

To Grow is Not Enough: Supplementary Material

Nash Rochman*

Department of Chemical and Biomolecular Engineering, Johns Hopkins University

Fangwei Si†

Department of Mechanical Engineering, Johns Hopkins University

Sean X. Sun‡

*Department of Mechanical Engineering and Biomedical Engineering,
Johns Hopkins University, Baltimore MD 21218*

*Electronic address: nashdeltarochman@gmail.com

†Electronic address: sifangwei@gmail.com

‡Electronic address: ssun@jhu.edu

I. IDEAL GROWTH-RATE IS NOT IMPROVED BY ALLOWING INCREASED NOISE

The growth of a population of cells in exponential phase satisfies the following ODE:

$$\frac{\partial N(t)}{\partial t} = rN(t) \quad (1)$$

where the growth-rate, r , may be rewritten as a statistical average:

$$r = \langle f(\phi, P(\tau)) \rangle_{\phi} \quad (2)$$

$P(\tau)$ is the division time distribution, ϕ is a variable parametrized over \mathbb{R} which carries any environmental dependence (e.g. temperature, nutrient concentration), $N(t)$ is the number of cells at time t , f is an undetermined function, and $\langle \rangle_{\phi}$ indicates averaging over all possible environmental states. The solution to (1) is of course the familiar exponential growth expression:

$$N(t) = N_0 \exp(rt) \quad (3)$$

where N_0 is the initial number. For the simplified case where every cell divides in time $\tau \equiv \mu$ (such that $N(\mu) = 2N_0$) we may find $r = \frac{\ln(2)}{\mu}$. In the treatment of the real case for arbitrary $P(\tau)$, we rewrite $f(P(\tau), \phi)$ the following way:

$$f(\phi, P(\tau)) = \int_0^{\infty} g(\tau, \phi) P(\tau) d\tau \quad (4)$$

Where $g(\tau, \phi)$ is the effective growth-rate of an individual cell, which unrestricted will divide in time τ , subject to additional environmental constraints under conditions ϕ . For example, suppose ϕ specifies that the environment is not rich enough to sustain the nutrient uptake required for a cell to divide in under 20 minutes. In this case, $g(30, \phi) = \frac{1}{30}$, but $g(10, \phi) = 0$ as the cell will die before it divides. Now consider that each environmental state is found with some probability $P(\phi)$. Averaging over these states yields:

$$\langle f(\phi, P(\tau)) \rangle_{\phi} = \int_0^{\infty} P(\phi) \left[\int_0^{\infty} g(\tau, \phi) P(\tau) d\tau \right] d\phi \quad (5)$$

Rearranging:

$$\langle f(\phi, P(\tau)) \rangle_{\phi} = \int_0^{\infty} \int_0^{\infty} g(\tau, \phi) P(\tau) P(\phi) d\tau d\phi \quad (6)$$

We may note $g(\tau, \phi)P(\tau)P(\phi)$ is always positive and $g(\tau, \phi)$ is bounded (i.e. growth-rate is always finite). Further considering only the case where $g(\tau, \phi)$, $P(\tau)$, $P(\phi)$ are continuous and $\int_0^{\infty} g(\tau, \phi)P(\tau)d\tau$ is continuous as a function of ϕ , we can freely switch the order of integration[1]:

$$\langle f(\phi, P(\tau)) \rangle_{\phi} = \int_0^{\infty} \int_0^{\infty} g(\tau, \phi) P(\tau) P(\phi) d\phi d\tau = \int_0^{\infty} \left[\int_{\phi} P(\phi) g(\tau, \phi) d\phi \right] P(\tau) d\tau \equiv \int_0^{\infty} h(\tau) P(\tau) d\tau \quad (7)$$

Now we know three things about $h(\tau)$: $h(0) = 0$ - a cell cannot divide in 0 time, $h(\infty) = 0$ - a cell “dividing in infinite time” does not grow, and further that for some $\tau = \tau'$, $h(\tau') > 0$ - or else there is no growth at all. Therefore we can also say that there exists some $\tau = \tau^*$ where $h(\tau^*)$ attains its maximum value which is greater than zero. This implies that the maximum may be achieved when the distribution attains a delta function:

$$\max(r(P(\tau))) \Rightarrow P(\tau) = \delta(\tau, \tau^*) \quad (8)$$

Thus given any cell cycle duration distribution with finite width, there exists a narrower one which attains the same growth-rate or greater.

II. DERIVATION OF THE TRANSITION PROBABILITY $M(\tau \rightarrow \tau')$

Please note that some of this material is also presented in the main text. We repeat it here to maintain continuity of the material. We want to take into account two sources of regulation for cell cycle duration - optimization of protein synthesis rates to the current environment (instantaneous information), and the maintenance of proteome similarity to the mother cell (inherited information). Let us begin with the shorter argument - the comparison of proteome composition between mother and daughter cells.

A. Construction: Inherited Information

At the beginning of a cell cycle, a cell’s proteome is entirely inherited from proteins present in the mother cell. Due to partition noise, the probability of inheriting m proteins from a mother cell containing N proteins is given by the symmetric binomial distribution $P(m) = \frac{N!}{m!(N-m)!} \frac{1}{2^N}$. For N very large, this distribution is well approximated by a Gaussian distribution of mean $\frac{N}{2}$ and variance $\frac{N}{4}$. These inherited proteins constitute roughly half of the cell’s final proteome (or the proteome of the cell at the time of its division) though this fraction differs based on the number of proteins inherited. Furthermore, there exists a map between the proteome of a cell and its cycle duration. The greater number of proteins shared between the mother and daughter cell, the smaller the difference between their cycle durations. As this inheritance is roughly normally distributed, we approximate the cycle duration τ' (of the daughter cell) to be normally distributed with mean τ (the cycle duration of the mother cell) and some unknown variance σ_2^2 :

$$M_2(\tau \rightarrow \tau') = A \exp \left[-\frac{1}{2\sigma_2^2} (\tau' - \tau)^2 \right] \quad (9)$$

B. Construction: Instantaneous Information

Let us begin by describing the simplest construction of the constant-adder model: let us define a mean protein/DNA synthesis rate of k proteins/DNA per minute, and say that the cell must synthesize additional N proteins/DNA during its cell cycle. Here we denote the probability of observing N proteins at time t as $\pi(N, t)$. From a uni-directional random-walker model we get for $N \geq 1$:

$$\frac{\partial \pi(N, t)}{\partial t} = k [\pi(N-1, t) - \pi(N, t)]$$

for N equals zero: $\frac{\partial \pi(0, t)}{\partial t} = -k\pi(0, t)$. Thus (normalizing):

$$\pi(0, t) = k \exp(-kt)$$

Suppose:

$$\pi(N, t) = \frac{k^{N+1}}{N!} t^N e^{-kt}$$

Then:

$$\frac{\partial \pi(N+1, t)}{\partial t} = k \left[\frac{k^{N+1}}{N!} t^N e^{-kt} - \pi(N+1, t) \right]$$

Which implies:

$$\pi(N+1, t) = \frac{k * k^{N+1}}{(N+1)!} t^{N+1} e^{-kt}$$

Therefore, by induction:

$$\pi(N, t) = \frac{k^{N+1}}{N!} t^N e^{-kt}$$

We may note that this is simply a Gamma distribution:

$$\pi(N, t) = \frac{k^{N+1}}{\Gamma(N+1)} t^N e^{-kt} \quad (10)$$

with a mean of $\mu = \frac{N+1}{k}$ and variance $\sigma_1^2 = \frac{N+1}{k^2}$. Since k depends on the environmental condition ϕ , both μ and σ_1 depend on ϕ . As long as $N \gg 1$, this distribution is suitably symmetric, and we

may well approximate the transition probability with a Gaussian function (in terms of $t \equiv \tau'$) with the given mean and variance:

$$P(N, t) \approx M_1(\tau') \equiv \frac{1}{\sqrt{2\pi}\sigma_1} \exp \left[-\frac{1}{2\sigma_1^2} (\tau' - \mu)^2 \right] \quad (11)$$

The total transition probability from τ of the mother cell to τ' of the daughter cell should balance the inherited information with the process of protein/DNA synthesis. Therefore we expect

$$M(\tau \rightarrow \tau') \sim M_1(\tau')M_2(\tau \rightarrow \tau') \quad (12)$$

There is a problem with this construction, however, because taking a look at the product of our two transition probabilities (where the normalization is absorbed into the constant A):

$$M(\tau \rightarrow \tau') = A \exp \left[-\frac{1}{2\sigma_1^2} (\tau - \mu)^2 \right] \exp \left[-\frac{1}{2\sigma_2^2} (\tau' - \tau)^2 \right] \quad (13)$$

we can see that the most probable state for τ' is always between τ and μ regardless of the value of σ_1 or σ_2 . This contradicts the result from the constant adder model and experiment which predicts a weak negative correlation between adjacent generations at steady state[2, 3]. So in order to construct the most general form for the transition probability that can reproduce both response and steady-state dynamics we will also include some current state (τ) dependence in the term representing instantaneous information constructed below.

In what follows we will describe what may be the original “sizer”, or mass-accumulation, model described by Bremer and Chuang in 1981 [4]. We do not wish to imply that we believe the “sizer” model to be a good substitute for the “constant adder” that has proven so robust in recent studies [2, 5], but merely that when interpreted as a product of two weighted regulatory tendencies (mass-accumulation and proteome maintenance), the “constant adder” can be even more widely applied to cell populations experiencing variable environmental conditions. We begin by breaking down the cell cycle into three smaller stages. Consider the cell cycle to be composed of three periods τ^A , before the initiation of DNA replication, τ^B , during replication, and τ^C after replication:

$$\tau = \tau^A + \tau^B + \tau^C \quad (14)$$

Bremer and Chuang found that the time between associated points in replication from generation to generation (e.g. the time between the initiation of replication in the mother cell and daughter cell or termination of replication between cells) were highly conserved. This implies (where generation n is denoted τ_n):

$$\tau^B = \text{constant} \equiv B \quad (15)$$

and furthermore:

$$\tau_{n-1}^C + \tau_n^A = \text{constant} \equiv D \quad (16)$$

So we can write down an expression for the n^{th} division time in terms of the preceding $n-1$ division time:

$$\tau_n = \tau_n^A + \tau_n^B + \tau_n^C = (D - \tau_{n-1}^C) + B + \tau_n^C = (D + B) + \tau_n^C - \tau_{n-1}^C \quad (17)$$

Now, $D + B$ is simply the mean the total cycle time (relabeling it μ). With this we have:

$$\tau_n = \mu + \tau_n^C - \tau_{n-1}^C \quad (18)$$

Continuing, we know τ_n^C is just some fraction α_1 of the total mean and similarly τ_{n-1}^C is some fraction α_2 of the total division time τ_{n-1} :

$$\tau_n = \mu + \alpha_1\mu - \alpha_2\tau_{n-1} \quad (19)$$

Approximating the fraction of the cell cycle accounted for by region C as constant across each generation, i.e. $\alpha_1 = \alpha_2 \equiv \alpha$ yields:

$$\tau_n = \mu + \alpha\mu - \alpha\tau_{n-1} = \mu + \alpha(\mu - \tau_{n-1}) \quad (20)$$

Therefore, on average, we have

$$\langle \tau_n \rangle = \mu + \alpha\mu - \alpha\tau_{n-1} = \mu + \alpha(\mu - \langle \tau_{n-1} \rangle) \quad (21)$$

Similarly, we want an expression for the variance:

$$\sigma^2(\tau_n) = \sigma^2(\mu + \tau_n^C - \tau_{n-1}^C) = \sigma^2(\mu - \tau_{n-1}^C) + \sigma^2(\tau_n^C) = \sigma^2(\tau_n^C) \equiv \sigma_1^2 \quad (22)$$

Applying the stochastic protein synthesis argument above restricted to period C yields:

$$M_1(\tau_{n-1} \rightarrow \tau_n) \approx A \exp \left[\frac{1}{2\sigma_1^2} (\tau_n + \alpha\tau_{n-1} - (1 + \alpha)\mu)^2 \right] \quad (23)$$

We have provided physiological motivation for a weighted averaging of the current state and the mean of the total distribution; however, as we are primarily interested in a generalizable phenomenological model and are unable to retrieve the value for α from experiment, we will use the symmetric average $\alpha = 1$ during data fitting for simplicity. Though this parameter selection loses physiological significance, the trends observed do not meaningfully change when a smaller alpha

value is used. Putting these two considerations (instantaneous and inherited information) together we have constructed our desired transition probability:

$$M(\tau_{n-1} \rightarrow \tau_n) = A \exp \left[-\frac{1}{2\sigma_1^2} (\tau_n + \alpha\tau_{n-1} - (1 + \alpha)\mu)^2 \right] \exp \left[-\frac{1}{2\sigma_2^2} (\tau_n - \tau_{n-1})^2 \right] \quad (24)$$

We note that the preceding derivation of the transition probability is heuristic, and was partially motivated by connections with previous work. We are primarily interested in this approximate Gaussian model because it admits the analytic treatment below. However, the proposed model does have strong experimental support from recent quantitative data. For example, the model predicts a correlation in fluctuation of cell cycle durations: $\langle \delta\tau(n)\delta\tau(n') \rangle$. There exists an analytic result for this correlation function (see below). The comparison with data from multiple experiments in the Mother Machine are shown in Fig. S1. We also have directly measured the transition probability in constant environmental conditions. The measured transition probability compares well with our model (Fig. S1).

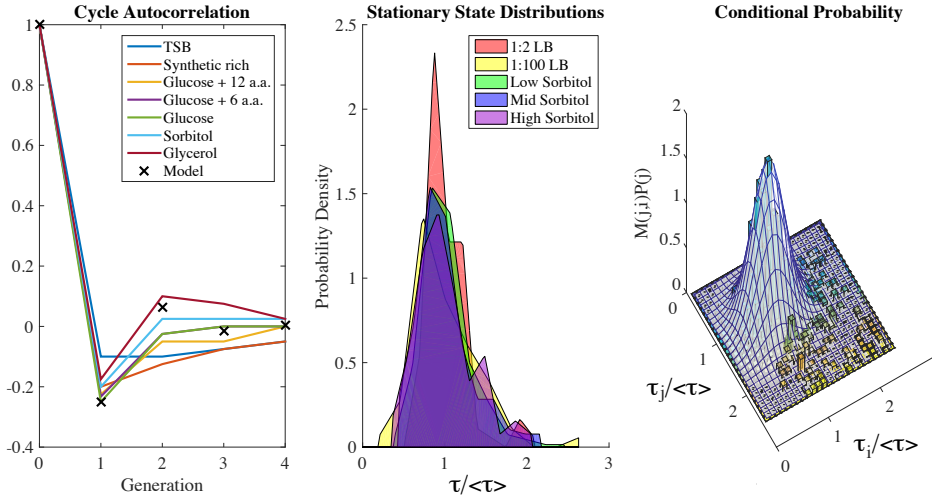


Figure S1: (A) Experimentally measured cell cycle duration fluctuation correlation function $C(n) = \langle \delta\tau(n')\delta\tau(n' + n) \rangle / \langle \delta\tau^2 \rangle$ from Ref. [2] (lines) and our model correlation function derived from the transition probability in Eq. (36) and Eq. (62) with $\Pi = -1/4$ and $\alpha = 1$ (crosses). The function shows a negative correlation after 1 generation, consistent with our model. (B) The cell cycle duration distributions displayed in Figure 1 of the main text scaled by the mean value from each condition. One may note the shape conservation of the mean scaled distributions. (C) The experimentally measured conditional probability $M(\tau_j / \langle \tau \rangle \rightarrow \tau_i / \langle \tau \rangle) P(\tau_j / \langle \tau \rangle)$ from 1500 data points (collected for this study) is compared with the proposed model in Eq. (36). Here the distributions from all environmental conditions studied are scaled by the mean of each distribution and compiled. From these results, our proposed model is able to capture both the correlation function as well as explicit transition probabilities between difference cell cycle durations.

C. Rewriting in reduced notation and normalization

It now remains to find the normalization of this form as well as the derivation of the stationary distribution and analytic treatment of the response when μ is perturbed. Before we move forward, we will introduce some reduced notation to aid in the calculation of the desired quantities. In what follows, we will use the following notation for pre-normalized Gaussian functions:

$$\exp \left[-\gamma (x - \mu)^2 \right] \equiv G(x, \mu, \gamma) \quad (25)$$

where γ is a width parameter. In addition, all integration will be conducted over \mathbb{R} :

$$\int \equiv \int_{-\infty}^{\infty} \quad (26)$$

Thus in this notation:

$$\int G(x, \mu, \gamma) dx = \sqrt{\frac{\pi}{\gamma}} \quad (27)$$

Introducing the change of variables defined below:

$$a \equiv \frac{\alpha^2}{2\sigma_1^2}, b \equiv \frac{1}{2\sigma_2^2}, x \equiv \tau' - \mu, y \equiv \tau - \mu \quad (28)$$

yields the following:

$$M(y \rightarrow x) = AG \left(y, \frac{-x}{\alpha}, a \right) G(y, x, b) \quad (29)$$

Additionally, we will find it convenient to label the following quantity:

$$\Pi \equiv \left[\frac{b - \frac{a}{\alpha}}{b + \frac{a}{\alpha^2}} \right] \quad (30)$$

We are going to need to compute the products of many Gaussian functions and may utilize the following identity [6]:

$$G \left(x, \mu_1, \frac{1}{2\sigma_1^2} \right) G \left(x, \mu_2, \frac{1}{2\sigma_2^2} \right) = G \left(x, \frac{\mu_1\sigma_2^2 + \mu_2\sigma_1^2}{\sigma_2^2 + \sigma_1^2}, \frac{\sigma_2^2 + \sigma_1^2}{2\sigma_2^2\sigma_1^2} \right) G \left(\mu_1, \mu_2, \frac{1}{2(\sigma_2^2 + \sigma_1^2)} \right) \quad (31)$$

which becomes simpler in terms of $\gamma_i \equiv \frac{1}{2\sigma_i^2}$:

$$G(x, \mu_1, \gamma_1) G(x, \mu_2, \gamma_2) = G \left(x, \frac{\mu_1\gamma_1 + \mu_2\gamma_2}{\gamma_1 + \gamma_2}, \gamma_1 + \gamma_2 \right) G \left(\mu_1, \mu_2, \frac{\gamma_1\gamma_2}{\gamma_1 + \gamma_2} \right) \quad (32)$$

To compute the normalization constant we will use a delta function as the initial distribution $\rho(y)$ for convenience:

$$1 = \int \left[\int M(y \rightarrow x) \rho(y) dy \right] dx = A \int \left[\int G \left(y, -\frac{x}{\alpha}, a \right) G(y, x, b) \delta(y - y') dy \right] dx \quad (33)$$

Evaluating the inner integral and rearranging yields:

$$1 = A \int G \left(y', -\frac{x}{\alpha}, a \right) G(y', x, b) dx = AG \left(y', 0, (1 + \alpha)^2 \frac{ab}{a + \alpha^2 b} \right) \int G \left(x, y'\Pi, b + \frac{a}{\alpha^2} \right) dx \quad (34)$$

finally, evaluation of the remaining integral yields:

$$A = \left[\sqrt{\frac{\pi}{b + \frac{a}{\alpha^2}}} G \left(y', 0, (1 + \alpha)^2 \frac{ab}{a + \alpha^2 b} \right) \right]^{-1} \quad (35)$$

This leaves us with the reduced form for the transition probability:

$$M(y \rightarrow x) = \sqrt{\frac{b + \frac{a}{\alpha^2}}{\pi}} G \left(x, y'\Pi, b + \frac{a}{\alpha^2} \right) \quad (36)$$

D. Identification of the stationary state

We want to solve for the distribution $P(y)$ satisfying the relationship:

$$P(x) = \int M(y \rightarrow x) P(y) dy \quad (37)$$

We will show that this distribution is in fact also Gaussian and will begin with the trial distribution:

$$\rho(y) = G(y, m, c) \quad (38)$$

where c is an unknown Gaussian width parameter for the steady state, and m is the unknown Gaussian center. Using Eq. (36) and Eq. (37), we are left to solve for $\rho(x)$:

$$\rho(x) = \int \sqrt{\frac{b + \frac{a}{\alpha^2}}{\pi}} G\left(x, y\Pi, b + \frac{a}{\alpha^2}\right) \sqrt{\frac{c}{\pi}} G(y, m, c) dy \quad (39)$$

Rearranging yields:

$$\rho(x) = \sqrt{\frac{c}{\pi}} \sqrt{\frac{b + \frac{a}{\alpha^2}}{\pi}} G\left(x\Pi^{-1}, m, \frac{(b - \frac{a}{\alpha})^2 c}{(b - \frac{a}{\alpha})^2 + c(b + \frac{a}{\alpha^2})}\right) \int G\left(y, \dots, \frac{(b - \frac{a}{\alpha})^2}{b + \frac{a}{\alpha^2}} + c\right) dy \quad (40)$$

Integration and simplification yields:

$$\rho(x) = \sqrt{\frac{c(b + \frac{a}{\alpha^2})^2}{\pi \left((b - \frac{a}{\alpha})^2 + c(b + \frac{a}{\alpha^2}) \right)}} G\left(x, m\Pi, \frac{c(b + \frac{a}{\alpha^2})^2}{(b - \frac{a}{\alpha})^2 + c(b + \frac{a}{\alpha^2})}\right) \quad (41)$$

We have indeed shown that if a steady state exists, it must be Gaussian. We are now left with recurrence relations on the mean (where the subscripts indicate the generation number, or the number of applications of M):

$$m_{n+1} = m_n\Pi \quad (42)$$

and the width parameter:

$$c_{n+1} = \frac{c_n \left(b + \frac{a}{\alpha^2}\right)^2}{\left(b - \frac{a}{\alpha}\right)^2 + c_n \left(b + \frac{a}{\alpha^2}\right)} \quad (43)$$

To find the steady state mean (m^*) we solve:

$$m^* = m^*\Pi$$

We may note $|\Pi| < 1$ whenever $0 < a, b$ and $\alpha \leq 1$. Thus we see $m^* = 0$. Similarly for the width parameter we must solve:

$$c^* = \frac{c^* \left(b + \frac{a}{\alpha^2}\right)^2}{\left(b - \frac{a}{\alpha}\right)^2 + c^* \left(b + \frac{a}{\alpha^2}\right)} \quad (44)$$

Simplification yields:

$$\left(b - \frac{a}{\alpha}\right)^2 + c^* \left(b + \frac{a}{\alpha^2}\right) = \left(b + \frac{a}{\alpha^2}\right)^2 \quad (45)$$

and

$$c^* = \frac{\left(b + \frac{a}{\alpha^2}\right)^2 - \left(b - \frac{a}{\alpha}\right)^2}{\left(b + \frac{a}{\alpha^2}\right)} = \frac{2ab(1 + \alpha) + a^2\left(\frac{1}{\alpha^2} - 1\right)}{\alpha^2 b + a} \quad (46)$$

Thus we have shown that a stationary solution does exist and that it is in fact Gaussian with the mean and variance identified above:

$$P(x) = AG\left(x, 0, \frac{2ab(1 + \alpha) + a^2\left(\frac{1}{\alpha^2} - 1\right)}{\alpha^2 b + a}\right) \quad (47)$$

E. Stability of the stationary state

We want to verify the global asymptotic stability of the stationary state. We will do this in two steps: first verifying the mean, and then the width parameter. As was shown above, we have the following recurrence relation for the mean:

$$m_{n+1} = m_n \Pi \quad (48)$$

Stability requires that the sequence m_n converges to the stationary mean of zero. Thus the stationary mean is stable since the Π is less than one. Now we will treat the width parameter. As with the mean, stability requires that the sequence c_n converges to the stationary width parameter of c^* . First we want to show that c_n converges. To do so, we'll show that the sequence is monotone in a bounded set and thus is guaranteed to converge by the monotone sequence theorem. We start with the recursion relation from above.

$$c_{n+1} = \frac{c_n \left(b + \frac{a}{\alpha^2}\right)^2}{\left(b - \frac{a}{\alpha}\right)^2 + c_n \left(b + \frac{a}{\alpha^2}\right)} \quad (49)$$

This implies (after simplification):

$$c_{n+2} - c_{n+1} = \left[\frac{\left(\left(b + \frac{a}{\alpha^2}\right)\left(b - \frac{a}{\alpha}\right)\right)^2}{\left(\left(b - \frac{a}{\alpha}\right)^2 + c_{n+1}\left(b + \frac{a}{\alpha^2}\right)\right)\left(\left(b - \frac{a}{\alpha}\right)^2 + c_n\left(b + \frac{a}{\alpha^2}\right)\right)} \right] (c_{n+1} - c_n) \quad (50)$$

and for the object of interest:

$$\text{sign}(c_{n+2} - c_{n+1}) = \text{sign}(c_{n+1} - c_n) \quad (51)$$

Therefore our width parameters are monotonic. Evaluating the bounds, we know that the supremum and infimum of the sequence are contained within the minimum and maximum of the function:

$$f(x) = \frac{x \left(b + \frac{a}{\alpha^2}\right)^2}{\left(b - \frac{a}{\alpha}\right)^2 + x \left(b + \frac{a}{\alpha^2}\right)} \quad (52)$$

We see that:

$$-\infty < 0 = \min(f(x)) \leq \inf(c^n) \leq \sup(c^n) \leq \max(f(x)) = \left(b + \frac{a}{\alpha^2}\right) < \infty \quad (53)$$

which verifies the sequence is bounded. Since the sequence converges, we may take the limit:

$$\lim_{n \rightarrow \infty} c_{n+1} = \lim_{n \rightarrow \infty} \frac{c_n \left(b + \frac{a}{\alpha^2}\right)^2}{\left(b - \frac{a}{\alpha}\right)^2 + c_n \left(b + \frac{a}{\alpha^2}\right)} \quad (54)$$

and labelling the limit $\lim_{n \rightarrow \infty} c_n = \lim_{N \rightarrow \infty} c_{n+1} = c^*$, we see that the sequence converges to the stationary value. We have shown that there exists a unique stationary state which is globally attractive for any arbitrary Gaussian initial state. We may note that any normalizable function, excluding special functions, can be written as a sum of Gaussian functions and so we have shown that there exists exactly one stationary state which is globally attractive.

F. Calculation of the Autocorrelation Function

We may also calculate the autocorrelation function for constant environmental conditions, generated from M analytically. Restating M :

$$M(y \rightarrow x) = G\left(y, \frac{-x}{\alpha}, a\right) G(y, x, b) \quad (55)$$

which rearranged and simplified becomes:

$$G\left(x, -\alpha y, \frac{a}{\alpha^2}\right) G(x, y, b) = AG\left(x, y, \frac{a}{\alpha^2} + b\right) \quad (56)$$

With this we may write down the mean value of the n^{th} generation beginning at the initial state x :

$$\int_{-\infty}^{\infty} x' P(x' = y, t = n' + n | x' = x, t = n') dx' = x \Pi^n \quad (57)$$

where P in the integrand is the probability of a trajectory starting at x and ending at y after n generations, which from the Chapman-Kolmogorov relation is the repeated application of the transition probability M over n generations. We now turn to the definition of the autocorrelation function:

$$C(n) = \frac{\sum_{n'=0}^M (\tau(n') - \mu) (\tau(n'+n) - \mu)}{\sum_{n'=0}^M (\tau(n') - \mu) (\tau(n') - \mu)} = \frac{\sum_{n'=0}^M x(n') x(n'+n)}{\sum_{n'=0}^M x(n') x(n')} \quad (58)$$

Noting that these sums may be rewritten as integrals yields:

$$\sum_{n'=0}^M x(n') x(n'+n) = \int_{-\infty}^{\infty} xP(x) \left[\int_{-\infty}^{\infty} x'P(x' = y, t = n' + n | x' = x, t = n') dx' \right] dx \quad (59)$$

and

$$\sum_{n'=0}^M x(n') x(n') = \int_{-\infty}^{\infty} xP(x) x dx \quad (60)$$

where $P(x)$ is the initial stationary distribution. We may use our result in (56) to simplify (60):

$$\int_{-\infty}^{\infty} xP(x) \left[\int_{-\infty}^{\infty} x'P(x' = y, t = n' + n | x' = x, t = n') dx' \right] dx = \int_{-\infty}^{\infty} xP(x) x \Pi^n dx \quad (61)$$

Thus we have:

$$C(n) = \frac{\Pi^n \int_{-\infty}^{\infty} xP(x) x dx}{\int_{-\infty}^{\infty} xP(x) x dx} = \Pi^n \quad (62)$$

This correlation function is compared with experimental results in Fig. S1.

III. EXPERIMENTAL METHODS

All experiments were conducted with *E. coli* strain WM4419 (derived from wild type MG1655). PDMS devices modelled after the ‘‘Mother Machine’’[3] were constructed and sonicated in IPA and DI water before bonding to glass to remove PDMS waste from the channels. Cells were grown overnight at 37C in 20 mL of LB media. Devices were subjected to oxide plasma treatment for 10 minutes at 70 watts to decrease the hydrophobicity of the microchannel surfaces and subsequently coated with a 10 mg/mL BSA solution. A solid cell pellet was produced via centrifugation of the cultured cells at 5000 rpm for 5 minutes at 4C. The suspension was discarded and the cell pellet was mixed with 1.5 mL LB media to produce a concentrated solution that was then injected into the device by syringe. The cells were then loaded into the microchannels through centrifugation of the devices at 5000 rpm for 40 min at 4C. The device was then prepared for time-lapse microscopy on a stage kept at approximately 33C in an incubating box (Pathology Devices). Phase contrast microscopy was conducted with a 100X oil objective on a Nikon TE2000 (Nikon corp.). Note that for processing, septum formations were determined by eye due to poor image resolution.

During microscopy, 5 different growth conditions were tested: 1:2 LB:autoclaved water solutions with the addition of 0.40 g/mL sorbitol (high hypertonic stress), 0.33 g/mL sorbitol (medium hypertonic stress), 0.22 g/mL sorbitol (low hypertonic stress), and 0 g/mL sorbitol (control) as well

as a 1:100 LB:autoclaved water solution with the addition of 0.07 g/mL sorbitol (low nutrient stress). In addition 0.001 g/mL BSA was incorporated into each growth condition to avoid further cell adhesion. Eight pairs of media were tested where the cells were allowed to reach stable exponential growth in the first condition before the growth medium was switched and the cells relaxed into the stationary distribution corresponding to the second condition. The eight pairs were: (1) low nutrient to control, (2) control to low nutrient, (3) low hypertonicity to control, (4) control to low hypertonicity, (5) medium hypertonicity to control, (6) control to medium hypertonicity, (7) high hypertonicity to control, (8) control to high hypertonicity. A relatively slow flow rate of approximately 40 microliters per hour was maintained throughout the experiment apart from the initial flow of fresh media into the device where a high flow rate of approximately 400 microliters per hour was utilized to clear the device of excess cells and immediately following the medium switch for some experiments. Even at the slowest flow rate used, it was confirmed through fluorescence imaging that transfer of the media is fast relative to the timescale of the cellular response and diffusion of the media within the tubing did not cause any issues. The medium was switched through manual clips: two syringes were connected with a forked joint with one clip for each syringe. Only one syringe (with its clip open) was placed on the syringe pump at any time while the other syringe was set aside with its clip closed.

A total of 6510 cell divisions were analyzed for the construction of Figure 3. There were a minimum of 374 and a maximum of 1804 divisions (mean of 813.75) observed for each experiment. A representative fraction (min. 68, max. 265., mean 132.5) of these divisions observed at the beginning and end of each experiment where the environment is constant were selected for the construction of the distributions displayed in Figure 1 (of the main text). These distributions are well described by shifted gamma distributions.

IV. DATA PROCESSING, FITTING, AND ANALYSIS

A. Processing

After the completion of each experiment, only channels that contained healthy cells throughout the switch of the growth medium and past the point of relaxation into the final stationary state were indexed. For each of these channels, every septum formation for the mother cell (the cell at the bottom of of the channel) was recorded with the time elapsed from the beginning of microscopy. Once all septum formations were recorded, cell cycle durations τ were calculated as the time elapsed

between pairs of sequential septum formations and each corresponding division event was recorded at the time of the second septum formation. Once the cell cycle durations were indexed, we calculated the final stationary distribution by constructing a normalized histogram of a representative ensemble of cell cycles. We then calculated all the relevant statistical information (mean, variance, etc.) from this distribution. In total, this information yields $\rho(\tau, n)$: the probability that the n^{th} cell cycle (relative to the start of microscopy) will be of duration τ_n - a measure of cell state as a function of discrete time. To analyze the response dynamics we wanted a measure of cell state that was continuous in time so we constructed $\rho(\tau, t)$: the probability that a cell is currently within a cycle of duration τ at time t . To construct this object we first defined the cell state “trajectory function” $F(t)$ for each cell as follows, letting s_n represent the time of the N^{th} septum formation:

$$F(t) = \tau_n, s_n \leq t < s_{n+1} \quad (63)$$

We may see that $\rho(\tau, t) = \rho(F(t))$. We used the median of the trajectory distribution to analyze the response dynamics and fit the model to the data collected. A figure summarizing the data processing is shown below:

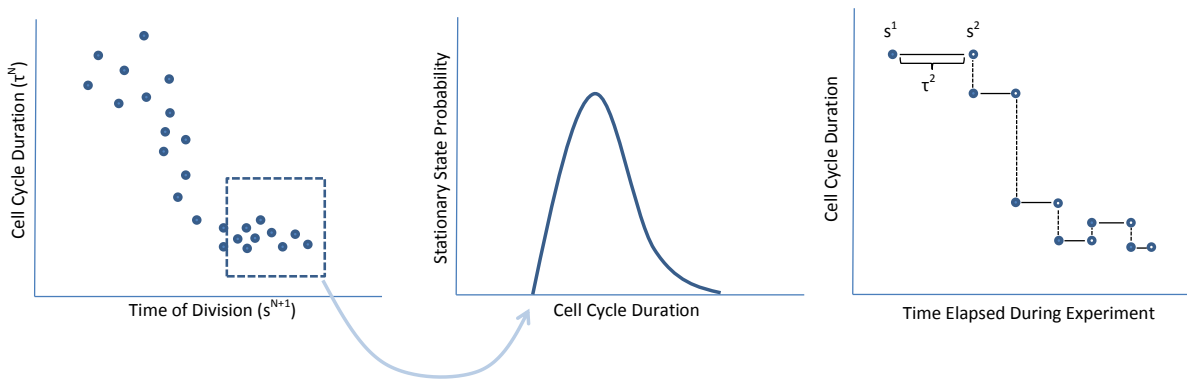


Figure S2: Data Processing: (from left to right) indexing cell cycle durations, calculating the stationary distribution, and constructing the cell trajectory functions.

We would like to acknowledge an issue with this method of processing: during the step change experiments which transition from rich media to poor media, the cells seem to respond before the shock. This is due to the fact that cells that enter their cycle shortly before the switch attain a large τ since they spend most of their cycle in the poor media; however, this cycle began before the switch leading to an increase in $\langle \tau \rangle$ before the media changes.

B. Fitting

We have a transition probability of the form:

$$M(y \rightarrow x) = M(y \rightarrow x) = \sqrt{\frac{b + \frac{a}{\alpha^2}}{\pi}} G\left(x, y\Pi, b + \frac{a}{\alpha^2}\right) \quad (64)$$

which when applied on an arbitrary Gaussian initial distribution, yields Gaussian distributions corresponding to the n^{th} generation with mean:

$$\langle \tau(n) \rangle \equiv m_n + \mu_f = \mu_f + \Pi^n m \quad (65)$$

where μ_f is the mean of the final distribution. Eventually, when environment is constant, we reach a stationary distribution with a width parameter of:

$$c^* = \frac{2ab(1 + \alpha) + a^2\left(\frac{1}{\alpha^2} - 1\right)}{\alpha^2 b + a} \quad (66)$$

We need to relate these quantities to experimental observables to find expressions for a and b . This is quite straightforward fitting to only the stationary distributions. Looking back at our form for the autocorrelation function we find:

$$C(n) = \Pi^n \quad (67)$$

The constant adder model predicts that $\Pi \approx \frac{-1}{4}$ [2] so we may solve for a and b in terms of c^* , the width parameter for the stationary state, and Π :

$$a = c^* \left[\frac{1 + \alpha \left(\frac{1 + \Pi}{1 - \Pi}\right)}{\left(\frac{1}{\alpha^2} - 1\right) + \frac{2(1 + \alpha)}{\alpha} \left(\frac{1 + \Pi}{1 - \Pi}\right)} \right] = \frac{c^*}{2(1 + \Pi)} \text{ for } \alpha = 1 \quad (68)$$

and

$$b = \frac{c^*}{\alpha} \left(\frac{1 + \Pi}{1 - \Pi}\right) \left[\frac{1 + \alpha}{\left(\frac{1}{\alpha^2} - 1\right)(1 - \Pi) + 2\left(\frac{1}{\alpha} + 1\right)\left(1 + \frac{\Pi}{\alpha}\right)} \right] = \frac{c^*}{2(1 - \Pi)} \text{ for } \alpha = 1 \quad (69)$$

For the purposes of comparing to the main text, we may restate these relationships in terms of σ_1 , σ_2 , coefficient of variation (CV), and μ . Looking at the case where $\alpha = 1$:

$$\Pi = \left(\frac{\left(\frac{\sigma_1}{\sigma_2}\right)^2 - 1}{\left(\frac{\sigma_1}{\sigma_2}\right)^2 + 1} \right) \quad (70)$$

$$\sigma_1 = \sqrt{2(1 + \Pi)} CV \mu \quad (71)$$

$$\sigma_2 = \sqrt{2(1 - \Pi)} CV \mu \quad (72)$$

$$CV = \frac{\sigma_1}{\sqrt{2(1 + \Pi)}\mu} = \frac{\sigma_2}{\sqrt{2(1 - \Pi)}\mu} \quad (73)$$

In fitting the model parameters this way, we have relied on our hypothesis that M is completely determined by the current environment ϕ . Additionally, we want to fit a and b directly to the response curve both for direct comparison with experiment and to probe the validity of this assumption. Fitting this way is a little more challenging. We still identify c^* as the width parameter for the stationary state, after the switch, and will arrive at the same expressions; however, we do not know Π a priori. Instead we must fit it to the experimental response curve. While constructing fitting rules, we must keep in mind that the model predicts mean division time $\langle \tau \rangle \equiv m + \mu_f$ in discrete generations where generation zero refers to the last generation at steady state with μ_i and a, b before the application of the new transition probability and generation one refers to the first generation after application. In the real system, however, the cell cycles are not phase locked and the identification of generations “zero” and “one” is poorly defined. This problem did not come up in the discussion of data processing because we constructed the cell trajectories $F(t)$ which are defined in real time. (We have checked that cell cycle phase is uniform and cells are not synchronized in the Mother Machine.) We can generate matching trajectories from the model results by assuming that the cell cycle phase distribution is uniform. This means that if a population of cells with a fixed average cell cycle duration $\langle \tau \rangle = m + \mu_f$ is monitored beginning at time t_0 , the probability density for observing the first division is uniform from t_0 to $t_0 + m + \mu_f$.

If we let t_0 represent the time of the growth medium switch we may construct the average cell cycle duration trajectories $G_k(t)$ for an ensemble of M cells (indexed by k) as follows:

$$G_k(t) = \begin{cases} m + \mu_f & -(m + \mu_f) \leq t < t_0 + \left(\frac{2(k-1)}{M} - 1\right) (m + \mu_f) \\ m_n + \mu_f & t_0 + \left(\frac{2(k-1)}{M} - 1\right) m + \sum_{i=1}^{n-1} m_i \leq t < t_0 + \left(\frac{2(k-1)}{M} - 1\right) m + \sum_{i=1}^n m_i \end{cases} \quad (74)$$

This ensemble represents a population of idealized cells that always divide in time $m_n + \mu_f$ for their n^{th} generation relative to the application of the new transition probability but differ in the phase of the cell cycle. $G_k(t)$ is uniquely defined given m_n and the time of the medium switch. Further we know that given μ_i and μ_f taken from the experimental values for the median cell state (i.e. $median(F(t))$) at the time of the switch and the end of the experiment, the m_n are uniquely defined up to the constant Π :

$$m_n = median(F(t_\infty)) + \Pi^n (median(F(t_0)) - median(F(t_\infty))) \quad (75)$$

Given this expression, we found the value for Π corresponding to the least squared error between $median(F(t))$ and $G_k(t)$. Given both Π and c^* we may solve for a and b as we did for the stationary case above. A figure illustrating how we find the best Π for the data is shown below:

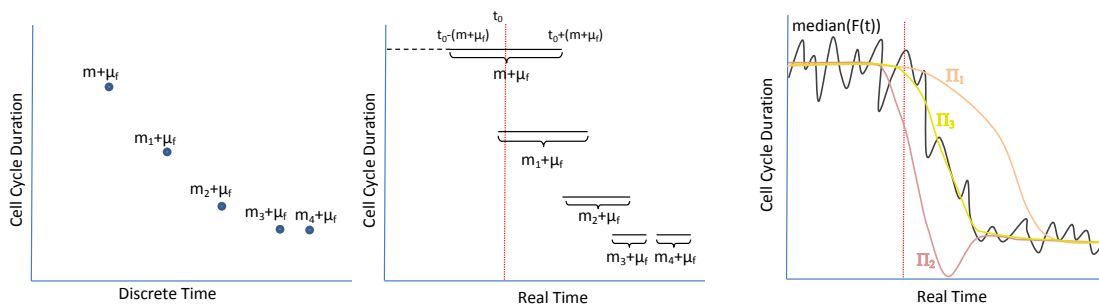


Figure S3: Data Fitting: (from left to right) predicted mean cell cycle durations indexed by generation (discrete time), conversion to real time, and fitting the parameter Π via a least squared error match of the set of theoretical real time curves $G(t)$ to the experimental curve $median(F(t))$.

C. Analysis

Here we display the curves used for data fitting and discuss the trends observed across all experiments. In Fig. S4 and S5, all single cell trajectories $F(t)$ are plotted in yellow along with the $median(F(t))$ displayed in black and the model result in blue. Below in Fig. S4, we compare experimental and model results where a and b were fit using only the final stationary distribution.

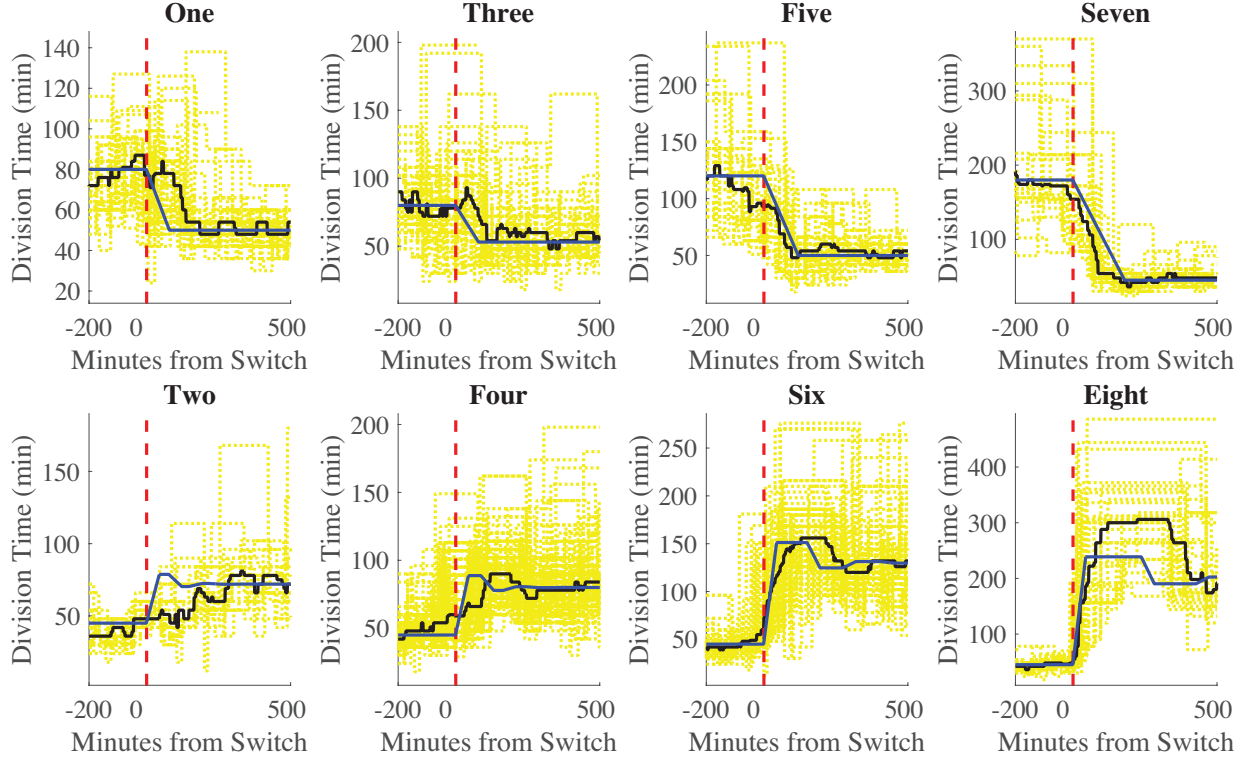


Figure S4: Single cell trajectories $F(t)$ are plotted in yellow along with the $median(F(t))$ displayed in black and the model fit to the final stationary distribution in blue. The plots correspond to the following: Experiment 1, increase nutrient; Experiment 2, decrease nutrient; Experiment 3, decrease osmolarity I; Experiment 4, increase osmolarity I; Experiment 5, decrease osmolarity II; Experiment 6, increase osmolarity II; Experiment 7, decrease osmolarity III; Experiment 8, increase osmolarity III.

In Fig. S5, we compare experimental and model results where a and b were fit directly to the full response curve. Note, this is the fit that corresponds to Figure 3 in the main text. We see the response curves corresponding to the model fit to only the stationary distributions (Fig. S4) are fairly consistent with those returned from a least-squared error fitting.

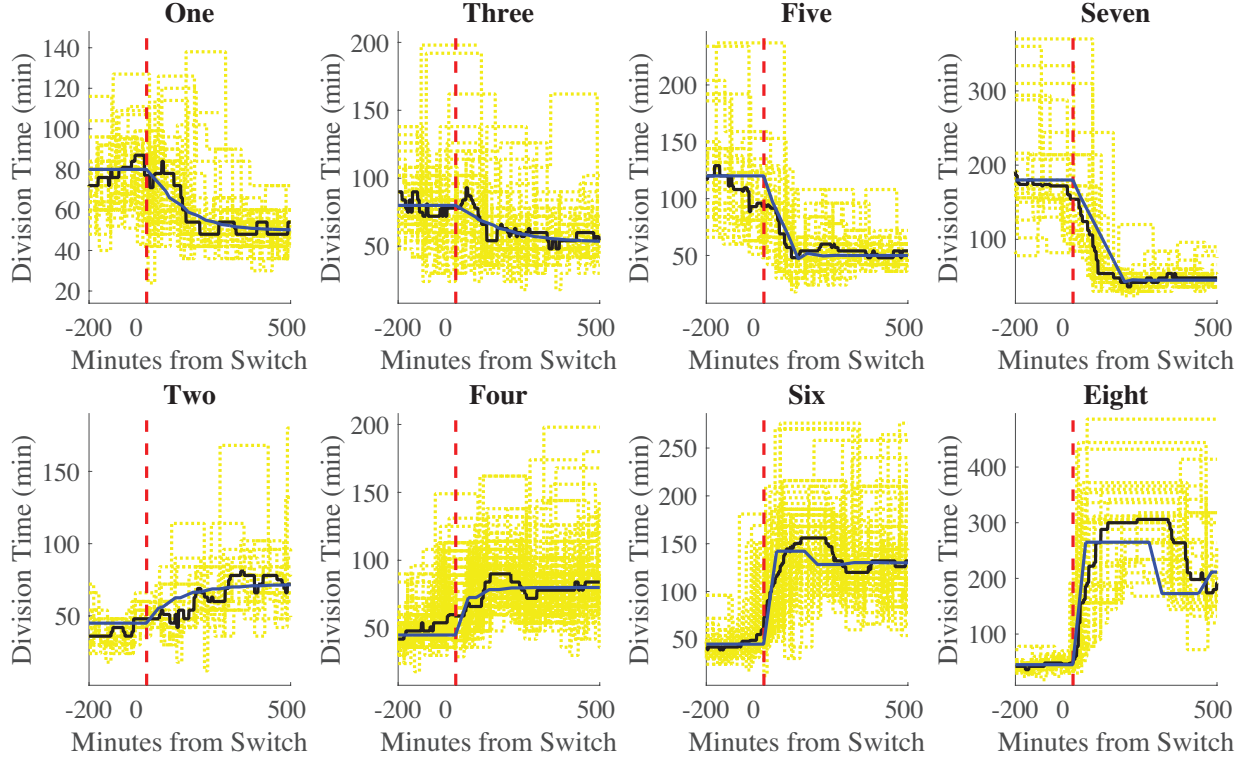


Figure S5: Single cell trajectories $F(t)$ are plotted in yellow along with the $median(F(t))$ displayed in black and the least-squared error model fit in blue. The plots correspond to the following: Experiment 1, increase nutrient; Experiment 2, decrease nutrient; Experiment 3, decrease osmolarity I; Experiment 4, increase osmolarity I; Experiment 5, decrease osmolarity II; Experiment 6, increase osmolarity II; Experiment 7, decrease osmolarity III; Experiment 8, increase osmolarity III.

Examining the cell responses to environmental changes, the trend across the relaxation experiments where the environmental stress is removed (odd numbered experiments - top row) is quite clear: as the variance (and mean) of the distributions before the environmental change increase, the response times remain roughly constant and thus the response speed significantly increases. For the shock experiments where the stress is induced (even numbered experiments - bottom row), as the mean of the final distribution increases the response time marginally increases and the overshoot of the final mean significantly increases.

Additionally, we show a comparison between the experimental and model (least-squared error) cell cycle duration distributions for the first three generations after the media switch in Fig. S6 and S7. Here "Generation 1" is defined to be the collection of cell cycles for which the first septum formation occurred before the media switch and the second septum formation occurred after the media switch.

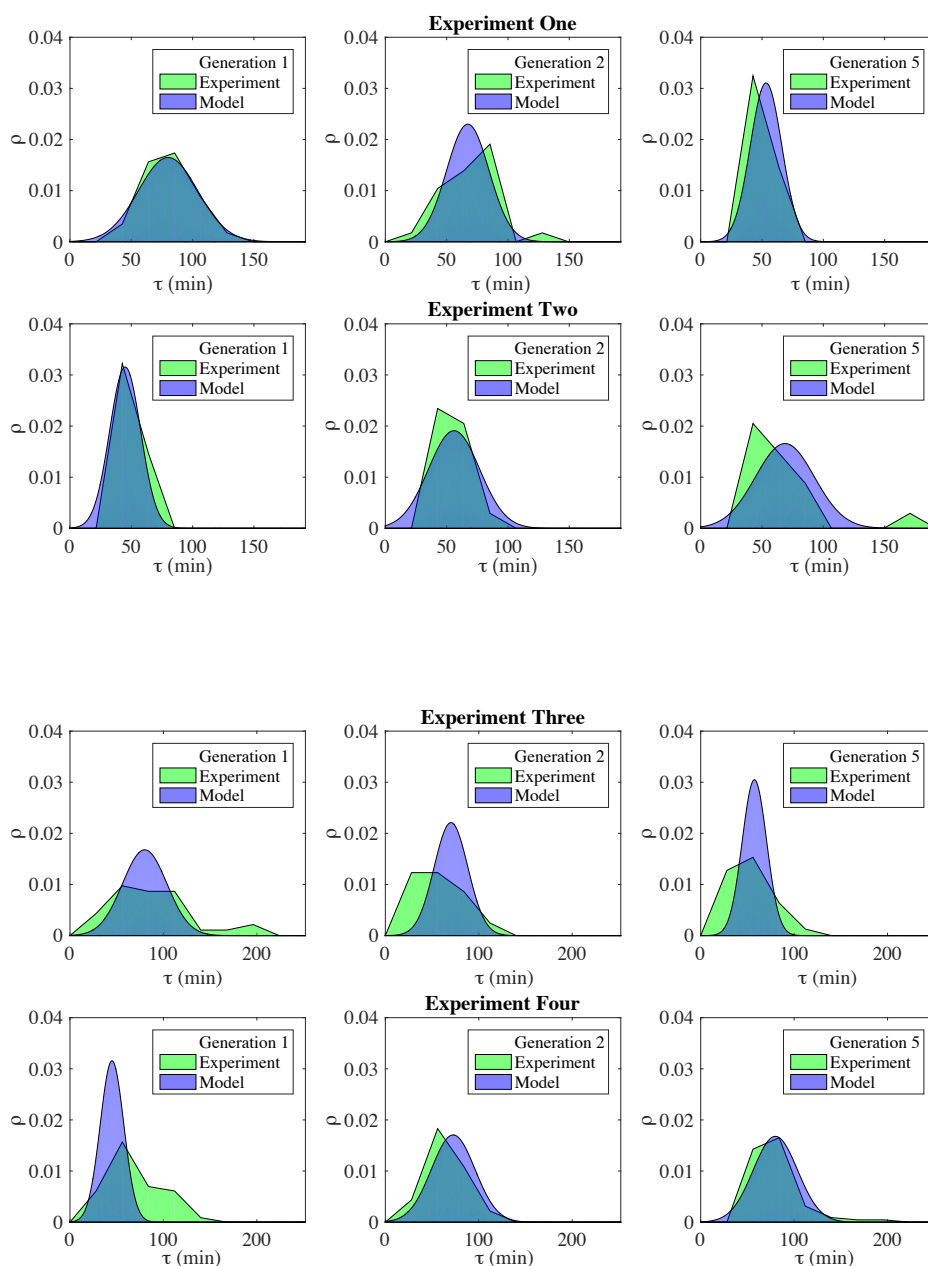


Figure S6: A comparison of the experimental and model cell cycle duration distributions for the first three generations after the media switch. Here "Generation 1" is defined to be the collection of cell cycles for which the first septum formation occurred before the media switch and the second septum formation occurred after the media switch. Experiments 1-4 are shown: Experiment 1, increase nutrient; Experiment 2, decrease nutrient; Experiment 3, decrease osmolarity I; Experiment 4, increase osmolarity I.

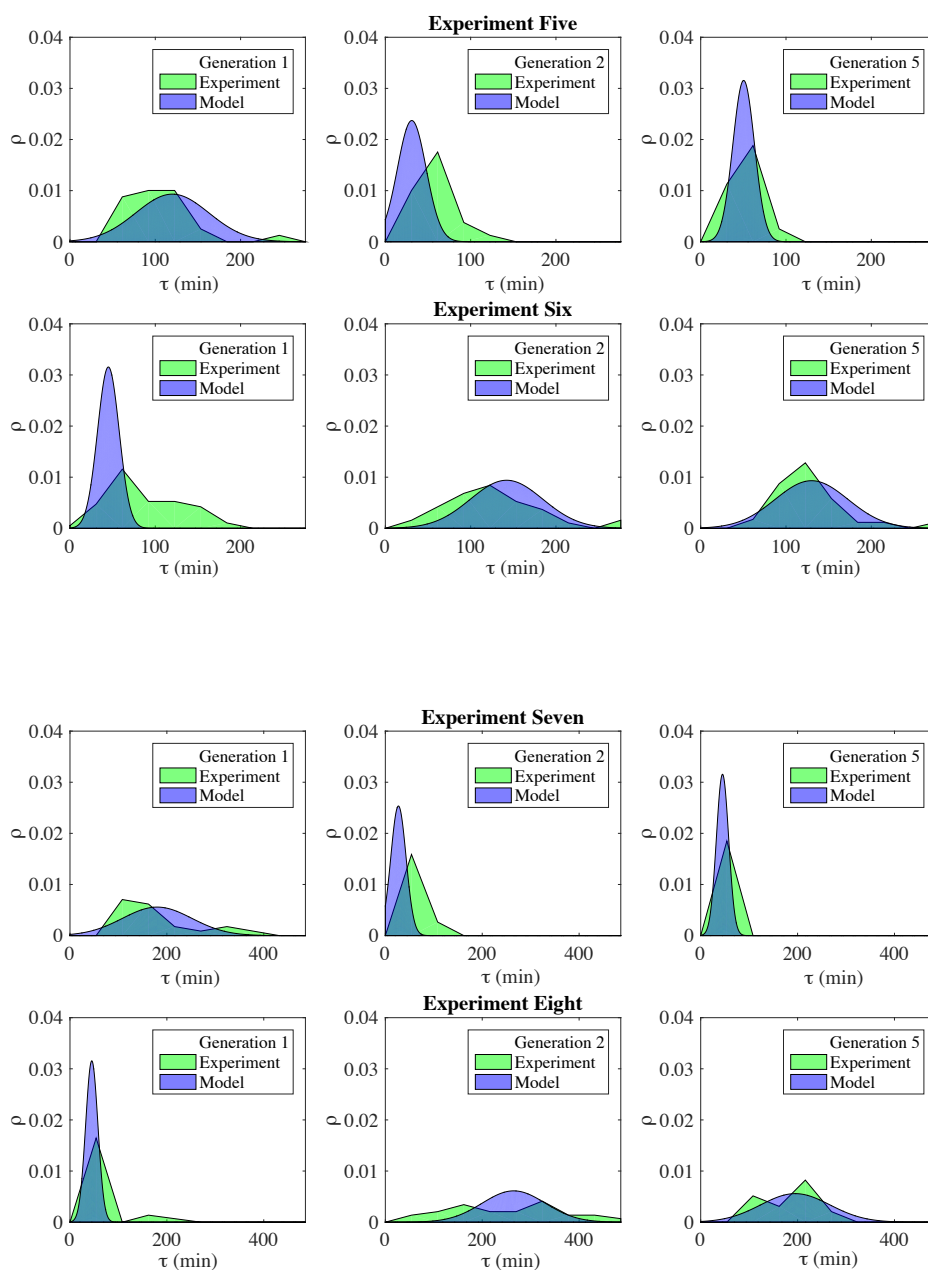


Figure S7: A comparison of the experimental and model cell cycle duration distributions for the first three generations after the media switch. Here "Generation 1" is defined to be the collection of cell cycles for which the first septum formation occurred before the media switch and the second septum formation occurred after the media switch. Experiments 5-8 are shown: Experiment 5, decrease osmolarity II; Experiment 6, increase osmolarity II; Experiment 7, decrease osmolarity III; Experiment 8, increase osmolarity III.

In addition to our theoretical fit and the discussion of the trends in the variance values across experiments, we also desired an empirical method of comparison (shown to the right in Fig. 4 of the main text). Response speed and absolute response time proved to be poor measures as nearly half of the data is non-monotonic and, as stated above, the response time varied little across the relaxation experiments. Instead we defined the following measure Δ : the higher of the initial or final mean division time subtracted from mean division time of the total distribution over the first $\lambda = 500$ minutes succeeding the environment shift, i.e.

$$\Delta \equiv \left\{ \frac{1}{N_{cells}} \sum_{N_{cells}} \left[\frac{1}{\lambda} \int_0^\lambda F(t) dt \right] \right\} - \max(\mu_i, \mu_f) \quad (76)$$

The limit of 500 minutes was selected as it is the minimum time within which all responses are complete. This is a measure of the efficiency of response. A Δ value greater than zero means during response, cells were forced to divide even slower than the slowest dividing stationary state. A Δ value less than zero means during response, cells were able to primarily divide faster than the slowest stationary state. Fig. S8 displays the Δ calculation and comparing the trends across variance and shift magnitude (the magnitude of the difference between the initial and final mean division times).

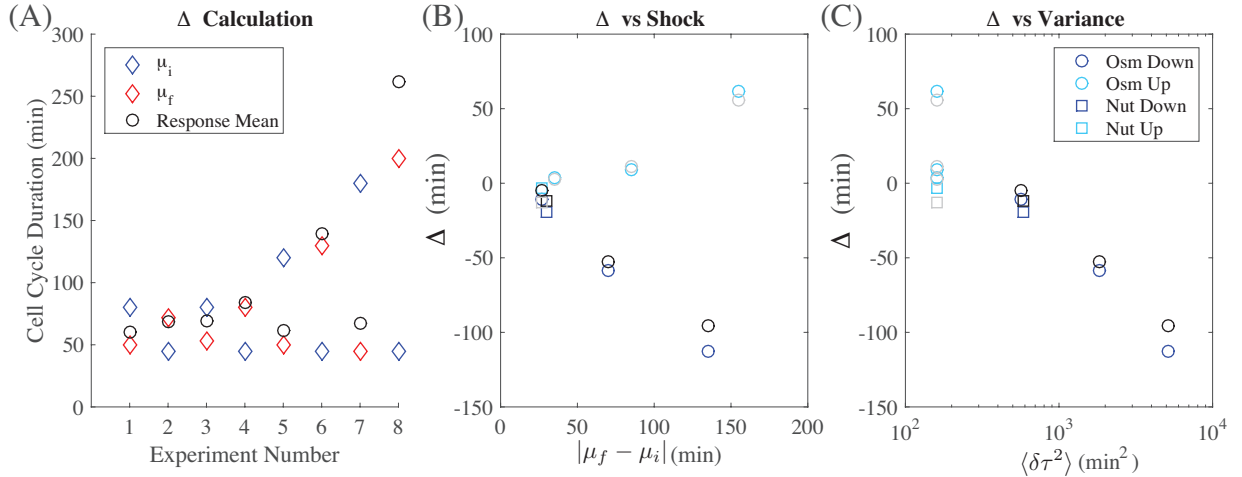


Figure S8: Calculation of Δ from Eq. (76). (A) The mean division time during response, $\frac{1}{N_{cells}} \sum_{cells} \frac{1}{\lambda} \int_0^\lambda F(t) dt$, is plotted as a black circle against the initial mean in blue and the final mean in red for each experiment. (B) The Δ values plotted against the environmental shift magnitude $|\mu_f - \mu_i|$. (C) The Δ values plotted against variance, $\langle \delta\tau^2 \rangle$, before the environmental shock. For (B) and (C) the grayscale symbols indicate alternative Δ values calculated with $\lambda = 250$. We may note that the trends are conserved across $\lambda = 250$ and $\lambda = 500$.

Three of the four shock experiments display a Δ value greater than zero indicating an ineffi-

cient response where the final mean was overshoot. As expected Δ increases with increasing shift magnitude. The four relaxation experiments all display negative Δ values which increase in magnitude, and thus improve response efficiency, as variance of the steady state distribution before the environmental change increases.

V. APPENDIX I: SUPPLEMENTARY MICROFLUIDIC EXPERIMENTS

Here we present data collected from a different strain (WM4584 derived from MG1655) in a larger version of the "mother machine" where cells are less confined within the microchannels. Imaging was also conducted on a different microscope (a Zeiss Axio Observer). We show this data to explicitly display ergodicity of the culture in even the slowest growth condition tested. We also show the repeat of experiment eight for the alternate strain in the alternate device and note that the response is qualitatively different for this strain in the modified mechanical environment in the same media. Figure S9 displays the results of the shock experiment (8) control to high hypertonicity.

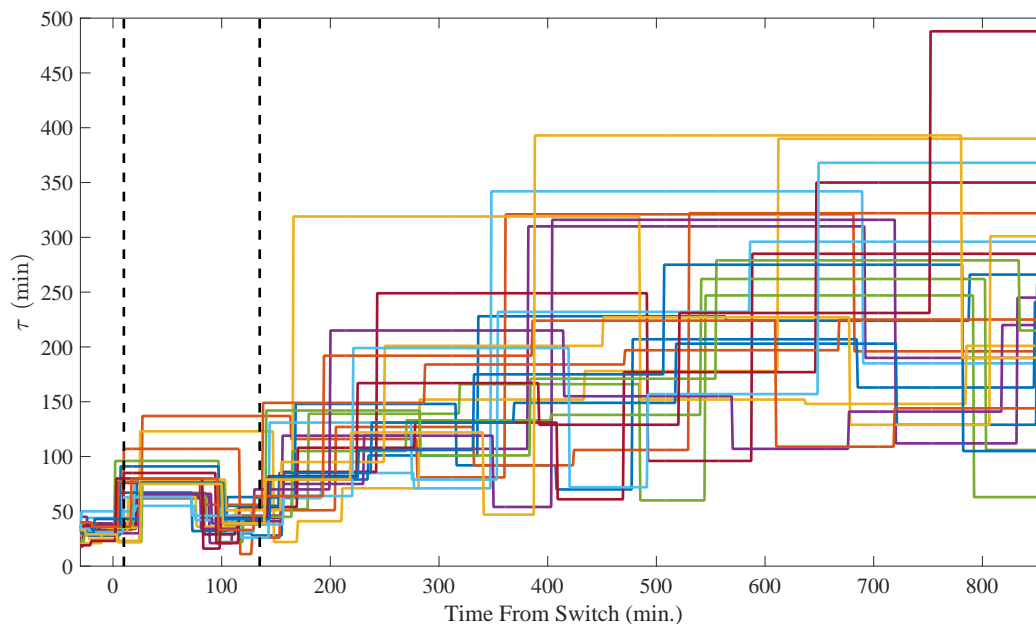


Figure S9: Media switch from control (1:2 LB:autoclaved water solution) to high hypertonic stress (+0.33 g/mL sorbitol). Both solutions contain 0.001 g/mL BSA

Figure S10 displays the autocorrelation function for the ensemble grown under high osmotic

stress as well as an overlay of the ensemble distribution with a sample of individual cell histograms. The autocorrelation function displays no long term memory, and combined with the single-cell distributions, confirms that the CCDD is ergodic even for this slowest-growth state tested.

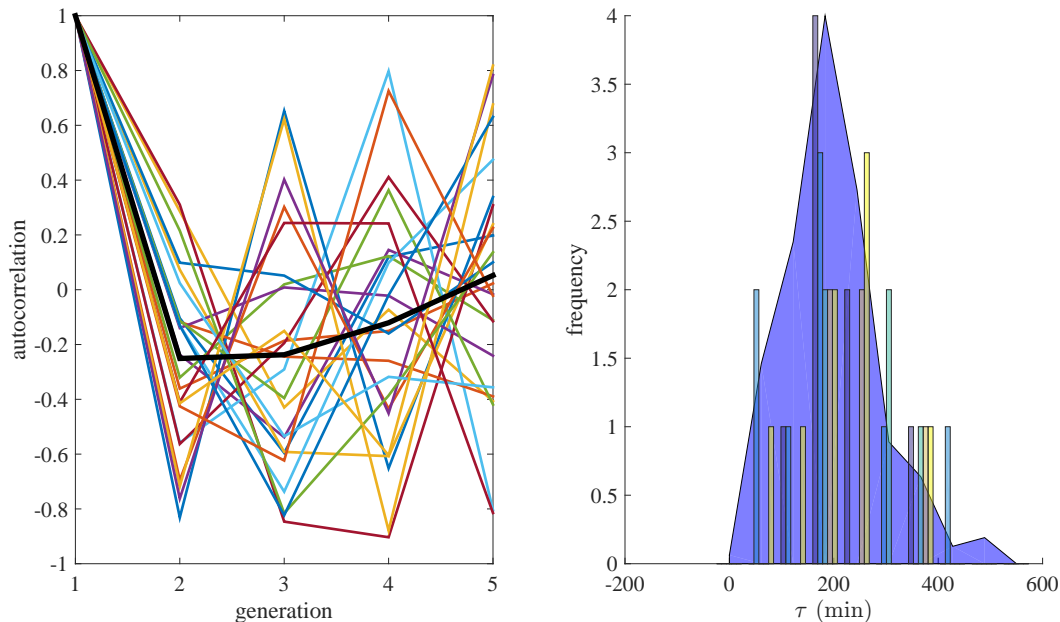


Figure S10: On the left, the autocorrelation function, $C(n) = \langle \delta\tau(n')\delta\tau(n' + n) \rangle / \langle \delta\tau^2 \rangle$, and on the right, the scaled ensemble distribution overlaid with representative single-cell distributions (histogrammed).

VI. APPENDIX II: SUPPLEMENTARY BULK EXPERIMENTS

In an effort to reproduce similar results in bulk, and to test temperature variation response, two bulk experiments were conducted. For each experiment cells were cultured overnight at 37C in 20 mL of LB media and in the morning a sample was removed from the original culture chamber. This sample was grown in a bulk culture of 1:2 LB:autoclaved water solution with the addition of 0.001 g/mL BSA until the optical density (OD) reached thirty percent at which time the new culture was diluted by a factor of three reducing the OD to ten percent. At the time of dilution, the temperature of the culture was also switched and at that second temperature, the cells were observed until the optical density reached forty percent. The temperature increase experiment was conducted with a culture grown at 22C and switched to 37C. The temperature decrease experiment was the reverse (37-22C). The bulk doubling times were calculated simply by finding the time required for the

optical density to double. Precisely, the bulk doubling time at time t was calculated as the time that elapsed between t and the time at which the optical density was half its current value i.e.:

$$\text{Doubling Time}(t) = t - t', OD(t') = \frac{1}{2}OD(t) \quad (77)$$

To calculate the doubling times near the time of the switch, the OD curve was expatiated backwards in time in accordance with the expected growth curve for a culture with the known initial doubling time τ_0 :

$$OD(t - \tau_0) = \frac{1}{2}OD(0) \exp\left(\frac{\ln(2)}{\tau_0}t\right), 0 \leq t < \tau_0 \quad (78)$$

A figure displaying the OD curves (where the expatiated portions are plotted as dashed lines) and doubling times are shown below:

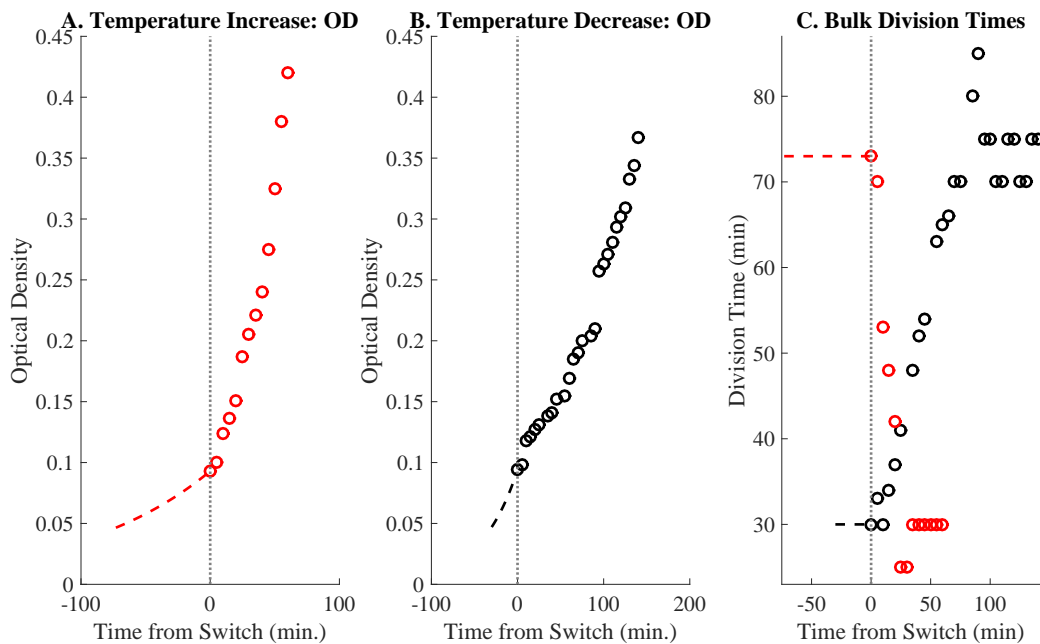


Figure S11: (A) OD curve for the temperature increase experiment - at time 0 the temperature is increased from 22C to 37C. (B) OD curve for the temperature decrease experiment - at time 0 the temperature is decreased from 37C to 22C. C.) bulk doubling time curves. The portions of the curves expatiated backwards in time from the temperature switch are displayed as dashed lines.

It is important to note that the physiological changes during temperature variation may be very different from those during nutrient or osmotic variation. Additionally in these bulk experiments, unlike the microfluidic work of the main text, cell-cell signaling is possible which may play an important role; and even without signaling, the instability of the environment (variable and growing OD) may lead to time varying mean cell behavior (e.g. mean division time, mean cell size, etc.) unlike what has been observed in the stable microfluidic environments. It has been shown that cell density, which is non-uniform and time varying in bulk conditions, plays an important role in modulating cell metabolism [7]. Expression levels of important proteins is a strong function of OD. Nonetheless, one can observe that the relaxation response (temperature increase) is faster than the shock (temperature decrease) which mirrors the microfluidic results.

It may also be interesting to note that during the bulk investigation, cells were cultured in the nutrient poor media (1:100 LB:autoclaved water solution with the addition of 0.07 g/mL sorbitol and 0.001 g/mL BSA) until the OD no longer varied in an attempt to probe the cell density at which the culture transitioned to stationary phase. We found that growing on the poor media, cells reached stationary phase at an OD ≈ 0.1 which is about 25% the value corresponding to growth on the rich media. This result, while perhaps unsurprising, merits discussion due to the practical concern that many spectrophotometers cannot accurately measure OD below 0.1 making a measurement of the bulk growth rate impossible using the method described in this section. This highlights another benefit of utilizing the “mother machine”: this microfluidic device not only enabled precise time-varying control of the environment impractical on the bulk scale, but also the study of stable growth conditions that might not be adequately characterized by spectrophotometry.

VII. REFERENCES

-
- [1] Chaudhry MA, Zubair SM. On a class of incomplete gamma functions with applications. CRC press; 2010.
 - [2] Taheri-Araghi S, Bradde S, Sauls JT, Hill NS, Levin PA, Paulsson J, et al. Cell-Size Control and Homeostasis in Bacteria. *Current Biology*. 2015;25(3):385-91.
 - [3] Wang P, Robert L, Pelletier J, Dang WL, Taddei F, Wright A, et al. Robust Growth of *Escherichia coli*. *Current biology*. 2010;20(12):1099–1103.
 - [4] Bremer H, Chuang L. The cell cycle in *Escherichia coli* *Journal of theoretical biology*. 1981;88(1):47–81.

- [5] Campos M, Surovtsev IV, Kato S, Paintdakhi A, Beltran B, Ebmeier SE, et al. A Constant Size Extension Drives Bacterial Cell Size Homeostasis. *Cell*. 2014;159(6):1433–1446.
- [6] Bromiley P. Products and convolutions of Gaussian probability density functions. *Tina-Vision Memo*. 2003;3.
- [7] Kalinin Y, Neumann S, Sourjik V, Wu M. Responses of *Escherichia coli* bacteria to two opposing chemoattractant gradients depend on the chemoreceptor ratio. *Journal of bacteriology*. 2010;192(7):1796–1800.