Supporting Materials and Methods

Growth conditions and Media. Cells were grown at 37ºC in M9 minimal media with succinate as the main carbon source. Overnight cultures were prepared in the absence of TMG or in the presence of 100 μM TMG to yield populations of cells uninduced or fully induced for *lac* expression, respectively. Cells from the overnight cultures were diluted into fresh media containing intermediate TMG levels and maintained at low density ($0.001 < OD_{600} < 0.005$) to prevent TMG depletion throughout the experiment. At specified time points, a portion of each culture was removed and prepared for imaging; fresh, pre-warmed media with the appropriate TMG level was added to dilute the remaining culture so that the cell density of each imaged sample was the same ($OD_{600} = 0.005$) for all time points. Samples were concentrated and prepared for imaging by filtration, centrifugation, resuspension in 1.25 μl of the appropriate growth media, and placement on a microscope slide.

Bacterial Strains. Dynamic population distributions of *lac* promoter activity were gathered using *Escherichia coli* MUK21 (1), in which the *gfp* gene is placed under the control of a wild-type *lac* promoter and chromosomally inserted. Published steady-state distributions from the ERT113 strain (1) are further analyzed here for both noise measurements and partitioning analysis. Strain ERT113 was constructed by transforming MUK21 cells with a plasmid containing the red-fluorescent protein HcRed under the control of the *gat* promoter.

Fluorescence Microscopy. Measurements of GFP fluorescence in dynamic (pre-steady-state) cell populations were obtained using a Nikon TE300 microscope equipped with automatic stage and focus, and a Micromax 1024B CCD camera (Roper Scientific), all controlled by MetaMorph software (Universal Imaging). Steady-state measurements were gathered as previously described (1). Fluorescence values for cells are corrected by subtracting background fluorescence measured in a region of the field of view devoid of cells. Cell boundaries were determined by auto-thresholding phase contrast images, and GFP intensity was averaged over this area. Mean fluorescence levels are assumed to be representative of reporter concentration and are calculated by dividing the total intensity of the cell by the area in pixels of the cell in the phase contrast image. These numbers are then normalized so the induced population average of the mean fluorescence is 100 for both GFP and RFP.

Derivation of Deterministic Model. We modify a deterministic model for the lactose uptake network (1) as shown below:

$$\frac{R}{R_{\mathrm{T}}} = \frac{1}{1 + \frac{X^2}{X_0^{\,2}}} \qquad \textbf{[6]}$$

$$\frac{dY}{dt} = \frac{N_{\mathrm{LacY}} / \tau_{\mathrm{y}}}{1 + \frac{R}{R_0}} - \frac{Y}{\tau_{\mathrm{y}}} \qquad \textbf{[7]}$$

$$\frac{dG}{dt} = \frac{N_{\mathrm{GFP}} / \tau_{\mathrm{g}}}{1 + \frac{R}{R_0}} - \frac{G}{\tau_{\mathrm{g}}} \qquad \textbf{[8]}$$

$$\frac{dX}{dt} = BY + \Lambda TMG_{\mathrm{external}} - \frac{X}{\tau_{\mathrm{x}}} \qquad \textbf{[9]}$$

The fraction of active LacI tetramers (R) as a function of internal TMG (X) is modeled by a Hill function with Hill coefficient of two. (Eq. **6**). R_T is the total concentration of LacI tetramers, and X_0 represents the half-saturation point. The rate of production of LacY molecules (Y) and GFP molecules (G) in the presence of LacI (R) is also a Hill function with Hill coefficient of one (Eqs. **7** and **8**). Here R_0 is the half-saturation concentration of active LacI (R) while N_{LacY} and N_{GFP} are the equilibrium number of LacY and GFP molecules, respectively, in fully induced cells. B represents the active uptake of TMG per LacY molecule, while Λ represents the passive, LacY-independent uptake of external TMG (Eq. **9**). Y, G, and X are all assumed to undergo first-order decay with time constants τ_y, τ_g and τ_x, respectively. To derive the simplified equations (Eqs. **1-3**) shown in the main text, Eq. **6** is substituted into Eqs. **7** and **8**, and the following new parameters (Eq. **10**) are defined:

$$g \equiv \frac{\alpha}{N_{\mathrm{GFP}}} G$$

$$y \equiv \frac{\alpha}{N_{\mathrm{LacY}}} Y$$

$$x \equiv \frac{X}{X_0}$$

$$\rho \equiv \frac{R_{\mathrm{T}}}{R_0} + 1$$

$$\beta \equiv \frac{N_{\mathrm{LacY}}\tau_{\mathrm{x}}}{\alpha X_0} B$$

$$\lambda \equiv \frac{\tau_{\mathrm{x}}}{X_0} \Lambda \qquad \textbf{[10]}$$

Effect of Induction on Doubling Time. It has been suggested that in order to model stochastic transitions in the lactose uptake network, one must include effects of TMG induction on doubling time (2). We find that induction with TMG has no statistically significant change on growth rate; therefore, we ignore this effect in our model.

Estimating deterministic parameters: τ_x and λ. To estimate τ_x, the time constant of internal TMG decay, we examine a population of cells that has been grown for 24 hours in media containing 50 μM TMG, which is sufficient to force all cells into the fully induced state of high GFP expression. These cells were then transferred into media lacking TMG at $t = 0$. GFP measurements were made at $t = 4$, 6, 8 and 10 hours (Fig. 5*a*), and the average concentration of GFP was measured at each time point with an error estimated from the population standard deviation. Because there is no external TMG in the media, the cells will cease production of GFP when the internal concentration of TMG becomes sufficiently dilute. Given that GFP is not actively degraded, its level should fall exponentially based on the doubling time. Therefore, we fit an exponential curve to the mean GFP concentration at each time point, using the least-squares method. The resulting best-fit curve was extrapolated to $t = 0$ and used to determine the approximate time at which GFP production ceased (Fig. 5*c*). Our extrapolation indicates that GFP production in induced cells begins to decay within 10-20 minutes following removal of external TMG, suggesting that $\tau_x << \tau_{1/2}$. Thus, we assume Eq. **1** is in equilibrium when compared to Eqs. **2** and **3**.

We expect that uninduced cells switched into media with a high TMG concentration will transition quickly to the induced state, minimizing stochastic effects. In this situation the deterministic model should provide a reasonable estimate of this behavior at the population level. Therefore, we arrive at an estimate of λ, the rate of TMG leakage, by fitting the deterministic model to experimental data in which we observed cells transitioning from an uninduced state to a fully induced state. Cells initially grown in absence of TMG were switched into media containing 55 μM of TMG, and measurements were taken at $t = 1$, 2, 3 and 4 hours (Fig. 5*b*). The population average GFP level was determined from this distribution with error bars set by the standard deviation. The parameter λ was then varied from 0 to 1.0 in 0.0025 intervals, and the χ^2 between the predictions of the deterministic model and the population average was

calculated at each point (Fig. 5*d*). The best-fit value was found to be $\lambda = 0.06$, with an 80% confidence range of (0.03, 0.12). Because α, β, and ρ were determined with $\lambda = 0$, a non-zero value of λ will change the position of the monostable-bistable transition and in principle require the refitting of all other parameters. However, we find that the lower monostable-bistable boundary shifts only from 3.5 to 3.4 μM TMG, which is smaller than the precision at which the boundary was originally determined. The position of the upper monostable-bistable boundary is changed as well, but the behavior of the stochastic model still matches the experiments in this region indicating that refitting is unnecessary.

Measurement of Noise around Steady State. We show the main sources of noise in Fig. 1*b*. Each gene has a source of intrinsic noise, which is related to mRNA burst size, *b*, and protein number, *N*, by the relation $\eta_{\text{int}}^2 = \frac{b+1}{< N >}$. Because GFP expression and RFP expression share one source of noise, the correlation between the levels of these molecules will depend on this noise source and no others. To derive this relation, we can use the Langevin formalism in the same manner in which Pedraza *et al.* (3) have applied it to noise propagation in a synthetic gene cascade. Here we treat the rate of change of each molecule as having two components: continuous terms due to rates in the deterministic model and stochastic terms due to intrinsic noise. This is made more precise by writing the corresponding Langevin equations for the two extrinsic sources and the two reporters in induced cells, where LacI and LacY noise is not transmitted.

$$\begin{aligned} \tau\dot{E} &= 1 - E + \mu_{\text{E}} \\ \tau\dot{G} &= E\alpha_{\text{G}} - G + \mu_{\text{G}} \\ \tau\dot{R} &= E\alpha_{\text{R}} - R + \mu_{\text{R}} + \mu_{\text{P}} \end{aligned} \qquad \textbf{[11]}$$

Here E represents the collective effect of global noise factors (for example, CRP, ribosome and RNA polymerase concentrations) and is scaled to an equilibrium value of one. G and R are the number of GFP and RFP molecules, while α_G and α_R represent the equilibrium amount of GFP and RFP in induced cells. We assume that observed drops in the levels of these molecules result from cell division, so τ is the characteristic decay time due to dilution. The following relations define the noise terms accounting for intrinsic fluctuations of proteins.

$$\begin{aligned} &< \mu_{\text{G,R}} >= 0 \\ &< \mu_{\text{G,R}}(t)\mu_{\text{G,R}}(t+\Delta t) >= 2\tau\alpha_{\text{G,R}}(b_{\text{GFP,RFP}}+1)\delta(\Delta t) \end{aligned} \qquad \textbf{[12]}$$

We include a term μ_P to account for noise introduced by fluctuations in plasmid number, which we assume is uncorrelated with other sources of noise. Fluctuations of the global factors may be defined in a manner similar to that for intrinsic protein noise, but we do not *a priori* know the magnitude of these fluctuations:

$$< \mu_{\mathrm{E}} >= 0$$
$$< \mu_{\mathrm{E}}(t)\mu_{\mathrm{E}}(t+\Delta t) >=< \mu_{\mathrm{E}}^2 > \delta(\Delta t) \qquad \textbf{[13]}$$

We want to use Eq. **11** to derive the noise properties of our fluorescent reporters, which can be accomplished by calculating deviations from steady state values: $\delta E \equiv E - 1$, $\delta G \equiv G - \alpha_{\mathrm{G}}$, $\delta R \equiv R - \alpha_{\mathrm{R}}$. Substituting these calculations in Eq. **11**, Fourier transforming, squaring, and inverse Fourier transforming yields the following result for the second moment of δG and the correlation between δG and δR:

$$< \delta G^2 >= \alpha_{\mathrm{G}}(b_{\mathrm{GFP}} + 1) + \frac{\alpha_{\mathrm{G}}}{\alpha_{\mathrm{R}}} < \delta G \delta R > \qquad \textbf{[14]}$$

Eq. **14** relates the total fluctuations around steady state in GFP to a contribution from intrinsic fluctuations and a contribution from global fluctuations, meaning that we can separate the total noise in GFP ($\eta_{\mathrm{g-total}}$) into intrinsic ($\eta_{\mathrm{g-int}}$) and global (η_{global}) components by the relation

$$\eta_{\mathrm{g-total}}^2 = \eta_{\mathrm{g-int}}^2 + \eta_{\mathrm{global}}^2 \qquad \textbf{[15]}$$

Here, the individual noise contributions are related to the reporter fluctuations:

$$\eta_{\mathrm{g-total}}^2 = \frac{< \delta G^2 >}{\alpha_{\mathrm{G}}^2} = \frac{< G^2 > - < G >^2}{< G >^2}$$
$$\eta_{\mathrm{global}}^2 = \frac{< \delta G \delta R >}{\alpha_{\mathrm{G}}\alpha_{\mathrm{R}}} = \frac{< GR > - < G >< R >}{< G >< R >} \qquad \textbf{[16]}$$

The biochemical parameters α_{G} and b_{GFP} can be obtained from

$$\eta^2_{\text{g-int}} = \frac{b_{\text{GFP}} + 1}{\alpha_{\text{G}}} = \eta^2_{\text{g-total}} - \eta^2_{\text{global}} \qquad \textbf{[17]}$$

By assuming the mean fluorescence in a single cell is proportional to the concentration of molecules in that cell, we can measure total and extrinsic noise for a population using Eq. **16**, and the intrinsic noise is given by Eq. **17**. Surprisingly, we find that even in fully induced cells the fluctuations in GFP are coming almost entirely from intrinsic noise (Fig. 4*b*).

Measurement of the Number of Molecules. Decomposing the noise in a gene into intrinsic and extrinsic components still does not characterize fluctuations with sufficient detail to build a microscopic simulation; the numbers of relevant proteins in each cell is still needed. We estimate this number using a method similar to that introduced by Rosenfeld *et al.* (4), where GFP fluorescence is compared between dividing cells.

The process of cell division can be viewed as a binomial process where each molecule of GFP is randomly and independently assigned to one of the two daughter cells. Letting N_1 and N_2 be the number of molecules in the first and second daughter, respectively, we assume that the following relations fully determine the statistics of this process: $N_{\text{Pair}} = N_1 + N_2$, $< N_1 >= \frac{N_{\text{Pair}}}{2}$, and $< N_1^2 >= \frac{N_{\text{Pair}}(1 + N_{\text{Pair}})}{4}$, where averages are taken over the cell population. Furthermore, we assume that the mean fluorescence values, gfp_1 and gfp_2, in both daughter cells are directly proportional to the number of GFP molecules. Based on these assumptions, N_{GFP} can be estimated as shown below from the fluctuations of intensity between dividing cells without requiring details of photon flux or quantum efficiency.

$$\frac{< (gfp_1 - gfp_2)^2 >}{< gpf_1 >< gfp_2 >} = \frac{< (N_1 - N_2)^2 >}{< N_1 >< N_2 >} = \frac{4 < N_1^2 > -4N_{\text{Pair}} < N_1 > + N_{\text{Pair}}^2}{N_{\text{Pair}}^2 / 4} = \frac{4}{N_{\text{Pair}}} \qquad \textbf{[18]}$$

Eq. **18** relates the distribution of cellular fluorescence values in the daughter cells to the total number of molecules present in the mother. Because the fluorescence of an average cell is somewhere between the undivided cell (N_{Pair}) and the daughter cells ($N_{\text{Pair}}/2$), we assume that $N_{\text{GFP}} = 3/4N_{\text{Pair}}$. Finally we measure gfp_1 and gfp_2 for a population of 70 pairs of fully induced cells in 30 μM TMG and find that N_{GFP}=790±210 molecules with error bars estimated by bootstrapping.

Stochastic Model Details. Because mRNA production is a large source of noise, we build a model where noisy events are dominated by mRNA processes and use protein levels to determine the instantaneous state of the system. The three main events in the model are mRNA production/degradation, protein degradation, and global noise.

We assume the number of proteins produced from an individual mRNA to be chosen from an exponential distribution by treating the decay of mRNA as a random Poisson process and assuming that the number of proteins translated is proportional to the lifetime. Because the lifetime of an mRNA is very short in relation to the timescales associated with fluctuations in protein level ($\tau_{1/2}$) we condense the three events (production, translation and degradation) into a single 'burst' event. We model this process by production and immediate decay of an mRNA whose net effect was the addition of a random number of new proteins to the system. We quantify the rate at which these bursts occur by dividing the rate of protein production by the mean number of proteins produced from an mRNA.

We model the loss of protein levels by random Poisson decay of individual proteins at a rate commensurate with that caused by dilution. This adds noise to the system that is not inherent in cell growth and ignores noise due to the partitioning process. However, the noise difference between the decay and dilution processes should not be important because the noise from mRNA bursts is dominant in our experiments. This could become relevant for dynamics that are heavily influenced by decay processes when there are few if any mRNA burst events.

Global noise is included as a multiplicative term on the production of GFP and LacY. To include this effect, the random walk process *E(t)*, described in Eqs. **11** and **13**, is simulated and included in the relevant mRNA production rates. Here the value $<\mu_E^2>$ is chosen so that the simulated value of η_{global} matches that measured experimentally. The statistics on the process are thus given by $\langle E \rangle = 1$ and $\langle \delta E(t)\delta E(t+\Delta t)\rangle = 2\eta_{\text{global}}^2 \exp(-\Delta t/\tau_{1/2})$. This is then numerically simulated by a discrete random walk, and the mRNA production rates of relevant proteins are multiplied by the resultant factor $E(t)$.

Further Experimental Verification of Model. To verify that the model is functioning properly, growth of cells from both ON and OFF histories is simulated for the equivalent of 20 h in a range of external concentrations of TMG. As can be seen in Fig. 6 the cells largely remain in their original states through the shaded region, whereas transitions between the two states occur more frequently past the edges of this region. The hysteretic effect is similar to published measurements of single cells (1), indicating

that the model is capturing many of the essential properties of the biological network. The greater than expected transition rate from OFF-to-ON in the deterministically bistable region could be due to increased sensitivity to parameter error in this region, or an overestimation of LacY noise. In Fig. 7 we compare the model results to experimental data for cells grown in 6 μM and 9 μM TMG, both in the deterministically bistable region, and we find reasonable agreement between these data.

To further test our model in the bistable region, we calculate the OFF-to-ON transition rate when cells are switching stochastically. This can be done by fitting the equation $f_0 e^{-\gamma t}$ to the fraction of cells remaining in the OFF state at time t, $f_{OFF}(t)$. The results for the switching rate γ(TMG) from this fitting are shown in Fig. 8 for both experimental data and simulations in the range of 3 - 60 μM TMG. The transition rate for both curves is near or equal to zero until approximately 9 μM TMG, which coincides with the upper boundary of the deterministically bistable region. Above 9μM TMG, the model and experimental transition rates rise together as extracellular TMG increases.

Parameter Robustness. In order to quantify the sensitivity of the model's output to variations in the deterministic and noise parameters, we construct a cost function that compares two population distributions. The cost function should return small values for nearly identical distributions and large values for very different distributions. We set $H(G, T, t)$ to be the fraction of initially uninduced cells in the bin centered at G units of GFP fluorescence after t hours of growth in T μM TMG. This is similar to the curves shown in Fig. 2. We define a cost function, X, similar to a χ^2 error, and evaluate it on the logarithmic-normalized histograms:

$$X = \Sigma_T \Sigma_t \Sigma_G \, log^2(H_{\text{experiment}}(G,T,t)/ H_{\text{model}}(G,T,t))$$

The sums are confined to the histograms shown in Figs. 3 *b-d*, and any terms with $H = 0$ were ignored in the sum. To estimate the sensitivity of the model to parameter error, each parameter was individually varied, and the cost between the model predictions and experimental measurements was calculated. We vary α, β, ρ, λ, N_{GFP} and b_{GFP} by the calculated errors shown in Table 2, while N_{LacY}, b_{LacY}, N_{LacI} and b_{LacI} are each varied by a factor of two because these parameter values were not measured directly. X_{high} and X_{low} represent the cost functions generated from simulations in which individual parameters are set to the upper and lower error boundaries, respectively. These calculations are shown for each parameter in Table 3. When all parameters are given by the values

indicated in Table 1 the cost is found to be $X_0 = 67.02$, which is close to the cost between two histograms generated from replicates of the same experiment: $X_{\text{Experimental}} = 45.50$.

Error on the parameters α, λ, and $\tau_{1/2}$ has the greatest impact on the model's predictions when compared to experiments. Even though the highest value of X is more than twice as large as the lowest, the qualitative predictions remain similar throughout the range of parameters. Variations in each parameter slightly change the shape of the simulated histograms. For example, increased (decreased) values of λ created a higher (lower) rate of OFF to ON transitions. While the shapes of each peak in the bimodal histogram of transitioning cells remains similar, the relative magnitudes of the peaks are changed. For changes in α, however, the peak of OFF cells decays with a rate similar to experiments, but the simulated transitioning cells either produce GFP too quickly or too slowly for the histogram to closely match experiments. The noise parameters N and b seem to affect prediction accuracy the least. Decreasing b_{LacY} even causes better agreement than the value estimated through GFP noise, suggesting that noise in the LacY levels might have been over-estimated.

1. Ozbudak, E. M., Thattai, M., Lim, H. N., Shraiman, B. I. & Van Oudenaarden, A. (2004) *Nature* **427,** 737-740.
2. Vilar, J. M., Guet, C. C. & Leibler, S. (2003) *J Cell Biol* **161,** 471-476.
3. Pedraza, J. M. & van Oudenaarden, A. (2005) *Science* **307,** 1965-1969.
4. Rosenfeld, N., Young, J. W., Alon, U., Swain, P. S. & Elowitz, M. B. (2005) *Science* **307,** 1962-1965.