Temporal precision of molecular events with regulation and feedback

Shivam Gupta,[1] Sean Fancher,[1,2] Hendrik C. Korswagen,[3] and Andrew Mugler[1,*]

[1]*Department of Physics and Astronomy, Purdue University, West Lafayette, Indiana 47907, USA*

[2]*Department of Physics and Astronomy, University of Pennsylvania, Philadelphia, Pennsylvania 19104, USA*

[3]*Hubrecht Institute, Royal Netherlands Academy of Arts and Sciences and University Medical Center Utrecht, 3584 CT Utrecht, Netherlands*

(Received 11 October 2019; accepted 8 June 2020; published 26 June 2020)

Cellular behaviors such as migration, division, and differentiation rely on precise timing, and yet the molecular events that govern these behaviors are highly stochastic. We investigate regulatory strategies that decrease the timing noise of molecular events. Autoregulatory feedback increases noise. Yet we find that in the presence of regulation by a second species, autoregulatory feedback decreases noise. To explain this finding, we develop a method to calculate the optimal regulation function that minimizes the timing noise. The method reveals that the combination of feedback and regulation minimizes noise by maximizing the number of molecular events that must happen in sequence before a threshold is crossed. We compute the optimal timing precision for all two-node networks with regulation and feedback, derive a generic lower bound on timing noise, and discuss our results in the context of neuroblast migration during *Caenorhabditis elegans* development.

DOI: 10.1103/PhysRevE.101.062420

I. INTRODUCTION

Precise timing is crucial for many biological processes including, cell division [1–3], cell differentiation [4], cell migration [5], embryonic development [6,7], and cell death [8]. Ultimately, the timing of these processes is governed by the timing of molecular events inside the cell. However, these events are inherently stochastic. Cells use regulatory networks to reduce this stochasticity, but the effects of particular regulatory features on timing precision remain poorly understood.

We recently demonstrated that the time at which an accumulating molecular species crosses an abundance threshold is more precise if that species is regulated by a second species with its own stochastic dynamics [9]. In contrast, it was recently demonstrated that if the species is instead regulated by itself (feedback), then the crossing time is less precise [10]. It is difficult to predict in advance the effect of a particular regulatory feature on timing precision. Moreover, even if the effect of a regulatory feature is known for specific parameters, a systematic method of determining whether the effect is generic or of determining the optimal regulatory function is currently lacking.

Feedback cannot be generically harmful to timing precision, as feedback is common in many important timing processes. In yeast, the cyclin proteins that cross an abundance threshold to initiate the cell cycle [3] are subject to positive feedback [1,11,12]. In *Caenorhabditis elegans*, the mig-1 protein that crosses an abundance threshold to terminate migration in QR neuroblasts [5] has been found in experiments on the sister QL lineage to be subject to feedback via Wnt signaling [13]. This raises the question of why feedback is observed in key timing processes if it has previously been shown to decrease timing precision.

Here we investigate the combined effect of regulation and feedback on timing precision. We develop a gradient-descent approach to find the globally optimal regulation function for a given network topology that minimizes the timing noise. We find that, despite the fact that feedback generically increases timing noise when it acts alone, feedback decreases timing noise when it acts in combination with regulation by an external species. We explain the mechanisms behind this counterintuitive result, derive a generic lower bound on the timing noise, and discuss the relevance of our results to the timing of neuroblast migration in *C. elegans*.

II. RESULTS

Consider a molecular species Y that is produced over time and first reaches a molecule-number threshold y_* at a particular time t_* on average [Fig. 1(a)]. In what follows we assume that Y can only be produced, not degraded, and that when regulated, the regulator of Y can only be produced or degraded, not both. This simplifying assumption limits the possible paths that the system can take through molecule-number space, but the methods we outline below are readily generalized to species that are both produced and degraded, and we comment on this case in Sec. III.

Stochasticity in the accumulation process leads to variability in the crossing time t. The timing noise is given by the variance σ_t^2. For unregulated production of Y, the time between each production event is exponentially distributed with mean t_*/y_* and variance $(t_*/y_*)^2$. Because the production events are independent, the variances add, giving a total variance of $\sigma_t^2 = y_*(t_*/y_*)^2 = t_*^2/y_*$. Therefore, we focus on the scaled variance $\sigma_t^2 y_*/t_*^2$, whose value is 1 for unregulated production.

First, we investigate the effect of feedback on timing precision using a simple example: we suppose that the production rate of Y is not a constant but rather is a simple sigmoidal

*amugler@purdue.edu

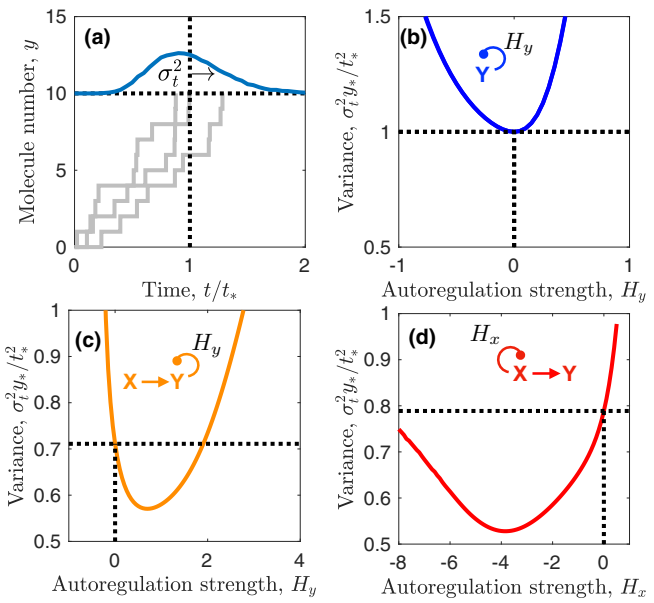

FIG. 1. Feedback increases timing precision in the presence but not absence of regulation. (a) A species Y crosses a molecule-number threshold y_* at mean time t_* with timing variance σ_t^2. (b) Feedback increases the variance. However, in the presence of regulation by a second species X, feedback on either (c) Y or (d) X can decrease the variance. Parameters are $K_y = 2.5$ in (b); $\alpha_0 t_* = 10$, $H_x = -0.5$, $H_{xy} = -H_y$, $K_x = 15$, $K_y = 5$, and $K_{xy} = 6$ in (c); $\alpha_0 t_* = 10$, $H_y = 4$, $K_x = 10$, and $K_y = 7.5$ in (d); and $y_* = 10$ throughout. (a) is generated by Gillespie's simulation [14]; (b)–(d) are generated by solving the master equation numerically [9].

function of the current number of molecules y,

$$\beta(y) = \beta_0\{1 + \tanh[H_y(y/K_y - 1)]\}, \tag{1}$$

where positive (negative) H_y corresponds to positive (negative) feedback, $|H_y|$ is the maximum steepness, K_y is the molecule number at which β is half-maximal, and β_0 is set to ensure that the average time at which y first reaches y_* is t_*. We calculate the variance σ_t^2 from the master equation by matrix inversion [9]. In Fig. 1(b) we see that when there is no feedback ($H_y = 0$), the variance satisfies $\sigma_t^2 y_*/t_*^2 = 1$ and that either positive or negative feedback increases the variance. This result is consistent with previous findings for a species that does not degrade [10], and it has an intuitive explanation: a sequence of time-ordered stochastic events is most precisely timed if the mean times for each event to occur are equal, but feedback makes these times unequal.

Next, we investigate the interplay of feedback and regulation by introducing a second species, X, that is produced at a constant rate α_0. The Y production rate $\beta(x, y)$ is now a function of both molecule numbers x and y. We find that if it is a simple sum $\beta(x, y) = f_1(x) + f_2(y)$ or product $\beta(x, y) = f_1(x)f_2(y)$, then feedback continues to generically increase the timing variance, but if we include a coupling term $\beta(x, y) = f_1(x)f_2(y)f_3(xy)$, the situation is different. Specifically, Fig. 1(c) shows the case where

$$\begin{aligned}\beta(x, y) = \beta_0\{&1 + \tanh[H_x(x/K_x - 1)]\}\\ &\times \{1 + \tanh[H_y(y/K_y - 1)]\}\\ &\times \{1 + \tanh[H_{xy}(xy/K_{xy}^2 - 1)]\}.\end{aligned} \tag{2}$$

We see that with no feedback ($H_y = 0$) we have $\sigma_t^2 y_*/t_*^2 < 1$, which demonstrates that regulation by a second species increases the timing precision as found previously [9]. However, now we also see that with positive feedback ($H_y > 0$), the variance can be even lower. Together with Fig. 1(b), this result implies that although feedback increases timing noise in the absence of regulation, it can decrease timing noise in the presence of regulation.

Similarly, we investigate the case where the feedback occurs on X, not Y. We take the production rates of x and y to be

$$\alpha(x) = \alpha_0\{1 + \tanh[H_x(x/K_x - 1)]\}, \tag{3}$$

$$\beta(x) = \beta_0\{1 + \tanh[H_y(x/K_y - 1)]\}, \tag{4}$$

respectively. We see in Fig. 1(d) that with negative feedback ($H_x < 0$) the variance is lower than when there is no feedback ($H_x = 0$), again implying that feedback can reduce timing noise when coupled to regulation.

To understand this effect, we develop a gradient-descent method to find the optimal regulation that minimizes the timing variance. The regulation is specified by the X and Y production rates $\alpha(x, y)$ and $\beta(x, y)$, respectively, which each depend on the molecule numbers x and y in general but whose dependencies will later be restricted to consider particular feedback topologies. The probability of first reaching $y = y_*$ at time t is $P(t) = \sum_{\{\vec{s}\}} P(t|\vec{s})P(\vec{s})$, where

$$P(\vec{s}) = \prod_{i=0}^{S-1} \frac{r_i}{k_i}, \tag{5}$$

$$P(t|\vec{s}) = \left(\prod_{i=0}^{S-1} \int_0^\infty dt_i k_i e^{-k_i t_i}\right) \delta\left(t - \sum_{j=0}^{S-1} t_j\right). \tag{6}$$

In Eq. (5), $P(\vec{s})$ is the probability of taking a path $\vec{s}$ from $(x_0, y_0) = (0, 0)$ to $(x_S, y_S) = (x_S, y_*)$ for any nonnegative x_S, where S is the length of the path. Each step i takes the system out of state (x_i, y_i) with rate $k_i = \alpha(x_i, y_i) + \beta(x_i, y_i)$ and into a new state with probability r_i/k_i, where the new state is either $(x_i + 1, y_i)$ with $r_i = \alpha(x_i, y_i)$ or $(x_i, y_i + 1)$ with $r_i = \beta(x_i, y_i)$. In Eq. (6), $P(t|\vec{s})$ is the probability that traversing the given path $\vec{s}$ takes a time t. The first term integrates over all values of each step's transition time t_i, which is exponentially distributed with rate k_i, and the second term ensures that the sum of these transition times is t. From $P(t)$ we calculate the moments (see Appendix A), of which the first two are

$$\langle t \rangle = \sum_{\{\vec{s}\}} P(\vec{s}) \sum_{i=0}^{S-1} \frac{1}{k_i}, \tag{7}$$

$$\langle t^2 \rangle = \sum_{\{\vec{s}\}} P(\vec{s}) \left[\left(\sum_{i=0}^{S-1} \frac{1}{k_i^2}\right) + \left(\sum_{j=0}^{S-1} \frac{1}{k_j}\right)^2\right]. \tag{8}$$

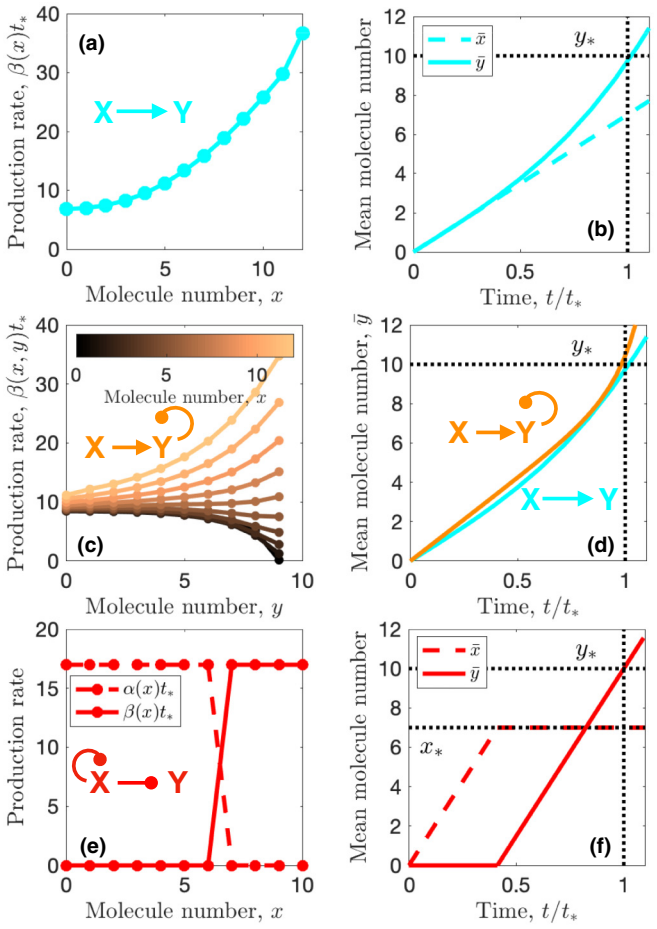

FIG. 2. Optimal regulation functions that minimize timing variance. (a) Without feedback, X activates Y, (b) allowing $\bar{y}$ to accelerate before crossing y_*. (c) With feedback on Y, X acts as a "timer" for Y, allowing Y to self-repress at early times and self-activate at late times and (d) providing further, late-phase acceleration of $\bar{y}$. (e) With feedback on X, it represses itself and activates Y sharply, (f) resulting in kinked dynamics where $\bar{x}$ and $\bar{y}$ growth are separated in time. Parameters are $\alpha_0 t_* = 7$ in (a)–(d), $x_* = 7$ in (e) and (f), and $y_* = 10$ throughout.

The optimal regulation function minimizes $\langle t^2 \rangle$ at fixed $\langle t \rangle = t_*$. Therefore, defining a vector $\vec{\gamma}$ whose components are all components of both the $\alpha(x, y)$ and $\beta(x, y)$ matrices, we initialize $\vec{\gamma}$ to satisfy $\langle t \rangle = t_*$ and update it as

$$\vec{\gamma}^{(n+1)} = \vec{\gamma}^{(n)} - \epsilon \vec{u}. \tag{9}$$

Here $\epsilon \ll 1$, and $\vec{u}$ is calculated such that $\vec{u} \cdot \nabla_\gamma \langle t^2 \rangle$ is maximized with respect to the constraints $\vec{u} \cdot \nabla_\gamma \langle t \rangle = 0$ and $|u|^2 = 1$. To summarize this numerical method, we initialize the rates $\vec{\gamma}$, calculate the gradient of the timing variance $\sigma_t^2 = \langle t^2 \rangle - \langle t \rangle^2$ [Eqs. (7) and (8)] with respect to $\vec{\gamma}$ at each point in state space, and update the rates such that $\nabla_{\vec{\gamma}} \sigma_t^2$ is maximized while keeping the target condition fixed by $\nabla_{\vec{\gamma}} \langle t \rangle = 0$. After many iterations this method converges to the minimum variance.

First, we apply this method to the case where X regulates Y with no feedback. Thus, we fix $\alpha = \alpha_0$ and optimize $\beta(x)$. Figure 2(a) shows the result, and we see that the optimal $\beta(x)$ is an increasing function of x (i.e., X activates Y). The reason, clear from the mean dynamics in Fig. 2(b), is that as x increases over time, $\beta(x)$ increases over time, which causes y to accelerate. The acceleration allows $\bar{y}$ to cross y_* with a large slope, reducing the uncertainty of the crossing time. We observed this effect previously with Hill-function activation [9], but the optimal regulation function was unknown.

Next, we keep $\alpha = \alpha_0$, but we allow feedback on Y and find the optimal $\beta(x, y)$. Figure 2(c) shows the result, and we see that the optimal $\beta(x, y)$ depends on y, confirming that feedback is beneficial in the presence of regulation. Specifically, we see that $\beta(x, y)$ decreases with y (negative feedback) when x is small and increases with y (positive feedback) when x is large. These two properties are also exhibited by Eq. (2) with $H_x < 0$, $H_y > 0$, and $H_{xy} < 0$, as in Fig. 1(c). The first property ensures that Y is not prematurely activated at early times when x is small. The second property provides an additional acceleration of y at late times when x is large. Thus, X acts as a "timer" for Y, allowing Y to apply self-amplification only at late times. This strategy has two advantages, as seen in Fig. 2(d): (i) it increases the slope of $\bar{y}$ at crossing, beyond the slope without feedback, and (ii) it allows the acceleration to begin at a $\bar{y}$ value that is already close to y_*, thus reducing trajectory-to-trajectory variability caused by prolonged self-amplification [10].

Finally, we consider the case where feedback acts on X instead of Y. Here, to provide a reasonable constraint on $x(t)$, we introduce a bound x_* and restrict $\alpha(x)$ such that $\bar{x}(t) \leqslant x_*$ over the range $0 \leqslant t \leqslant t_*$. The optimal regulation functions $\alpha(x)$ and $\beta(x)$ are shown in Fig. 2(e). We see that X represses itself and activates Y and that both regulation functions have a sharp transition when $x = x_*$. We see in Fig. 2(f) that the resulting dynamics are sharply kinked.

To understand the sharp nature of the optimal solution in Figs. 2(e) and 2(f), we investigate our optimization scheme [Eqs. (5)–(9)] analytically. The analytic version of Eq. (9) is $0 = \gamma_i \partial_{\gamma_i} (\langle t^2 \rangle - \lambda \langle t \rangle)$, where the Lagrange multiplier λ enforces $\langle t \rangle = t_*$ and the factor of γ_i in front enforces $\gamma_i > 0$ (see Appendix B). By inserting Eqs. (7) and (8) into this condition, we show in Appendix B that it is satisfied when (i) α and β are such that all possible paths $\vec{s}$ to reach $y = y_*$ have the same length S and (ii) all transition rates along each of these paths are equal. Each such set of equal-length, constant-velocity paths is a local optimum, and the global optimum that minimizes the timing variance is the set for which (iii) the path length S is as large as possible. More generally, if only property (ii) is satisfied, we show in Appendix C that the timing variance satisfies

$$\frac{\sigma_t^2}{t_*^2} = \frac{\sigma_S^2}{\langle S \rangle^2} + \frac{1}{\langle S \rangle}, \tag{10}$$

where $\langle S \rangle$ and σ_S^2 are the mean and variance of the path lengths, weighted by the path probabilities $P(\vec{s})$. Clearly, the variance is minimized when $\sigma_S^2 = 0$ and $\langle S \rangle$ is as large as possible, consistent with properties (i) and (iii) above, respectively.

Now we can understand why the optimal solution in Figs. 2(e) and 2(f) looks the way it does. The sharp nature of the regulation functions ensures that at early times only x changes, and at late times only y changes, confining the

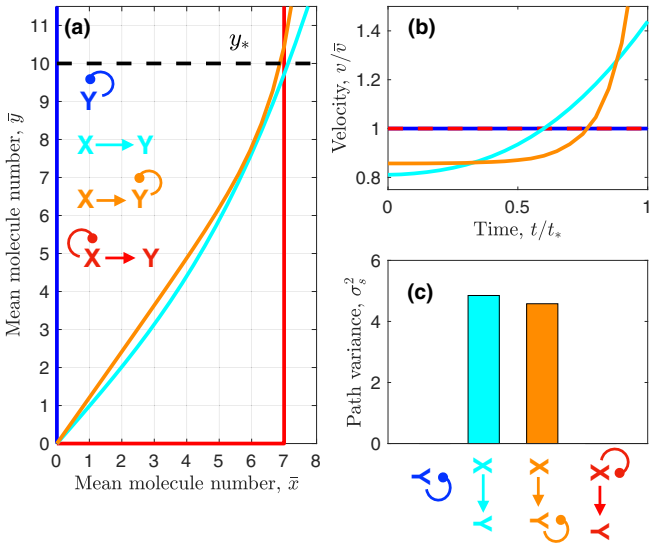

FIG. 3. Properties that minimize timing variance: (a) large path length S, (b) constant velocity $v(t)$ along the path, and (c) small path length variance σ_S^2. Parameters are as in Fig. 2.

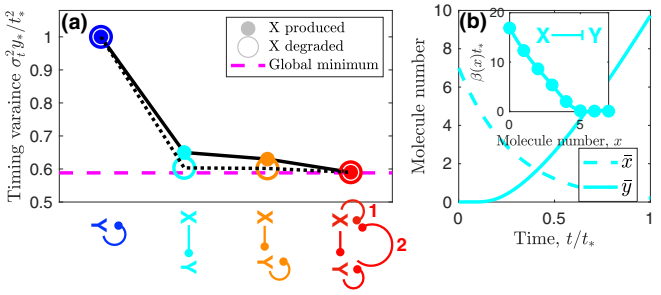

FIG. 4. (a) Ranking of timing variance for all one- and two-node networks. The global minimum is $\sigma_t^2 y_*/t_*^2 = y_*/(x_* + y_*)$. In the red network, link 1, link 2, or both are required. Parameters are as in Fig. 2. (b) Mean dynamics and regulation function (inset) for the case when X is degraded. Here $\alpha_0 t_* = 3.5$.

stochastic dynamics to only one possible path in (x, y) space [property (i)]. The values of α and β, when they are nonzero, are constant and equal to each other, ensuring that the velocity along this path is constant [property (ii)]. Finally, both x and y attain their maximal values x_* and y_*, ensuring that the path is as long as possible [property (iii)].

Indeed, Fig. 3 shows the optimal solutions for all of the networks considered thus far in terms of these three properties. Specifically, Fig. 3(a) shows the mean dynamics in (x, y) space; Fig. 3(b) shows the velocity $v(t) = \sqrt{(d\bar{x}/dt)^2 + (d\bar{y}/dt)^2}$ along this path, normalized by its time average $\bar{v} = t_*^{-1} \int_0^{t_*} dt\ v(t)$, and Fig. 3(c) shows the variance σ_S^2 in the path length across all paths. With only Y and no X (blue), there is only one possible path [Fig. 3(a)], and therefore, $\sigma_S^2 = 0$ [Fig. 3(c)]. The optimal solution has constant velocity along the path [Fig. 3(b)], which is achieved with no feedback. When X regulates Y (cyan and orange), the mean path extends into the (x, y) plane [Fig. 3(a)], which increases its length and thus lowers the timing variance. However, it also makes the velocity nonconstant [Fig. 3(b)] and allows for many possible paths such that $\sigma_S^2 > 0$ [Fig. 3(c)]. Only upon allowing X to also regulate itself (red) does the path become as long as possible [Fig. 3(a)], constant velocity [Fig. 3(b)], and unique [Fig. 3(c)].

The minimal values of the timing variance for the networks are shown by the solid circles in Fig. 4(a). We see that the single species Y achieves the standard $\sigma_t^2 y_*/t_*^2 = 1$ (blue), regulation by X lowers the variance (cyan), feedback on Y lowers it further (orange), and regulation of X lowers it to the global minimum given by Eq. (10) with $\sigma_S^2 = 0$ and $\langle S \rangle = x_* + y_*$, namely, $\sigma_t^2 y_*/t_*^2 = y_*/(x_* + y_*)$. Because the results in Fig. 4(a) are minima, it does not matter in the last case whether the regulation of X is by X itself (red link 1), by Y (red link 2), or by both; the optimal regulation functions will produce the red path in Fig. 3 regardless.

Thus far we have considered only the scenario where X is produced over time. However, X could alternatively be degraded over time [9]. In the cases where X is unregulated (cyan and orange), this corresponds to replacing its production propensity α_0 (for $x \to x + 1$) with a degradation propensity $\alpha_0 x$ (for $x \to x - 1$). The resulting minimal values of the timing variance are shown by the open circles in Fig. 4(a), and we see that they are lower than the corresponding values when X is produced over time (solid circles). The reason, illustrated for the case where X regulates Y in Fig. 4(b), is that when X is produced over time, it increases linearly [Fig. 2(b), dashed line], whereas when X is degraded over time it decreases exponentially [Fig. 4(b), dashed line]. The curvature of the exponential begins to approximate the kinked dynamics of the globally optimal solution [Fig. 2(f), dashed line]. Specifically, X is most dynamic at early times [Fig. 4(b), dashed line], and Y is produced only once x drops below a particular value [Fig. 4(b), inset], allowing it to be most dynamic at late times [Fig. 4(b), solid line]. Thus, even without feedback, the nonlinear dynamics of a degraded regulator allow its target to more closely approach the globally optimal timing precision.

III. DISCUSSION

We have developed a gradient-descent approach that provides the optimal regulation functions for a given network topology that minimize the timing noise of a threshold-crossing event. The approach has revealed that feedback reduces timing noise in the presence but not absence of regulation because the combination of the two increases the number of transitions that must happen sequentially in molecular state space. More generally, our work suggests a perspective in which noise is not minimized by finding the right network topology, but rather by finding the right combination of regulation functions that produce a path through state space that is as long, steady, and unique as possible. We anticipate that this perspective applies broadly to biological processes where timing is crucial. Although our work is limited to two-species networks, whereas most biochemical networks are much larger, our approach is straightforward to generalize to larger and more complex networks given sufficient computational power.

Our findings generalize previously known results. It was previously found that Hill-function autoregulation increases noise [10]. Here we have shown that in fact, any autoregulation function increases noise, as it results in unequal production rates at different points in state space, which is suboptimal. We previously found that Hill-function regulation by an upstream species reduces noise [9], but we lacked a procedure to determine whether Hill-function optimization is optimal. Here we have computed the optimal regulation function that results in minimal noise, and we see that it is not necessarily Hill-like but rather concave up [Fig. 2(a)].

We have considered networks in which species can be produced or degraded but not both. However, species that can be both produced and degraded are ubiquitous in biochemical networks. We have checked for small system sizes in the case where X regulates Y without feedback that adding degradation to either X or Y does not significantly change the optimal regulation function from that in Fig. 2(a), but it does increase the timing variance. The increase in the timing variance makes sense because degradation introduces a much larger set of possible paths for the molecule numbers to take through state space, as now one of the species can go down in molecule number as well as up. Consistent with our results herein, a larger path number should lead to higher path stochasticity and therefore larger threshold-crossing noise. This finding also suggests that if the degradation rate is also optimized over, the optimal degradation rate would be zero and therefore that our results on theoretical optimality apply also when degradation is included. We leave a more comprehensive investigation of simultaneous production and degradation to future work.

We find that timing noise is minimized by following a single deterministic path through state space, which is likely unrealistic for biochemical reactions. However, this result is, nonetheless, useful because it demonstrates that under ideal conditions any network will have a fundamental minimum timing variance. Here we provide the value of that variance for the simple cases of one- and two-species networks. This is important because it provides a bound to which an actual network may (or may not) come close, even if that network does not take a single deterministic path through state space. Moreover, the more realistic examples in Fig. 1 demonstrate that significant reductions in timing noise are possible due to regulation and feedback even without the globally optimal regulation functions. We anticipate that our general insights will serve as guiding predictions for experimental investigations, such as the findings that in order to increase timing precision species should be only produced or degraded but not both and that different species should change molecule number at separate times.

Can our results be compared to specific experimental systems? Our findings suggest that a cellular process for which timing precision is important should be governed by a molecular network with both multistep regulation and feedback, particularly one in which every species is subject to regulation as in Fig. 4(a) (red). An experimental example in which timing precision is particularly well studied is neuroblast migration in developing *C. elegans* larvae. Here the QR neuroblast produces a protein called mig-1 that crosses an abundance threshold to terminate migration; overproduction causes undermigration and vice versa [5]. It was recently discovered in the sister QL lineage that mig-1 is subject to both regulation and negative feedback via canonical Wnt signaling [13]. Specifically, mig-1 activates one or more Wnt signaling factors, which in turn repress mig-1. These interactions form a network of the red type in Fig. 4(a) (with link 2), where X is the Wnt factor and Y is mig-1, which is precisely the class of networks that we predict achieve the globally minimum timing noise. We anticipate that other biological processes where timing precision is paramount will be governed by interaction networks in this class.

Finally, our work has connections to other active areas of research and other biological systems. In principle, confining a stochastic system to a single deterministic path in state space should come at a large thermodynamic cost. Although we do not consider this cost here, the connection between noise reduction and energy expenditure is a rich and active field [15–17], beginning with the seminal example of kinetic proofreading [18]. Furthermore, our finding that timing precision is maximized by systems that move in only one direction (i.e., production or degradation but not both), whose species progress one at a time, is reminiscent of molecular motors on filaments. Systems of motors may exhibit these properties mechanically, via ratcheted motion and steric hinderance, and therefore may be promising examples of precise biological timers [19–22].

ACKNOWLEDGMENTS

This work was supported by Human Frontier Science Program Grant No. RGP0030/2016 and Simons Foundation Grant No. 376198.

APPENDIX A: CALCULATION OF THE MOMENTS OF THE FIRST PASSAGE TIME

Using Eqs. (5) and (6) of the main text, we write the first passage time distribution as

$$\begin{aligned} P(t) &= \sum_{\{\vec{s}\}} P(\vec{s})P(t|\vec{s}) \\ &= \sum_{\{\vec{s}\}} \left(\prod_{i=0}^{S-1} \frac{r_i}{k_i}\right)\left(\prod_{j=0}^{S-1}\int_0^\infty dt_j k_j e^{-k_j t_j}\right)\delta\left(t - \sum_{\ell=0}^{S-1} t_\ell\right) \\ &= \sum_{\{\vec{s}\}} \left(\prod_{i=0}^{S-1}\int_0^\infty dt_i r_i e^{-k_i t_i}\right)\delta\left(t - \sum_{j=0}^{S-1} t_j\right). \end{aligned} \tag{A1}$$

The nth moment is

$$\begin{aligned} \langle t^n\rangle &= \int_0^\infty dt\; t^n P(t) \\ &= \int_0^\infty dt\; t^n \sum_{\{\vec{s}\}} \left(\prod_{i=0}^{S-1}\int_0^\infty dt_i r_i e^{-k_i t_i}\right)\delta\left(t - \sum_{j=0}^{S-1} t_j\right) \\ &= \sum_{\{\vec{s}\}} \left(\prod_{i=0}^{S-1}\int_0^\infty dt_i r_i e^{-k_i t_i}\right)\left(\sum_{j=0}^{S-1} t_j\right)^n. \end{aligned} \tag{A2}$$

Specifically, the first and second moments are

$$\begin{aligned}\langle t\rangle &= \sum_{\{\vec{s}\}}\left(\prod_{i=0}^{S-1}\int_0^\infty dt_i r_i e^{-k_i t_i}\right)\left(\sum_{j=0}^{S-1} t_j\right)\\ &= \sum_{\{\vec{s}\}}\left(\prod_{i=0}^{S-1}\frac{r_i}{k_i}\right)\sum_{j=0}^{S-1}\frac{1}{k_j}\\ &= \sum_{\{\vec{s}\}} P(\vec{s})\sum_{j=0}^{S-1}\frac{1}{k_j}\end{aligned} \tag{A3}$$

and

$$\begin{aligned}\langle t^2\rangle &= \sum_{\{\vec{s}\}}\left(\prod_{i=0}^{S-1}\int_0^\infty dt_i r_i e^{-k_i t_i}\right)\left(\sum_{j=0}^{S-1} t_j\right)^2\\ &= \sum_{\{\vec{s}\}}\left(\prod_{i=0}^{S-1}\int_0^\infty dt_i r_i e^{-k_i t_i}\right)\left(\sum_{j=0}^{S-1} t_j^2 + \sum_{j=0}^{S-2}\sum_{\ell=j+1}^{S-1} 2t_j t_\ell\right)\\ &= \sum_{\{\vec{s}\}}\left(\prod_{i=0}^{S-1}\frac{r_i}{k_i}\right)\sum_{j=0}^{S-1}\sum_{\ell=j}^{S-1}\frac{2}{k_j k_\ell}\\ &= \sum_{\{\vec{s}\}}\left(\prod_{i=0}^{S-1}\frac{r_i}{k_i}\right)\left[\left(\sum_{j=0}^{S-1}\frac{1}{k_j^2}\right)+\left(\sum_{j=0}^{S-1}\frac{1}{k_j}\right)^2\right],\\ &= \sum_{\{\vec{s}\}} P(\vec{s})\left[\left(\sum_{j=0}^{S-1}\frac{1}{k_j^2}\right)+\left(\sum_{j=0}^{S-1}\frac{1}{k_j}\right)^2\right],\end{aligned} \tag{A4}$$

as in Eqs. (7) and (8) of the main text, where the last line in each case recalls Eq. (5) from the main text.

APPENDIX B: ANALYTIC MINIMIZATION OF TIMING VARIANCE USING LAGRANGE MULTIPLIERS

To find the minimum variance when the mean is fixed to be t^*, we utilize Lagrange multipliers. Because the variance is a function of only the first and second moments and is monotonically increasing with the second moment, finding the minimum of the variance with a fixed mean is equivalent to finding the minimum of the second moment with a fixed mean. Thus, the set of r_ℓ values which produces the minimum variance is the set which solves

$$0 = \frac{\partial}{\partial r_\ell}\left(\langle t^2\rangle - \lambda\langle t\rangle\right) \tag{B1}$$

for Lagrange multiplier λ.

However, Eq. (B1) raises an issue. Assume that $x_* = y_* = 1$. In this case, there are only three possible rates α_{xy} and β_{xy}, namely, α_{00}, β_{00}, and β_{10}. There are also only two possible paths: $\vec{s}_1 = [\{0,0\},\{0,1\}]$ and $\vec{s}_2 = [\{0,0\},\{1,0\},\{1,1\}]$. Putting these rates and paths into Eqs. (A3) and (A4) yields

$$\begin{aligned}\langle t\rangle &= \frac{\beta_{00}}{\alpha_{00}+\beta_{00}}\frac{1}{\alpha_{00}+\beta_{00}} + \frac{\alpha_{00}}{\alpha_{00}+\beta_{00}}\frac{\beta_{10}}{\beta_{10}}\left(\frac{1}{\alpha_{00}+\beta_{00}}+\frac{1}{\beta_{10}}\right)\\ &= \frac{1}{\alpha_{00}+\beta_{00}}\left(1+\frac{\alpha_{00}}{\beta_{10}}\right)\end{aligned} \tag{B2}$$

and

$$\begin{aligned}\langle t^2\rangle &= \frac{\beta_{00}}{\alpha_{00}+\beta_{00}}\frac{2}{(\alpha_{00}+\beta_{00})^2}\\ &\quad+\frac{\alpha_{00}}{\alpha_{00}+\beta_{00}}\frac{\beta_{10}}{\beta_{10}}\left(\frac{2}{(\alpha_{00}+\beta_{00})^2}\right.\\ &\quad\left.+\frac{2}{(\alpha_{00}+\beta_{00})\beta_{10}}+\frac{2}{\beta_{10}^2}\right)\\ &= \frac{2}{(\alpha_{00}+\beta_{00})^2}\left(1+\frac{\alpha_{00}}{\beta_{10}}+\frac{\alpha_{00}(\alpha_{00}+\beta_{00})}{\beta_{10}^2}\right).\end{aligned} \tag{B3}$$

By putting Eqs. (B2) and (B3) into Eq. (B1) and solving the resulting system of equations, one finds that some rates must be negative or even undefined depending on the order in which they are solved. Since negative rates are unphysical, we can enforce positivity by making the substitutions $\alpha_{xy} = \exp(a_{xy})/t^*$ and $\beta_{xy} = \exp(b_{xy})/t^*$ and finding the minimum variance in (a_{xy}, b_{xy}) space rather than $(\alpha_{xy}, \beta_{xy})$ space. This procedure can be done without ever leaving $(\alpha_{xy}, \beta_{xy})$ space by noting that $\partial/\partial a = (\partial\alpha/\partial a)\partial/\partial\alpha = \alpha(\partial/\partial\alpha)$ and, similarly, that $\partial/\partial b = \beta(\partial/\partial\beta)$. This allows Eq. (B1) to be rewritten as

$$0 = r_\ell\frac{\partial}{\partial r_\ell}\left(\langle t^2\rangle - \lambda\langle t\rangle\right). \tag{B4}$$

Putting Eqs. (B2) and (B3) into Eq. (B4) yields two possible solutions to the resulting equations: $[\beta_{00}, \alpha_{00}, \beta_{10}] = [1/t_*, 0, \beta_{10}]$ with $\sigma^2 = t_*^2$ for any value of β_{10} and $[\beta_{00}, \alpha_{00}, \beta_{10}] = [0, 2/t_*, 2/t_*]$ with $\sigma^2 = t_*^2/2$. One important note is the fact that when $\alpha_{00} = 0$, only the $\vec{s}_1$ path is available, while when $\beta_{00} = 0$, only the $\vec{s}_2$ path is available. Thus, the variance is seen to be extremized when only one possible path is available and all rates along that path are equal. Additionally, the longer path yields a smaller variance.

This can be seen to be a simple case of a larger trend. For any possible values of x_* and y_* it is possible to choose a set of reaction rates such that there is only one possible path through (x, y) space. When this is done, the product terms in Eqs. (A3) and (A4) become identically 1 since $r_i = k_i$ must be true along the one possible path. All other paths will have $r_i = 0$ for some i and will thus not contribute. This allows Eq. (B4) to be easily calculated for any r_ℓ that is in the single possible path,

$$\begin{aligned}0 &= r_\ell\frac{\partial}{\partial r_\ell}\left[\left(\sum_{i=0}^{S-1}\frac{1}{r_i^2}\right)+\left(\sum_{i=0}^{S-1}\frac{1}{r_i}\right)^2-\lambda\left(\sum_{i=0}^{S-1}\frac{1}{r_i}\right)\right]\\ &= \frac{\lambda}{r_\ell}-\frac{2}{r_\ell^2}-\frac{2}{r_\ell}\left(\sum_{i=0}^{S-1}\frac{1}{r_i}\right).\end{aligned} \tag{B5}$$

Equation (B5) is true for all r_ℓ along the single path if and only if all r_ℓ along that path have the same value, which, from the restriction that the mean first passage time must be t^* and Eq. (A3), means $r_\ell = S/t^*$. Putting these values back into Eq. (A4) then allows the variance to be simply calculated to be $\sigma^2 = t^{*2}/S$.

Equation (B4) must hold for all off-path reactions as well. This can be seen to be true by noting that for all other paths at least one r_i must be zero in the product term. If $\ell \neq i$, this fact

is not changed, and that path will still have zero contribution. If $\ell = i$, then the r_ℓ in front of the derivative operator will still force that path to have zero contribution since no k_i can be zero. Similarly, if r_ℓ is not a reaction that occurs at any state along the one possible path, then the derivative will cause it to vanish since the contribution from the one possible path does not depend on rates that exist in other states, while if r_ℓ is a zero rate that exists at a state in the one possible path, then the factor of r_ℓ in front of the derivative will cause the whole expression to vanish. Thus, choosing a set of reaction rates such that there is a single possible path and all rates along that path are equal is a solution to Eq. (B4) for all r_ℓ. Additionally, since $\sigma^2 = t^{*2}/S$, the longer that path is, the smaller the variance will be.

APPENDIX C: DERIVATION OF THE LOWER BOUND ON TIMING VARIANCE

If all rates are the same, $k_i = k$, then Eqs. (A3) and (A4) become

$$\langle t \rangle = \sum_{\{\vec{s}\}} P(\vec{s}) \frac{S}{k} = \frac{\langle S \rangle}{k} \tag{C1}$$

and

$$\langle t^2 \rangle = \sum_{\{\vec{s}\}} P(\vec{s}) \left(\frac{S}{k^2} + \frac{S^2}{k^2} \right) = \frac{\langle S \rangle}{k^2} + \frac{\langle S^2 \rangle}{k^2}. \tag{C2}$$

We then have

$$\begin{aligned} \frac{\sigma_t^2}{\langle t \rangle^2} &= \frac{\langle t^2 \rangle - \langle t \rangle^2}{\langle t \rangle^2} \\ &= \frac{k^2}{\langle S \rangle^2} \left(\frac{\langle S \rangle}{k^2} + \frac{\langle S^2 \rangle}{k^2} - \frac{\langle S \rangle^2}{k^2} \right) \\ &= \frac{1}{\langle S \rangle} + \frac{\sigma_S^2}{\langle S \rangle^2}, \end{aligned} \tag{C3}$$

as in Eq. (10) of the main text. Equation (C3) implies that the system does not need to be restricted to only one path, but rather to any set of paths of the same length. Then, $\sigma_S^2 = 0$ still, and a local minimum in the variance is still obtained.

We summarize the results of this and Appendix B by establishing three rules which state that the variance in first passage time is minimized when (1) all possible paths have the same length, (2) the rate at which the system moves through state space is as constant as possible, and (3) the path length through state space is maximized.

[1] J. M. Bean, E. D. Siggia, and F. R. Cross, Mol. Cell **21**, 3 (2006).
[2] I. Nachman, A. Regev, and S. Ramanathan, Cell **131**, 544 (2007).
[3] B. L. Schneider, J. Zhang, J. Markwardt, G. Tokiwa, T. Volpe, S. Honey, and B. Futcher, Mol. Cell. Biol. **24**, 10802 (2004).
[4] K. Carniol, P. Eichenberger, and R. Losick, J. Biol. Chem. **279**, 14860 (2004).
[5] R. A. Mentink, T. C. Middelkoop, L. Rella, N. Ji, C. Y. Tang, M. C. Betist, A. van Oudenaarden, and H. C. Korswagen, Dev. Cell **31**, 188 (2014).
[6] H. Meinhardt, *Models of Biological Pattern Formation* (Academic, Cambridge, MA, 1982).
[7] D. E. Tufcea and P. François, Biophys. J. **109**, 1724 (2015).
[8] J. Roux, M. Hafner, S. Bandara, J. J. Sims, H. Hudson, D. Chai, and P. K. Sorger, Mol. Syst. Biol. **11**, 803 (2015).
[9] S. Gupta, J. Varennes, H. C. Korswagen, and A. Mugler, PLoS Comput. Biol. **14**, e1006201 (2018).
[10] K. R. Ghusinga, J. J. Dennehy, and A. Singh, Proc. Natl. Acad. Sci. USA **114**, 693 (2017).
[11] L. Dirick and K. Nasmyth, Nature (London) **351**, 754 (1991).
[12] F. R. Cross and A. H. Tinkelenberg, Cell **65**, 875 (1991).
[13] N. Ji, T. C. Middelkoop, R. A. Mentink, M. C. Betist, S. Tonegawa, D. Mooijman, H. C. Korswagen, and A. van Oudenaarden, Cell **155**, 869 (2013).
[14] D. T. Gillespie, J. Phys. Chem. **81**, 2340 (1977).
[15] U. Seifert, Rep. Prog. Phys. **75**, 126001 (2012).
[16] P. R. ten Wolde, N. B. Becker, T. E. Ouldridge, and A. Mugler, J. Stat. Phys. **162**, 1395 (2016).
[17] P. Mehta, A. H. Lang, and D. J. Schwab, J. Stat. Phys. **162**, 1153 (2016).
[18] J. J. Hopfield, Proc. Natl. Acad. Sci. USA **71**, 4135 (1974).
[19] O. Campàs, Y. Kafri, K. B. Zeldovich, J. Casademunt, and J.-F. Joanny, Phys. Rev. Lett. **97**, 038101 (2006).
[20] P. C. Bressloff and J. M. Newby, Rev. Mod. Phys. **85**, 135 (2013).
[21] S. Iyer-Biswas and A. Zilman, in *Advances in Chemical Physics*, edited by S. A. Rice and A. R. Dinner, Vol. 160 (Wiley Online Library, 2016), p. 261.
[22] P. J. Mlynarczyk and S. M. Abel, Phys. Rev. E **99**, 022406 (2019).