Vol 467|9 September 2010|doi:10.1038/nature09333

ARTICLES

Fundamental limits on the suppression of molecular fluctuations

Ioannis Lestas[1], Glenn Vinnicombe[1] & Johan Paulsson[2]

Negative feedback is common in biological processes and can increase a system's stability to internal and external perturbations. But at the molecular level, control loops always involve signalling steps with finite rates for random births and deaths of individual molecules. Here we show, by developing mathematical tools that merge control and information theory with physical chemistry, that seemingly mild constraints on these rates place severe limits on the ability to suppress molecular fluctuations. Specifically, the minimum standard deviation in abundances decreases with the quartic root of the number of signalling events, making it extremely expensive to increase accuracy. Our results are formulated in terms of experimental observables, and existing data show that cells use brute force when noise suppression is essential; for example, regulatory genes are transcribed tens of thousands of times per cell cycle. The theory challenges conventional beliefs about biochemical accuracy and presents an approach to the rigorous analysis of poorly characterized biological systems.

Life in the cell is a complex battle between randomizing and correcting statistical forces: births and deaths of individual molecules create spontaneous fluctuations in abundances[1–4]—noise—and many control circuits have evolved to eliminate, tolerate or exploit the noise[5–8]. The net outcome is difficult to predict because each control circuit in turn consists of probabilistic chemical reactions. For example, negative feedback loops can compensate for changes in abundances by adjusting the rates of synthesis or degradation[7], but such adjustments are only certain to suppress noise if the individual deviations immediately and surely affect the rates[5]. Even the simplest transcriptional autorepression, by contrast, involves gene activation, transcription and translation, introducing intermediate probabilistic events that can randomize or destabilize control. Negative feedback may thus either suppress or amplify fluctuations depending on the exact mechanisms, reaction steps and parameters[9]—details that are difficult to characterize at the single-cell level and that differ greatly from system to system. This raises the fundamental questions of to what extent biological noise is inevitable and to what extent it can be controlled. Perhaps evolution could simply favour networks—however elaborate or ingeniously designed—that enable cells to homeostatically suppress any disadvantageous noise, or maybe the nature of the mechanisms imposes inherent constraints that cannot be overcome.

Control is limited by information loss

To address this issue without oversimplifying or guessing at the complexity of cells, we consider a chemical species, X_1, that affects the production of a second species, X_2, which in turn indirectly controls the production of X_1 through an arbitrarily complicated reaction network with any number of components, nonlinear reaction rates or spatial effects (Fig. 1). For generality, we only specify three of the chemical events of the larger network:

$$x_1 \xrightarrow{u(x_2(-\infty,t))} x_1+1$$
$$x_1 \xrightarrow{x_1/\tau_1} x_1-1 \qquad (1)$$
$$x_2 \xrightarrow{f(x_1)} x_2+1$$

Here x_1 and x_2 are respectively the numbers of X_1 and X_2 molecules per cell, the birth and death rates are probabilistic reaction intensities, τ_1 is the average lifetime of X_1 molecules, f is a specified rate function and the unspecified control network allows the birth rate, u, to be dynamically and arbitrarily set by the full time history of X_2 values. Death events for X_2 are omitted because the results we derive rigorously hold for all types and rates of X_2 degradation mechanisms, as long as they do not depend on X_1. The generality of u and f allows X_1 to represent many different biological species: a messenger RNA with X_2 as the corresponding protein, a protein with X_2 as either its own mRNA or an mRNA downstream in the control pathway, an enzyme with X_2 as a product or a self-replicating DNA with X_2 as a replication control molecule.

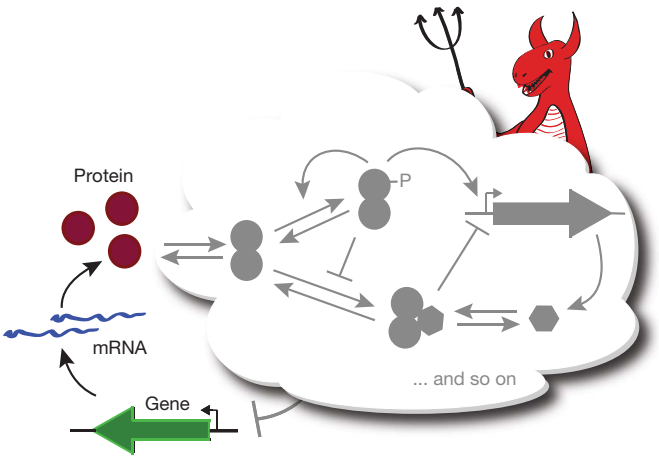

Figure 1 | Schematic of optimal control networks and information loss. Biological networks can be overwhelmingly complex, with numerous feedback loops and signalling steps. Predictions about noise then rely on quantitative estimates for how every probabilistic reaction rate responds to every type of perturbation. To investigate bounds on behaviour, most of the network is here replaced by a 'control demon' representing a controller that is optimized over all possible network topologies, rates and mechanisms. The bounds are then calculated in terms of the few specified features.

[1]Department of Engineering, University of Cambridge, Cambridge CB2 1PZ, UK. [2]Department of Systems Biology, Harvard University, Boston, Massachusetts 02115, USA.

The arbitrary birth rate, u, represents a hypothetical 'control demon' that knows everything about past and present values of x_2 and uses this information to minimize the variance in x_1. This corresponds to an optimal reaction network capable of any type of time-integration, frequency-based control, spatially extended dynamics or other exotic action. The sole restriction is that the control system depends on x_1 only through the third reaction in equation (1), which is an example of a common chemical signalling relay where a concentration determines a rate. Because individual X_2 birth events are probabilistic, some information about X_1 is therefore inevitably and irrecoverably lost and the current value of X_1 cannot be perfectly inferred from the X_2 time series. Specifically, the number of X_2 birth events in a short time period is on average proportional to $f(x_1)$, with a statistical uncertainty that depends on the average number of events. If x_1 remained constant, the uncertainty could be arbitrarily reduced by integrating over a longer time, but because it keeps changing randomly on a timescale set by τ_1, integration can only help so much. The problem is thus equivalent to determining the strength of a weak light source by counting photons: each photon emission is probabilistic, and if the light waxes and wanes, counts from the past carry little information about the current strength. The otherwise omniscient control demon thus cannot know the exact state of the component it is trying to control.

We next quantify how finite signalling rates restrict noise suppression, without linearizing or otherwise approximating the control systems, by analytically deriving a feedback-invariant upper limit on the mutual information[10] between X_1 and X_2—an information-theoretic entropic measure of the degree to which knowledge of one variable reduces uncertainty about another—and derive lower bounds on variances in terms of this limit. We use a continuous stochastic differential equation for the dynamics of species X_1, an approximation that makes it easier to extend the results to other contexts and processes, but keep the signalling and control processes discrete. This theory (summarized in Box 1 and detailed in the Supplementary Information) allows us to calculate fundamental lower bounds on variances.

Noise is limited by the quartic root of the signal rate

When the rate of X_2 production is proportional to X_1, that is, $f = \alpha x_1$, for example when X_1 is a template or enzyme producing X_2, the strict lower bound on the (squared) relative standard deviation created by the loss of information follows

$$\frac{{\sigma_1}^2}{\langle x_1\rangle^2} \geq \frac{1}{\langle x_1\rangle}\frac{2}{1+\sqrt{1+4N_2/N_1}} \approx \begin{cases} 1/N_1 & \text{for } N_2 < N_1 \\ 1/\sqrt{N_1 N_2} & \text{for } N_2 > N_1 \end{cases} \quad (2)$$

where angle brackets denote population average and $N_1 = \langle u\rangle\tau_1 = \langle x_1\rangle$ and $N_2 = \alpha\langle x_1\rangle\tau_1$ are the numbers of birth events of X_1 and X_2 made on average during time τ_1. Thus, no control network can significantly reduce noise when the signal, X_2, is made less frequently than the controlled component. When the signal is made more frequently than the controlled component, the minimal relative standard deviation (square root of equation (2)) at most decreases with the quartic root of the number of signal birth events. Reducing the standard deviation of X_1 tenfold thus requires that the signal is made at least 10,000 times more frequently. This makes it hard to achieve high precision, and practically impossible to achieve extreme precision; and this is true even for the slowest changing X_1 in the cell, relative to which signals may be made faster.

Systems with nonlinear amplification before the infrequent signalling step are also subject to bounds. For arbitrary nonlinear encoding where f is an arbitrary functional of the whole x_1 time history—corresponding to a second control demon between X_1 and X_2—the quartic-root limit turns into a type of square-root limit (Box 1 and Supplementary Information). However, gene regulatory functions typically saturate at full activation or leak at full repression, as for the generalized Hill function $f = \nu(K_1 + {x_1}^h)/(K_2 + {x_1}^h)$ with $K_1 < K_2$, where ν is the maximal rate and h is a Hill coefficient. Here X_1 may be an activator or repressor and X_2 may be an mRNA encoding either X_1 or a downstream protein. Without linearizing f or restricting the control demon, an extension of the methods above (Supplementary Information) reveals similar quartic root bounds as in equation (2), with the difference that N_2 is replaced by $\gamma N_{2,\max}$ where γ is of the order of one for a wide range of biologically relevant parameters (Supplementary Information) and $N_{2,\max} = \nu\tau_1 = N_2\nu/\langle f\rangle$. Cells can then produce many fewer signal molecules without reducing the information transfer, depending on the maximal rate increase, $\nu/\langle f\rangle$, but the quartic-root effect still strongly dampens the impact on the noise limit. If X_2 is an mRNA, $N_{2,\max}$ is also limited

Box 1 | Outline of the underlying theory

Statistical uncertainties and dependencies are often measured by variances and correlation coefficients, but both uncertainty and dependence can also be defined purely in terms of probabilities (p_i), without considering the actual states of the system. The Shannon entropy, $H(X) = \sum_i p_i \ln(p_i)$, measures inherent uncertainty rather than how different the outcomes are, and the mutual information between random variables, $I(X_1;X_2) = H(X_1) - H(X_1|X_2)$, measures the degree to which knowledge of one variable reduces entropic uncertainty in another, regardless of how their outcomes may correlate[10,29]. Despite the fundamental differences between these measures, however, there are several points of contact that can be used to predict limits on stochastic behaviour.

First, because imperfectly estimating the state of a system fundamentally restricts the ability to control it (Supplementary Information), there is a strict bound on variances whenever there is incomplete mutual information between the signal, X_2, and the controlled variable, X_1. We quantify the bound by means of Pinsker's non-anticipatory epsilon entropy[30], a rarely used information-theoretic concept that exploits the fact that the transmission of information in a feedback system must occur in real time. This shows (Supplementary Information) how an upper bound on the mutual information, $I(X_1;X_2)$, that is, a limited Shannon capacity in the channel from X_1 to X_2, imposes a lower bound on the mean squared estimation error, $E(X_1 - \hat{X}_1)^2$, where the 'estimator', $\hat{X}_1$, is an arbitrary function of the discrete signal time series and the X_1 dynamics at equilibrium is described by a stochastic differential equation. Because the capacity of the molecular channels we consider is not increased by feedback, this results in a lower limit on the variance of X_1 (in terms of the channel capacity, C) that holds for arbitrary feedback control laws: ${\sigma_1}^2/\langle x_1\rangle \geq (1 + C\tau_1)^{-1}$.

Second, the Shannon capacity is potentially unlimited when information is sent over point-process 'Poisson channels'[31], $x_2 \xrightarrow{f} x_2 + 1$, as in stochastic reaction networks where a controlled variable affects the rate of a probabilistic signalling event. However, infinite capacity requires that the rate $f(x_1)$ is unrestricted and, thus, that X_1 is unrestricted—contrary to the purpose of control. Here we consider two types of restriction. First, if the rate has an upper limit, $f_{\max}$, it follows[32] that $C = K\langle f\rangle$, where $K = \ln(f_{\max}/\langle f\rangle)$. The channel capacity then equals the average intensity multiplied by the natural logarithm of the effective dynamic range, $f_{\max}/\langle f\rangle$, and the noise bound obeys ${\sigma_1}^2/\langle x_1\rangle^2 \geq 1/(N_1(KN_2 + 1))$. This holds for any nonlinear function $f(x_1)$, but for specific functions restricting the variance in x_1 can further reduce the capacity. For example, we analytically show that the capacity of the generic Poisson channel subject to mean and variance constraints obeys $C = \langle f\rangle\ln(1 + {\sigma_f}^2/\langle f\rangle^2)$. Having less noise in x_1 will reduce the variance in f and thereby make it harder to transmit the information that is fundamentally required to reduce noise. Combining this expression for the channel capacity with the feedback limit above reveals strict limits beyond which no improvements can be made: any further reduction in the variance would require a higher mutual information, which is impossible to achieve without instead increasing the variance. When f is linear in x_1, this produces the result in equation (2). Analogous calculations allow us to derive capacity and noise results when f is a Hill function, or for processes with bursts, extrinsic noise, parallel channels and cascades (Supplementary Information). Finite channel capacities are the only fundamental constraints considered here, so at infinite capacity perfect noise suppression is possible by construction.

because transcription events tend to be relatively rare even for fully expressed genes.

Many biological systems show much greater fluctuations due to upstream sources of noise, or sudden 'bursts' of synthesis[4,11,12]. If X_1 molecules are made or degraded in bursts (of size b_1, averaged over births and deaths), there is much more noise to suppress, and if X_2 signal molecules are produced in bursts (of size b_2), each independent burst only counts as a single signalling event in terms of the Shannon information transfer, and

$$\frac{{\sigma_1}^2}{\langle x_1\rangle^2} \geq \frac{b_1}{\langle x_1\rangle}\frac{2}{1+\sqrt{1+4(N_2/b_2)/(N_1/b_1)}} \quad (3)$$

The effective average number of molecules or events is thus reduced by the size of the burst, which can increase the noise limits greatly in many biological systems. The effect of slower upstream fluctuations in turn depends on their timescales, on how they affect the system and on whether or not the control system can monitor the source of such noise directly. If noise in the X_1 birth rate is extrinsic to X_1 but not directly accessible by the controller, the predicted noise suppression limits can follow similar quartic-root principles for both fast and slow extrinsic noise, whereas for intermediate timescales the power-law exponent is between 3/8 and 1/4 (Fig. 2 and Supplementary Information).

Information losses in cascades

Signalling in the cell typically involves numerous components that change in probabilistic events with finite rates. Information about upstream states is then progressively lost at each step, much like in a game of 'broken telephone', in which messages are imperfectly whispered from person to person. If each signalling component, X_{i+1}, decays exponentially and is produced at rate $\alpha_i x_i$, an extension of the theory (Supplementary Information) shows that if a control demon monitors X_{n+1} and controls X_1, N_2 above is replaced by

$$N_{\text{eff}} = \left(\sum_{j=2}^{n+1} {N_j}^{-1}\right)^{-1} \quad (4)$$

where N_j is the average number of birth events (or bursts, as in equation (3)) of species j during time period τ_1. Information transfer in cascades is thus limited by the components made in the lowest numbers, and because the total average number of birth events over the n steps obeys $N_{\text{tot}} \geq n^2 N_{\text{eff}}$, a five-step linear cascade requires at least 25 times more birth events as a single-step mechanism to maintain the same capacity to suppress noise. This effect of information loss is superficially similar to noise propagation where variation in inputs causes variation in outputs, but although both effects reflect the probabilistic nature of infrequent reactions, the governing principles are very different. In fact, the mechanisms for preventing noise propagation—such as time averaging or kinetic robustness to upstream changes[6]—cause a greater loss of information, and mechanisms that minimize information losses—such as all-or-nothing nonlinear effects[13]—instead amplify noise. Large variation in signalling intermediates is thus not necessarily a sign of reduced precision but could reflect strategies to minimize information loss, which in turn allow tighter control of downstream components.

The rapid loss of information in cascades also suggests another trade-off: effective control requires a combination of appropriately nonlinear responses and small information losses, but nonlinear amplification in turn requires multiple chemical reactions with a loss of information at each step. The actual bounds may thus be much more restrictive than predicted above, where the assumption of Hill functions or arbitrary control networks conceals this trade-off. One of the greatest challenges in the cell may be to generate appropriately nonlinear reaction rates without losing too much mutual information between upstream and downstream components.

Parallel signal and control systems can instead improve noise suppression, because each signalling pathway contributes independent information about the upstream state. However, for a given total number of signalling events, parallel control cannot possibly reduce noise below the limits above: the loss of information is determined only by the total frequency of the signalling events, not their physical nature. The analyses above in fact implicitly allow for arbitrarily parallel control, with f interpreted as the total rate of making control molecules affected directly by X_1 (Supplementary Information).

Systems selected for noise suppression

The results above suggest that it is exceedingly difficult to suppress molecular noise efficiently. At first glance, this seems to be contradicted by a wealth of biological counterexamples: although molecules are often present in low numbers and signalling cascades where one component affects the rates of another are ubiquitous, many processes are extremely precise. That does not appear possible if the limits apply universally. However, the transmission of chemical information is not fundamentally limited by the number of molecules present at any

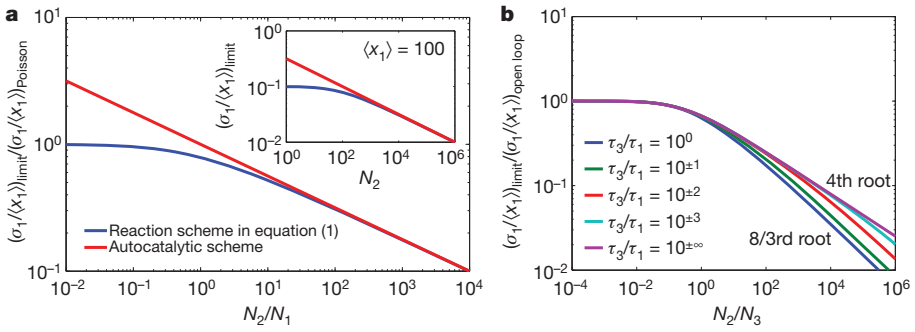

Figure 2 | Hard limits on standard deviations. a, Intrinsic noise (equation (1)): the lower limit on the relative standard deviation normalized by that of a Poisson distribution, as a function of the ratio N_2/N_1. The blue curve corresponds to the reaction scheme in equation (1), and the red curve corresponds to the autocatalytic scheme described above equation (5). The quartic root is the strongest relative response along either curve, and at low relative signalling frequencies the limit is an even more damped function of N_2/N_1. Inset, the same lower limit for an average of 100 X_1 molecules, as a function of N_2. **b**, Extrinsic noise: X_1 is made at rate $x_3 u$, where X_3 is born with constant probability and decays exponentially with rate $1/\tau_3$, and intrinsic birth and death noise in X_1 is ignored. For $\tau_3 \ll \tau_1$ and $\tau_3 \gg \tau_1$, the quartic-root asymptotic still applies, essentially because the process mimics a one-variable random process in both cases. At intermediate timescales, the N_2 dependence is less strict and $\tau_3 = \tau_1$ produces an asymptotic power-law exponent of 3/8 rather than 1/4, partly supporting previous[6,16] conclusions that extrinsic noise is slightly easier to suppress. However, many actual control systems may find intermediately slow noise the hardest to eliminate, and any predictions about suppressing extrinsic noise will depend on the properties of that noise. The predicted extrinsic noise limit is also a conservative estimate, and the actual magnitude of the noise limit may be slightly higher (Supplementary Information).

given time, but by the number of chemical events integrated over the timescale of control (that is, by N_2 rather than $\langle x_2 \rangle$). Also, most processes that have been studied quantitatively in single cells do in fact show large variation, and the anecdotal view of cells as microscopic-yet-precise largely comes from a few central processes where cells can afford a very high number of chemical events at each step, often using post-translational signalling cascades. Just as gravity places energetic and mechanistic constraints on flight but does not confine all organisms to the surface of the earth, the rapid loss of information in chemical networks places strong constraints on molecular control circuits but does not make any level of precision inherently impossible.

It can also be tempting to dismiss physical constraints simply because life processes seem to be doing 'just fine' despite them. For example, many cellular processes operate with a great deal of stochastic variation, and central pathways seem able to achieve sufficiently high precision. But such arguments are almost circular. The existence of flight does not make gravity irrelevant, nor do winged creatures simply fly sufficiently well. The challenges are instead to understand the trade-offs involved, for example to predict what performances are selectively advantageous given the associated costs, and how small fitness differences can be and still be selectively relevant.

To illustrate the biological consequences of imperfect signalling, we consider systems that must suppress noise for survival and must relay signals through gene expression, where chemical information is lost owing to infrequent activation, transcription and translation. The best-characterized examples are the homeostatic copy-number control mechanisms of bacterial plasmids that reduce the risk of plasmid loss at cell division. These have been described much like the example above, with X_1 as plasmids and X_2 as plasmid-expressed inhibitors[5], except that plasmids self-replicate with rate $u(t)x_1$ and therefore are bound by the quartic-root limit for all values of N_1 and N_2 (Fig. 2 and Supplementary Information). To identify the mechanistic constraints when X_1 production is directly inhibited by X_2, rather than by a control demon that is infinitely fast and that delivers the optimal response to every perturbation, we consider a closed toy model:

$$\begin{array}{ccc} x_1 \xrightarrow{x_1 u(x_2)} x_1+1 & & x_2 \xrightarrow{x_1 R_2^+(x_2)} x_2+1 \\ x_1 \xrightarrow{x_1/\tau_1} x_1-1 & \text{and} & x_2 \xrightarrow{R_2^-(x_2)} x_2-1 \end{array} \quad (5)$$

Here X_1 degradation with arbitrary rate function R_2^- is a proxy for partitioning at cell division, and the rate, $x_1 R_2^+$, of making X_2 is proportional to X_1 because each plasmid copy encodes a gene for X_2. We then use the logarithmic gains[6,14] $H_{12} = -\partial \ln(u)/\partial \ln(x_2)$ and $H_{22} = \partial \ln(R_2^-/R_2^+)/\partial \ln(x_2)$ to quantify the percentage responses in rates to percentage changes in levels without specifying the exact rate functions. Parameter H_{12} is similar to a Hill coefficient of inhibition, and H_{22} determines how X_2 affects its own rates, increasing when it is negatively autoregulated and decreasing when it is degraded by saturated enzymes. The ratio H_{12}/H_{22} is thus a total gain, corresponding to the eventual percentage response in u to a percentage change in x_1. With τ_2 as the average lifetime of X_2 molecules, stationary fluctuation–dissipation approximations[6,15] (linearizing responses; Supplementary Information) then give

$$\frac{{\sigma_1}^2}{\langle x_1 \rangle^2} = \overbrace{\frac{1}{\langle x_1 \rangle}\left(\frac{H_{22}}{H_{12}} + \frac{\tau_2}{\tau_1}\frac{1}{H_{22}}\right)}^{\text{Noise from low } X_1 \text{ numbers}} + \overbrace{\frac{1}{\langle x_2 \rangle}\frac{H_{12}}{H_{22}}\frac{\tau_2}{\tau_1}}^{\text{Noise from low } X_2 \text{ numbers}} \geq \underbrace{\frac{2}{\sqrt{N_1 N_2}}}_{\text{Lower limit to total noise}}$$

where the limit holds for all H_{ij} and τ_i (Supplementary Information). This reflects a classic trade-off in control theory: higher total gain suppresses spontaneous fluctuations in X_1 but amplifies the transmitted fluctuations from X_2 to X_1. Numerical analysis confirms that even a Hill-type inhibition function u can get close to the limit (not shown), and, thus, that direct inhibition can do almost as well as a control demon. However, the parameter requirements can be extreme: the signal molecules must be very short lived, and the optimal gain, $(H_{12}/H_{22})_{\text{opt}} \approx \sqrt{N_2/N_1}$, may be so high that introducing any delays or 'extrinsic' fluctuations[6,16] would destabilize the dynamics. Regardless of the inhibition control network, plasmids thus need to express inhibitors at extraordinarily high rates, and generate strongly nonlinear feedback responses without introducing signalling cascades. Most plasmids indeed take these strategies to the extreme, for example by transcribing control genes tens of thousands of times per cell cycle using several gene copies and some of the strongest promoters known. Some plasmids also eliminate many of the cascade steps inherent in gene expression, using small regulatory RNAs, and still create highly nonlinear responses using proofreading-type mechanisms (Fig. 3a). Others partly avoid indirect control by ensuring that the plasmid copies themselves prevent each others' replication (Fig. 3b), or suppress noise without closing control loops[17,18] by changing the Poisson nature of the X_1 and X_2 chemical events (equation (1)). Although such schemes may have limited effects on variances[11], some plasmids seem to take advantage of them[5].

Outlook

Several recent studies have generalized control-theoretic notions[19,20] or applied them to biology[21,22]. Others have demonstrated physical limits on the accuracy of cellular signalling[13,23–25], for example by

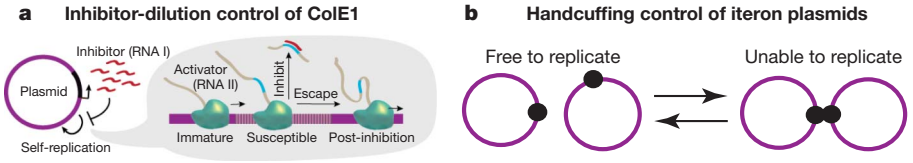

Figure 3 | Plasmid replication control. a, Plasmid ColE1 expresses an inhibitor that prevents replication, similarly to the self-replication model in the main text with X_1 as plasmid and X_2 as inhibitor. Because plasmids are under selection for noise suppression, the theory predicts it must maximize expression rates and minimize the length of signalling cascades while still achieving 'cooperative' nonlinear effects in the control loop. ColE1 indeed expresses a short-lived antisense RNA inhibitor (RNA I), tens of thousands of times per cell cycle (~10 Hz), that directly and irreversibly blocks the maturation of a constitutively synthesized sense-RNA replication pre-primer (RNA II)[5]—eliminating both the translation step and binding and unbinding to genes, and making it energetically and mechanistically possible to produce inhibitors at such high rates. ColE1 could also create strongly nonlinear control kinetics by exploiting kinetic proofreading in RNA II elongation[5,27]. Many unrelated plasmids similarly express antisense inhibitors at high rates, avoid cascades and use multistep inhibition kinetics. **b**, Plasmids such as P1, F and pSC101 use 'handcuffing' mechanisms, where repeated DNA sequences (iterons) bind each other and prevent replication[28]. This can achieve similar homeostatic dynamics as monomer–dimer equilibria where a higher fraction of molecules are in dimer form at higher abundance. Using DNA itself as inhibitor, this could eliminate the need for indirect signalling altogether, but because the mechanisms seem incapable of strongly nonlinear corrections[28], most such plasmids use additional control systems that go through gene expression and thus are subject to information loss. Plasmids also commonly use counteracting loops, where replication inhibitors also auto-inhibit their own synthesis—a counterintuitive strategy that in fact can improve control greatly (increasing H_{22} for a given, high H_{21} in equation (4)).

using fluctuation–dissipation approximations to predict estimation errors associated with a constant number of diffusing molecules hitting a biological sensor[26]. The latter show that the minimal relative error decreases with the square root of the number of events, regardless of detection mechanism. Some studies have also analysed the information transfer capacity of open-loop molecular systems[25] or have extracted valuable insights from Gaussian small-noise approximations. Here we extend these works by developing exact mathematical methods for arbitrarily complex and nonlinear real-time feedback control of a dynamic process of noisy synthesis and degradation. In such systems, the minimal error decreases with the quartic root of the integer number of signalling events, making a decent job 16 times harder than a half-decent job. This perhaps explains why there is so much biochemical noise—correcting it would be too costly—but also constrains other aspects of life in the cell. For example, the noise levels may increase or decrease along signalling cascades, depending on the kinetic details at each step, but information about upstream states is always progressively and irreversibly lost. Although it is tempting to believe that large reaction networks are capable of almost anything if the rates are suitably nonlinear, the opposite perspective may thus be more appropriate: having more steps where one component affects the rates of another creates more opportunities for information to be lost and fundamentally prevents more types of behaviour. Until the development of detailed models that predict what single cells actually do, which require every probabilistic chemical step to be well characterized, fusing control and information theory with stochastic kinetics thus provides a useful starting point: predicting what cells cannot do.

Received 1 November 2009; accepted 8 July 2010.

1. Ozbudak, E. M., Thattai, M., Kurtser, I., Grossman, A. D. & van Oudenaarden, A. Regulation of noise in the expression of a single gene. *Nature Genet.* **31,** 69–73 (2002).
2. Elowitz, M. B., Levine, A. J., Siggia, E. D. & Swain, P. S. Stochastic gene expression in a single cell. *Science* **297,** 1183–1186 (2002).
3. Newman, J. R. *et al.* Single-cell proteomic analysis of *S. cerevisiae* reveals the architecture of biological noise. *Nature* **441,** 840–846 (2006).
4. Golding, I., Paulsson, J., Zawilski, S. M. & Cox, E. C. Real-time kinetics of gene activity in individual bacteria. *Cell* **123,** 1025–1036 (2005).
5. Paulsson, J. & Ehrenberg, M. Noise in a minimal regulatory network: plasmid copy number control. *Q. Rev. Biophys.* **34,** 1–59 (2001).
6. Paulsson, J. Summing up the noise in gene networks. *Nature* **427,** 415–418 (2004).
7. Dublanche, Y., Michalodimitrakis, K., Kuemmerer, N., Foglierini, M. & Serrano, L. Noise in transcription negative feedback loops: simulation and experimental analysis. *Mol. Syst. Biol.* **2,** 41 (2006).
8. Barkai, N. & Shilo, B. Z. Variability and robustness in biomolecular systems. *Mol. Cell* **28,** 755–760 (2007).
9. Maxwell, J. On governors. *Proc. R. Soc. Lond.* **16,** 270–283 (1867).
10. Cover, T. M. & Thomas, J. A. *Elements of Information Theory* 2nd edn (Wiley, 1991).
11. Pedraza, J. M., &. Paulsson, J. Effects of molecular memory and bursting on fluctuations in gene expression. *Science* **319,** 339–343 (2008).
12. Cai, L., Friedman, N. & Xie, X. S. Stochastic protein expression in individual cells at the single molecule level. *Nature* **440,** 358–362 (2006).
13. Tkacik, G., Callan, C. G. Jr & Bialek, W. Information flow and optimization in transcriptional regulation. *Proc. Natl Acad. Sci. USA* **105,** 12265–12270 (2008).
14. Savageau, M. A. Parameter sensitivity as a criterion for evaluating and comparing the performance of biochemical systems. *Nature* **229,** 542–544 (1971).
15. Keizer, J. *Statistical Thermodynamics of Nonequilibrium Processes* (Springer, 1987).
16. Singh, A. & Hespanha, J. P. Optimal feedback strength for noise suppression in autoregulatory gene networks. *Biophys. J.* **96,** 4013–4023 (2009).
17. Korobkova, E. A., Emonet, T., Park, H. & Cluzel, P. Hidden stochastic nature of a single bacterial motor. *Phys. Rev. Lett.* **96,** 058105 (2006).
18. Doan, T., Mendez, A., Detwiler, P. B., Chen, J. & Rieke, F. Multiple phosphorylation sites confer reproducibility of the rod's single-photon responses. *Science* **313,** 530–533 (2006).
19. Martins, N. C., Dahleh, M. A. & Doyle, J. C. Fundamental limitations of disturbance attenuation in the presence of side information. *IEEE Trans. Automat. Contr.* **52,** 56–66 (2007).
20. Martins, N. C. & Dahleh, M. A. Feedback control in the presence of noisy channels: "bode-like" fundamental limitations of performance. *IEEE Trans. Automat. Contr.* **53,** 1604–1615 (2008).
21. El-Samad, H., Kurata, H., Doyle, J. C., Gross, C. A. & Khammash, M. Surviving heat shock: control strategies for robustness and performance. *Proc. Natl Acad. Sci. USA* **102,** 2736–2741 (2005).
22. Yi, T. M., Huang, Y., Simon, M. I. & Doyle, J. Robust perfect adaptation in bacterial chemotaxis through integral feedback control. *Proc. Natl Acad. Sci. USA* **97,** 4649–4653 (2000).
23. Bialek, W. & Setayeshgar, S. Cooperativity, sensitivity, and noise in biochemical signaling. *Phys. Rev. Lett.* **100,** 258101 (2008).
24. Gregor, T., Tank, D. W., Wieschaus, E. F. & Bialek, W. Probing the limits to positional information. *Cell* **130,** 153–164 (2007).
25. Walczak, A. M., Mugler, A. & Wiggins, C. H. A stochastic spectral analysis of transcriptional regulatory cascades. *Proc. Natl Acad. Sci. USA* **106,** 6529–6534 (2009).
26. Bialek, W. & Setayeshgar, S. Physical limits to biochemical signaling. *Proc. Natl Acad. Sci. USA* **102,** 10040–10045 (2005).
27. Tomizawa, J. Control of ColE1 plasmid replication: binding of RNA I to RNA II and inhibition of primer formation. *Cell* **47,** 89–97 (1986).
28. Das, N. *et al.* Multiple homeostatic mechanisms in the control of P1 plasmid replication. *Proc. Natl Acad. Sci. USA* **102,** 2856–2861 (2005).
29. Shannon, C. E. A mathematical theory of communication. *Bell Syst. Tech. J.* **27,** 379–423, 623–656 (1948).
30. Gorbunov, A. K. & Pinsker, M. S. Nonanticipatory and prognostic epsilon entropies and message generation rates. *Probl. Inf. Transm.* **9,** 184–191 (1973).
31. Kabanov, Y. The capacity of a channel of the Poisson type. *Theory Probab. Appl.* **23,** 143–147 (1978).
32. Davis, M. H. A. Capacity and cut-off rate for Poisson type channels. *IEEE Trans. Inf. Theory* **26,** 710–715 (1980).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements This research was supported by the BBSRC under grant BB/C008073/1, by the National Science Foundation grants DMS-074876-0 and CAREER 0720056, and by grants GM081563-02 and GM068763-06 from the National Institutes of Health.

Author Contributions I.L., G.V. and J.P. contributed equally, and all conceived the study, derived the equations and wrote the paper.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to G.V. (gv@eng.cam.ac.uk) or J.P. (johan_paulsson@hms.harvard.edu).