

Hysteresis and noise floor in gene expression optimised for persistence against lethal events

Pavol Bokes ^{*1} and Abhyudai Singh ^{†2}

¹Department of Applied Mathematics and Statistics, Comenius University, Bratislava 84248, Slovakia

²Department of Electrical and Computer Engineering, University of Delaware, Newark, Delaware 19716, USA

July 19, 2024

Abstract

Bacterial cell persistence, crucial for survival under adverse conditions like antibiotic exposure, is intrinsically linked to stochastic fluctuations in gene expression. Certain genes, while inhibiting growth under normal circumstances, confer tolerance to antibiotics at elevated expression levels. The occurrence of antibiotic events lead to instantaneous cellular responses with varied survival probabilities correlated with gene expression levels. Notably, cells with lower protein concentrations face higher mortality rates. This study aims to elucidate an optimal strategy for protein expression conducive to cellular survival. Through comprehensive mathematical analysis, we determine the optimal burst size and frequency that maximise cell proliferation. Furthermore, we explore how the optimal expression distribution changes as the cost of protein expression to growth escalates. Our model reveals a hysteresis phenomenon, characterised by discontinuous transitions between deterministic and stochastic optima. Intriguingly, stochastic optima possess a noise floor, representing the minimal level of fluctuations essential for optimal cellular resilience.

Keywords: stochastic gene expression, phenotypic heterogeneity, antibiotic tolerance, antibiotic persistence, antibiotic resistance

1 Introduction

Noisy gene expression constitutes a significant source of phenotypic variability within populations of genetically identical cells [1, 2]. This heterogeneity arises from the inherent discreteness of transcription and/or translation reactions [3, 4]. Mathematical modeling of these processes elucidates the relationship between kinetic gene expression parameters and the resulting population distribution of protein concentration [5–7]. Notably, even highly abundant proteins display a basal level of noise, commonly known as the noise floor [8, 9].

^{*}Corresponding Author: pavol.bokes@fmph.uniba.sk

[†]absingh@udel.edu

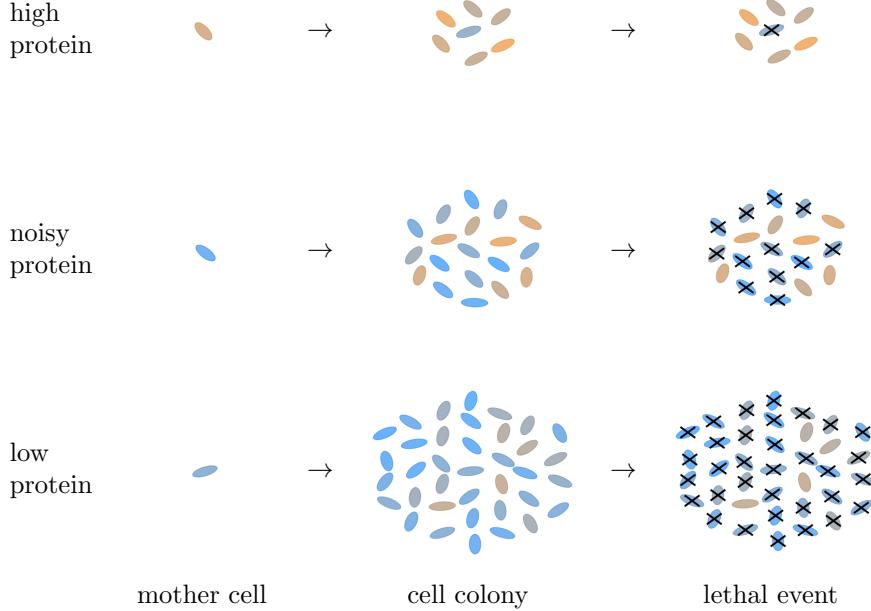


Figure 1: High protein expression imposes a cost to cell growth but provides protection at lethal events. Low protein expression enables growth but is sensitive to lethal events. The aim is to find optimal protein expression parameters (burst size and frequency) that maximise the expected colony size post lethal event.

Stochastic gene expression is effectively modeled by a framework rooted in the concept of bursty protein synthesis [10, 11]. Experimental evidence has documented gene expression bursts at both transcriptional and translational levels [12, 13]. This bursty model accurately predicts a gamma distribution of protein concentration within cellular populations [14, 15]. Notably, the shape and scale parameters of this distribution directly correspond to the mean frequency and size of bursts, providing valuable insights into the gene expression dynamics [16, 17].

Phenotypic variability leads to the persistence of clonal cells following treatment [18, 19]. This phenomenon entails a fractional killing of the cell population, wherein a portion of sensitive cells succumbs while a persistent fraction survives [20]. Notably, persistent cells are not characterized as resistant mutants and may retain sensitivity to subsequent treatments [21, 22]. Bacterial persistence has been correlated with heightened toxin expression and diminished cell growth rates [23–25] (see Figure 1). Notably, protein-dependent suppression of cell growth limits protein dilution and constitutes a positive feedback loop on protein expression [26, 27]. Trade-off between growth and persistence is also observed in cancer cells [28, 29].

Evolutionary pressures drive the selection of gene expression strategies that maximize species growth. Neglecting population heterogeneity may result in the

Notation	Interpretation
y	random process representing intercellular protein concentration
$n(y, t)$	number of cells with protein concentration y at time t
$r(y)$	cell growth rate as function of protein concentration y
a	burst frequency — the rate of occurrence of protein bursts
b	the mean size of exponentially distributed protein bursts
$x = ab$	steady state mean protein concentration
k	cost to cell growth rate per unit protein concentration
T	mean waiting time until a lethal event
$\mu = kT$	protein burden on growth at the time of a lethal event
$s(y)$	survival probability as function of protein concentration
s_0	probability that a cell without protein survives a lethal event
s_∞	survival probability of a cell with an infinite amount of protein
h	Hill coefficient for the survival probability function
$f(x, b)$	cost of protein expressed at mean x in bursts of mean size b

Table 1: Model parameters and variables. Without loss of generality, unit time is chosen so that $r(0) = 1$; unit protein concentration is set so that $s(1) = (s_0 + s_\infty)/2$.

adoption of suboptimal strategies [30, 31]. Achieving optimal gene expression involves assessing the trade-off between the cost of protein production and the benefits of protein-dependent growth [32] such as improved precision at high protein levels [33]. Other investigations have focused on maximizing mutual information between a transcription factor and the expression level of its target gene [34, 35]. Additionally, synthetic gene regulatory circuits have been proposed to sense and respond to environmental changes, adjusting their activity for optimal performance [36].

This paper centers on elucidating the role of a protein whose expression confers antibiotic tolerance while hampering cell growth. Section 2 presents the formulation of the bursty gene expression model alongside the associated cost function for the protein. In Section 3, we identify the critical points of this cost function and classify their type. Subsequently, in Section 4, we conduct a bifurcation analysis of these critical points with respect to a parameter measuring the burden that the protein places on cell growth. Finally, in Section 5, we offer a comprehensive summary and interpretation of our findings.

2 Model formulation

We consider a model for stochastic gene expression in a growing population of cells. We assume that a protein is produced in each cell in bursts as per a Poisson process with burst frequency a . In each burst, a random amount of protein is produced, which we assume to be exponentially distributed with mean size b . Cells proliferate with rate $r = r(y)$ which depends on the cellular concentration y of the protein. Cell growth continually dilutes the protein concentration as per the non-linear decay equation $dy/dt = -yr(y)$. The number $n(y, t)$ of cells with protein concentration y at time t satisfies an integro-partial-differential

equation

$$\frac{\partial n(y, t)}{\partial t} = \frac{\partial(yr(y)n(y, t))}{\partial y} + r(y)n(y, t) + \frac{a}{b} \int_0^y e^{-(y-z)/b} n(z, t) dz - an(y, t). \quad (1)$$

This population balance equation holds in a suitable time interval and for $y > 0$; the point $y = 0$ is a natural boundary. The four terms on the right-hand side of (1) quantify the change in population composition due to protein dilution, cell growth, burst influx, and burst efflux, respectively.

A single-cell version of the equation without the growth term was introduced in [14] and subsequently studied by other authors (e.g. [37, 38]) including us [39, 40]. We and co-authors extended the model by protein-dependent growth and analysed differences between the single-cell and population models in [41, 42]. The difficulty introduced by the growth term can be recognised by integrating (1) over all admissible protein concentration values $y \in (0, \infty)$. Doing so shows that the total cell count

$$N(t) = \int_0^\infty n(y, t) dy \quad (2)$$

satisfies

$$\frac{dN(t)}{dt} = \int_0^\infty r(y)n(y, t) dy. \quad (3)$$

If the growth rate $r(y)$ is independent of the protein level y , equation (3) is closed and immediately implies exponential growth. Otherwise, (3) is not closed and subtler methods are required to calculate the principal eigenvalue [42].

Here we circumvent this difficulty by assuming that the effect of the protein is relatively small:

$$r(y) = 1 - ky, \quad \text{where } ky \ll 1. \quad (4)$$

Without loss of generality, the growth rate is set to one in the absence of the protein.

After an initial transient, solutions to (1) are approximately equal to

$$n(y, t) \approx N(0)e^{(1-kab)t}p(y), \quad (5)$$

where

$$p(y) = \frac{y^{a-1}e^{-y/b}}{\Gamma(a)b^a} \quad (6)$$

is the normalised protein distribution (the gamma distribution). We note that the gamma distribution is independent on the burden k imposed by the protein. The effect of the burden is in slowing down the protein growth by a factor proportional to the steady state protein mean ab . These results are consistent with the exact analysis in [42]. We emphasise that this simplicity is due to the assumption of a small burden in (4). Without this assumption, the differentials in growth rates at different protein concentrations will tilt the protein probability distribution toward faster growing cells [42].

Cells that express the protein proliferate slower but have a better chance of survival of stress events (treatment). The probability of surviving treatment in the presence of y protein is modelled by a sigmoid function

$$s(y) = \frac{s_0 + s_\infty y^h}{1 + y^h}, \quad (7)$$

with $0 \leq s_0 < s_\infty$ and $h > 1$ assumed. The protein concentration at which the survival probability is equal to the average value $(s_0 + s_\infty)/2$ is thereby fixed to $y = 1$ without loss of generality. The proportion of cells surviving treatment is given by the integral

$$E(s(y)) = \int_0^\infty s(y)p(y)dy, \quad (8)$$

where the protein probability density function $p(y)$ depends on the gene expression parameters a and b as specified in (6).

We assume that treatment events occur as per a Poisson process with mean interarrival times T . Over a typical period from one treatment to another, the population grows by the factor of

$$\frac{N(T)}{N(0)} \approx g = e^{(1-kab)T} E(s(y)). \quad (9)$$

The fittest cells maximise g with respect to the parameters of the gene expression model. We apply on the dependent variable a strictly decreasing transformation $f = T - \ln g$. Our goal is therefore to minimise with respect to the protein mean $x = ab$ and the mean burst size b the *cost function*

$$f = \mu x - \ln E(s(y)). \quad (10)$$

The product $\mu = kT$ is henceforth referred to as *protein burden rate*. Dividing $s(y)$ by s_∞ is equivalent to adding a constant term $\ln s_\infty$ to the cost function; therefore, $s_\infty = 1$ can be assumed without loss of generality.

We parametrise the cost function by the protein mean $x = ab > 0$ and burst size $b > 0$. On the boundary of the parameter domain, the following asymptotic behaviour is identified (cf. Figure 2):

$$p(y) \rightarrow \delta(y - x), \quad b \rightarrow 0, \quad x = O(1), \quad (11)$$

$$p(y) \rightarrow \delta(y), \quad x \rightarrow 0, \quad b = O(1), \quad (12)$$

where $\delta(y)$ denotes the Dirac delta function. Regime (11) corresponds to a frequently expressed ($a = x/b \rightarrow \infty$) protein which deviates little from its mean value x . Regime (12) corresponds to a protein expressed rarely ($a = x/b \rightarrow 0$) in random bursts of mean size b . Both regimes will play a role in the optimisation problem.

The above asymptotics allow to extend the distribution (6) to the closure $x \geq 0$ and $b \geq 0$ of the parameter domain. For $x \geq 0$ and $b = 0$, the limit in (11) gives a deterministic protein expression state; increasing b from zero to a small positive value while keeping x unchanged makes a small, nearly Gaussian, stochastic perturbation to the deterministic expression. For $x = 0$ and $b \geq 0$, the limit (12) gives the non-expression state of the protein; increasing x from zero to a small positive value while keeping b unchanged makes a highly skewed and non-Gaussian perturbation of the non-expression state. These types of perturbations will be useful for characterising the character of the deterministic/non-expression states in the optimisation problem.

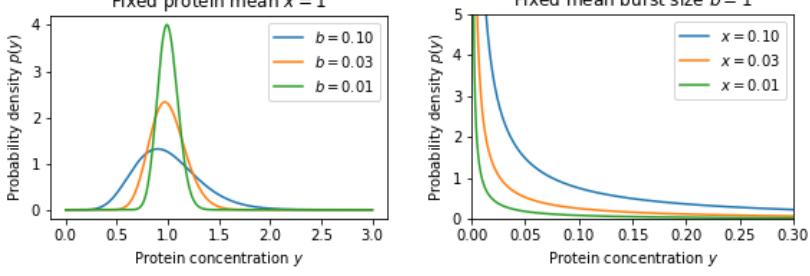


Figure 2: Protein distribution (6) for varied values of protein mean x and expected burst size b

3 Critical points of the cost function

We want to minimise the cost function $f = f(x, b)$ as function of protein mean x and burst size b within the positive quadrant $x \geq 0$, $b \geq 0$. More generally, we are also interested in the critical points of $f(x, b)$, which in addition to local minimisers also include maximisers and saddle points. Leaving aside the non-expression state, critical points can be deterministic ($b = 0$) or stochastic ($b > 0$).

3.1 Deterministic critical points

Inserting (11) into (8) gives $E(s(y)) = s(x)$ for $b = 0$. The cost of expressing the protein deterministically at level x is thus equal to

$$f(x, 0) = \mu x - \ln s(x). \quad (13)$$

By a deterministic critical point we understand a point $x > 0$ and $b = 0$ at which the cost is stationary:

$$\frac{\partial f}{\partial x}(x, 0) = \mu - \frac{s'(x)}{s(x)} = 0. \quad (14)$$

The character of the critical point is determined by the partial derivatives

$$\frac{\partial f}{\partial b}(x, 0) = -\frac{xs''(x)}{2s(x)}, \quad (15)$$

$$\frac{\partial^2 f}{\partial x^2}(x, 0) = \frac{(s'(x))^2 - s''(x)s(x)}{s^2(x)}. \quad (16)$$

Equation (16) is obtained by differentiating (14) with respect to x . If (16) is positive (negative), then perturbing the critical point by a small deterministic value increases (decreases) the cost. Equation (15) follows from a linear noise approximation (Appendix A). If (15) is positive (negative), then adding a small amount of stochastic noise increases (decreases) the cost. Clearly, the critical point is a point of (i) minimum if $\frac{\partial^2 f}{\partial x^2} > 0$ and $\frac{\partial f}{\partial b} > 0$; (ii) maximum if $\frac{\partial^2 f}{\partial x^2} < 0$ and $\frac{\partial f}{\partial b} < 0$; (iii) saddle if $\frac{\partial^2 f}{\partial x^2} \times \frac{\partial f}{\partial b} < 0$.

3.2 Stochastic critical points

By a stochastic critical point we understand a point $x > 0$ and $b > 0$ at which

$$\frac{\partial f}{\partial x} = 0, \quad \frac{\partial f}{\partial b} = 0. \quad (17)$$

This is further distinguished depending on the Hessian matrix

$$H = \begin{pmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x \partial b} \\ \frac{\partial^2 f}{\partial x \partial b} & \frac{\partial^2 f}{\partial b^2} \end{pmatrix}. \quad (18)$$

A stochastic critical point is a point of (i) minimum if H is positive definite; (ii) maximum if H is negative definite; (iii) saddle if H is indefinite.

4 Bifurcation analysis

We are interested in the problem of parameter continuation, i.e. determining how the coordinates and the character of critical points evolve as a bifurcation parameter changes. We use the protein burden rate μ as the bifurcation parameter and denote by $\mu_1 < \mu_2 < \dots$ its bifurcation values. We also denote by x_1, x_2, \dots and b_1, b_2, \dots the coordinates of the bifurcation points. The relatively simpler situation is treated first which occurs when survival is impossible without any protein ($s_0 = 0$); the effects of introducing a basal chance of survival ($s_0 > 0$) are studied thereafter. The latter case is referred to as “leaky” and the former as “tight”.

4.1 Tight case ($s_0 = 0$)

By (14) and (7), deterministic critical points satisfy

$$\frac{h}{x(1+x^h)} = \mu, \quad b = 0. \quad (19)$$

The left-hand side of (19) is a decreasing function of x , approaching infinity as $x \rightarrow 0$ and diminishing to zero as $x \rightarrow \infty$. Therefore, there exists a unique deterministic critical point $x = x(\mu)$, $b = 0$ for any given $\mu > 0$. This is shown in the bifurcation diagrams of Figure 3 in red colour.

The character of the deterministic critical point is determined by the signs of (15) and (16). Monotonicity of (19) implies that (16) is positive. Depending on the sign of (15), the deterministic critical point can be a minimum or a saddle point. Equation (15) simplifies to

$$\frac{\partial f}{\partial b}(x, 0) = \frac{((h+1)x^h - h+1)h}{2x(1+x^h)^2}. \quad (20)$$

Equating the numerator of (20) to zero and inserting into (19) give the explicit bifurcation values (cf. Figure 3)

$$\mu_2 = \frac{(h-1)^{\frac{1}{h}+1}}{2(h+1)^{\frac{1}{h}}}, \quad x_2 = \left(\frac{h-1}{h+1}\right)^{\frac{1}{h}}, \quad b_2 = 0. \quad (21)$$

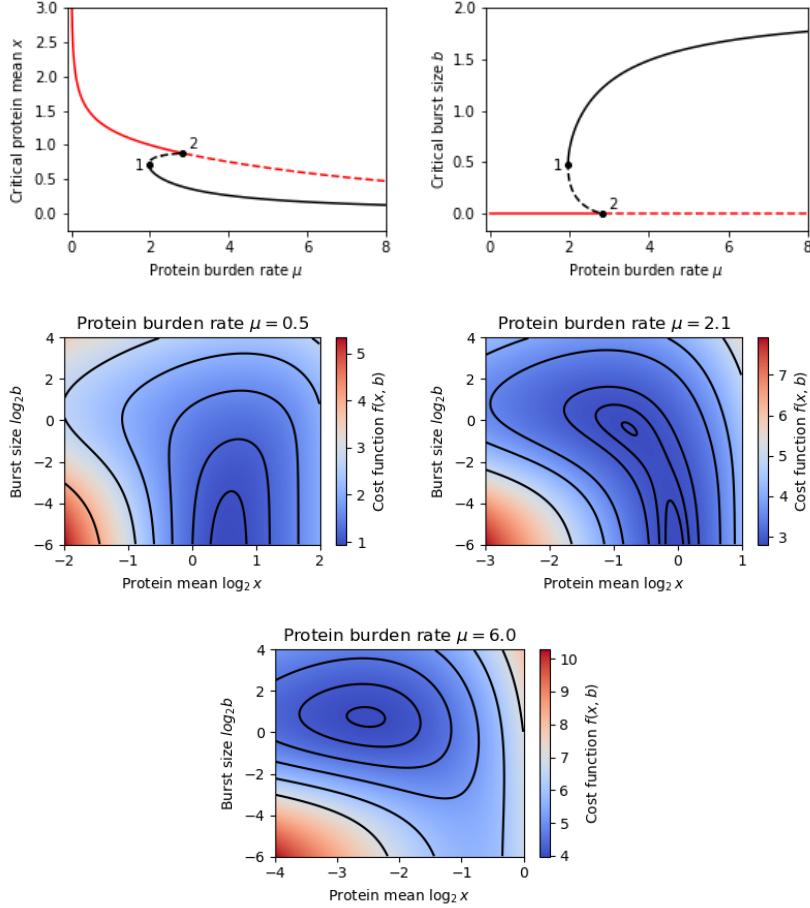


Figure 3: *Top panels:* The μ -dependence of the x - and b -coordinates of the cost function minimisers (full lines) and saddle-points (dashed lines) in the tight case ($s_0 = 0$). Markers indicate the bifurcation points. *Central and bottom panels:* The cost function $f(x, b)$ over a range of (logarithmically scaled) protein means x and burst sizes b for selected values of the protein burden rate μ . *Other parameter values:* $s_\infty = 1$, $h = 4$.

The critical point is a minimum if $\mu < \mu_2$ (i.e. $x > x_2$) and a saddle point if $\mu > \mu_2$ (i.e. $x < x_2$).

In addition to the deterministic critical point, we can also have up to two stochastic critical points for any value of μ . The μ -dependence is determined numerically using the method of Appendix B. The overall bifurcation picture as follows (cf. Figure 3): for $\mu < \mu_1 = 1.97$ the only critical point is the deterministic minimum. A typical heat map of $f(x, b)$ is depicted in central left panel. At $\mu = \mu_1$, a saddle-node bifurcation occurs at $(x_1, b_1) = (0.71, 0.48)$. For $\mu_1 < \mu < \mu_2 = 2.84$, a stochastic minimum and a stochastic saddle point exist in addition to the deterministic minimum. A typical heat map of $f(x, b)$ is depicted in central right panel. At $\mu = \mu_2$, the stochastic saddle coalesces with the deterministic minimum in a transcritical bifurcation. For $\mu > \mu_2$, there is a deterministic saddle and a stochastic minimum. A typical heat map of $f(x, b)$ is depicted in the bottom panel.

4.2 Leaky case ($s_0 > 0$)

If $s_0 > 0$, the function $s'(x)/s(x)$ is non-monotonous, and equation (14) for deterministic critical points has two solutions if $\mu < \mu_3$, one solution if $\mu = \mu_3$, and no solution if $\mu > \mu_3$, where μ_3 is a point of a saddle-node bifurcation. The μ -dependence of the deterministic critical points is shown in red colour in Figure 4.

If $s_0 > 0$, we also need to consider the trivial critical point with $x = 0$ (no protein expression) which has a finite cost

$$f(0, b) = -\ln s_0. \quad (22)$$

The character of the trivial critical point is determined by the sign of

$$\frac{\partial f}{\partial x}(0, b) = \mu - \frac{s_\infty - s_0}{s_0 b} \int_0^\infty \frac{y^{h-1} e^{-y/b}}{1 + y^h} dy. \quad (23)$$

Formula (23) follows from an asymptotic evaluation of the cost function (Appendix A). In Appendix A, we also show that there exists a bifurcation value μ_4 such that (i) if $\mu < \mu_4$, then $\partial f(0, b)/\partial x$ changes sign in the interval $0 \leq b < \infty$; (ii) if $\mu > \mu_4$, then $\partial f(0, b)/\partial x > 0$ for all $b \geq 0$. Therefore, we classify the trivial critical point as a saddle point if $\mu < \mu_4$ and a minimum if $\mu > \mu_4$.

The overall bifurcation behaviour is as follows (cf. Figure 4): for $\mu < \mu_1 = 1.63$, there are two deterministic critical points (a maximum and a minimum) and the trivial critical point, which is a saddle point. A typical heat map of $f(x, b)$ is depicted in the central left panel. At $\mu = \mu_1 = 1.63$, a saddle-node bifurcation occurs at $(x_1, b_1) = (0.67, 0.57)$. It is the same kind of bifurcation that is observed in the tight case, but the bifurcation values are perturbed due to the leakiness. For $\mu_1 < \mu < \mu_2 = 2.48$, a stochastic minimum and a stochastic saddle point exist in addition to the aforementioned deterministic critical points. A typical heat map of $f(x, b)$ is depicted in the central right panel. At $\mu = \mu_2$, the stochastic saddle point and the deterministic minimum collapse into a deterministic saddle point. Again, the same kind of transcritical bifurcation occurred in the tight case, but at a slightly different value of the bifurcation parameter. For $\mu_2 < \mu < \mu_3 = 4.1$, there is a stochastic minimum, a deterministic maximum, a deterministic saddle point, and the trivial saddle

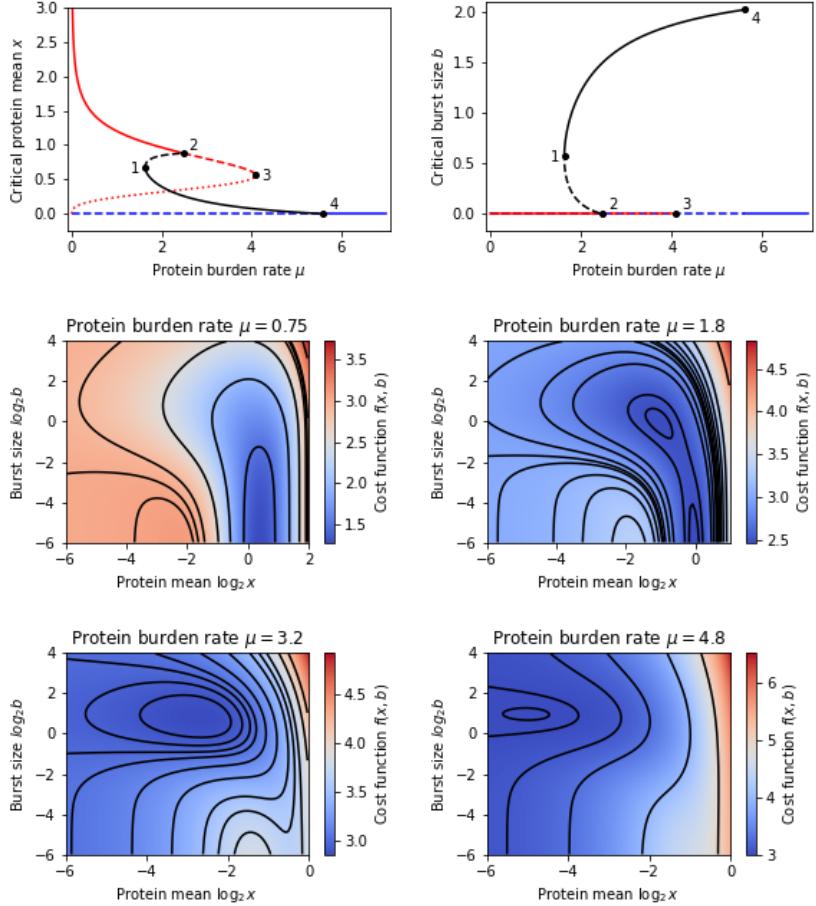


Figure 4: *Top panels:* The μ -dependence of the x - and b -coordinates of the cost function minimisers (full lines), saddle-points (dashed lines), and maximisers (dotted line) in the leaky case ($s_0 = 0.05$). Markers indicate the bifurcation points. *Central and bottom panels:* The cost function $f(x, b)$ over a range of (logarithmically scaled) protein means x and burst sizes b for selected values of the protein burden rate μ . *Other parameter values:* $s_\infty = 1$, $h = 4$.

point. A typical heat map of $f(x, b)$ is depicted in the bottom left panel. At $\mu = \mu_3$, the deterministic saddle point and maximum coalesce in a saddle-node bifurcation. This bifurcation does not have an analogue in the tight case. For $\mu_3 < \mu < \mu_4 = 5.61$, we have the trivial saddle point and a stochastic minimum. A typical heat map of $f(x, b)$ is depicted in the bottom right panel. At $\mu = \mu_4$, the stochastic minimum coalesces with the trivial saddle point in a transcritical bifurcation. This bifurcation also does not have an analogue in the tight case. For $\mu > \mu_4$, the only critical point is the trivial one, which is a minimum.

5 Conclusion

We considered a stochastically expressed protein which confers tolerance to antibiotic treatment but places a burden on the cell machinery and inhibits cell growth. The population growth observed in the presence of such a protein is quantified by a population balance equation. The underlying assumption behind the use of the equation is that the population density is continuous; it does not account for discrete population effects such as extinction and fixation [43]. In addition to the continuity assumption, we assumed that (i) the cost of protein to growth is relatively small and (ii) the waiting times till treatment are relatively large. Owing to these assumptions, the population distribution can be assumed to have equilibrated at a gamma distribution with shape and scale directly linked to the protein burst frequency and the burst size. Without these simplifying assumptions, we expect that the analysis would have to rely on a numerical solution of the integro-partial-differential population balance equation.

Subject to the aforementioned simplifications, the overall effect of the protein on the cell growth is quantified by a tractable cost function f . The cost function depends on the growth parameters (protein burden rate μ ; survival probabilities s_0 and s_∞ ; the Hill constant h) and the gene expression parameters (protein mean x ; mean burst size b). Gene expression can be stochastic ($x > 0$ and $b > 0$), deterministic ($x > 0$, $b = 0$), or turned off ($x = 0$, $b \geq 0$).

The cost is minimised as function $f = f(x, b)$ of the gene expression parameters for a given set of growth parameters. Depending on how cells without protein respond to treatment, we differentiated between the tight case ($s_0 = 0$) and the leaky case ($s_0 > 0$). Optimal gene expression can be stochastic or deterministic; in the leaky case it can be optimal not to express any protein.

The main result of our analysis is the description of how optimal gene expression parameters change in response to changing protein burden rate μ . This is visualised in bifurcation diagrams in Figure 3 (tight case) and Figure 4 (leaky case). In addition to minima, the bifurcation diagrams also show the μ -dependence of other types of critical points of the cost function, namely saddle points and maxima, and highlight bifurcation points $\mu_1 < \mu_2 < \mu_3 < \mu_4$, at which the critical points coalesce and change type.

Increasing protein burden rate μ leads to a decrease in the mean value x of the optimal protein expression. In the tight case, there is a single deterministic optimum if $\mu < \mu_1$, coexisting deterministic and stochastic optima if $\mu_1 < \mu < \mu_2$, and a single stochastic optimum if $\mu > \mu_2$. Thus, making protein more costly is associated with lower expression levels as well as the transition toward a stochastic regime of expression. In the leaky case, the aforementioned holds with one caveat: it is optimal not to express protein if $\mu > \mu_4$. Thus, if the cost

of expressing the protein is too high, it is optimal to switch it off despite the protection it confers.

Deterministic and stochastic optima present contrasting strategies to use the protein to survive treatment. In the former, the protein is maintained at high levels throughout the population, so as achieve a high probability of survival in all cells. In the latter, the population expresses the protein in random bursts, with a small subpopulation of cells being sufficiently protected at any given time. The bifurcation diagrams reveal hysteresis in transitions between stochastic and deterministic optima. Consequently, we observe an all-or-none property in gene expression noise: optima are either deterministic or highly stochastic; indeed, stochastic optima are characterised by large burst sizes and low mean values, leading to large coefficients of variation $CV^2 = b/x$. Our results thus suggest the existence of a noise floor for resistance driven by stochastic gene expression.

Appendix A. Asymptotics of the cost function

The limit of $b \rightarrow 0$ with x fixed (linear noise approximation). Taylor expanding the survival probability around the protein mean gives:

$$s(y) = s(x) + s'(x)(y - x) + \frac{s''(x)}{2}(y - x)^2 + \dots$$

Taking the expectation with respect to the protein distribution gives

$$E(s(y)) = s(x) + \frac{s''(x)}{2} \text{Var}(y) + \dots = s(x) + \frac{s''(x)}{2} xb + o(b), \quad (24)$$

in which we substituted the variance of the gamma distribution [44]; higher-order central moments of the gamma distribution are $o(b)$. Taking the logarithm of (24) yields

$$f(x, b) = \mu x - \ln s(x) - \frac{xs''(x)b}{2s(x)} + o(b), \quad (25)$$

in which the $O(1)$ term is equal to the cost of a deterministic expression (13) and the $O(b)$ term gives the cost derivative (15) due to a stochastic perturbation of the deterministic expression.

The limit of $x \rightarrow 0$ with b fixed. This is equivalent to taking $a = x/b \rightarrow 0$.

$$\begin{aligned} E(s(y)) &= s_0 + (s_\infty - s_0)E\left(\frac{y^h}{1+y^h}\right) \\ &= s_0 + \frac{s_\infty - s_0}{b^a \Gamma(a)} \int_0^\infty \frac{y^{h+a-1} e^{-y/b}}{1+y^h} dy \\ &= s_0 + a(s_\infty - s_0) \int_0^\infty \frac{y^{h-1} e^{-y/b}}{1+y^h} dy + o(a). \end{aligned} \quad (26)$$

Taking the logarithm of (26) leads to

$$f(x, b) = -\ln s_0 + x \left(\mu - \frac{s_\infty - s_0}{s_0 b} \int_0^\infty \frac{y^{h-1} e^{-y/b}}{1+y^h} dy \right) + o(x), \quad (27)$$

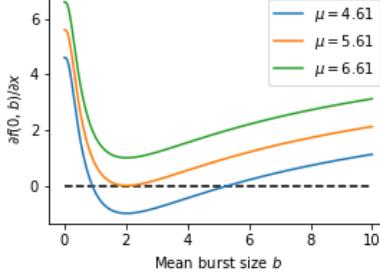


Figure 5: The cost derivative $\partial f(0, b)/\partial x$ has zero, one, or two zeros depending on the burden rate μ . Parameter values are $h = 4$, $s_0 = 0.05$, $s_\infty = 1$.

in which the $O(1)$ term is equal to the cost of the trivial state (22) and the $O(x)$ term gives the cost derivative (23) due to a stochastic perturbation of the trivial state. The derivative as function of mean burst size b can change sign depending on the burden rate μ (Figure 5).

Appendix B. Parameter continuation

We provide a specialised algorithm for the μ -continuation of stochastic critical points of the cost function (10). It is easier to implement and is expected to be more efficient than general approaches [45, 46].

Inserting (6) into (10) and simplifying give

$$f(x, b) = \mu x + \ln \Gamma(a) + a \ln b - \ln \int_0^\infty s(y) y^{a-1} e^{-y/b} dy. \quad (28)$$

The burst frequency a in (28) is understood to be a function of the protein mean x and the burst size b :

$$a = \frac{x}{b}, \quad \frac{\partial a}{\partial x} = \frac{1}{b}, \quad \frac{\partial a}{\partial b} = -\frac{x}{b^2}. \quad (29)$$

Differentiating (28) gives

$$\frac{\partial f}{\partial x} = \mu + \frac{1}{b} \left(\psi(a) + \ln b - \frac{\int_0^\infty s(y) y^{a-1} e^{-y/b} \ln y dy}{\int_0^\infty s(y) y^{a-1} e^{-y/b} dy} \right), \quad (30)$$

$$\frac{\partial f}{\partial b} = \frac{1}{b^2} \left((1 - \psi(a) - \ln b) ab - \frac{\int_0^\infty s(y) y^{a-1} e^{-y/b} (y - ab \ln y) dy}{\int_0^\infty s(y) y^{a-1} e^{-y/b} dy} \right), \quad (31)$$

where $\psi(a) = d \ln \Gamma(a)/da$ is the digamma function. The bifurcation parameter μ features only as an additive term in (30) and is absent from (31). This simplicity facilitates an easy calculation of a bifurcation diagram as described below.

The critical point equations (17) are solved in the extended domain $(x, b, \mu) \in \mathbb{R}_+^3$ (gene expression and bifurcation parameters) in the following steps:

1. Pick a value $b > 0$.

2. Solve numerically in $a = a(b)$ the critical equation $\partial f / \partial b = 0$, i.e.

$$(1 - \psi(a) - \ln b)ab - \frac{\int_0^\infty s(y)y^{a-1}e^{-y/b}(y - ab \ln y)dy}{\int_0^\infty s(y)y^{a-1}e^{-y/b}dy} = 0. \quad (32)$$

3. Set $x = x(b) = a(b)b$.

4. Use the second critical equation $\partial f / \partial x = 0$ to calculate $\mu = \mu(b)$ for the given pair of $a = a(b)$ and b , i.e. set

$$\mu = \frac{1}{b} \left(\frac{\int_0^\infty s(y)y^{a-1}e^{-y/b} \ln y dy}{\int_0^\infty s(y)y^{a-1}e^{-y/b} dy} - \psi(a) - \ln b \right). \quad (33)$$

5. Perturb the value $b \rightarrow b + \varepsilon$ and go to Step 2.

Equation (32) is solved practically using the Newton method, which requires an initial guess for the solution $a = a(b)$. In the first iteration, we assume that such a guess is available (e.g. from a visual inspection of the cost function). In the second iteration, we use the solution $a(b)$ from the first iteration as the initial guess for the new solution $a(b + \varepsilon)$, which itself becomes the initial guess in the next iteration, etc.

The above procedure constructs a curve $(x, b, \mu) = (x(b), b, \mu(b))$ parametrised by $b > 0$ in the extended domain of the cost function; $(x(b), b)$ is thereby a critical point of the cost function for $\mu = \mu(b)$. Plotting $x(b)$ and b against $\mu(b)$ for a range of b values visualises how stochastic critical points (x, b) depend on the bifurcation parameter μ ; such plots are shown in the main text.

References

- [1] B. M. C. Martins and J. Locke, “Microbial individuality: how single-cell heterogeneity enables population level strategies,” *Current Opinion in Microbiology*, vol. 24, pp. 104–112, 2015.
- [2] C. Schwall, T. Loman, B. Martins, S. Cortijo, C. Villava, V. Kusmartsev, T. Livesey, T. Saez, and J. Locke, “Tunable phenotypic variability through an autoregulatory alternative sigma factor circuit,” *Molecular Systems Biology*, vol. 17, 2021.
- [3] D. Jones, R. Brewster, and R. Phillips, “Promoter architecture dictates cell-to-cell variability in gene expression,” *Science*, vol. 346, pp. 1533–1536, 2014.
- [4] M. Niepel, S. Spencer, and P. Sorger, “Non-genetic cell-to-cell variability and the consequences for pharmacology,” *Current Opinion in Chemical Biology*, vol. 13, pp. 556–561, 2009.
- [5] M. Sturrock, S. Li, and V. Shahrezaei, “The influence of nuclear compartmentalisation on stochastic dynamics of self-repressing gene expression,” *Journal of Theoretical Biology*, vol. 424, pp. 55–72, 2017.
- [6] K. Iida, N. Obata, and Y. Kimura, “Quantifying heterogeneity of stochastic gene expression,” *Journal of Theoretical Biology*, vol. 465, pp. 56–62, 2019.

- [7] R. Rudnicki and A. Tomski, “On a stochastic gene expression with pre-mRNA, mRNA and protein contribution,” *Journal of Theoretical Biology*, vol. 387, pp. 54–67, 2015.
- [8] R. D. Dar, B. S. Razooky, L. S. Weinberger, C. D. Cox, and M. L. Simpson, “The low noise limit in gene expression,” *PLoS one*, vol. 10, no. 10, p. e0140969, 2015.
- [9] J. Jedrak and A. Ochab-Marcinek, “Contributions to the noise floor in gene expression in a population of dividing cells,” *Scientific Reports*, vol. 10, no. 1, p. 13533, 2020.
- [10] N. Almeira and S. Risau-Gusman, “Role of transcriptional bursts in cellular oscillations,” *Journal of Theoretical Biology*, vol. 426, pp. 49–56, 2017.
- [11] J. Tyson, T. Laometsachit, and P. Kraikivski, “Modeling the dynamic behavior of biochemical regulatory networks,” *Journal of Theoretical Biology*, vol. 462, pp. 514–527, 2019.
- [12] V. Elgart, T. Jia, A. Fenley, and R. Kulkarni, “Connecting protein and mRNA burst distributions for stochastic models of gene expression,” *Physical Biology*, vol. 8, p. 046001, 2011.
- [13] H. Ochiai, T. Hayashi, M. Umeda, M. Yoshimura, A. Harada, Y. Shimizu, K. Nakano, N. Saitoh, Z. Liu, T. Yamamoto, T. Okamura, Y. Ohkawa, H. Kimurâ, and I. Nikaido, “Genome-wide kinetic properties of transcriptional bursting in mouse embryonic stem cells,” *Science Advances*, vol. 6, 2020.
- [14] N. Friedman, L. Cai, and X. Xie, “Linking stochastic dynamics to population distribution: an analytical framework of gene expression,” *Physical Review Letters*, vol. 97, p. 168302, 2006.
- [15] J. C. J. Ray, “Survival of phenotypic information during cellular growth transitions,” *ACS Synthetic Biology*, vol. 5, no. 8, pp. 810–816, 2016.
- [16] R. Dar, B. Razooky, A. Singh, T. Trimeloni, J. McCollum, C. Cox, M. Simpson, and L. Weinberger, “Transcriptional burst frequency and burst size are equally modulated across the human genome,” *Proceedings of the National Academy of Sciences*, vol. 109, pp. 17454–17459, 2012.
- [17] D. Nicolas, B. Zoller, D. Suter, and F. Naef, “Modulation of transcriptional burst frequency by histone acetylation,” *Proceedings of the National Academy of Sciences*, vol. 115, pp. 7153–7158, 2018.
- [18] N. Q. Balaban, J. Merrin, R. Chait, L. Kowalik, and S. Leibler, “Bacterial persistence as a phenotypic switch,” *Science*, vol. 305, no. 5690, pp. 1622–1625, 2004.
- [19] E. B. Gunnarsson, S. De, K. Leder, and J. Foo, “Understanding the role of phenotypic switching in cancer drug resistance,” *Journal of Theoretical Biology*, vol. 490, p. 110162, 2020.

- [20] F. Puccioni, J. Pausch, P. Piho, and P. Thomas, “Noise-induced survival resonances during fractional killing of cell populations,” *arXiv preprint 2402.19045*, 2024.
- [21] S. M. Shaffer, M. C. Dunagin, S. R. Torborg, E. A. Torre, B. Emert, C. Krepler, M. Beqiri, K. Sproesser, P. A. Brafford, M. Xiao, *et al.*, “Rare cell variability and drug-induced reprogramming as a mode of cancer drug resistance,” *Nature*, vol. 546, no. 7658, pp. 431–435, 2017.
- [22] T. Hossain, A. Singh, and N. C. Butzin, “Escherichia coli cells are primed for survival before lethal antibiotic stress,” *Microbiology Spectrum*, vol. 11, no. 5, pp. e01219–23, 2023.
- [23] E. Rotem, A. Loinger, I. Ronin, I. Levin-Reisman, C. Gabay, N. Shores, O. Biham, and N. Q. Balaban, “Regulation of phenotypic variability by a threshold-based mechanism underlies bacterial persistence,” *Proceedings of the National Academy of Sciences*, vol. 107, no. 28, pp. 12541–12546, 2010.
- [24] J. Feng, D. A. Kessler, E. Ben-Jacob, and H. Levine, “Growth feedback as a basis for persister bistability,” *Proceedings of the National Academy of Sciences*, vol. 111, no. 1, pp. 544–549, 2014.
- [25] K. T. Rahman, R. Amaratunga, A. Singh, T. Hossain, and N. C. Butzin, “An isogenic E. coli population gives rise to multiple persister phenotypes,” *bioRxiv preprint 10.1101/2023.12.09.570944*, 2023.
- [26] O. Patange, C. Schwall, M. Jones, C. Villava, D. A. Griffith, A. Phillips, and J. C. Locke, “Escherichia coli can survive stress by noisy growth modulation,” *Nature Communications*, vol. 9, no. 1, p. 5333, 2018.
- [27] I. Zabaikina, Z. Zhang, C. Nieto, P. Bokes, and A. Singh, “Amplification of noisy gene expression by protein burden: An analytical approach,” in *2023 American Control Conference (ACC)*, pp. 2861–2866, IEEE, 2023.
- [28] D. A. Kessler and H. Levine, “Phenomenological approach to cancer cell persistence,” *Physical Review Letters*, vol. 129, no. 10, p. 108101, 2022.
- [29] G. Harmange, R. A. R. Hueros, D. L. Schaff, B. Emert, M. Saint-Antoine, L. C. Kim, Z. Niu, S. Nellore, M. E. Fane, G. M. Alicea, *et al.*, “Disrupting cellular memory to overcome drug resistance,” *Nature Communications*, vol. 14, no. 1, p. 7130, 2023.
- [30] T. Mora and A. M. Walczak, “Effect of phenotypic selection on stochastic gene expression,” *Journal of Physical Chemistry B*, vol. 117, no. 42, pp. 13194–13205, 2013.
- [31] D. Lunz, J. F. Bonnans, and J. Ruess, “Optimal control of bioproduction in the presence of population heterogeneity,” *Journal of Mathematical Biology*, vol. 86, no. 3, p. 43, 2023.
- [32] T. Kalisky, E. Dekel, and U. Alon, “Cost–benefit theory and optimal design of gene regulation functions,” *Physical Biology*, vol. 4, no. 4, p. 229, 2007.

- [33] J. Hausser, A. Mayo, L. Keren, and U. Alon, “Central dogma rates and the trade-off between precision and economy in gene expression,” *Nature Communications*, vol. 10, no. 1, p. 68, 2019.
- [34] G. Tkačik, C. G. Callan Jr, and W. Bialek, “Information flow and optimization in transcriptional regulation,” *Proceedings of the National Academy of Sciences*, vol. 105, no. 34, pp. 12265–12270, 2008.
- [35] G. Tkačik and A. M. Walczak, “Information transmission in genetic regulatory networks: a review,” *Journal of Physics: Condensed Matter*, vol. 23, no. 15, p. 153102, 2011.
- [36] A. Gyorgy, A. Menezes, and M. Arcak, “A blueprint for a synthetic genetic feedback optimizer,” *Nature Communications*, vol. 14, no. 1, p. 2554, 2023.
- [37] J. Jedrak and A. Ochab-Marcinek, “Influence of gene copy number on self-regulated gene expression,” *Journal of Theoretical Biology*, vol. 408, pp. 222–236, 2016.
- [38] M. A. Vághy, I. Otero-Muras, M. Pájaro, and G. Szederkényi, “A kinetic finite volume discretization of the multidimensional PIDE model for gene regulatory networks,” *Bulletin of Mathematical Biology*, vol. 86, no. 2, pp. 1–26, 2024.
- [39] P. Bokes and A. Singh, “Protein copy number distributions for a self-regulating gene in the presence of decoy binding sites,” *PLoS one*, vol. 10, p. e0120555, 2015.
- [40] P. Bokes, “Heavy-tailed distributions in a stochastic gene autoregulation model,” *Journal of Statistical Mechanics: Theory and Experiment*, vol. 2021, no. 11, p. 113403, 2021.
- [41] I. Zabaikina, P. Bokes, and A. Singh, “Joint distribution of protein concentration and cell volume coupled by feedback in dilution,” in *International Conference on Computational Methods in Systems Biology* (J. Pang and J. Niehren, eds.), vol. 14137, pp. 253–268, Springer, Cham, 2023.
- [42] Z. Zhang, I. Zabaikina, C. Nieto, Z. Vahdata, P. Bokes, and A. Singh, “Stochastic gene expression in proliferating cells: Differing noise intensity in single-cell and population perspectives,” *bioRxiv preprint 10.1101/2024.06.28.601263*, 2024.
- [43] M. Assaf and B. Meerson, “WKB theory of large deviations in stochastic populations,” *Journal of Physics A: Mathematical and Theoretical*, vol. 50, no. 26, p. 263001, 2017.
- [44] P. Bokes and A. Singh, “Controlling noisy expression through auto regulation of burst frequency and protein stability,” in *Hybrid Systems Biology. HSB 2019. Lecture Notes in Computer Science* (M. Češka and N. Paoletti, eds.), vol. 11705, Springer, Cham, 2019.
- [45] E. J. Doedel, “Lecture notes on numerical analysis of nonlinear equations,” in *Numerical Continuation Methods for dynamical systems*, pp. 1–49, Springer, 2007.

[46] E. J. Doedel and B. E. Oldeman, “AUTO-07P: continuation and bifurcation software for ordinary differential equations,” *Concordia University, Montreal*, 2007.