First-passage time approach to controlling noise in the timing of intracellular events

Khem Raj Ghusinga\textsuperscript{a}, John J. Dennehy\textsuperscript{b}, and Abhyudai Singh\textsuperscript{a,c,d,1}

\textsuperscript{a}Department of Electrical and Computer Engineering, University of Delaware, Newark, DE 19716; \textsuperscript{b}The Graduate Center, City University of New York, New York, NY 10016; \textsuperscript{c}Department of Biomedical Engineering, University of Delaware, Newark, DE 19716; and \textsuperscript{d}Department of Mathematical Sciences, University of Delaware, Newark, DE 19716

Edited by Charles S. Peskin, New York University, New York, NY, and approved November 29, 2016 (received for review June 4, 2016)

In the noisy cellular environment, gene products are subject to inherent random fluctuations in copy numbers over time. How cells ensure precision in the timing of key intracellular events despite such stochasticity is an intriguing fundamental problem. We formulate event timing as a first-passage time problem, where an event is triggered when the level of a protein crosses a critical threshold for the first time. Analytical calculations are performed for the first-passage time distribution in stochastic models of gene expression. Derivation of these formulas motivates an interesting question: Is there an optimal feedback strategy to regulate the synthesis of a protein to ensure that an event will occur at a precise time, while minimizing deviations or noise about the mean? Counterintuitively, results show that for a stable long-lived protein, the optimal strategy is to express the protein at a constant rate without any feedback regulation, and any form of feedback (positive, negative, or any combination of them) will always amplify noise in event timing. In contrast, a positive feedback mechanism provides the highest precision in timing for an unstable protein. These theoretical results explain recent experimental observations of single-cell lysis times in bacteriophage $\lambda$. Here, lysis of an infected bacterial cell is orchestrated by the expression and accumulation of a stable $\lambda$ protein up to a threshold, and precision in timing is achieved via feedforward rather than feedback control. Our results have broad implications for diverse cellular processes that rely on precise temporal triggering of events.

first-passage time | event timing | stochastic gene expression | feedback control | single cell

Timing of events in many cellular processes, such as cell-cycle control (1–3), cell differentiation (4, 5), sporulation (6, 7), apoptosis (8, 9), development (10, 11), temporal order of gene expression (12–14), and so on, depend on regulatory proteins reaching critical threshold levels. Triggering of these events in single cells is influenced by fluctuations in protein levels that arise naturally due to noise in gene expression (15–18). Increasing evidence shows considerable cell-to-cell variation in timing of intracellular events among isogenic cells (19–21), and it is unclear how noisy expression generates this variation. Characterization of control strategies that buffer stochasticity in event timing are critically needed to understand reliable functioning of diverse intracellular pathways that rely on precision in timing.

Mathematically, noise in the timing of events can be investigated via the first-passage time (FPT) framework, where an event is triggered when a stochastic process (single-cell protein level) crosses a critical threshold for the first time. There is already a rich tradition of using such FPT approaches to study timing of events in biological and physical sciences (22–26). Following this tradition, exact analytical expression for the FPT distribution is computed in experimentally validated and commonly used stochastic models of gene expression. These results provide insights into how expression parameters shape statistical fluctuations in event timing.

To investigate control mechanisms for buffering noise in timing, we consider feedback regulation in protein synthesis, where the transcription rate varies arbitrarily with the protein count. Such feedback can be implemented directly through autoregulation of gene promoter activity by its own protein (27–29) or indirectly via intermediate states (30). It is important to point out that although the effects of such feedback loops on fluctuations in protein copy number are well studied (29, 31–33), their impacts on stochasticity in event timing have been overlooked. We specifically formulate the problem of controlling precision in event timing as follows: What optimal form of feedback regulation ensures a given mean time to an event, while minimizing deviations or noise about the mean? It turns out that for a minimal model of stochastic gene expression, this optimization problem can be solved analytically, providing counterintuitive insights. For example, a negative feedback regulation is found to amplify noise in event timing and the optimal form of feedback is to not have any feedback at all. The robustness of these results is analyzed in the context of different noise mechanisms, such as intrinsic versus extrinsic noise in transcription/translation machinery (34–37). Finally, we discuss in detail how our results explain recent experimental observations of single-cell lysis times in bacteriophage $\lambda$, where precision in timing is obtained without any feedback regulation.

Stochastic Model Formulation

Consider a gene that is switched on at time $t = 0$ and begins to express a timekeeper protein. The intracellular event of interest is triggered once the protein reaches a critical level in the cell. We describe a minimal model of gene expression that assumes transcription in bursts and incorporates feedback regulation by considering the transcription rate as a function of the protein level.

Significance

Understanding how randomness in the timing of intracellular events is buffered has important consequences for diverse cellular processes, where precision is required for proper functioning. To investigate event timing in noisy biochemical systems, we develop a first-passage time framework in which an event is triggered when a regulatory protein accumulates up to a critical level. Formulas quantifying event-timing fluctuations in stochastic models of protein synthesis with feedback regulation are developed. Formulas shed counterintuitive insights into regulatory mechanisms essential for scheduling an event at a precise time with minimal fluctuations. These results uncover various features in the biochemical pathways used by phages to lyse individually infected bacterial cells at an optimal time, despite stochastic expression of lysis proteins.

Author contributions: J.J.D. and A.S. designed research; K.R.G. and A.S. performed research; K.R.G. analyzed data; and K.R.G. and A.S. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

1To whom correspondence should be addressed. Email: absingh@UDel.Edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1609012114/-/DCSupplemental.
(Fig. 1). More precisely, if \( x(t) \in \{0, 1, \ldots \} \) denotes the level of a protein in a single cell at time \( t \), then the gene is transcribed at a Poisson rate \( k_i \) when \( x(t) = i \). Any arbitrary form of feedback can be realized by appropriately defining \( k_i \) as a function of \( i \). For example, increasing (decreasing) \( k_i \)'s correspond to a positive (negative) feedback loop in protein production, and a fixed transcription rate implies no feedback. The translation burst approximation is based on assuming short-lived mRNAs, that is, each mRNA degrades with a constant rate \( \gamma \). The time evolution of \( x(t) \) is described through the following probabilities of occurrences of burst and decay events in the next infinitesimal time \( (t, t + dt) \):

\[
\mathbb{P}(x(t + dt) = i + 1 | x(t) = i) = k_i dt, \quad [2a]
\]

\[
\mathbb{P}(x(t + dt) = i - 1 | x(t) = i) = \gamma dt. \quad [2b]
\]

Note that in this representation of gene expression as a bursty birth–death process, the mRNA transcription rate \( k_i \) is the burst arrival rate, whereas the rate at which proteins are translated from an mRNA determines the mean protein burst size \( b \). Next, we formulate event timing through the FPT framework.

**Computing Event Timing Distribution**

The time to an event is the FPT for \( x(t) \) to reach a threshold \( X \) starting from a zero initial condition \( x(0) = 0 \) (Fig. 1). It is mathematically described by the following random variable:

\[
T := \min\{t : x(t) \geq X | x(0) = 0\}, \quad [3]
\]

and can be interpreted as the time taken by a random walker to first reach a defined point. For the bursty birth–death process in Eq. 2, the probability density function (pdf) of the FPT is given by

\[
f_T(t) = \sum_{i=0}^{X-1} k_i \left( \frac{b}{b+1} \right)^i \left( 1 - \frac{b}{b+1} \right)^{X-i} p_i(t), \quad [4]
\]

where \( p_i(t) = \mathbb{P}(x(t) = i) \) (SI Appendix, section S1).

The FPT pdf in Eq. 4 can be compactly written as product of two vectors:

\[
f_T(t) = U^T \mathbf{P}(t), \quad [5a]
\]

\[
U = \begin{bmatrix} k_0 \left( \frac{b}{b+1} \right)^X k_1 \left( \frac{b}{b+1} \right)^{X-1} \cdots k_{X-1} \left( \frac{b}{b+1} \right) \end{bmatrix}^T, \quad [5b]
\]

\[
\mathbf{P}(t) = \begin{bmatrix} p_0(t) & p_1(t) & \cdots & p_{X-1}(t) \end{bmatrix}^T. \quad [5c]
\]

Here, the dynamics of \( \mathbf{P}(t) \) can be written as a linear system

\[
\frac{d}{dt} \mathbf{P}(t) = \mathbf{A} \mathbf{P}(t) = \mathbf{U}^T \mathbf{P}(t), \quad [6]
\]

where \( \mathbf{U}(0) = 1 \cdots 0 \) is vector of probabilities at \( t = 0 \) that follows from \( x(0) = 0 \). Although this pdf provides complete characterization of the event timing, we are particularly interested in the lower-order statistical moments of FPT. Next, we exploit the structure of matrix \( \mathbf{A} \) to obtain analytical formulas for the first-and second-order moments of FPT.

**Moments of the FPT**

From Eq. 8, the \( m \)-th order centered moment of the FPT is given by

\[
\langle T^m \rangle = \mathbf{U}^T \left( \int_0^\infty e^{\mathbf{A}t} dt \right) \mathbf{P}(0) = (-1)^{m+1} m! \mathbf{U}^T (\mathbf{A}^{-1})^{m+1} \mathbf{P}(0). \quad [9b]
\]

Here, in computing the above integral, we used the fact that the matrix \( \mathbf{A} \) is full-rank with negative eigenvalues (SI Appendix, section S2). One can also find explicit formulas for the first two moments as series summations in terms of event threshold \( X \), mean burst size \( b \), protein decay rate \( \gamma \), and transcription rates \( k_i, 0 \leq i \leq X - 1 \) (SI Appendix, section S3).

Analysis of the FPT moments in some limiting cases gives important insights (SI Appendix, section S4). For the simplest case of a stable protein (\( \gamma = 0 \)) and a constant transcription rate (no feedback; \( k_i = k \)), the moment expressions simplify to

\[
\langle T \rangle = \frac{1}{k} \left( \frac{X}{b} + 1 \right) \approx \frac{X}{bk}, \quad CV_T^2 = \frac{b^2 + X + 2bX}{(b+X)^2} \approx \frac{1 + 2b}{X}, \quad [10]
\]

where \( CV_T^2 \) represents the noise in FPT as quantified by its coefficient of variation squared (variance/mean^2; \( T^2 / \langle T \rangle^2 - 1 \)). The approximate formulas in Eq. 10 are valid for a high event
threshold compared with the mean protein burst size \( (X/b \gg 1) \). The mean FPT formula can be interpreted as the time taken to reach \( X \) with an accumulation rate \( bk \). Further, \( X/b \) represents the average number of burst events required for the protein level to cross the threshold, and increasing \( X/b \) leads to noise reduction through more efficient averaging of the bursty process. One can also gain important insights, such as, the noise in FPT is invariant of the transcription rate \( k \). Therefore, \( \langle T \rangle \) and \( CV_2 \) can be independently tuned—increasing the event threshold and/or reducing the burst size will lower the noise level. Once \( CV_2 \) is sufficiently reduced, \( k \) can be altered to obtain a desired mean event timing.

Interesting features of the FPT statistics are revealed when \( \gamma \neq 0 \) (unstable protein) is considered with a constant transcription rate \( k = k \). In this case, the expressions of FPT moments are quite involved (SI Appendix, section S4), and we investigate the effect of various parameters numerically. It turns out that changing the event threshold leads to a U-shaped profile for \( CV_2 \), where noise in event timing first decreases and then increases (Fig. 2). Intuitively, when \( \gamma \neq 0 \), the protein level approaches a steady-state level \( x_{ss} = b \gamma/k \). When the event threshold \( X \) is sufficiently below \( x_{ss} \), \( CV_2 \) reduces with increasing \( X \), similar to the \( 0 \) case. As \( X \) approaches \( x_{ss} \), the protein trajectories start saturating and crossing the threshold becomes a noise-driven event. This results in an increase in \( CV_2 \) and ultimately leads to \( CV_2 \rightarrow 1 \) as \( X \gg x_{ss} \). Recall that the coefficient of variation of an exponentially distributed random variable is exactly equal to one. Thus, when \( X \) is much larger than \( x_{ss} \), the timing process becomes memoryless, yielding exponentially distributed FPTs. The minimum value of \( CV_2 \) is achieved at an intermediate threshold level \( X \approx x_{ss}/2 \) for a birth–death process \( (b \to 0) \), and the dip in the U-shape shifts to the right as the protein expression becomes more bursty (Fig. 2). Another interesting point to note is that whereas \( b \) increases \( b \) in event timing when \( X \ll x_{ss} \), it has a contrasting effect when \( X \gg x_{ss} \), where increasing \( b \) can sometimes reduce \( CV_2 \). Next, we explore how feedback regulation of the transcription rate affects noise in timing, for a given \( X \) and \( b \).

**Optimal Feedback Strategy**

Having derived the FPT moments, we investigate optimal forms of transcriptional feedback that schedule an event at a given time with the lowest \( CV_2 \). Because \( \langle T \rangle \) is assumed to be fixed, minimizing \( CV_2 \) is equivalent to minimizing \( \langle T^2 \rangle \). Thus, the problem mathematically corresponds to a constraint optimization problem: Find transcription rates \( k_0, k_1, \ldots, k_{X-1} \) that minimize \( \langle T^2 \rangle \) for a fixed \( \langle T \rangle \). We first consider a stable protein whose half-life is much longer than the event timescale, and, hence, degradation can be ignored \( (\gamma = 0) \).

**Optimal Feedback for a Stable Protein.** When the protein of interest does not decay \( (\gamma = 0) \), the expressions for the FPT moments take much simpler forms:

\[
\langle T \rangle = \frac{1}{k_0} + \frac{1}{b} \sum_{i=0}^{X-1} \frac{1}{k_i},
\]

\[
\langle T^2 \rangle = \frac{2}{b^2} \left( \frac{\tau_0}{bk_0} + \sum_{i=0}^{X-1} \frac{\tau_i}{k_i} \right), \quad \tau_i := \frac{b}{k_i} + \sum_{j=i+1}^{X-1} \frac{1}{k_j}.
\]

Note that, in Eq. 11a, the contribution of \( k_0 \) (transcription rate when there is no protein) differs from the other transcription rates \( k_i, i \in \{1, 2, \ldots, X-1\} \). For instance, when the event threshold is large compared with the mean burst size \( (X/b) \), then the term \( 1/k_0 \) can be ignored and \( \langle T \rangle \approx \sum_{i=1}^{X-1} \frac{1}{bk_i} \). In contrast, if the burst size is large \( (b \gg X) \) then \( \langle T^2 \rangle \approx 1/k_0 \), because a single burst event starting from zero protein molecules is sufficient for threshold crossing. A similar observation for different contributions of \( k_i \) can be made about Eq. 11b.

It turns out that, for these simplified formulas, the problem of minimizing \( \langle T^2 \rangle \) given \( \langle T \rangle \) can be solved analytically using the method of Lagrange multipliers (SI Appendix, section S5). The optimal transcription rates are given by

\[
k_0 = \frac{1 + b}{1 + 2b} \frac{2b + X}{\langle T \rangle}, \quad k_i = \frac{1 + 2b}{1 + b} k_0, \quad 1 \leq i \leq X - 1,
\]

and all rates are equal to each other except for \( k_0 \). Intuitively, the difference for \( k_0 \) comes from the fact that it contributes differently to the FPT moments compared with other rates. Note that for a small mean burst size \( (b \ll 1) \), \( k_0 = \frac{b}{X} \), whereas \( k_i = k_0/2 \) for a sufficiently large \( b \). Despite this slight deviation in \( k_0 \), for the purposes of practical implementation, the optimal feedback strategy in this case is to have a constant transcription rate (i.e., no feedback in protein expression).

We tested the above result for a more complex stochastic gene expression model that explicitly includes mRNA dynamics via Monte Carlo simulations (Fig. 3). For ease of implementation, the feedbacks are assumed to be linear:

\[
k_i = c_1 + c_2 i, \quad i \in \{0, 1, \ldots, \},
\]

where \( c_2 = 0 \) represents no feedback, and \( c_2 > 0 (c_2 < 0) \) denotes a positive (negative) feedback. In agreement with Eq. 12, a no-feedback strategy outperforms negative/positive feedbacks in terms of minimizing noise in FPT around a given mean event time. The qualitative shape of trajectories in Fig. 3 is determined by the feedback strategy used, with no feedback resulting in linear time evolution of protein levels. This provides an intriguing geometric interpretation of our results—an approximate linear path from zero protein molecules at \( t = 0 \) to \( X \) molecules at time \( \langle T \rangle \) provides the highest precision in timing. Next, we discuss the optimal feedback strategy when protein degradation is taken into consideration.

**Optimal Feedback for an Unstable Protein.** Now consider the scenario where protein degradation cannot be ignored over the
event timescale ($\gamma \neq 0$). Unfortunately, the expressions of the FPT moments are too convoluted for the optimization problem to be solved analytically (SI Appendix, section S3, Eq. S3.10), and the effect of different feedbacks is investigated numerically.

We implement the feedbacks using physiologically relevant Hill functions, where the transcription rates for a negative feedback mechanism take the following form:

$$k_i = \frac{k_{\text{max}}}{1 + (c_i)^H}, \quad i \in \{0, 1, \ldots\}. \quad [14]$$

Here $H$ denotes the Hill coefficient, $k_{\text{max}}$ corresponds to the maximum transcription rate, and $c_i$ characterizes the negative feedback strength, with $c_i = 0$ representing no feedback (29, 43). Similarly, a positive feedback is assumed to take the following form:

$$k_i = k_{\text{max}} \left( r + (1 - r) \frac{(c_i)^H}{1 + (c_i)^H} \right) = k_{\text{max}} \frac{r + (c_i)^H}{1 + (c_i)^H}. \quad [15]$$

Note that an additional parameter $r \in (0, 1)$, referred to as the basal strength, is introduced in Eq. 15. This is to ensure that the transcription rate in protein absence $k_0 = k_{\text{max}}r > 0$, and this is necessary to prevent protein levels from getting stuck at zero molecules.

To find the optimal feedback mechanism, our strategy is as follows: For given $r$ and $H$, choose a certain feedback strength $c_i$ in Eq. 14/15, appropriately tune $k_{\text{max}}$ for the desired mean event timing, and explore the corresponding noise in FPT as measured by its coefficient of variation squared $CV_T^2$. Counter-intuitively, results show that for a given value of $\gamma$, a negative feedback loop in gene expression has the highest $CV_T^2$, and its performance deteriorates with increasing feedback strength (Fig. 4, Top). In contrast, $CV_T^2$ first decreases with increasing strength of the positive feedback and then increases after an optimal feedback strength is crossed (Fig. 4, Top). Thus, when the protein is not stable, precision in timing is attained by having a positive feedback in protein synthesis with an intermediate strength.

Next we explore how the minimal achievable noise in event timing, for a fixed $\langle T \rangle$, varies with the protein decay rate $\gamma$. Our analysis shows that for a given basal strength $r$, the minimum $CV_T^2$ obtained via positive feedback increases monotonically with $\gamma$, and $CV_T^2 \to 1$ as $\gamma$ becomes large (Fig. 4, Bottom). A couple of interesting observations can be made from Fig. 4, Bottom: (i) The difference in $CV_T^2$ for optimal feedback and no feedback is indistinguishable when the protein is stable ($\gamma = 0$) or highly unstable ($\gamma \to \infty$); (ii) for a range of intermediate protein half-lives the optimal feedback strategy provides better reduction of $CV_T^2$, as compared with no feedback regulation (which also corresponds to minimum $CV_T^2$ obtained via a negative feedback); (iii) lowering the basal strength $r$ results in better performance in terms of noise suppression; and (iv) a linear feedback based on Hill functions and provides significantly lower levels of $CV_T^2$ for high protein decay rates. It is also worth pointing out that the qualitative shape of curves in Fig. 4, Bottom does not change for different values of event threshold $X$ or mean burst size $b$ (SI Appendix, section S6).

Why is positive feedback the optimal control strategy for ensuring precision in event timing? One way to understand this result is to consider the linear feedback form Eq. 13, in which case the mean protein levels evolve according to the following ordinary differential equation:

$$\frac{dx(t)}{dt} = b(c_1 + c_2 x) - \gamma x, \quad x(0) = 0. \quad [16]$$

Recall the geometric argument presented in Fig. 3, where an approximately linear path for the protein to reach the prescribed threshold in a given time provides the highest precision in event timing. Whereas no feedback ($c_2 = 0$) and negative feedback ($c_2 < 0$) in Eq. 16 will create nonlinear protein trajectories, choosing a positive value $c_2 \approx \gamma/b$ results in linear $x(t)$, and hence minimal noise in event timing. Indeed, our detailed stochastic analysis shows that the optimal feedback strength that minimizes $CV_T^2$ in the stochastic model is qualitatively similar to $c_2 \approx \gamma/b$ (SI Appendix, section S6).

Discussion

We have systematically investigated ingredients essential for precision in timing of biochemical events at the level of single cells. Our approach relies on modeling event timing as the FPT for a stochastically expressed protein to cross a threshold level. This framework was used to uncover optimal strategies for synthesizing the protein that ensures a given mean time to event triggering (threshold crossing) with minimal fluctuations around the mean. The main contributions and insights can be summarized as follows: (i) analytical calculations for the FPT in stochastic models of gene expression, with and without feedback regulation are performed; (ii) if the protein half-life is much longer than the timescale of the event, the highest precision in event
Another important model aspect is geometrically distributed protein burst size, which results from the assumption of exponentially distributed mRNA lifetimes. We have also explored the scenario of perfect memory in the mRNA degradation process, which results in an mRNA lifetime distribution given by the delta function. In this case, the protein burst size is Poisson and the optimal feedback strategy is fairly close to having no feedback for a stable protein (SI Appendix, section S8). Next, we discuss the biological implications of our findings in the context of phage λ’s lysis times (i.e., the time taken by the virus to destroy infected bacterial cells).

**Connecting Theoretical Insights to λ Lysis Times.** Phage λ has recently emerged as a simple model system for studying event timing at the level of single cells (19, 20). After infecting *Escherichia coli*, λ expresses a protein, holin, which accumulates in the inner membrane. When holin reaches a critical threshold concentration, it undergoes a structural transformation, forming holes in the membrane (45). Subsequently the cell lysis and phage progeny are released into the surrounding medium.

Because hole formation and cell rupture are nearly simultaneous, lysis timing depends on de novo expression and accumulation of holin in the cell membrane up to a critical threshold (45). Data reveal precision in the timing of lysis—individual cells infected by a single virus lyse on average at 65 min, with an SD of 3.5 min, implying a coefficient of variation of ≈5% (SI Appendix, section S9). Such precision is expected given the existence of an optimal lysis time (46–49). Intuitively, if λ lysis is early then there are no viral progeny. In contrast, if λ lysis is late then the infected cell could die before lysis is effected, trapping the virus with it.

The threshold for lysis is reported to be a few thousand holin molecules (50). Moreover, the holin mean burst size (average number of holins produced in a single mRNA lifetime) is estimated as $b \approx 1–3$ (50). Based on our FPT moment calculations in Eq. 10, such a small protein burst size relative to the event threshold will yield a tight distribution of lysis times. Interestingly, Eq. 10 provides insights for engineering mutant λ that lyse, on average, at the same time as the wild type, but with much higher noise. This could be done by lowering the threshold $X$ for lysis through mutations in the holin amino acid sequence (20), and also reducing the holin mRNA transcription rate $k$ or mean protein burst size $b$ by reducing the translation rate so as to keep the same mean lysis time. Notably, the holin proteins are long-lived and do not degrade over relevant timescales (51); therefore, λ’s lysis system with no known feedback in holin expression provides better suppression of lysis-time fluctuation compared with any feedback regulated system.

**Additional Mechanism for Noise Buffering.** The surprising ineffectiveness of feedback control motivates the need for other mechanisms to buffer noise in event timing. Intriguingly, λ uses feedforward control to regulate the timing of lysis that is implemented through two proteins with opposing functions: holin and antiholin (52, 53). In the wild-type virus both proteins are expressed in a 2:1 ratio (for every two holins there is one anti-holin) from the same mRNA through a dual start motif. Anti-holin binds to holin and prevents holin from participating in hole formation, creating an incoherent feedforward circuit. Synthesis of antiholin leads to a lower burst size for active holin molecules and increases the threshold for the total number of holins needed for lysis—both factors functioning to lower the noise in event timing. Consistent with this prediction, variants of λ lacking antiholin are experimentally observed to exhibit much higher intercellular variation in lysis times compared with the wild-type virus (20, 54) (SI Appendix, section S9). Succinctly put, λ encodes several regulatory mechanisms (low holin burst size,

![Fig. 4. For an unstable protein, positive feedback provides the lowest noise in event timing for a fixed mean FPT. (Top) Noise in timing ($CV^2$) as a function of the feedback strength $c$ for different control strategies. The value of $k_{max}$ is changed in Eq. 14/15 so as to keep $(T) = 40$ mins fixed. The performance of the negative feedback worsens with increasing feedback strength. In contrast, positive feedback with an optimal value of $c$ provides the highest precision in event timing. Other parameters used are $\gamma = 0.05 \text{ min}^{-1}$, $X = 500 \text{ molecules}$, $H = 1$, $b = 2$, and for positive feedback $r = 0.05$. (Bottom) The minimum value of $CV^2$ obtained via positive feedback increases monotonically with the protein degradation rate. A smaller basal promoter strength $r = 0.01$ in Eq. 15 gives better noise suppression than a larger value $r = 0.05$. For comparison purposes, $CV^2$ obtained without any feedback ($c = 0$), and a linear feedback with $c_1$ and $c_2$ in Eq. 13 chosen to minimize $CV^2$ for a given $(T) = 40$ mins are also shown. The parameter values used are $X = 500 \text{ molecules}$, $H = 1$, and $b = 2$.](image-url)
no feedback regulation, and feedforward control) to ensure that single infected cells lyse at an optimal time, despite the stochastic expression of lysis proteins.

These results illustrate the utility of the FPT framework for characterizing noise in the timing of intracellular events and motivate alternate formulations of the timing problem that might involve additional constraints such as fixing the cost of protein production (SI Appendix, section S10). Exploring these constraints in more detail will be an important avenue for future research. Finally, analytical results and insights obtained here have broader implications for timing phenomenon in chemical kinetics, ecological modeling, and statistical physics.
Supporting Information for "A first–passage time approach to controlling noise in the timing of intracellular events"

Khem Raj Ghusinga, John J. Dennehy, Abhyudai Singh 10.1073/pnas.1609012114

S1. Computation of first-passage time distribution

Eq. (2) of the main text describes the time evolution of protein count \( x(t) \) by the following bursty birth-death process

\[
P(x(t+dt) = i + B | x(t) = i) = k_i dt, \]
\[
P(x(t+dt) = i - 1 | x(t) = i) = i \gamma dt,
\]

where \( B \) denotes the protein burst size which is assumed to follow a geometric distribution with mean \( b \). We are interested in computing the first-passage (FPT) described by the following random variable

\[
T := \min\{ t : x(t) \geq X | x(0) = 0 \}.
\]

Fig. S1. Illustration of a bursty birth-death process for computing the first-passage time. States \([0, 1, \ldots, X]\) represent the protein population counts, and arrows represent transition between states due to burst and decay events. The destination of a forward jump (a birth event) is decided by the burst size while each degradation event reduces the protein count by one. The process terminates when the protein level reaches the absorbing-state \( X \) and the first-passage time is recorded.

One can note that if the protein does not decay, then \( x(t) \) accumulates over time and the FPT distribution could be obtained by observing

\[
P(x(t) \geq X) = P(T \leq t).
\]

However, with protein degradation, the FPT calculation needs careful consideration so as to avoid counting multiple crossings of the threshold. To this end, we construct an equivalent bursty birth-death process on a finite state-space \([0, 1, \ldots, X]\), where the states represent the protein count (Fig. S1). All states denoting \( x(t) \geq X \) are combined into a single absorbing state \( X \). In this model, the probability of the protein level reaching \( X \) in the small time window \((t, t+dt)\) is the probability that \( x(t) = i \) and a jump of size \( X - i \) or larger occurs in \((t, t+dt)\). That is, we have

\[
f_T(t) dt = P(T \in (t, t+dt))
\]
\[
= \sum_{i=0}^{X-1} k_i P(B \geq X - i) P(x(t) = i) dt,
\]

where \( f_T(t) \) is the first-passage time probability density function. Note that Eq. (S1.4b) is valid for any distribution of the burst size \( B \). When the burst size \( B \) is geometrically distributed, we can use the following relation in Eq. (S1.4b)

\[
P(B \geq X - i) = \left( \frac{b}{b+1} \right)^{X-i}.
\]
Defining \( p_i(t) := P(x(t) = i) \), the FPT pdf can be written as product of two vectors
\[
    f_T(t) = U^\top P(t),
\]
\[
    U = \begin{bmatrix}
        k_0 \left( \frac{b}{b+t} \right)^X \\
        k_1 \left( \frac{b}{b+t} \right)^{X-1} \\
        \vdots \\
        k_{X-1} \left( \frac{b}{b+t} \right)
    \end{bmatrix},
\]
\[
    P(t) = \begin{bmatrix}
        p_0(t) & p_1(t) & \cdots & p_{X-1}(t)
    \end{bmatrix}^\top.
\]  

To obtain \( P(t) \), we write the Chemical Master Equation (CME) corresponding to the equivalent bursty-birth death process in Fig. S1
\[
    p_0(t) = -k_0 \frac{b}{b+t} p_0(t) + \gamma p_1(t),
\]
\[
    p_i(t) = - \left( k_i \frac{b}{b+1} + i \gamma \right) p_i(t) + (i+1) \gamma p_{i+1}(t) + \sum_{n=0}^{i-1} k_n \frac{b^{i-n}}{(b+1)^{n+1}} p_n(t), \quad 1 \leq i \leq X - 2,
\]
\[
    p_{X-1}(t) = -(k_{X-1} \frac{b}{b+1} + (X-1) \gamma) p_{X-1}(t) + \sum_{n=0}^{X-2} k_n \frac{b^{X-n}}{(b+1)^{X-n+1}} p_n(t).
\]  

The CME can be conveniently written as the following linear dynamical system
\[
    \dot{P}(t) = AP(t),
\]
where an \( i^{th} \) row and \( j^{th} \) column element \( a_{ij} \) of the matrix \( A \) is given by
\[
    a_{ij} = \begin{cases} 
        0, & j > i + 1 \\
        (i - 1) \gamma, & j = i + 1 \\
        (-k_i \frac{b}{b+1} - (i - 1) \gamma), & j = i \\
        k_i \frac{b^{i-j}}{(b+1)^{i-j}}, & j < i 
    \end{cases},
\]  
i, j \in \{1, \ldots, X\}. Solving Eq. (S1.8a) and using Eq. (S1.6a) yields the following pdf for the first-passage time
\[
    f_T(t) = U^\top P(t) = U^\top \exp(At) P(0),
\]  

where \( P(0) = \begin{bmatrix} 1 & 0 & \cdots & 0 \end{bmatrix}^\top \) is vector of probabilities at \( t = 0 \) that follows from \( x(0) = 0 \).  

**S2. On some properties of the matrix \( A \)**

In this section, we discuss some properties of the matrix \( A \) given in Eq. (S1.8b). One can note that \( A \) is an “almost” lower triangular matrix: all entries above its first superdiagonal are zero. This kind of matrices are known as lower Hessenberg matrices. Additionally, \( A \) satisfies two important properties that it is a Hurwitz matrix, and its inverse can be explicitly obtained. We’ll see in the next section how these properties allow one to compute expressions of FPT moments. Below we give discuss these two properties in detail.

**A is a Hurwitz matrix.** The matrix \( A \) is given by
\[
    A = \begin{bmatrix}
        -\frac{bk_0}{k_0} & -\left( \frac{b}{b+1} + \gamma \right) & \cdots & 0 \\
        \frac{b^{X-2}k_0}{(b+1)^{X-2}} & \frac{b^{X-3}k_0}{(b+1)^{X-3}} & \cdots & (X-1) \gamma \\
        \vdots & \vdots & \ddots & \vdots \\
        \frac{b^X}{(b+1)^X} & \frac{b^{X-1}k_0}{(b+1)^{X-1}} & \cdots & \left( \frac{b}{b+1} + (X-1) \gamma \right)
    \end{bmatrix}.
\]  

In order to prove that \( A \) is a Hurwitz matrix, we show that the following two conditions hold true [1, pp. 48–49]:

1. The diagonal elements \( a_{ii} < 0 \) for \( i = 1, 2, \cdots, X \),
2. \( \max_{1 \leq j \leq X} \sum_{i=1}^{X} \left| \frac{a_{ij}}{a_{jj}} \right| < 1 \).
Therefore the inverse of the matrix

$$\sum_{j=1}^{X} a_{ij} = \frac{(j-1)\gamma}{bk_{j+1} + (j-1)\gamma} + \frac{bk_{j-1}}{b+1} \sum_{i=j+1}^{X} \frac{(b_{i/j})^{i-j}}{bk_{i/j} + (j-1)\gamma} \quad \text{[S2.2a]}$$

$$= \frac{(j-1)\gamma}{bk_{j+1} + (j-1)\gamma} + \frac{bk_{j-1}}{b+1} \left(1 - \left(\frac{b_{i/j}}{b+1}\right)^{X-j}\right) \quad \text{[S2.2b]}$$

$$= \frac{bk_{j-1}}{b+1} \left(1 - \left(\frac{b_{i/j}}{b+1}\right)^{X-j}\right) + (j-1)\gamma$$

$$< 1. \quad \text{[S2.2c]}$$

Thus, the matrix $A$ is Hurwitz, i.e., the eigenvalues of $A$ have negative real part.

**Determining inverse of the matrix $A$.** Let us use $A_0$ to denote the matrix $A$ when $\gamma = 0$. The lower triangular matrix $A_0$ is given by

$$A_0 = \begin{bmatrix}
-bk_0 & 0 & \cdots & 0 \\
-bk_1 & -bk_2 & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
-bk_{X-1}k_0 & -bk_{X-1}k_1 & \cdots & -bk_{X-1}
\end{bmatrix} \quad \text{[S2.3]}
$$

We claim that the inverse of $A_0$ is given by the following matrix

$$A_0^{-1} = -\frac{1}{b} \begin{bmatrix}
\frac{b+1}{k_0} & 0 & \cdots & 0 \\
\frac{b+1}{k_1} & \frac{b+1}{k_2} & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
\frac{b+1}{k_{X-1}} & \frac{b+1}{k_{X-2}} & \cdots & \frac{b+1}{k_{X-1}}
\end{bmatrix} \quad \text{[S2.4]}
$$

This claim can be quickly verified by multiplying the matrices which results in identity matrix. Next, to determine $A^{-1}$, we observe that when $\gamma \neq 0$, the matrix $A$ can be written as

$$A = A_0 + A_e, \quad \text{[S2.5]}$$

where $A_e$ is given by

$$A_e = \begin{bmatrix}
0 & \gamma & \cdots & 0 & 0 \\
0 & -\gamma & \cdots & 0 & 0 \\
0 & 0 & \ddots & 0 & 0 \\
\vdots & \vdots & \ddots & \ddots & \vdots \\
0 & 0 & \cdots & -(X-2)\gamma & (X-1)\gamma \\
0 & 0 & \cdots & 0 & -(X-1)\gamma
\end{bmatrix}. \quad \text{[S2.6]}
$$

Therefore the inverse of the matrix $A$ can be written as

$$A^{-1} = (A_0 + A_e)^{-1} = (I + A_0^{-1} A_e)^{-1} A_0^{-1}. \quad \text{[S2.7]}$$
As we have already determined the expression of $\mathbf{A}_0^{-1}$ in Eq. (S2.4), the expression of $\mathbf{A}_0^{-1}\mathbf{A}_e$ can be determined as

$$ -\frac{1}{b} \begin{bmatrix} \frac{b+1}{k_0} & 0 & \cdots & 0 & 0 \\ \frac{b+1}{k_1} & \frac{b+1}{k_1} & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \frac{1}{k_{X-2}} & \frac{1}{k_{X-2}} & \cdots & \frac{b+1}{k_{X-2}} & 0 \\ \frac{1}{k_{X-1}} & \frac{1}{k_{X-1}} & \cdots & \frac{b+1}{k_{X-1}} & \frac{b+1}{k_{X-1}} \end{bmatrix} \begin{bmatrix} 0 & \gamma & \cdots & 0 & 0 \\ 0 & -\gamma & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & -(X-2)\gamma & (X-1)\gamma \\ 0 & 0 & \cdots & 0 & -(X-1)\gamma \end{bmatrix} $$

[S2.8a]

Thus, the matrix $\mathbf{I} + \mathbf{A}_0^{-1}\mathbf{A}_e$ is a bidiagonal matrix with its diagonal elements $1 + \frac{(i-1)\gamma}{k_{i-1}} = \frac{k_{i-1} + (i-1)\gamma}{k_{i-1}}$ for $i = 1, 2, \cdots, X$. The super diagonal elements are given by $-\frac{i\gamma}{k_{i-1}}$ for $j = 1, 2, \cdots, X - 1$. Using the result for inverse of a bidiagonal matrix [2], we can write an element at the $i^{th}$ row and the $j^{th}$ column of inverse of $\mathbf{E} := \mathbf{I} + \mathbf{A}_0^{-1}\mathbf{A}_e$ as follows

$$ e_{i,j} = \begin{cases} 0, & \text{if } i > j, \\ \frac{k_{i-1}}{k_i + \gamma} \prod_{l=i}^{j-1} \frac{k_l}{k_{l+1} + \gamma}, & \text{if } i = j, \\ \frac{k_{i-1} + (i-1)\gamma}{k_{i-1}}, & \text{if } i < j. \end{cases} $$

[S2.9]

Alternatively, in matrix form

$$ \mathbf{E}^{-1} = \begin{bmatrix} 1 & \frac{k_1}{k_1 + \gamma} & \cdots & \frac{k_{X-1}}{k_{X-1} + \gamma} \\ \frac{k_1}{k_1 + \gamma} & \frac{k_1}{k_1 + \gamma} & \cdots & \frac{k_{X-1}}{k_{X-1} + \gamma} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \frac{k_1}{k_1 + \gamma} & \cdots & 1 \end{bmatrix} \cdot (\frac{k_1}{k_1 + \gamma} \prod_{l=2}^{X-1} (\frac{k_l}{k_{l+1} + \gamma})). $$

[S2.10]

We can compute $\mathbf{A}^{-1}$ by calculating $\mathbf{E}^{-1}\mathbf{A}_0^{-1}$. Here, we do not give explicit form of $\mathbf{A}^{-1}$ as it is not required for calculations in this document.

**S3. Expression of $m^{th}$ moment of first-passage time**

In this section, we make use of the properties discussed in the previous section to determine the moments of the first-passage time. As discussed earlier, the distribution of first-passage time (FPT) is given by Eq. (S1.9). Using this, a moment of $m^{th}$ order can be calculated as

$$ \langle T^m \rangle = \int_0^\infty \int_0^\infty t^m \mathbf{U}^T \exp(\mathbf{A}t) \mathbf{P}(0) dt $$

[S3.1a]

$$ = \mathbf{U}^T \left( \int_0^\infty t^m \exp(\mathbf{A}t) dt \right) \mathbf{P}(0). $$

[S3.1b]

Let us consider $\int_0^\infty t^m \exp(\mathbf{A}t) dt$. Integrating by parts

$$ \int_0^\infty t^m \exp(\mathbf{A}t) dt = t^m \mathbf{A}^{-1} \exp(\mathbf{A}t) \bigg|_0^\infty - m \mathbf{A}^{-1} \int_0^\infty t^{m-1} \exp(\mathbf{A}t) dt. $$

[S3.2]

As the matrix $\mathbf{A}$ is a Hurwitz matrix, i.e., eigenvalues of $\mathbf{A}$ have negative real parts, the first term in the above expression goes to zero. Using $a_m$ as a notion to represent $\int_0^\infty t^m \exp(\mathbf{A}t) dt$, we can write the following recursive relationship

$$ a_m = -m \mathbf{A}^{-1} a_{m-1}. $$

[S3.3]
Thus, \( a_m = (-1)^m m! (A^{-1})^m a_0 \). Further \( a_0 \) can be computed as

\[
a_0 = \int_0^\infty \exp(At) dt = A^{-1} \exp(A) \bigg|_0^\infty = -A^{-1}.
\]  

[S3.4]

Therefore \( a_m = (-1)^{m+1} m! (A^{-1})^{m+1} \). Substituting this in Eq. (S3.1b) gives the following for a general \( m^{th} \) moment of FPT

\[
(T^m) = (-1)^m m! U^T (A^{-1})^{m+1} P(0).
\]  

[S3.5]

Calculation of \( U^T A^{-1} \). As we saw in equation Eq. (S3.5), calculation of the moments have a term of the form \( U^T A^{-1} \). Here we provide the calculation of this term.

Consider two matrices \( G \) and \( H \) such that \( A = GH \) where \( G \) is a \( X \times X - 1 \) matrix

\[
G = \gamma \begin{bmatrix}
-1 & 0 & 0 & \cdots & 0 & 0 \\
1 & -2 & 0 & \cdots & 0 & 0 \\
0 & 2 & -3 & \cdots & 0 & 0 \\
0 & 0 & 3 & \cdots & 0 & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
0 & 0 & 0 & \cdots & X - 2 & -(X - 1) \\
0 & 0 & 0 & \cdots & 0 & X - 1
\end{bmatrix},
\]  

[S3.6a]

while \( H \) is a \( X - 1 \times X \) matrix

\[
H = \begin{bmatrix}
0 & -1 & 0 & \cdots & 0 & 0 \\
0 & 0 & -1 & \cdots & 0 & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
0 & 0 & 0 & \cdots & -1 & 0 \\
0 & 0 & 0 & \cdots & 0 & -1
\end{bmatrix}.
\]  

[S3.6b]

Using the matrix inversion lemma, \( A^{-1} \) can be written as

\[
A^{-1} = (A_0 + GH)^{-1},
\]  

[S3.7a]

\[
= A_0^{-1} - A_0^{-1} G (I + HA_0^{-1} G)^{-1} HA_0^{-1}.
\]  

[S3.7b]

The expression \( U^T A_0^{-1} G \) can be computed as below

\[
U^T A_0^{-1} G = -\begin{bmatrix}
-1 & 0 & \cdots & 0 \\
1 & -2 & \cdots & 0 \\
0 & 2 & \cdots & 0 \\
0 & 0 & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & -(X - 1) \\
0 & 0 & \cdots & X - 1
\end{bmatrix},
\]  

[S3.8a]

\[
= -\begin{bmatrix}
0 & 0 & \cdots & 0
\end{bmatrix}.
\]  

[S3.8b]

Therefore, we can conclude that \( U^T A^{-1} \) is in fact equal to \( U^T A_0^{-1} \) which could be calculated by multiplying \( U^T \) and \( A_0^{-1} \).
where the terms of 20 Khem Raj Ghusinga, John J. Dennehy, Abhyudai Singh 10.1073/pnas.1609012114 to these formulas.

\[
\langle T \rangle = \frac{1}{b} + \frac{1}{b} \sum_{i=1}^{X} \frac{k_{i-1}}{k_{i-1} + (i - 1)\gamma} \left( \frac{1}{k_{i-1}} + \sum_{i=1}^{X} \frac{1}{k_{i-1}} \prod_{l=i+1}^{i} k_{l} (b + 1)l\gamma^{l} \frac{b}{bk_{l-1}} \right),
\]

[S3.10a]

\[
\langle T^2 \rangle = \frac{2}{b^2} \sum_{i=1}^{X} \frac{k_{i-1}}{k_{i-1} + (i - 1)\gamma} \left( \frac{b\eta_{i} + \sum_{j=1}^{i} \eta_{j}}{k_{i-1}} + \sum_{i=1}^{X} \frac{b\eta_{i} + \sum_{j=1}^{i} \eta_{j}}{k_{i-1}} \prod_{l=i+1}^{i} k_{l} (b + 1)l\gamma \frac{b}{bk_{l-1}} \right),
\]

[S3.10b]

where the terms \( \eta_{i} \) are given by

\[
\eta_{i} = \frac{b}{b_{0}} \delta_{i-1} + \frac{k_{i-1}}{k_{i-1} + (i - 1)\gamma} \left( \frac{1}{k_{i-1}} + \sum_{i=1}^{X} \frac{1}{k_{i-1}} \prod_{l=i+1}^{i} k_{l} (b + 1)l\gamma \frac{b}{bk_{l-1}} \right),
\]

[S3.10c]

with \( \delta_{i-1} \) denoting the Kronecker delta which is one if \( i = 1 \) and zero otherwise. Below we show detailed calculations that lead to these formulas.

Mean FPT The mean FPT’s expression can be written as

\[
\langle T \rangle = U^T A^{-1} P(0) = U^T A_0^{-1} E^{-1} A_0^{-1} P(0).
\]

[S3.11]

The expression of \( A_0^{-1} P(0) \) is just the first column of \( A_0 \). Therefore

\[
E^{-1} A_0^{-1} P(0) = -\frac{1}{b} \begin{bmatrix}
1 & 0 & \cdots & \prod_{i=1}^{X-1} \left( \frac{k_{i+1}}{k_{i+1} + (b + 1)\gamma} \right) \\
0 & \frac{k_{1+1}}{k_{1+1} + \gamma} & \cdots & \prod_{i=2}^{X-1} \left( \frac{k_{i+1}}{k_{i+1} + (b + 1)\gamma} \right) \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & \frac{1}{k_{X-1} + (X-1)\gamma} \\
\end{bmatrix}
\]

[S3.12a]

\[
= -\frac{1}{b} \begin{bmatrix}
\frac{k_{1+1}}{k_{1+1} + \gamma} + \frac{k_{1+1}}{k_{2+1} + \gamma} \cdots \prod_{i=2}^{X-1} \left( \frac{k_{i+1}}{k_{i+1} + (b + 1)\gamma} \right) \\
0 & \frac{k_{2+1}}{k_{2+1} + \gamma} & \cdots & \prod_{i=3}^{X-1} \left( \frac{k_{i+1}}{k_{i+1} + (b + 1)\gamma} \right) \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & \frac{1}{k_{X-1} + (X-1)\gamma} \\
\end{bmatrix}
\]

[S3.12b]
Since the vector $\mathbf{U}^\top \mathbf{A}^{-1} = -[1 \ 1 \ 1 \ \cdots \ 1]$, $\mathbf{U}^\top \mathbf{A}^{-1} \mathbf{E}^{-1} \mathbf{A}^{-1} \mathbf{P}(0)$ is essentially negative sum of the elements of $\mathbf{E}^{-1} \mathbf{A}^{-1} \mathbf{P}(0)$. Therefore we have the expression of mean FPT is given by

$$
\langle T \rangle = \frac{1}{k_0} + \frac{1}{b} \sum_{i=1}^{X} \frac{k_{i-1}}{k_{i-1} + (i - 1) \gamma} \left( \frac{1}{k_{i-1}} + \sum_{j=i+1}^{X} \frac{1}{k_{j-1}} \prod_{l=1}^{j-1} \frac{k_{l}}{k_{l} + l \gamma} \right).
$$

[S3.13]

**Second order moment**

The second order moment is given by

$$
\langle T^2 \rangle = -2 \mathbf{U}^\top \mathbf{A}^{-1} \mathbf{E}^{-1} \mathbf{A}^{-1} \mathbf{P}(0) = -2 \mathbf{U}^\top \mathbf{A}^{-1} \mathbf{E}^{-1} \mathbf{A}^{-1} \mathbf{E}^{-1} \mathbf{A}^{-1} \mathbf{P}(0).
$$

[S3.14]

Let us use the notation $\eta_i$ defined as

$$
\eta_i := \frac{k_{i-1}}{k_{i-1} + (i - 1) \gamma} \left( \frac{1}{k_{i-1}} + \sum_{j=i+1}^{X} \frac{1}{k_{j-1}} \prod_{l=1}^{j-1} \frac{k_{l}}{k_{l} + l \gamma} \right) + \frac{b}{k_0} \delta_{i-1},
$$

[S3.15]

where $\delta_{i-1}$ denotes the Kronecker delta which is one if $i = 1$ and zero otherwise.

Using Eq. (S3.12b), we can write

$$
\mathbf{E}^{-1} \mathbf{A}^{-1} \mathbf{P}(0) = -\frac{1}{b} \begin{bmatrix}
\eta_1 \\
\eta_2 \\
\vdots \\
\eta_{X-2} \\
\eta_{X-1}
\end{bmatrix},
$$

[S3.16]

Therefore

$$
\mathbf{A}^{-1} \mathbf{E}^{-1} \mathbf{A}^{-1} \mathbf{P}(0) = \left( \frac{1}{b} \right)^2 \begin{bmatrix}
\frac{b+1}{k_0} & 0 & \cdots & 0 & 0 \\
\frac{b+1}{k_1} & 0 & \cdots & 0 & 0 \\
\vdots & \vdots & \ddots & \vdots & \vdots \\
\frac{b+1}{k_{X-1}} & \frac{1}{k_{X-1}} & \cdots & \frac{1}{k_{X-1}} & 0 \\
\frac{b+1}{k_{X-1}} & \frac{1}{k_{X-1}} & \cdots & \frac{1}{k_{X-1}} & 0 \\
\frac{b+1}{k_{X-1}} & \frac{1}{k_{X-1}} & \cdots & \frac{1}{k_{X-1}} & 0 \\
\end{bmatrix} \begin{bmatrix}
\eta_1 \\
\eta_2 \\
\vdots \\
\eta_{X-2} \\
\eta_{X-1}
\end{bmatrix}.
$$

[S3.17a]

Using the notion $\xi_i = \sum_{j=1}^{i-1} \eta_i + (b + 1) \eta_i$, we can write $\mathbf{E}^{-1} \mathbf{A}^{-1} \mathbf{E}^{-1} \mathbf{A}^{-1} \mathbf{P}(0)$ as

$$
\left( \frac{1}{b} \right)^2 \begin{bmatrix}
1 & 0 & \cdots & 0 & 0 \\
\frac{k_1}{k_{1} + \gamma} & \frac{k_2}{k_{1} + \gamma} & \cdots & \frac{k_{X-1}}{k_{1} + \gamma} & 0 \\
0 & \frac{k_2}{k_{2} + \gamma} & \cdots & \frac{k_{X-2}}{k_{2} + \gamma} & 0 \\
0 & 0 & \cdots & \frac{k_{X-2}}{k_{X-1} + (X-1) \gamma} & 0 \\
0 & 0 & \cdots & \frac{k_{X-2}}{k_{X-1} + (X-1) \gamma} & 0 \\
\end{bmatrix} \begin{bmatrix}
\xi_1 \\
\xi_2 \\
\vdots \\
\xi_{X-2} \\
\xi_{X-1}
\end{bmatrix}.
$$

[S3.18]

As $\mathbf{U}^\top \mathbf{A}^{-1}$ gives the negative summation of the elements of the column vector it pre-multiplies to, the second order moment of FPT can be given by following explicit formula

$$
\langle T^2 \rangle = 2 \left( \frac{1}{b} \right)^2 \sum_{i=1}^{X} \frac{k_{i-1}}{k_{i-1} + (i - 1) \gamma} \left( \xi_i + \sum_{r=i+1}^{X} \xi_r \prod_{l=i}^{r-1} \left( \frac{k_{l}}{k_{l} + l \gamma} \right) \right).
$$

[S3.20]
S4. Effect of model parameters on FPT statistics

Having developed the expressions for first two moments of FPT, we can investigate how different model parameters affect the mean and noise in FPT. Perhaps simplest possible case is to consider a constant transcription rate, and that protein does not degrade. For this case, the FPT moments in Eq. (S3.10) lead to

$$\langle T \rangle = \frac{1}{k} \left( \frac{X}{b} + 1 \right) \approx \frac{X}{bk} \text{CV}^2_T = \frac{b^2 + X + 2bX}{(b + X)^2} \approx \frac{1 + 2b}{X},$$

where CV^2_T represents the noise in FPT as quantified by its coefficient of variation squared (variance/mean^2; \( \langle T^2 \rangle / \langle T \rangle^2 - 1 \)). There formulas have been discussed in main text in detail.

The next possibility is to study the case when protein degradation is considered along with constant transcription rate, i.e., \( k_i = k \). For this case, the first two FPT moments are given by

$$\langle T \rangle = \frac{1}{k} + \frac{1}{b} \sum_{i=1}^{X} \frac{1}{k + (i - 1)\gamma} \left( 1 + \sum_{j=i+1}^{X} \prod_{l=i}^{j-1} \left( \frac{l\gamma}{k + l\gamma} b + 1 \right) \right),$$

$$\langle T^2 \rangle = \frac{2}{b^2} \sum_{i=1}^{X} \frac{1}{k + (i - 1)\gamma} \left( \xi_i + \sum_{r=i+1}^{X} \sum_{l=1}^{r-1} \xi_r \prod_{l=1}^{r-1} \left( \frac{l\gamma}{k + l\gamma} b + 1 \right) \right),$$

where

$$\xi_i = \sum_{j=1}^{i-1} \eta_j + (b + 1)\eta_i,$$

$$\eta_i = \frac{1}{k + (i - 1)\gamma} \left( 1 + \sum_{j=i+1}^{X} \prod_{l=i}^{j-1} \left( \frac{l\gamma}{k + l\gamma} b + 1 \right) \right) + \frac{b}{k} \delta_{i-1}.$$

Some insights obtained from these formulas are discussed at length in the main text. Next we discuss some additional insights obtained from these results.

**Independent tuning of mean FPT and noise in FPT.** An important result that can be seen from these formulas is that the noise only depends upon the ratio \( k/\gamma \) and not on their individual values. To show this, let us rewrite the formulas with \( \theta := k/\gamma \)

$$\langle T \rangle = \frac{1}{\theta} \left( \frac{1}{\theta} + \frac{1}{b} \sum_{i=1}^{X} \frac{1}{\theta + (i - 1)} \left( 1 + \sum_{j=i+1}^{X} \prod_{l=i}^{j-1} \left( \frac{l\theta}{\theta + l} b + 1 \right) \right) \right),$$

$$\langle T^2 \rangle = \frac{2}{b^2} \sum_{i=1}^{X} \frac{1}{\theta + (i - 1)} \left( \xi_i + \sum_{r=i+1}^{X} \sum_{l=1}^{r-1} \xi_r \prod_{l=1}^{r-1} \left( \frac{l\theta}{\theta + l} b + 1 \right) \right),$$

where

$$\xi_i = \sum_{j=1}^{i-1} \eta_j + (b + 1)\eta_i,$$

$$\eta_i = \frac{1}{\theta + (i - 1)} \left( 1 + \sum_{j=i+1}^{X} \prod_{l=i}^{j-1} \left( \frac{l\theta}{\theta + l} b + 1 \right) \right) + \frac{b}{\theta} \delta_{i-1}.$$

Thus, the noise CV^2_T can be written as

$$CV^2_T = \frac{\langle T^2 \rangle}{\langle T \rangle^2} - 1 = \frac{2}{b^2} \sum_{i=1}^{X} \frac{1}{\theta + (i - 1)} \left( \xi_i + \sum_{r=i+1}^{X} \sum_{l=1}^{r-1} \xi_r \prod_{l=1}^{r-1} \left( \frac{l\theta b + 1}{\theta + l} \right) \right)^2 - 1,$$

which depends only on \( \theta = k/\gamma \), and not on the values of \( k \) and \( \gamma \) themselves. This result has an important implication that if the noise is in desirable range for a given a set of parameters \( k, b, \gamma, X \), then the mean FPT can be independently tuned by varying degradation rate and keeping \( k/\gamma \) constant.
Effect of model parameters for constant production and degradation. Using the formulas in Eq. (S4.2), we can investigate how different parameters impact mean and noise in FPT. As discussed in the main text, fixing other parameters and varying the event threshold results in a U-shape profile for noise in FPT (Fig. 2 in main text, Fig. S2 (left)). Furthermore, because the noise is minimized at a certain value of $X$, the cell might choose this value and vary $k/\gamma$ constant to get a required mean FPT (Fig. S2 (right)). The mean burst size $b$ can also be varied while keeping the steady-state protein level same by appropriately changing $k/\gamma$. It is seen that when $X$ is smaller than or comparable to the steady-state protein level, increasing $b$ increases noise in timing. However, when $X$ is higher than the steady-state protein level, the noise in timing is decreasing by increasing $b$ (Fig. 2 in the main text).

Remark: The result that $CV^2$ is independent of individual values of $k$ and $\gamma$ can be straightforwardly extended to the FPT moments with feedback (Eq. (S3.13), Eq. (S3.20)). For this purpose, we define rescaled transcription rates $K_i = k_i/k_0$, and the ratio $\theta := k_0/\gamma$. Then, the FPT moments can be rewritten as

$$
\langle T \rangle = \frac{1}{\gamma} \left( \frac{1}{\theta} + \frac{1}{b} \sum_{i=1}^{X} \frac{\theta K_{i-1}}{\theta K_{i-1} + (i-1)} \left( \frac{1}{\theta K_{i-1}} + \frac{1}{\theta K_{i-1}} \prod_{l=i}^{j-1} \frac{K_l}{\theta K_l + (b+1)l} \right) \right), \tag{S4.5a}
$$

$$
\langle T^2 \rangle = \frac{1}{\gamma^2} \left( \frac{2}{b^2} \sum_{i=1}^{X} \frac{\theta K_{i-1}}{\theta K_{i-1} + (i-1)} \left( \frac{b \overline{\pi}_i + \sum_{j=i}^{X} \overline{\pi}_j}{\theta K_{i-1}} + \sum_{r=i+1}^{X} \frac{b \overline{\pi}_r + \sum_{j=r}^{X} \overline{\pi}_j}{\theta K_{r-1}} \prod_{l=i}^{r-1} \frac{K_l}{\theta K_l + (b+1)l} \right) \right), \tag{S4.5b}
$$

where the terms $\overline{\pi}_i$ are given by

$$
\overline{\pi}_i = \frac{b}{\theta} \delta_{i-1} + \frac{\theta K_{i-1}}{\theta K_{i-1} + (i-1)} \left( \frac{1}{\theta K_{i-1}} + \frac{1}{\theta K_{i-1}} \prod_{l=i}^{j-1} \frac{K_l}{\theta K_l + (b+1)l} \right). \tag{S4.5c}
$$

It can be seen that the noise only depends upon $k_0/\gamma$ and not on their individual values. While one could have defined $\theta$ as scaling some other transcription rate $k_i$ by $\gamma$ and made similar observation, the significance of using $k_0/\gamma$ can be understood by considering Hill function forms of the feedbacks. As in the main text Eq. (16), let us consider a positive feedback implemented as

$$
k_i = k_{\text{max}} \frac{r + (ci)^H}{1 + (ci)^H}, \tag{S4.6}
$$

where $k_{\text{max}}$ is the maximum transcription rate, $k_{\text{max}}r$ is the basal transcription rate with $0 < r < 1$ representing the basal strength, $c$ is the feedback strength, and $H$ is the Hill coefficient. In this case, we have

$$
k_0 = k_{\text{max}}r, \quad K_i = \frac{r + (ci)^H}{r(1 + (ci)^H)}. \tag{S4.7}
$$
Typically \( r \) is a constant small number less than one. Thus, it can be observed that the noise must only depend only on \( k_{\text{max}}/\gamma \) and not on their individual values. One can make similar argument to see that in the negative feedback case (Eq. (15) in main text)

\[
k_i = \frac{k_{\text{max}}}{1 + (ci)^p},
\]

as well the noise depends only on \( k_{\text{max}}/\gamma \). These results collectively show that the noise in FPT can be tuned by choosing a value of \( k_{\text{max}}/\gamma \) and the mean FPT can be independently tuned by choosing a value of \( \gamma \) without affecting the noise.

### S5. Optimal feedback when protein does not degrade

As mentioned in the main text, our objective is to find optimal feedback strategy that minimizes \( \langle T^2 \rangle \) such that \( \langle T \rangle \) is fixed. For calculation purposes, we will denote this constraint as \( \langle T \rangle = t_{\text{opt}} \). Let \( m \) represents the Lagrange’s multiplier, then we define the following objective function

\[
\phi := \langle T^2 \rangle + m \left( \langle T \rangle - t_{\text{opt}} \right).
\]

The optimization problem is solved in two steps. First, we determine the critical points. Second, we find the critical point corresponding to a global minimum.

Determining the critical points requires the following system of equations to be solved

\[
\frac{\partial \langle T^2 \rangle}{\partial k_i} = m \frac{\partial \langle T \rangle}{\partial k_i}, 0 \leq i \leq X - 1,
\]

\[
\langle T \rangle = t_{\text{opt}}.
\]

The expressions of the first-two moments of FPT when protein does not degrade can be obtained by substituting \( \gamma = 0 \) in the moment expressions. These are given by

\[
\langle T \rangle = \frac{1}{k_0} + \frac{1}{b} \sum_{i=0}^{X-1} \frac{1}{k_i},
\]

\[
\langle T^2 \rangle = \frac{2}{b^2} \left( \frac{\tau_0}{bk_0} + \sum_{i=0}^{X-1} \frac{\tau_i}{k_i} \right), \text{ where } \tau_i := \frac{b}{k_i} + \sum_{j=1}^{X-1} \frac{1}{k_j}.
\]

The optimization problem in Eq. (S5.2) requires calculation of the first order derivatives of \( \langle T \rangle \) and \( \langle T^2 \rangle \). The derivatives of \( \langle T \rangle \) with respect to \( k_i \)'s are given by

\[
\frac{\partial \langle T \rangle}{\partial k_i} = \frac{b+1}{b} \left( -\frac{1}{k_i^2} \right), \quad \frac{\partial \langle T \rangle}{\partial k_0} = \frac{1}{b} \left( -\frac{1}{k_0^2} \right), 1 \leq i \leq X - 1.
\]

Similarly, the derivative of \( \langle T^2 \rangle \) are

\[
\frac{\partial \langle T^2 \rangle}{\partial k_0} = 2 \left( \frac{1}{b} \right)^2 \left( \frac{2(b+1)^2}{k_0^2} - \frac{b+1}{k_0} \sum_{j=1}^{X-1} \frac{1}{k_j} \right),
\]

\[
\frac{\partial \langle T^2 \rangle}{\partial k_i} = 2 \left( \frac{1}{b} \right)^2 \left( \frac{b+1}{k_0} + \frac{2(b+1)}{k_i} + \sum_{j=1}^{X-1} \frac{1}{k_j} \right) \left( -\frac{1}{k_i^2} \right), 1 \leq i \leq X - 1.
\]

Substituting these expressions and assuming \( k_0 \neq 0, k_i \neq 0 \), the system of equations to be solved becomes

\[
\frac{2}{b} \left( \frac{2(b+1)}{k_0} + \sum_{j=1}^{X-1} \frac{1}{k_j} \right) = m,
\]

\[
\frac{2}{b} \left( \frac{b+1}{k_0} + \frac{2(b+1)}{k_i} + \sum_{j=1}^{X-1} \frac{1}{k_j} \right) = m, 1 \leq i \leq X - 1,
\]

\[
\frac{1}{b} \left( \frac{b+1}{k_0} + \sum_{j=1}^{X-1} \frac{1}{k_j} \right) = t_{\text{opt}}.
\]
We have calculated the critical point for the optimization problem. However, it needs to be checked whether it's a minimum or maximum. For this purpose, we consider the bordered Hessian as follows.

\[
D_{\phi} = \begin{bmatrix}
\frac{\partial^2 \phi}{\partial m^2} & \frac{\partial^2 \phi}{\partial m \partial k_0} & \frac{\partial^2 \phi}{\partial m \partial k_i} & \cdots & \frac{\partial^2 \phi}{\partial m \partial k_{X-1}} \\
\frac{\partial^2 \phi}{\partial m \partial k_0} & \frac{\partial^2 \phi}{\partial k_0 \partial m} & \frac{\partial^2 \phi}{\partial k_0 \partial k_1} & \cdots & \frac{\partial^2 \phi}{\partial k_0 \partial k_{X-1}} \\
\frac{\partial^2 \phi}{\partial m \partial k_i} & \frac{\partial^2 \phi}{\partial k_i \partial m} & \frac{\partial^2 \phi}{\partial k_i \partial k_1} & \cdots & \frac{\partial^2 \phi}{\partial k_i \partial k_{X-1}} \\
\cdots & \cdots & \cdots & \cdots & \cdots \\
\frac{\partial^2 \phi}{\partial m \partial k_{X-1}} & \frac{\partial^2 \phi}{\partial k_{X-1} \partial m} & \frac{\partial^2 \phi}{\partial k_{X-1} \partial k_1} & \cdots & \frac{\partial^2 \phi}{\partial k_{X-1} \partial k_{X-1}}
\end{bmatrix}.
\]  

[S5.8]

We will show that all the principal minors of this matrix are negative. To start with, let us first determine the second order derivatives of (T).

\[
\frac{\partial^2 \langle T \rangle}{\partial k_0^2} = \frac{2(b+1)}{bk_0^2}, \quad \frac{\partial^2 \langle T \rangle}{\partial k_i^2} = \frac{2}{bk_i^2}, 1 \leq i \leq X-1; \quad \frac{\partial^2 \langle T \rangle}{\partial k_i \partial k_j} = 0, 0 \leq i, j \leq X-1, i \neq j.
\]  

[S5.9]

Similarly, the derivatives for \( \langle T^2 \rangle \) are given by

\[
\frac{\partial^2 \langle T^2 \rangle}{\partial k_0^2} = 2 \left( \frac{1}{b} \right)^2 \left( \frac{6(b+1)^2}{k_0^4} + \frac{2(b+1)}{k_0^2} \sum_{j=1}^{X-1} \frac{1}{k_j^2} \right),
\]  

[S5.10a]

\[
\frac{\partial^2 \langle T^2 \rangle}{\partial k_i^2} = 2 \left( \frac{1}{b} \right)^2 \left( \frac{2(b+1)}{k_0 k_i^3} + \frac{6(b+1)}{k_i^4} + \frac{2}{k_i^2} \sum_{j=1}^{X-1} \frac{1}{k_j^2} \right), 1 \leq i \leq X-1,
\]  

[S5.10b]

\[
\frac{\partial^2 \langle T^2 \rangle}{\partial k_i \partial k_0} = \frac{\partial^2 \langle T^2 \rangle}{\partial k_0 \partial k_i} = 2 \left( \frac{1}{b} \right)^2 b + \frac{1}{k_0^2 k_i^2}, 1 \leq i \leq X-1,
\]  

[S5.10c]

\[
\frac{\partial^2 \langle T^2 \rangle}{\partial k_i \partial k_j} = 2 \left( \frac{1}{b} \right)^2 \frac{1}{k_i^2 k_j^2}, 1 \leq i, j \leq X-1, i \neq j.
\]  

[S5.10d]

We can now determine the elements of the bordered Hessian matrix in equation Eq. (S5.8) computed at the solution given by equations Eq. (S5.7)

\[
\frac{\partial^2 \phi}{\partial m^2} = \frac{\partial^2 \left( \langle T^2 \rangle + m \left( \langle T \rangle - t_{opt} \right) \right)}{\partial m^2} = 0,
\]  

[S5.11a]

\[
\frac{\partial^2 \phi}{\partial m \partial k_0} = \frac{\partial^2 \phi}{\partial k_0 \partial m} = \frac{\partial \langle T \rangle}{\partial k_0} = \frac{b+1}{b} \left( -\frac{1}{k_0^2} \right),
\]  

[S5.11b]

\[
\frac{\partial^2 \phi}{\partial m \partial k_i} = \frac{\partial^2 \phi}{\partial k_i \partial m} = \frac{\partial \langle T \rangle}{\partial k_i} = \frac{1}{b} \left( -\frac{1}{k_i^2} \right).
\]  

[S5.11c]

\[
\frac{\partial^2 \phi}{\partial k_0^2} = \frac{\partial^2 \left( \langle T^2 \rangle + m \left( \langle T \rangle - t_{opt} \right) \right)}{\partial k_0^2} = \frac{4(b+1)^2(10b+2X+3)}{(2b+1)b^2 k_0^4},
\]  

[S5.11d]

\[
\frac{\partial^2 \phi}{\partial k_i^2} = \frac{\partial^2 \left( \langle T^2 \rangle + m \left( \langle T \rangle - t_{opt} \right) \right)}{\partial k_i^2} = \frac{4(b+1)^2(9b+2X+4)}{(2b+1)^2 b^2 k_i^4}.
\]  

[S5.11e]

\[
\frac{\partial^2 \phi}{\partial k_0 k_i} = \frac{\partial^2 \phi}{\partial k_i k_0} = \frac{\partial^2 \left( \langle T^2 \rangle + m \left( \langle T \rangle - t_{opt} \right) \right)}{\partial k_0 \partial k_i} = 2 \left( \frac{1}{b} \right)^2 \frac{b+1}{k_0^2 k_i^2}.
\]  

[S5.11f]
\[
\frac{\partial^2 \phi}{\partial k_i \partial k_j} = \frac{\partial^2 \phi}{\partial k_i \partial k_j} = \frac{\partial^2 \left( \langle T^2 \rangle + m (T) - t_{opt} \right)}{\partial k_i \partial k_j} = 2 \left( \frac{1}{b} \right)^2 \frac{1}{k_i^2 k_j^2} \]

[S5.11g]

It can be noted elements of \( D_\phi \) are from a set six quantities. Defining

\[
q_1 := \frac{b+1}{b} \left( -\frac{1}{k_2^2} \right), \quad q_2 := \frac{1}{b} \left( -\frac{1}{k_2^2} \right),
\]

[S5.12a]

\[
q_3 := \frac{4(b+1)^2(10b+2X+3)}{(2b+1)b^2k_0^4}, \quad q_4 := \frac{2(b+1)^3}{(2b+1)^2b^2k_0},
\]

[S5.12b]

\[
q_5 := \frac{4(b+1)^4(9b+2X+4)}{(2b+1)^4b^4k_0^4}, \quad q_6 := \frac{2(b+1)^2}{(2b+1)^2b^2k_0^2},
\]

[S5.12c]

we can write \( D_\phi \) as

\[
D_\phi = \begin{bmatrix}
0 & q_1 & q_2 & \cdots & q_2 \\
q_1 & q_3 & q_4 & \cdots & q_4 \\
q_2 & q_4 & q_5 & \cdots & q_6 \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
q_2 & q_4 & q_6 & \cdots & q_6
\end{bmatrix}.
\]

[S5.13]

Let us denote by \( \mathcal{M}(n) \) the principal minor of the matrix \( D_\phi \) of size \( n \times n \). It can be easily seen that \( \mathcal{M}(1) = 0, \mathcal{M}(2) = 0 - q_1^2 \leq 0 \) and \( \mathcal{M}(3) = -q_2^2q_3 + 2q_1q_2q_4 - q_1q_6 \). For \( 4 \leq n \leq X \), we perform the following two elementary operations on \( D_\phi \):

- \( \text{col}_r = \text{col}_r - \text{col}_{r-1} \)
- \( \text{row}_r = \text{row}_r - \text{row}_{r-1} \)

for \( r = n, n-1, \ldots, 1 \). This yields

\[
\mathcal{M}(n) = 2(q_5 - q_6)\mathcal{M}(n-1) - (q_5 - q_6)^2\mathcal{M}(n-2), \quad 4 \leq n \leq X.
\]

[S5.14]

The solution to the above recursive equation is given by

\[
\mathcal{M}(n) = -(q_5 - q_6)^{n-3} \left( (n-2) \left( q_2^2q_3 - 2q_1q_2q_4 + q_1^2q_6 \right) + \left( q_1^2q_6 + q_1^2q_6 \right) \right).
\]

[S5.15]

It can be easily checked that \( \mathcal{M}(n) \) is negative because \( q_5 > q_6, q_2^2q_3 - 2q_1q_2q_4 + q_1^2q_6 > 0 \) and \( q_1^2q_6 + q_1^2q_6 > 0 \). This proves that the critical point indeed corresponds to a minimum.

**Stochastic Simulations considering explicit mRNA dynamics.** Note that the stochastic gene expression model considers production of proteins in bursts, which relies upon the assumption that each mRNA degrades instantaneously. Here, we relax this assumption and perform stochastic simulations to see whether no feedback is still an optimal strategy.

In the Fig. 3 of the main text, we performed simulations assuming a linear form of feedback and our results suggest that the result from the analysis of the bursty approximation holds in this case. As an extension to that, here in Fig. S3 we consider the feedbacks implemented via Hill functions as in Eq. (S4.6)/ Eq. (S4.8).

**S6. Optimal feedback strategy in presence of protein degradation**

In the previous section, we have derived analytical expression of the optimal feedback strategy that minimizes \( \langle T^2 \rangle \) such that \( \langle T \rangle \) is constant for a stable protein. When the protein decay is considered, the expressions of \( \langle T \rangle \) and \( \langle T^2 \rangle \) in Eq. (S3.13) and Eq. (S3.20) are too convoluted to solve for optimal transcription rates analytically, and we take a numerical approach to solve the optimization problem. For this purpose, we fix the threshold \( X = 10 \) molecules, and mean burst size \( b = 1 \) molecules. Using numerical solvers, we search the parameter space of the transcription rates \( k_i, i \in \{0, 1, \ldots, 9\} \) for various values of the protein degradation rate \( \gamma \). As shown in Fig. S4, when \( \gamma = 0 \), the transcription rates are equal except for the first one. This is consistent with the expressions in Eq. (S5.7a)-Eq. (S5.7b). Further, as the protein degradation rate is increased, the transcription rates first increase and then decrease, suggesting a mixed feedback strategy. To keep the results biologically meaningful, the optimal feedback strategy is investigated in the main text by considering feedbacks which follow Hill functions forms.

**Optimal feedback strategy as \( X \) or \( b \) are changed.** While in the main text, we showed results for a given threshold \( X = 500 \) molecules and mean burst size \( b = 2 \) molecules. Here, we provide results when \( X \) and \( b \) are changed (Fig. S5). It can be seen that decreasing \( X \) does not change the qualitative trend of minimum \( CV_T^2 \) as \( \gamma \) is varied. However, the overall noise level increases ((Fig. S5) (Left)). Similar effect is seen as the mean burst size \( b \) is increased ((Fig. S5) (Right)).
Fig. S3. No feedback provides minimum noise in timing for a stable protein. Protein trajectories obtained using the Stochastic Simulation Algorithm (SSA) for a stochastic gene expression model with negative feedback (left), no feedback (middle), and positive feedback (right) are shown. The threshold is assumed to be 500 molecules and the feedbacks are implemented using Hill function forms of the feedbacks as in Eq. (S4.6)/Eq. (S4.8). The maximum transcription rate \( k_{\text{max}} \) is taken such that the mean FPT is kept constant at 40 minutes while other model parameters are assumed to be fixed as: \( c = 0.01 \), \( H = 1 \), \( r = 0.05 \). The mRNA half-life is assumed to be 2.7 min, and proteins are translated from mRNAs at a rate \( 0.5 \text{ min}^{-1} \) (these correspond to a mean burst size \( b = 2 \)). Histograms on the top represent distribution of FPT from 10,000 simulations.

Fig. S4. Optimal feedback strategy for unstable protein. The optimal transcription rates obtained via numerical optimization for different values of degradation rate are shown. The event threshold is assumed to be 10 molecules, and the mean FPT is constrained to be 1 minute. The optimal transcription rates show a non-monotonic pattern as the protein count increases, suggesting presence of a mixed feedback strategy.

Protein trajectories for unstable protein. As shown in Fig. 4 of the main text, a positive feedback strategy gives minimum noise in FPT around a given mean time. Here we show the protein trajectories for different feedback strategies and compare them with those of a no feedback case in Fig. 3 of the main text/Fig. S3. It is seen that the protein trajectories (Fig. S6/Fig. S7) look similar to those obtained by a no feedback strategy in stable protein case (Fig. 3 of the main text/Fig. S3). Thus, it appears that when the protein is allowed to degrade, a positive feedback attempts to mimic the path taken by the no feedback strategy when protein were stable.

As discussed in the main text, one way to understand this result is to consider the linear feedback form

\[
k_i = c_1 + c_2 i
\]

in which case the mean protein levels evolve according to the following ordinary differential equation

\[
\frac{dx(t)}{dt} = b(c_1 + c_2 x) - \gamma x, \quad x(0) = 0.
\]

While no feedback \((c_2 = 0)\), and negative feedback \((c_2 < 0)\) in Eq. (S6.2) create nonlinear protein trajectories (Fig. S7, choosing a positive value

\[
c_2 \approx \frac{\gamma}{b}
\]

results in linear \(x(t)\), and hence minimal noise in event timing. Indeed, the optimal \(c_2\) multiplied by the mean burst size \(b\) takes values close to the degradation rate \(\gamma\) as shown in Fig. S8.
Fig. S5. A positive feedback remains to be optimal as the event threshold or the mean burst size is varied. Left: The minimum value of $CV^2$ obtained via positive feedback increases monotonically with the protein degradation rate. Decreasing the event threshold $X$ shifts the curves upwards without changing the qualitative shape. Right: Increasing the event threshold $X$ shifts the curves upwards without changing the qualitative shape. The parameters for the positive (optimal) feedback used are $r = 0.05$, and $H = 1$.

S7. Optimal feedback when protein does not degrade in presence of extrinsic noise

In this section, assuming that the protein does not degrade, we investigate how the optimal regulation strategy deviates from a no feedback in presence of a static extrinsic noise. We consider two possibilities here: one, the extrinsic noise affects the mean burst size; two, the extrinsic noise affects the transcription rate. For the first case, we assume that the mean burst size is drawn from an arbitrary positive-valued distribution. The second case is analyzed by assuming that a factor $Z$ multiplies with the transcription rates resulting in an effective transcription rate when $x(t) = i$ to be $k_i Z$.

**Optimal feedback when the mean burst size is drawn from a distribution.** Let us assume the mean burst size a random variable with probability density function $f_0(\beta)$. Thus, the number of proteins in a burst are geometrically distributed with mean $b$ where $b$ itself is a random variable. The mean and second order moment are given by

$$\langle T \rangle | b = \beta = \frac{1}{\beta} \left( \frac{\beta + 1}{k_0} + \sum_{i=1}^{X-1} \frac{1}{k_i} \right),$$  \hspace{1cm} [S7.1a]

$$\Rightarrow \langle T \rangle = \left( 1 + \frac{1}{b} \right) \frac{1}{k_0} + \left( \frac{1}{b} \right) \sum_{i=1}^{X-1} \frac{1}{k_i},$$  \hspace{1cm} [S7.1b]

$$\langle T^2 \rangle | b = \beta = \left( \frac{1}{\beta} \right)^2 \left( \frac{\beta + 1}{k_0} + \sum_{i=1}^{X-1} \frac{1}{k_i} \right) + \sum_{i=1}^{X-1} \frac{1}{k_i} \left( \frac{\beta + 1}{k_i} + \sum_{j=i+1}^{X-1} \frac{1}{k_j} \right),$$  \hspace{1cm} [S7.2a]

$$\Rightarrow \langle T^2 \rangle = \left( \frac{1}{b^2} \right) + 2 \left( \frac{1}{b} \right) + 1 \frac{1}{k_0} + \left( \frac{1}{b^2} \right) + \left( \frac{1}{b} \right) \left( \frac{1}{k_0} \sum_{i=1}^{X-1} \frac{1}{k_i} + \sum_{i=1}^{X-1} \frac{1}{k_i} \right)$$

$$+ \left( \frac{1}{b} \right) \sum_{i=1}^{X-1} \frac{1}{k_i} \sum_{j=i+1}^{X-1} \frac{1}{k_j},$$  \hspace{1cm} [S7.2b]

Defining $\alpha_1 = \left( \frac{1}{b} \right)$ and $\alpha_2 = \left( \frac{1}{b^2} \right)$, we have

$$\langle T \rangle = \left( 1 + \alpha_1 \right) \frac{1}{k_0} + \alpha_1 \sum_{i=1}^{X-1} \frac{1}{k_i},$$  \hspace{1cm} [S7.3a]

$$\langle T^2 \rangle = \left( 1 + 2\alpha_1 + \alpha_2 \right) \frac{1}{k_0} + \left( \alpha_1 + \alpha_2 \right) \left( \frac{1}{k_0} \sum_{i=1}^{X-1} \frac{1}{k_i} + \sum_{i=1}^{X-1} \frac{1}{k_i} \right) + \alpha_2 \sum_{i=1}^{X-1} \frac{1}{k_i} \sum_{j=i+1}^{X-1} \frac{1}{k_j},$$  \hspace{1cm} [S7.3b]
The derivatives of $\langle T \rangle$ with respect to $k_i$'s are given by

$$\frac{\partial \langle T \rangle}{\partial k_0} = (1 + \alpha_1) \left( - \frac{1}{k_0} \right); \quad \frac{\partial \langle T \rangle}{\partial k_i} = \alpha_1 \left( - \frac{1}{k_i} \right), \quad 1 \leq i \leq X - 1. \quad [S7.4]$$

Similarly, the derivative of $\langle T^2 \rangle$ are

$$\frac{\partial \langle T^2 \rangle}{\partial k_0} = 2 \left( \frac{1}{k_0^2} \right); \quad \frac{\partial \langle T^2 \rangle}{\partial k_i} = 2 \left( \frac{(1 + 2\alpha_1 + \alpha_2)}{k_i^2} \right), \quad 1 \leq i \leq X - 1. \quad [S7.5a]$$

$$\frac{\partial \langle T^2 \rangle}{\partial k_i} = 2 \left( \frac{\alpha_1 + \alpha_2}{k_i} \right) + 2 \left( \frac{(\alpha_1 + \alpha_2)}{k_i} \right) \sum_{j=1}^{X-1} \frac{1}{k_j}, \quad 1 \leq i \leq X - 1. \quad [S7.5b]$$

Fig. S6. Shape of protein trajectories for an unstable protein with feedbacks implemented via Hill functions. Protein trajectories obtained using the Stochastic Simulation Algorithm (SSA) for a stochastic gene expression model with negative feedback (left), no feedback (middle), and positive feedback (right) are shown. The threshold is assumed to be 500 molecules and the feedbacks are implemented using Hill function forms of the feedbacks as in Eq. (S4.6) and Eq. (S4.8). The maximum transcription rate $k_{\text{max}}$ is taken such that the mean FPT is kept constant at 40 minutes while other model parameters are assumed to be fixed as: $e = 0.001, H = 1, \gamma = 0.05, \alpha = 0.05 \text{min}^{-1}$. The mRNA half-life is assumed to be 2.7 min, and proteins are translated from mRNAs at a rate 0.5 min$^{-1}$ (these correspond to a mean burst size $b = 2$). Histograms on the top represent distribution of FPT from 10,000 simulations.

Fig. S7. Shape of protein trajectories for an unstable protein with feedbacks implemented via a linear form. Protein trajectories obtained using the Stochastic Simulation Algorithm (SSA) for a stochastic gene expression model with negative feedback (left), no feedback (middle), and positive feedback (right) are shown. The threshold is assumed to be 500 molecules and the feedbacks are implemented using a linear form $k_i = c_1 + c_2 i$. The parameter $c_1$ is taken such that the mean FPT is kept constant at 40 minutes while other model parameters are assumed to be fixed as: $c_2 = 0.05 \text{min}^{-1}$ (positive feedback), $c_2 = -0.05 \text{min}^{-1}$ (negative feedback), $\gamma = 0.05 \text{min}^{-1}$. The mRNA half-life is assumed to be 2.7 min, and proteins are translated from mRNAs at a rate 0.5 min$^{-1}$ (these correspond to a mean burst size $b = 2$). Histograms on the top represent distribution of FPT from 10,000 simulations.
Fig. S8. Linear increase in $c_2$ with degradation rate. Left: A linear feedback of the form $k_1 = c_1 + c_2 i$ is considered. As expected from Eq. (S6.3), $c_2 \approx \gamma/b$. The parameter values are taken to be $X = 500$ molecules, $b = 2$.

To find a critical point, we have to solve the following system of equations

$$
2 \left( \frac{2 (1 + 2 \alpha_1 + \alpha_2)}{k_0} - \frac{(\alpha_1 + \alpha_2)}{k_0} \sum_{j=1}^{X-1} \frac{1}{k_j} \right) = m (1 + \alpha_1) \left( -\frac{1}{k_0} \right)
$$

S7.6a

$$
2 \left( \frac{\alpha_1 + \alpha_2}{k_0} + \frac{2 (\alpha_1 + \alpha_2)}{k_1} + \alpha_2 \sum_{j=1}^{X-1} \frac{1}{k_j} \right) \left( -\frac{1}{k_1} \right) = m \alpha_1 \left( -\frac{1}{k_1} \right), 1 \leq i \leq X - 1,
$$

S7.6b

$$
(1 + \alpha_1) \frac{1}{k_0} + \alpha_1 \sum_{i=1}^{X-1} \frac{1}{k_i} = t_{\text{opt}}.
$$

S7.6c

Assuming that $k_0, k_1, \ldots \neq 0$, we get

$$
2 \left( \frac{2 (1 + 2 \alpha_1 + \alpha_2)}{k_0} + (\alpha_1 + \alpha_2) \sum_{j=1}^{X-1} \frac{1}{k_j} \right) = m (1 + \alpha_1)
$$

S7.7a

$$
2 \left( \frac{\alpha_1 + \alpha_2}{k_0} + \frac{2 (\alpha_1 + \alpha_2)}{k_1} + \alpha_2 \sum_{j=1}^{X-1} \frac{1}{k_j} \right) = m \alpha_1, 1 \leq i \leq X - 1,
$$

S7.7b

$$
(1 + \alpha_1) \frac{1}{k_0} + \alpha_1 \sum_{i=1}^{X-1} \frac{1}{k_i} = t_{\text{opt}}.
$$

S7.7c

Solution to above system of equations gives

$$
k_i = k_0 \xi, 1 \leq i \leq X - 1,
$$

S7.8a

$$
k_0 = \frac{X \alpha_1 ((1 + \alpha_1)^2 + 2 \alpha_2 - 3 \alpha_1^2 + \alpha_2 + 2 \alpha_1 \alpha_2 + \alpha_2^2 \alpha_2 + \alpha_1)}{(X (\alpha_2 - \alpha_1^2) + \alpha_1 (2 + 3 \alpha_1 + \alpha_2)) t_{\text{opt}}},
$$

S7.8b

$$
\xi = \frac{X (\alpha_2 - \alpha_1^2) + 3 \alpha_1^2 + 2 \alpha_1 + \alpha_1 \alpha_2}{3 \alpha_1^2 + \alpha_1 + \alpha_1 \alpha_2 - \alpha_2}.
$$

S7.8c

These equations reduce to our previous results of having a constant mean burst size $b$ when $\alpha_1 = 1/b, \alpha_2 = 1/b^2$ are used.

**Optimal regulation when extrinsic factor affects the transcription rate.** We consider an extrinsic factor $Z$ with a positive-valued arbitrary distribution $f_Z(z)$. This factor is assumed to be static, i.e., it does not vary over the time scale of the event of interest. Further we assume that it affects the transcription rates in a multiplicative fashion. The first-passage time mean in this case can be written as

$$
\langle T | Z = z \rangle = \frac{1}{z} \left( \frac{1}{k_0} + \frac{1}{b} \sum_{i=0}^{X-1} \frac{1}{k_i} \right) \Rightarrow \langle T \rangle = \frac{1}{z} \left( \frac{1}{k_0} + \frac{1}{b} X \sum_{i=0}^{X-1} \frac{1}{k_i} \right).
$$

S7.9
Likewise the second order moment can be written as

\[ \langle T^2 \rangle = \langle \frac{1}{z^2} \rangle \frac{2}{b^2} \left( \tau_{b_0} + \frac{X - 1}{k_i} \sum_{j=0}^{X-1} \frac{\tau_j}{k_j} \right), \quad \tau_i := \frac{b}{k_i} + \sum_{j=0}^{X-1} \frac{1}{k_j}. \]  

[S7.10]

Solving the constrained optimization problem of minimizing \( \langle T^2 \rangle \) constrained to \( \langle T \rangle = t_{opt} \) in this case simplifies to solving the following system of equations

\[ \frac{1}{(1/z^2)} \frac{2}{b} \left( \frac{2 (b + 1)}{k_0} + \sum_{j=1}^{X-1} \frac{1}{k_j} \right) = m, \quad [S7.11a] \]

\[ \frac{1}{(1/z^2)} \frac{2}{b} \left( \frac{b + 1}{k_0} + \frac{2 (b + 1)}{k_i} + \sum_{j=0}^{X-1} \frac{1}{k_j} \right) = m, 1 \leq i \leq X - 1, \quad [S7.11b] \]

\[ \frac{1}{z} \frac{1}{b} \left( \frac{b + 1}{k_0} + \sum_{j=1}^{X-1} \frac{1}{k_j} \right) = t_{opt}, \quad [S7.11c] \]

where \( m \) represents the Lagrange’s multiplier. Solution to these equations gives

\[ k_0 = \frac{(b + 1)(2b + X) \left\langle \frac{1}{z} \right\rangle}{(2b + 1)t_{opt}}, \quad [S7.12a] \]

\[ k_i = \frac{2b + 1}{b + 1} k_0 = \frac{(2b + X) \left\langle \frac{1}{z} \right\rangle}{bt_{opt}}, \quad 1 \leq i \leq X - 1. \quad [S7.12b] \]

**S8. Optimal feedback strategy when burst size is drawn from a Poisson distribution**

In the main paper, the mRNA degradation process is assumed to be memoryless (exponential), and consequently the burst of proteins is assumed to follow a geometric distribution. However, in the limit when the mRNA degradation process is deterministic, the burst size distribution becomes Poisson. For this reason, we investigate how the optimal feedback strategy changes in the case when the burst follows a Poisson distribution.

The probability density function of FPT can be computed by using Eq. (S1.4b). Here the burst size \( B \) follows a Poisson distribution given by

\[ P(B = j) = \frac{b^j}{j!} e^{-b}. \]  

[S8.1]

Here \( b \), as before, represents the mean burst size, i.e., the average number of protein molecules produced by one mRNA. One can compute the moments of first-passage time in the same manner as done for the geometric burst size case. It turns out that
the first two moments can be compactly written as

\[
\langle T \rangle = \sum_{i=0}^{X-1} \frac{\rho_i}{k_i},
\]

\[
\langle T^2 \rangle = 2 \sum_{i=0}^{X-1} \frac{\tau_i}{k_i},
\]

\[\text{[S8.2]}\]

where

\[
\tau_i = \sum_{j=i}^{X-1} \frac{\rho_{i-j}}{k_j}.
\]

\[\text{[S8.4]}\]

The coefficients \(\rho_i, i \in \{0, 1, 2, \ldots, X - 1\}\) are defined as

\[
\rho_0 = 1, \quad \rho_i = \frac{e^b}{e^b-1} \frac{b^i}{i!} \sum_{m=0}^{i-1} a[i, m] e^{-mb},
\]

\[\text{[S8.5]}\]

with \(a[r, s]\) represents an Eulerian number whose expression is

\[
a[r, s] = \sum_{i=0}^{r+1} (-1)^i \binom{r+1}{i} (s+1-i)^r.
\]

\[\text{[S8.6]}\]

We performed numerical optimization with respect to parameters \(k_i\)'s for threshold \(X = 10\) to see the form of the optimal feedback strategy such that \(\langle T^2 \rangle\) is minimized, with constraint \(\langle T \rangle = t_{opt} = 10\) minutes. The results show that while the optimal transcription rates are not equal (i.e., no feedback mechanism in strict sense), they are fairly close to each other for mean burst sizes of 1 and 3 molecules. For mean burst size of 0.6 molecules, the first transcription rate when protein count is zero comes out to be significantly higher than other transcription rates which are more or less close to each other (Fig. S9). These results suggest that while the optimal feedback strategy deviates from a no feedback strategy with the underlying distribution of the burst size, it appears to remain close to a no feedback strategy.

**S9. Supplemental figures on phage \(\lambda\) lysis**

In this section, we provide two figures on phage \(\lambda\) lysis to supplement the discussion in the main text. Fig. S10 shows that the single-cell lysis time data for a wild-type phage \(\lambda\) shows precision: lysis takes place at about 65 minutes with approximately 5% coefficient of variation. Fig. S11 depicts the holin-antiholin feedforward loop in the lytic pathway of a phage \(\lambda\). Interestingly for a mutant IN71 which does not have antiholin expression, the lysis time coefficient of variation is higher than the wild-type as shown in Fig. S10.

**Fig. S10.** Single-cell lysis time histograms for bacteriophage \(\lambda\) show precision in timing. Time taken by wild-type \(\lambda\), and \(\lambda\) mutant IN71, to lyse individual \(E.\ coli\) cells infected by a single copy of the virus. For wild-type \(\lambda\), the coefficient of variation of lysis time is less than 5%. The mutant IN71, does not express antiholin but also has other compensatory mutations to ensure same mean lysis times [4]. The mutant strain has a significantly higher noise in lysis timing as measured by the coefficient of variation (p-value = 0.004 from bootstrapping). Lysis times are normalized by their respective means. See [4] for details on material and methods.
Fig. S11. An incoherent feedforward circuit in phage λ’s lytic pathway. Left: Bacteriophage λ lysis the infected host cell by expressing a membrane protein, holin (H). The protein slowly accumulates on the cell membrane over time and forms holes when a critical concentration threshold is reached. The mRNA encoding holin also expresses antiholin (AH), which binds to holin and prevents it from participating in hole formation creating a feedforward circuit.

S10. Optimal feedback strategy when the steady-state protein level is fixed

Although achieving precision in timing in an important task, the cost of protein producing may be another factor to consider in deciding upon the optimal feedback strategy. One way to incorporate energy costs, is to find feedback strategies that minimize noise in event timing for a given mean FPT, with the added constraint that they also guarantee a given mean steady-state protein level.

Let $x_{ss}$ denote the steady-state protein level obtained from the ordinary differential equation description of gene expression. Then, constraining $x_{ss}$ to some specific value leads to constraining the ratio $k_{max}b/\gamma$. For instance, for positive feedback

$$k_{max} \frac{r + (cx_{ss})^H}{1 + (cx_{ss})^H} b - \gamma x_{ss} = 0$$

$$\Rightarrow k_{max} \frac{b}{\gamma} = x_{ss} \left(1 + (cx_{ss})^H\right) \left(r + (cx_{ss})^H\right).$$

Similarly, for negative feedback, we obtain

$$k_{max} \frac{b}{\gamma} = x_{ss} \left(1 + (cx_{ss})^H\right).$$

Note that if we choose fixed values of $H$, $r$, $b$, and $\gamma$, then the above constraint on $x_{ss}$ uniquely determines the maximal transcription rate $k_{max}$. Since the idea is to schedule the event around a given mean FPT, one needs to vary the event threshold $X$ for a given feedback strategy and feedback strength.

Fig. S12. Optimal feedback strategy when event threshold is varied to schedule event at a given time. The time-scale of event as compared to the protein half-life determines the optimal feedback strategy. Left: For events whose time-scale is smaller than protein half-life, a negative feedback provides lowest noise in timing. Middle: If time-scale of the event and protein half-life are comparable, both negative and positive feedback are not able to suppress noise for high feedback strengths. For low feedback strengths, a negative feedback gives slightly better attenuation than a no feedback. Right: For events whose time-scales are longer than protein half-life, a positive feedback gives minimum noise in timing.

Our preliminary investigation of this related, but different optimization problem, reveals interesting differences with optimal feedback strategies uncovered for the case when the steady-state was not fixed. For example, in the case of a stable protein ($\langle T \rangle \ll 1/\gamma$), a negative feedback provides the best suppression of $CV^2_T$ for a fixed mean FPT and $x_{ss}$ (Fig S12). Intuitively, for a fixed protein steady-state, the negative feedback has a much faster response time compared to no feedback or positive

Khem Raj Ghusinga, John J. Dennehy, Abhyudai Singh 10.1073/pnas.1609012114 19 of 20
feedback [3], and hence has to have a much higher event threshold $X$ to maintain the mean FPT. This leads to lower noise in event timing for negative feedback, as for a stable protein $CV_T^2$ is inversely proportional to $X$ (Fig. 2 in the main text). Intriguingly, in the case of an unstable protein $(T \gg 1/\gamma)$, positive feedback remains the optimal control strategy for buffering fluctuations in event timing, as was the case when the steady-state protein level was not fixed. Hence, depending upon the desired constraints and the parameter regimes, different feedback strategies might be optimal in terms of reducing noise in timing.

It is worth mentioning that there is a range of other possibilities that can realize the constrain on mean FPT. For example, one could vary $\gamma$, or $b$ instead of $X$ as done here. Furthermore, a combination of these parameters might also be varied. Clearly, a more systematic study is required to investigate these cases.

References