

D. S Hill, S. Martin, J.L. Armstrong, R. Flockhart, J.J. Tonison, D.G. Simpson, M.A. Birch-Machin, C.P.F. Redfern, and P.E. Lovat. Combining the endoplasmic reticulum stress–inducing

bioRxiv preprint doi: https://doi.org/10.1101/2020.03.13.987032; this version posted March 13, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

agents bortezomib and fenretinide as a novel therapeutic strategy for metastatic melanoma. *Clin. Cancer Res.*, 15(4):1192–1198, 2009.

K.S. Hoek, N.C. Schlegel, P. Brafford, A. Sucker, S. Ugurel, R. Kumar, B.L. Weber, K.L. Nathanson, D.J. Phillips, M. Herlyn, et al. Metastatic potential of melanomas defined by specific gene expression profiles with no braf signature. *Pigment Cell Research*, 19(4):290–302, 2006.

F. Jafarpour. Cell size regulation induces sustained oscillations in the population growth rate. *Phys. Rev. Lett.*, 122(11):118101, 2019.

F. Jafarpour, C.S. Wright, H. Gudjonson, J. Riebling, E. Dawson, K. Lo, A. Fiebig, S. Crosson, A.R. Dinner, and S. Iyer-Biswas. Bridging the timescales of single-cell and population dynamics. *Phys. Rev. X*, 8(2):021007, 2018.

K. Kuritz, W. Halter, and F. Allgöwer. Passivity-based ensemble control for cell cycle synchronization. In *Emerging Applications of Control and Systems Theory*, pages 1–13. Springer, 2018.

A.K Laird. Dynamics of tumour growth: comparison of growth rates and extrapolation of growth curve to one cell. *Br. J. Cancer*, 19(2):278, 1965.

M. Lang, S. Waldherr, and F. Allgöwer. Amplitude distribution of stochastic oscillations in biochemical networks due to intrinsic noise. *PMC biophysics*, 2(1):10, 2009.

F. Lévi. Chronotherapeutics: the relevance of timing in cancer therapy. *Cancer Causes Control*, 17(4):611–621, 2006.

J. Monod. The growth of bacterial cultures. *Annu. Rev. Microbiol.*, 3(1):371–394, 1949.

R.G. Morris and T. Rogers. Growth-induced breaking and unbreaking of ergodicity in fully-connected spin systems. *J. Phys. A.-Math. Theor.*, 47(34):342003, 2014.

bioRxiv preprint doi: https://doi.org/10.1101/2020.03.13.987032; this version posted March 13, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

R.L. Mort, R.J.H. Ross, K.J. Hainey, O.J. Harrison, M.A. Keighren, G. Landini, R.E. Baker, K.J. Painter, I.J. Jackson, and C.A. Yates. Reconciling diverse mammalian pigmentation patterns with a fundamental mathematical model. *Nat. Commun.*, 7, 2016.

J.D. Murray. *Mathematical biology: I. An introduction*, volume 17. Springer Science & Business Media, 2007.

R.B. Paris. Exponential asymptotics of the Mittag-Leffler function. *Proc. R. Soc., Ser. B, London*, 458(2028):3041–3052, 2002.

D. Pirjol, F. Jafarpour, and S. Iyer-Biswas. Phenomenology of stochastic exponential growth. *Phys. Rev. E*, 95(6):062406, 2017.

T.L. Riss, R.A. Moravec, A.L. Niles, S. Duellman, H.A. Benink, T.J. Worzella, and L. Minor. Cell viability assays. In *Assay Guidance Manual*. Eli Lilly & Company and the National Center for Advancing Translational Sciences, 2016.

A. Sakaue-Sawano, H. Kurokawa, T. Morimura, A. Hanyu, H. Hama, H. Osawa, S. Kashiwagi, K. Fukami, T. Miyata, H. Miyoshi, et al. Visualizing spatiotemporal dynamics of multicellular cell-cycle progression. *Cell*, 132(3):487–498, 2008.

M.J. Simpson, W. Jin, S.T. Vittadello, T.A. Tambyah, J.M. Ryan, G. Gunasingh, N.K. Haass, and S.W. McCue. Stochastic models of cell invasion with fluorescent cell cycle indicators. *Phys. A*, 510:375–386, 2018.

K.S.M. Smalley, P. Brafford, N.K. Haass, J.M. Brandner, E. Brown, and M. Herlyn. Up-regulated expression of zonula occludens protein-1 in human melanoma associates with n-cadherin and contributes to invasion and adhesion. *Am. J. Pathol.*, 166(5):1541–1554, 2005.

K.S.M. Smalley, R. Contractor, N.K. Haass, A.N. Kulp, G.E. Atilla-Gokcumen, D.S. Williams, H. Bregman, K.T. Flaherty, M.S. Soengas, E. Meggers, et al. An organometallic protein

bioRxiv preprint doi: https://doi.org/10.1101/2020.03.13.987032; this version posted March 13, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

kinase inhibitor pharmacologically activates p53 and induces apoptosis in human melanoma cells. *Cancer research*, 67(1):209–217, 2007a.

K.S.M. Smalley, R. Contractor, N.K. Haass, J.T. Lee, K.L. Nathanson, C.A. Medina, K.T. Flaherty, and M. Herlyn. Ki67 expression levels are a better marker of reduced melanoma growth following mek inhibitor treatment than phospho-erk levels. *Br. J. Cancer*, 96(3):445–449, 2007b.

L. Spoerri, K.A. Beaumont, A. Anfosso, and N.K. Haass. Real-time cell cycle imaging in a 3d cell culture model of melanoma. In *Methods Mol. Biol.*, pages 1612: 401–416. Springer, 2017.

C. Strässle, B. Sonnleitner, and A. Fiechter. A predictive model for the spontaneous synchronization of saccharomyces cerevisiae grown in continuous culture. i. concept. *J. Biotechnol.*, 7 (4):299–317, 1988.

C. Strässle, B. Sonnleitner, and A. Fiechter. A predictive model for the spontaneous synchronization of saccharomyces cerevisiae grown in continuous culture. ii. experimental verification. *J. Biotechnol.*, 9(3):191–208, 1989.

I.W. Taylor and P.J. Hodson. Cell cycle regulation by environmental pH. *J. Cell Physiol*, 121 (3):517–525, 1984.

S.T Vittadello, S.W. McCue, G. Gunasingh, N.K Haass, and M.J. Simpson. Mathematical models for cell migration with real-time cell cycle dynamics. *Biophys. J.*, 114(5), 2018.

S.T. Vittadello, S.W. McCue, G. Gunasingh, N.K. Haass, and M.J. Simpson. Mathematical models incorporating a multi-stage cell cycle replicate normally-hidden inherent synchronization in cell proliferation. *J. R. Soc. Interface*, 16(20190382), 2019.

S.J. Welsh, H. Rizos, R.A. Scolyer, and G.V. Long. Resistance to combination braf and mek inhibition in metastatic melanoma: where to next? *Eur. J. Cancer*, 62:76–85, 2016.

bioRxiv preprint doi: https://doi.org/10.1101/2020.03.13.987032; this version posted March 13, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

C.A. Yates, M.J. Ford, and R.L. Mort. A multi-stage representation of cell proliferation as a markov process. *Bull. Math. Biol.*, 79(12):2905–2928, 2017.