A molecular noise generator

Ting Lu¹, Michael Ferry², Ron Weiss¹,⁴ and Jeff Hasty²,³,⁴

¹ Department of Electrical Engineering, Princeton University, J-319 E-quad, Princeton, NJ 08544-5263, USA
² Department of Bioengineering, University of California, San Diego, 9500 Gilman Drive MC0412, La Jolla, CA 92039, USA
³ Institute of Nonlinear Science, University of California, San Diego, 9500 Gilman Drive MC0412, La Jolla, CA 92039, USA
E-mail: rweiss@princeton.edu and hasty@bioeng.ucsd.edu

Received 16 May 2008
Accepted for publication 21 July 2008
Published 13 August 2008
Online at stacks.iop.org/PhysBio/5/036006

Abstract
Recent studies have demonstrated that intracellular variations in the rate of gene expression are of fundamental importance to cellular function and development. While such ‘noise’ is often considered detrimental in the context of perturbing genetic systems, it can be beneficial in processes such as species diversification and facilitation of evolution. A major difficulty in exploring such effects is that the magnitude and spectral properties of the induced variations arise from some intrinsic cellular process that is difficult to manipulate. Here, we present two designs of a molecular noise generator that allow for the flexible modulation of the noise profile of a target gene. The first design uses a dual-signal mechanism that enables independent tuning of the mean and variability of an output protein. This is achieved through the combinatorial control of two signals that regulate transcription and translation separately. We then extend the design to allow for DNA copy-number regulation, which leads to a wider tuning spectrum for the output molecule. To gain a deeper understanding of the circuit’s functionality in a realistic environment, we introduce variability in the input signals in order to ascertain the degree of noise induced by the control process itself. We conclude by illustrating potential applications of the noise generator, demonstrating how it could be used to ascertain the robust or fragile properties of a genetic circuit.

Introduction
Stochasticity is a ubiquitous feature of the biological networks that govern gene expression, signal transduction and metabolism. Arising from the random nature of the chemical reactions involved in biological function and the inherent bulk variability of the cellular environment, noise profoundly influences cellular function and development [1–9]. For example, noise can destroy the fidelity of signals propagating through networks, perturb a system from its optimal steady state or even possibly drive cells to dramatically different phenotypic states, including pathological ones seen in cancer and the transition to the genetic competence state in Bacillus subtilis [10, 11]. While the detrimental consequences of noise are perhaps the most straightforward to appreciate, noise can also be beneficial to biological systems in certain circumstances. Stochastic fluctuations randomize the traits manifested from genetically identical cells, generating different molecule numbers and even distinct physiological states. Consequently, this variability offers a mechanism for a population to increase its phenotypic variety and may confer greater adaptability of a species to stressful or changing environments [5–7, 12].

The specific consequences of noise can lead to simultaneous yet opposite demands. For example, selection could drive noise to a minimum in situations where it is a ‘perturbation’ away from an ideal state, but on longer timescales variability could give a selective advantage to the same organism. Such varied demands present a challenge for cells that must adapt in a context-dependent manner.
to environmental variations that occur on many different timescales. In order to begin to approach this fascinating area of inquiry, recent studies have focused on the characterization of the variability in terms of either the network architecture or the underlying mechanistic properties of a given network component [13–20]. Perhaps the most common motif for controlling noise is the ubiquitous negative feedback loop, which has been examined in the context of shaping the inherent noise spectra and increasing the stability of cellular states [21, 22]. Mechanistic properties for noise control include the modulation of the fluctuation frequency of a promoter state and the cascade length [23, 24]. While most of the studies to date have focused on noise as a potentially undesirable perturbation, there is no reason a priori to expect that such negative effects will play any greater role than positive effects. Over the past decade, concepts such as stochastic resonance [25, 26] have arisen in the physical sciences as paradigms for how noise can be beneficial to processes involving signal detection and transduction.

In the context of gene regulation, there is a clear need for experimental approaches that would permit the systematic tuning of the variability of a given intracellular protein. To meet this need, we propose a synthetic gene circuit, termed a molecular noise generator, that allows for independent external control of both the spectral properties and mean level of a desired protein. We first introduce a dual-signal system with transcription and translation independently controlled by two external signals. This architecture allows for the mean and variance of the system’s output protein to be regulated separately through the combinatorial modulation of the input signals. In order to increase the tuning spectrum, we further propose a triple-signal generator that harnesses tunable copy number through a well-characterized plasmid control mechanism. We then analyze the noise sources from input ‘devices’ and signals that are inevitably introduced by the signal transduction processes required for controlling the generator. We conclude with an example illustrating potential applications of the noise generator in an experimental context.

**Experimental details**

The stochastic simulations in this study were performed with the Gillespie algorithm [27], and included the reactions for mRNAs, proteins and relative signaling molecules. In this paper, statistical results were obtained by repeating the stochastic simulation many times and averaging the trials.

The Langevin equations used in this paper were linearized around their steady states. These linearized equations were transformed into Fourier space and solved exactly. Noise power spectrum was integrated in Fourier space and inverse transformed back to the real space. See the detailed calculations in the supplementary data available at stacks.iop.org/PhysBio/5/036006/mmedia.

**Results and discussion**

To design an effective and useful noise generator we need to consider several aspects including tunability, simplicity and economy. A good generator will allow adjustment of its output mean and variance along wide tuning regimes, be simple to build, easy to manipulate and be isolated as much as possible from the host’s intracellular environment. Ideally, the device would be generic enough to generate statistics for any desired output protein while not imposing a large metabolic burden on the host. A good generator should satisfy, or be the best compromise, all of the above criteria.

A dual-signal generator with tunable mean and variance

Recent studies on stochastic gene expression have explored the various mechanisms contributing to variability in gene expression [5]. For example, in prokaryotes noise predominantly comes from translation due to mRNA bursting rather than transcription [28]. This translational noise dominance encourages us to propose a simple noise generator as indicated in figure 1(A). Two signals ($S_1$ and $S_2$) individually control the transcription and translation of a gene D and would facilitate possible tuning of both the mean and variance of the protein P. Experimentally, transcription can be regulated through the modulation of an inducible promoter, such as the Lac promoter, while small RNAs with partial or full complementarity to the target sequence are good candidates for the regulation of translation [29]. One example of small non-coding RNA molecules affecting translation is known as a riboregulator which can be designed to disrupt the translation of a target mRNA by hybridizing to the target and blocking ribosomal access [30, 31]. The binding of these
riboregulators can be controlled by small molecule effectors and thus provides a mechanism for the precise regulation of a target mRNA’s translation. We propose using a system analogous to that described by Bayer and Smolke [30]. In this system the expression of a RNA aptamer will be controlled by the pBAD promoter which is inducible by arabinose. Once the RNA aptamer is produced in the cell its binding to the target mRNA can be controlled by the level of the small molecule theophylline. The level of theophylline in the cell will be controllable by the experimenter and will represent our second input signal. We will thus be able to modulate the translation of our target mRNA through the concentration level of this small molecule. In the following we are going to use simple analytical tools as well as simulation to study the tunability, sensitivity and other properties of our proposed noise generator.

The dynamics of such a generator can be described by the following Langevin equations:

\[
\dot{M} = F_m - H_m + \xi_m, \quad \dot{P} = F_p - H_p + \xi_p,
\]

where \(M\) and \(P\) represent mRNA and protein, \(F_m\) and \(F_p\) are the production and degradation rates of the mRNA, and \(H_m\) and \(H_p\) are those of the protein, respectively. The last terms, \(\xi_m\) and \(\xi_p\), reflect stochastic fluctuations of the mRNA and the protein, respectively. For simplicity, we assume both are Gaussian white noise, i.e., \(\langle \xi_i(t) \xi_j(t') \rangle = 2 \delta_{i,j} \delta(t-t') (i, j = m, p)\). Different dosage responses allow multiple choices for the combinations of the signals’ response functions. Here we use Hill-type inductions for both of the signals, \(\dot{F}_m = f_m \xi_m\) with \(f_m(s_m) = c_1(1 + f_1 s_m^N)/(1 + s_m^N)\) and \(\dot{F}_p = f_p M\) with \(f_p(s_p) = c_2(1 + f_2 s_p^N)/(1 + s_p^N)\), where \(c_1\) and \(c_2\) are the prefactors, \(f_1\) and \(f_2\) are the folds changed upon a saturated induction, and \(N\) and \(\nu\) are the Hill coefficients. First-order degradation assumptions are further made for both of the molecules, which are \(\dot{H}_m = d_m M\) and \(\dot{H}_p = d_p P\), with \(d_m\) and \(d_p\) the corresponding degradation constants.

This generator produces protein \(P\) with a mean \(\bar{P} = \frac{f_m f_p}{d_m d_p}\) and mRNA \(M\) with a mean \(\bar{M} = \frac{f_m}{d_m}\). The mRNA and protein noise can be measured by variance as

\[
\sigma_m^2 = \frac{f_m}{d_m}, \quad \sigma_p^2 = \frac{f_m f_p}{d_m d_p} \left(1 + \frac{f_p}{d_p + d_m}\right)
\]

and also by the Fano factor [32], defined as the ratio of the variance to the mean, as

\[
\phi_m = 1, \quad \phi_p = \left(1 + \frac{f_p}{d_p + d_m}\right).
\]

Here the variances \(\sigma_m^2\) and \(\sigma_p^2\) quantify the absolute dispersions of the probability distributions, in other words, the absolute noise magnitudes, while the Fano factors, \(\phi_m\) and \(\phi_p\), measure the deviations of the mRNA and the protein distributions from a corresponding Poisson distribution and are in effect the relative noise magnitudes.

Both the mean and variance of the output protein can be adjusted by the signals \(S_1\) and \(S_2\) through the modulation of the Hill functions for the transcription \(f_m(s_1)\) and the translation \(f_p(s_2)\). Figure 2 explicitly shows the mRNA and protein noise profiles of this dual-signal generator. Panels (A) and (B) are the landscapes of the mRNA and the protein noise levels with respect to the input signals \(S_1\) and \(S_2\). The mRNA noise depends only on the signal \(S_1\), while the protein noise depends on both. Furthermore, the protein noise (color surface) in figure 2(B) is generally larger than that of a Poisson process with the same mean (white surface). The corresponding Fano factors (panels (C) and (D)) indicate that the mRNA has a simple Poisson distribution regardless of the input signals but the protein distribution can be much wider especially for a small input of \(S_2\). These are consistent with our design: the signal \(S_1\) regulates the mRNA noise level through transcription and also the protein noise since mRNA and protein fluctuations are coupled through translation. In contrast, the signal \(S_2\) only affects the protein production rate and has no effect on the mRNA. Panel (E) shows the contour plots for the protein noise (black lines filled with colors) and the mean (dashed white lines) with respect to the signals \(S_1\) and \(S_2\). The noise level varies along the while lines (the mean contour lines), which indicates the noise tunability for a fixed mean. Similarly, the mean level varies along the black lines (the noise contour lines), which shows the noise level can also be modulated by the two signals while the noise remains the same. Panels (F) and (G) show how the protein noise depends on the signals \(S_1\) and \(S_2\) individually (the other signal is fixed in each case). As shown, \(S_1\) modulates the noise magnitude but does not change the Fano factor, while \(S_2\) modulates both of them ((F), (G) and their insets). Panel (H) shows how the protein noise level varies for different input signals while its mean remains the same. To test these analytical results, we performed numerical simulations using the Gillespie algorithm, and the results (red squares in (F), (G) and (H)) show that the protein noise levels agree well with those from analytical calculations.

To better illustrate the tunability of the protein noise, we directly plotted and analyzed the mRNA and protein data from the simulations of the generator. Figures 3(A) and (E) are the mRNA–protein density plots for two different sets of signal inputs (corresponding to the two black squares indicated in figure 2(G)). The statistics for these two cases are presented in (B) and (F) for the mRNA distributions and (C) and (G) for the protein distributions. Clearly, these two cases have distinct variabilities in their protein distributions: one is quite wide and the other is much more narrow. This can also be verified by typical protein number trajectories for these two sets of inputs (D and H). These comparisons give a direct illustration of this generator’s noise tunability.

Expanding tuning space through DNA copy number control

One of the most desirable characteristics of a noise generator is its tunability. In reality, we often need both a specific mean and a specific variance. As a reasonable estimation, the Fano factor (relative noise magnitude) of our dual-signal design has a 50–100-fold range from its minimum to maximum and experiences a substantial change [28]. However, due to the combinatorial control of the generator, tuning the Fano factor affects the mean protein level simultaneously, which imposes...
Figure 2. Analysis of the dual-signal noise generator. (A) Noise (variance) profile of the mRNA with respect to the two controlling signals $S_1$ and $S_2$. (B) Noise profile of the protein with respect to the two signals. The white surface is generated using Poisson statistics for a comparison with this generator’s noise profile (color surface). (C) The Fano factor (relative noise) of the mRNA remains a constant unit regardless of the levels of the signals. (D) The Fano factor of the protein. The increase of the signal $S_2$ results in the reduction of the Fano factor. Similar to before the white surface is the profile obtained from Poisson statistics. Parameters for (A)–(D) are chosen as: $c_1 = 1$, $c_2 = 50$, $f_1 = 50$, $f_2 = 200$, $d_m = 0.5$, $d_p = 0.02$, $u = 1$ and $v = 1$. (E) Contour plots for the protein noise level (black lines filled with colors) and the mean level (white dashed lines) with respect to the two signals. (F) Dependence of the protein noise on the signal $S_1$ for a fixed value of the signal $S_2$. Inset is the corresponding Fano factor. All the parameter values are the same as before except the fixed signal $S_2 = 3$. (G) Dependence of the protein noise on the signal $S_2$ for a fixed value of the signal $S_1$. Inset is again the corresponding Fano factor. All the parameter values are the same as before except the fixed signal $S_1 = 3$. (H) Output noise for different levels of signals with the mean protein level fixed. The inset is the relationship between the signals $S_1$ and $S_2$ in order to achieve a fixed mean protein level. (F)–(H) The green curves and red squares correspond to the analytical and simulational results, while the dashed blue curves are the profiles from Poisson distributions.

Figure 3. Comparison of two simulated cases for the dual-signal noise generator. (A)–(D) and (E)–(H) are for two limits that have the same mean but different noise outputs, one large and the other small, corresponding to the two black squares indicated in figure 2(G). Here (A) and (E) are the mRNA–protein density plots using simulation data. (B) and (F) are the corresponding mRNA distributions and (C) and (G) are the protein distributions. (D) and (H) are typical output protein trajectories for each case.
constraints to the generator’s tuning space and hence leads us to search for a better design.

We now extend the design by adding a plasmid copy number control mechanism that allows gene copy number tunability. As shown in figure 1(B), we introduce a controllable plasmid inhibitor, such as RNA I inhibitor, to attune the copy number of a plasmid containing the interested gene [33]. The dynamics of such a copy number regulation component can be described by the following equations:

\[
\begin{align*}
D &= \beta_d e^{-R/K} D - \gamma_d D + \xi_d, \\
R &= \beta_r D - \gamma_r R + \xi_r,
\end{align*}
\]

where \(D\) and \(R\) are the concentrations of the plasmid (equivalent to the gene) and the plasmid inhibitor RNA I, \(\beta_d\) and \(\gamma_d\) are the rate coefficients for the plasmid production and degradation, \(\beta_r\) and \(\gamma_r\) are those for the inhibitor, respectively. \(\xi_d\) and \(\xi_r\) are again the corresponding Gaussian white noise terms that represent the stochastic nature of the regulation dynamics. Here we have adopted a ColE1 vector as an example for our study with the term \(e^{-R/K}\) in the plasmid production rate representing the negative regulation of the vector by the plasmid inhibitor [33]. Since the inhibitor level is further determined by its production and degradation rates, we are able to implement the adjustment of the DNA copy number by changing the inhibitor’s production rate \(\beta_r\). This is feasible by introducing another external signal, say \(S\), for control (same as before, it can be modeled with another Hill function as \(\beta_r = c_0(1 + f_0s^0)/(1 + s^0)\)). Experimentally, we can simply replace the original promoter in the vector with an inducible promoter to achieve this goal. With the modulation of this production rate, the mean copy number of the gene \(\bar{D} = \frac{\gamma^K}{\bar{\beta}} \log \left( \frac{\bar{\beta}}{f_0} \right)\) can vary typically from 5 to 50, which is a ten-fold range for the wild-type ColE1 vector we used here.

For some other vectors, copy number could even change from a single one to hundreds [33]. Such a variation confers a much wider spectrum for noise modulation to the new proposed generator.

Calculations combining the plasmid copy number mechanism (4) with the gene expression process (1) yield the final protein statistics of the new generator, specifically its mean and the Fano factor, as

\[
\mathcal{P} = \left[ \frac{\gamma^K \log(\beta_d/\gamma_d)}{d_m d_p} \right] f_m f_p \beta_r, \\
\phi_p = 1 + \frac{f_p}{d_m + d_p} + f_m f_p \left( \frac{C_1}{\beta_m} + C_2 \right),
\]

where \(C_1\) and \(C_2\) are constants independent of the signaling responses (see the supplementary data available at stacks.iop.org/PhysBio/5/036006/mmedia), \(f_m\) is the \(S_1\)-controlled transcriptional rate for a single copy of the gene, \(f_p\) is the translation rate tuned by the signal \(S_2\) and \(\beta_r\) is the \(S_0\)-regulated synthesis rate of the DNA inhibitor RNA I.

To illustrate the advantages of the triple-signal design, we have compared its tuning space with that of the dual-signal system and that of a gene regulated simply by an inducible promoter. We assume for simplicity that all of these cases are integrated into the same plasmid backbone but only the triple-signal design has the copy number tunability. Figure 4 shows that the mean protein level of the transcription-regulated gene can be tuned for a limited range but its Fano factor (relative noise) remains nearly constant. The mean protein and the Fano factor of the dual-signal design are both tunable in wider regimes but are constrained by the inability to achieve a large mean and a low Fano factor. The tuning space for the triple-signal design is dramatically expanded because of the inherent tuning flexibility due to gene copy number control.

The copy number control apparatus greatly enlarges the tunability of the noise generator but, as a tradeoff, also introduces an extra noise source from DNA copy number fluctuations to the final output. In principle, we could introduce a fourth signal to regulate the variability of the plasmid copy number. Additionally, this apparatus noise could be eliminated by integrating the vector into the host chromosome, which is achievable for the dual-signal design. Both the dual-signal and triple-signal designs therefore have their own advantages and drawbacks. A hybrid of these two systems, using a bacterial artificial chromosome, would combine both advantages: without an input signal the generator system would be maintained as a constant single copy regardless of the physiological state of a host; however, the copy number could be increased to a very high level upon signal induction [34].

![Figure 4. Tuning space for different noise generator designs.](image-url)

**Noise from signals and input apparatuses**

We have adopted idealized inputs and perfect signal transduction for analyzing the noise generator in the above sections. However, noise exists inevitably in signaling. It arises from noisy external inducers, random fluctuations of molecules engaged in transmission pathways, and intracellular and extracellular fluctuating microenvironments. All of these fluctuations accumulate and propagate to downstream molecules in a network which, in this case, is our noise generator.
As a simple but generic illustration of the noise contribution due to signaling, we equip the generator with a common transduction motif for the modulations of the reaction rates as requested in our designs [35]. Illustrated in figure 1(C), an inducible signal $S$ modulates the conformational change of the intermediate molecule $U$ (between $U_{↓}$ and $U_{↑}$). Different conformations have distinct production rates for the downstream molecule $X$ and, as a result, tuning the signal leads to a dramatic change for the level of the molecule $X$. In our designs, the RNA I inhibitor production, DNA transcription and mRNA translation can all be regulated by this type of input motifs. The variability of the intermediate molecule in such a device is given by (see the supplementary data available at stacks.iop.org/PhysBio/5/036006/mmedia)

$$\sigma_u^2 = \frac{\kappa_d U^2}{(\kappa_b S^0 + \kappa_d)(\kappa_b S^0 + \kappa_d + d_u)} \left( \frac{\sigma^2}{2} \right) + \frac{\kappa_d \bar{U}}{\kappa_b S^0 + \kappa_d}$$

where $\bar{U}$ is the mean value of the molecule $U$ in the state $U_{↓}$, $\bar{S}$ and $\sigma_u$ are the mean and variance of the signal $S$ and all other kinetic parameters are shown in figure 1(C). Interestingly, the above expression splits into two terms representing different noise origins. The first term is the contribution originating from the randomness of the signal $S$ itself, while the second term accounts for the intrinsic stochasticity of the underlying reaction in the input ‘device’. For the generator controlled by these devices, both the signal and device noises are further propagated through and contribute additional components to the final output. For example, the spectrum of the terminal noise equation (2) is shaped as (see the supplementary data available at stacks.iop.org/PhysBio/5/036006/mmedia)

$$\frac{\Delta |P(\omega)|^2}{P} = \frac{2d_p}{\omega^2 + d_p^2} + \frac{2\left( \frac{d_p}{\omega^2} \right) d_p d_m}{(\omega^2 + d_p^2)(\omega^2 + d_m^2)}$$

$$+ \frac{2\chi_1 u^{-1} \frac{d_m}{\omega} d_p d_m}{\omega^2 + d_m^2} \frac{\sigma_t^2}{\sigma_i^2} + \frac{2\chi_2 u^{-1} \frac{d_p}{\omega} d_p d_m}{\omega^2 + d_p^2} \frac{\sigma_t^2}{\sigma_i^2}$$

$$+ \frac{2\chi_1 u^{-1} \frac{d_m}{\omega} d_p d_m}{\omega^2 + d_m^2} \frac{\sigma_t^2}{\sigma_i^2} + \frac{2\chi_2 u^{-1} \frac{d_p}{\omega} d_p d_m}{\omega^2 + d_p^2} \frac{\sigma_t^2}{\sigma_i^2}$$

which includes the transcriptional and translational noise (the first two terms) that are exactly the same as those in the idealized case. In addition, the spectrum indicates extra variability contributions coming from the stochastic transduction of the noisy signals $S_1$ and $S_2$: the third and fifth terms are contributed by the fluctuations of the signals $S_1$ and $S_2$ themselves and the fourth and sixth terms are due to the random nature of signaling in the corresponding devices. Propagated from this multi-step cascade, these four
signaling-related noise components may have been amplified as is the case in some ultrasensitive signal transduction networks [35]. As a result, these extra noise contributions in a realistic generator might be considerable and need be taken into account for a quantitative prediction of the final noise. Further discussions about sensitivity, economy and the fine-tune ability of the generator are available in the supplementary data available at stacks.iop.org/PhysBio/5/036006/mmedia.

An application example: noise generator drives a bistable switch

Our designed generator offers a tool for tuning a protein’s variability, testing the robustness of synthetic gene circuits and even perturbing the intracellular environment. Here we give an application example to illustrate the potential value of our noise generator.

As indicated in figure 5(A), a self-regulated bistable switch makes proteins P and G. Protein P is an activator regulating itself and protein G serves as a reporter. We have a dual-signal generator making the same activator P as the switch, which leads to coupling between the generator and the switch. The overall activator pool thus contains a subset made by the generator. Noise from this generator is then pumped into and propagated through the switch. Due to its noise tunability, the generator pool subset has the ability to independently change its variability over a large range while its mean level stays constant. The kinetics of the switch, such as the transition frequency between two stable states, is then affected by the generator through the modulation of its noise magnitude.

To elucidate this noise effect of the coupled system a stochastic simulation was performed [27]. It is necessary to first find a parameter regime where the coupled positive-feedback switch has two stable states regardless of its noise level. By rescaling the system’s kinetic equation and analyzing its phase diagram, we are able to get an appropriate set of parameters for multistability as shown in the inset of figure 5(B) (see the supplementary data available at stacks.iop.org/PhysBio/5/036006/mmedia). We then performed the simulation for 10⁶ unit time and counted the number of transition events (from any state to the other) for each given noise level pumped from the generator. The result (figure 5(B)) clearly shows that the noise increment changes the switch’s kinetics and leads to a monotonic enhancement of the transition frequency of the switch. To further illustrate these results, we plotted typical trajectories for the reporter protein and the generator-made protein in figures 5(C)–(F). Comparing the trajectories for the small noise case ((C) and (D)) with those in the larger noise case ((E) and (F)) confirms our speculation that pumped noise from the generator tunes the transition frequency of the bistable switch.

Conclusion and outlook

Complementary to exploring and characterizing noise in gene expression, we studied the possible control and entrainment of noise in this paper. We designed two versions of the noise generator allowing for wide-range modulation of both the mean and the variance of a protein, which was confirmed and supported by our analytical and simulational studies.

The generator’s terminal noise arises from multiple sources that can be classified as intrinsic, extrinsic and external contributions. For example, noise due to transcription and translation are intrinsic components, fluctuations of DNA copy numbers are extrinsic sources and external noise results from noisy input signals. However, there are still some sources, such as cellular growth and division, that were not explicitly studied (but have been implicitly incorporated in degradation processes). These sources could be important in some situations [4]. In our study we have employed the Langevin equations to describe the stochastic dynamics, which works well when the mean molecule numbers are not too small. But a master equation description would be mathematically more accurate. Nevertheless, our analysis has already provided a good quantitative understanding of the noise profile and clearly supports our idea about designing a generator for producing controllable noise.

Due to multiple levels of regulation in gene expression, noise generators can certainly be designed using other mechanisms, such as post-translational protein–protein and protein–mRNA interactions, that were not discussed here [36]. Treated in a mathematical way, protein fluctuations in our two designs are more or less Gaussian noise. However, with an analogy to different random number generators in the computational sciences, output protein noise can have other statistical distributions by employing different regulatory mechanisms [37]. For example, slow modulation of promoter states can possibly result in output proteins obeying a binomial distribution.

This generator offers a way to perturb cellular regulatory dynamics and the intracellular environment. It would help us to better understand gene regulation and make it possible to test a gene circuit’s robustness in living cells.

Acknowledgments

We thank Matthew Bennett for stimulating discussions. This work is supported by National Institutes of Health Grants GM69811-01 and GM082168-01.

References


[27] Levine E, Zhang Z, Kuhlman T and Hwa T 2007 Quantitative characteristics of gene regulation by small RNA *PLoS Biol.* 5 e229


Supporting Information for A Molecular Noise Generator

Ting Lu\textsuperscript{1}, Michael Ferry\textsuperscript{2}, Ron Weiss\textsuperscript{1}, and Jeff Hasty\textsuperscript{2,3}

\textsuperscript{1} Department of Electrical Engineering, Princeton University, Princeton, NJ 08544, USA
\textsuperscript{2} Department of Bioengineering, Institute of Nonlinear Science, University of California, San Diego, La Jolla, CA 92093, USA

Noise output for the triple-signal design: Derivation of Eq. 5

We assume that the DNA copy number control dynamics Eq. 4 in the main text has a steady state satisfying

\[ \beta_d e^{-R/K} D - \gamma_d D = 0 \]
\[ \beta_r D - \gamma_r R = 0, \]

(8)

from which we have the mean values for the inhibitor and the DNA copy number as

\[ \bar{R} = K \log(\frac{\beta_d}{\gamma_d}) \]
\[ \bar{D} = \frac{\gamma_r K}{\beta_r} \log(\frac{\beta_d}{\gamma_d}) \]

(9)

By linearizing Eq. 4 around the above steady state, i.e. \( \Delta D = D - \bar{D} \) and \( \Delta R = R - \bar{R} \), we have the corresponding linear equations for small changes of the DNA copy number and the inhibitor as

\[ \Delta \dot{D} = -\frac{\beta_d \bar{D}}{K} e^{-R/K} \Delta R + \xi_d \]
\[ \Delta \dot{R} = \beta_r \Delta D - \gamma_r \Delta R + \xi_r, \]

(10)
where the Gaussian noise terms $\xi_d$ and $\xi_r$ follow the relationships: $\langle \xi_d(t) \rangle = 0$, $\langle \xi_r(t) \rangle = 0$, $\langle \xi_d(t) \xi_r(t') \rangle = 0$, $\langle \xi_d(t) \xi_d(t') \rangle = 2 \gamma_d \delta(t - t')$ and $\langle \xi_r(t) \xi_r(t') \rangle = 2 \gamma_r \delta(t - t')$. Fourier transform of the above equations together with the Gaussian noise properties for $\xi_d$ and $\xi_r$ yields the power spectrum of the copy number variability as

$$
\langle |\Delta D(\omega)|^2 \rangle = D \left\{ \frac{2}{\omega^2} \left[ \frac{\gamma_d^2}{\bar{m}} (\log(\frac{\bar{m}}{\gamma_d}))^2 + \gamma_d \omega^2 \right] \right\}
$$

To understand how DNA copy number fluctuations propagate to the protein level upon expression, we have analyzed the mRNA and protein kinetics which are similar to Eq. 1 in the main text

$$
\dot{M} = f_m D - d_m M + \xi_m
$$
$$
\dot{P} = f_p M - d_p P + \xi_p
$$

where $\xi_m$ and $\xi_p$ are again Gaussian noise: $\langle \xi_m(t) \rangle = 0$, $\langle \xi_p(t) \rangle = 0$, $\langle \xi_m(t) \xi_m(t') \rangle = 0$, $\langle \xi_m(t) \xi_p(t') \rangle = 2 d_m \dot{M} \delta(t - t')$ and $\langle \xi_p(t) \xi_p(t') \rangle = 2 d_p \dot{P} \delta(t - t')$.

Steady state calculation of the above equation gives the mean value of the protein as

$$
\bar{P} = \frac{f_m f_p D}{d_m d_p}.
$$

Given the average copy number from the previous calculation, we have an expression for the mean protein level as

$$
\bar{P} = \left[ \frac{\gamma_r K \log(\beta_d/\gamma_d)}{d_m d_p} \right] \frac{f_m f_p}{\beta_r}
$$

Furthermore, linearization of Eq. 12 leads to the variability of the final protein as

$$
\langle |\Delta P(\omega)|^2 \rangle = \frac{\langle |\xi_p(\omega)|^2 \rangle}{(d_p^2 + \omega^2) + \frac{f_p^2 \langle |\xi_m(\omega)|^2 \rangle}{(d_m^2 + \omega^2)^2}} + \frac{f_p^2 f_m^2 \langle |\Delta D(\omega)|^2 \rangle}{(d_m^2 + \omega^2)(d_m^2 + \omega^2)}
$$

where $\langle |\xi_m|^2 \rangle = 2 f_m D$ and $\langle |\xi_p|^2 \rangle = 2 f_p M$.

The variability of the protein Eq. 14 and the DNA copy number Eq. 11 further lead to the expression for the protein noise which includes variability due to transcription, translation, and the DNA copy number control. Inverse Fourier transform of the protein noise expression $\langle |\Delta P|^2 \rangle = \int_{-\infty}^{\infty} \frac{d\omega}{2\pi} \langle |\Delta P(\omega)|^2 \rangle$ gives Eq. 5 in the main text.

$$
\frac{\langle |\Delta P|^2 \rangle}{P} = \left( 1 + \frac{f_p}{d_p + d_m} \right) + f_m f_p \left( \frac{C_1}{\beta_r} + C_2 \right)
$$
where the coefficients $C_1$ and $C_2$ are given by

$$C_1 = 2d_md_p \int_{-\infty}^{+\infty} \frac{d\omega}{2\pi} \frac{\gamma_d^2 \gamma_r^2 \log^2 \left( \frac{\beta_r}{\gamma_d} \right)}{\left( \omega^2 - \gamma_d \gamma_r \log \left( \frac{\beta_r}{\gamma_d} \right) \right)^2 + \gamma_r^2 \omega^2} (d_m^2 + \omega^2) (d_p^2 + \omega^2)$$

$$C_2 = 2d_md_p \int_{-\infty}^{+\infty} \frac{d\omega}{2\pi} \frac{\gamma_d \gamma_r \omega^2}{\left( \omega^2 - \gamma_d \gamma_r \log \left( \frac{\beta_r}{\gamma_d} \right) \right)^2 + \gamma_r^2 \omega^2} (d_m^2 + \omega^2) (d_p^2 + \omega^2)$$

which are independent of the tunable parameters $f_m$, $f_p$, and $\beta_r$.

### Noise from Signals and Input Apparatuses: Derivation of Eq. 6

The signal $S$ regulates the production level of the molecule $X$ through the modulation of the intermediate molecule $U$. According to the corresponding chemical reaction $U \downarrow + \theta S \xrightarrow{\kappa_b} \kappa_d U \uparrow$, we have the kinetics for the intermediate molecule in the state $U \uparrow$ as

$$\dot{U} = \kappa_b \bar{s}\theta (N_u - U) - \kappa_d U + \xi_u,$$

where $N_u$ is the total number of molecules, $U$ is the number of molecules in the state $U \uparrow$, and $\xi_u$ is a Gaussian noise term. From this equation, we can derive the steady state value $\bar{U} = \frac{\kappa_b \bar{s}\theta N_u}{\kappa_b \bar{s}\theta + \kappa_d}$ and a corresponding linearized equation for a small change of the number of molecules in the $U \uparrow$ state as

$$\Delta \dot{U} = \left[ \theta \kappa_b \bar{s}\theta^{-1} (N_u - \bar{U}) \right] \Delta s - (\kappa_b \bar{s}\theta + \kappa_d) \Delta U + \xi_u,$$

from which we can further obtain an expression for the power spectrum of the fluctuations

$$\langle |\Delta U(\omega)|^2 \rangle = \left[ \theta \kappa_b \bar{s}\theta^{-1} (N_u - \bar{U}) \right]^2 \frac{\langle |\Delta s|^2 \rangle + \langle |\xi_u|^2 \rangle}{\omega^2 + (\kappa_b \bar{s}\theta + \kappa_d)^2}$$

where $\langle |\xi_u|^2 \rangle = 2\kappa_d \bar{U}$ is fluctuation of the intermediate molecule and $\langle |\Delta s|^2 \rangle$ represents signal noise.

As a reasonable approximation, we assume that the signal’s noise has a characteristic correlation time (rather than an idealized gaussian noise that has zero correlation time). Such a signal can be simulated by

$$\dot{s} = \bar{\dot{s}}d_s - sd_s + \xi_s$$

3
where $\bar{s}$ is the mean level of the signal, $d_s$ is the decay rate for the noise correlation and $\xi_s$ represents fluctuations of the signal. Linearization of the above equation results in the following relation

$$\langle |\Delta s(\omega)|^2 \rangle = \frac{2d_s\sigma_s^2}{\omega^2 + d_s^2}$$

where we have used the relation for the variance of the noise $\sigma_s^2 = \int_{-\infty}^{+\infty} \frac{d\omega}{2\pi} (|\Delta s(\omega)|^2)$.

We therefore have the power spectrum for $U_i$ as

$$\langle |\Delta U(\omega)|^2 \rangle = \frac{2\kappa^2_d d_s \theta^2 \bar{U}^2}{(\omega^2 + (\kappa_b \bar{s}^2 + \kappa_d)^2)} \left( \frac{\sigma_s^2}{\bar{s}^2} \right) + \frac{2\kappa_d \bar{U}}{\omega^2 + (\kappa_b \bar{s}^2 + \kappa_d)^2},$$

integration of which leads to Eq. 6 in the main text

$$\sigma_u^2 = \frac{\kappa^2_b \theta^2 \bar{U}^2}{(\kappa_b \bar{s}^2 + \kappa_d) (\kappa_b \bar{s}^2 + \kappa_d + d_s)} \left( \frac{\sigma_s^2}{\bar{s}^2} \right) + \frac{\kappa_d \bar{U}}{\kappa_b \bar{s}^2 + \kappa_d}$$  \quad (22)

**Noise from Signals and Input Apparatuses: Derivation of Eq. 7**

Let us start with a signal transduction cascade as shown in Fig 6. The production rate of $X_i$ is regulated by the signal $S_i$ through the configuration modulation of the intermediate molecule $U_i$. Using the same reasoning as in previous sections, we have the following kinetic equations for the molecule $U_i$ and $X_i$

$$\dot{U}_i = A_i^+(U_i, s_i) - A_i^-(U) + \xi_{ui}$$

$$\dot{X}_i = C_i^+(U_i, X_{i-1}) - C_i^-(X_i) + \xi_{xi}$$

where $A_i^+$ and $A_i^-$ are the production and degradation rates for $U_i$, $C_i^+$ and $C_i^-$ are those for $X_i$, $\xi_{ui}$ and $\xi_{xi}$ are again Gaussian noise. Linearization of these two equations around the steady state gives the power spectrum of the fluctuations of $U_i$ and $X_i$ as

$$\langle |\Delta \tilde{U}_i(\omega)|^2 \rangle = \frac{A_i^{+2}[|\Delta \tilde{u}_i|^2 + |\xi_{ui}|^2]}{(\omega^2 + (\bar{A}_{ui} - A_i^{+})^2)}$$

$$\langle |\Delta \tilde{X}_i(\omega)|^2 \rangle = \frac{C_i^{+2}[|\Delta \tilde{u}_i|^2 + |\xi_{ui}|^2]}{(\omega^2 + C_i^{-2})(\omega^2 + (\bar{A}_{ui} - A_i^{+})^2)} + \frac{C_i^{+2}[|\Delta \tilde{X}_{i-1}|^2 + |\xi_{xi}|^2]}{(\omega^2 + C_i^{-2})}$$

which covers the main result in the previous section (Eq. 21). Here the expression $W_{iz_i}^\pm$ is the derivative of $W_i^\pm$ respecting to $z_i$, i. e., $W_{iz_i}^\pm = \frac{\partial W_i^\pm}{\partial z_i}$ ($W = A, C$, $z = x, u, s$).
Figure 6: A generic module for signal input. The cascade consists of a series of central nodes $X_i$ ($i = 0, 1, \ldots$) and a regulatory input device between every node pair. This figure depicts a repeating unit of the whole cascade. The production rate of $X_i$ is regulated by the signal $S_i$ through the configuration modulation of the intermediate molecule $U_i$.

Our dual-signal noise generator consists of transcription and translation and is therefore a two-unit cascade of Fig. 6. With the iterative expression of $\Delta \tilde{X}_i(\omega)$ in Eq. 25, we easily derive the following expression for the two-unit cascade

$$
(|\Delta \tilde{X}_2(\omega)|)^2 = \frac{C_{2x1}^2 |\xi_{21}|^2}{(\omega^2+C_{2x2}^2)} + \frac{C_{2x2}^2 |\xi_{21}|^2}{(\omega^2+C_{2x2}^2)} + \frac{C_{2u2}^2 A_{+2u}^+ |\Delta \tilde{s}_2|^2}{(\omega^2+(A_{2u2}^- - A_{+2u}^+)^2)} + \frac{C_{2u2}^2 |\xi_{21}|^2}{(\omega^2+C_{2x2}^2)} \left(\frac{\omega^2+(A_{2u2}^- - A_{+2u}^+)^2}{(\omega^2+C_{2x2}^2)} \right) + \frac{C_{2u2}^2 |\xi_{21}|^2}{(\omega^2+C_{2x2}^2)} \left(\frac{\omega^2+(A_{2u2}^- - A_{+2u}^+)^2}{(\omega^2+C_{2x2}^2)} \right) \left(\frac{\omega^2+(A_{2u2}^- - A_{+2u}^+)^2}{(\omega^2+C_{2x2}^2)} \right).
$$

(26)

Now we introduce sensitivity into this context. Sensitivity is often measured by the derivative of the logarithm of an output respecting to that of its input signal, which is, $\chi_i = d \log(P)/d \log(s_i)$ for $S_i$ signaling here. Simple calculations show that the sensitivities of the system for the two input signals $S_1$ and $S_2$ are given by

$$
\chi_1 = \frac{d \log X_2}{d \log s_1} = \frac{s_1 dX_2 dX_1 dU_1}{X_2 dX_1 dU_1 ds_1} = \frac{s_1}{X_2} \frac{C_{2x1}^+ C_{1u1}^+ A_{1s1}^+}{C_{2x2}^- C_{1x1}^- (A_{1u1}^- - A_{1u1}^+)},
$$

(27)

and

$$
\chi_2 = \frac{d \log X_2}{d \log s_2} = \frac{s_2 dX_2 dU_2}{X_2 dU_2 ds_2} = \frac{s_2}{X_2} \frac{C_{2x2}^+ C_{2u2}^+}{C_{2x2}^- (A_{2u2}^+ - A_{2u2}^-)}.
$$

(28)
and used the relation metabolic burden, sensitivity, and signaling noise. In the devices, we need to choose appropriate Hill coefficients to reach a compromise among and relating these conceptual ideas to the Hill functions underlying the signal transduction required. Given the common rule that a higher Hill coefficient indicates a higher sensitivity for accurate modulation, a low sensitivity and a correspondingly high metabolic burden is makes fine-tuning difficult. Conversely if we desire a low level of signal and device noise to save its metabolic cost by using only a small amount of a signal’s change for a large noise variability shift. However, it also simultaneously incurs large signal and device noise which makes fine-tuning difficult. Conversely if we desire a low level of signal and device noise for accurate modulation, a low sensitivity and a correspondingly high metabolic burden is required. Given the common rule that a higher Hill coefficient indicates a higher sensitivity and relating these conceptual ideas to the Hill functions underlying the signal transduction in the devices, we need to choose appropriate Hill coefficients to reach a compromise among metabolic burden, sensitivity, and signaling noise.

Substituting the reactions rates $A_i^+ = \kappa_{b1}s_i^\theta_1(N_{u_i} - U_i)$, $A_i^- = \kappa_{d1}U_i$, $C_i^+ = d_{x_i}X_i$, and $C_i^- = c_i [(N_{u_i} - U_i) + f_iU_i]X_{i-1}$ we finally have the expression of Eq. 7 in main text.

$$\frac{\Delta P(\omega)}{P} = \frac{2d_p}{\omega^2 + d_p^2} + \frac{2\left(\frac{dF_m}{dm}\right) d_p d_m}{(\omega^2 + d_p^2)(\omega^2 + d_m^2)} + 2\chi_2 \frac{\sigma^2_{s_2}}{(\omega^2 + d_p^2)(\omega^2 + d_m^2)} \frac{d_p d_{u_2}}{s_2^2}$$

$$+ \frac{2\chi_1 \mu^{-1} \left(\frac{dF_p}{dm} \frac{dF_m}{dm}\right) d_p d_m d_{u_3}}{(\omega^2 + d_p^2)(\omega^2 + d_m^2)}$$

(29)

where we have adopted the corresponding $X_i$ in our generator: $X_0 = D$, $X_1 = M$ and $X_2 = P$ and used the relation $A_{i\theta_1} = \frac{\theta_1}{s_i} A_i^+$ in our calculations. Here the effective degradation rates for $U_1$ and $U_2$ are given by $d_{u_1} = \kappa_{d1} + \kappa_{b1}s_i^\theta_1$ and $d_{u_2} = \kappa_{d2} + \kappa_{b2}s_i^\theta_1$, the Hill coefficients are given by $\theta_1 = \mu$ and $\theta_2 = \nu$, and the transcription and translation follows $F_m = c_1(N_{u_1} - U_1 + f_1U_1)D$ and $F_p = c_2(N_{u_2} - U_2 + f_2^{-1}U_2)M$.

The signaling feature in controlling the generator further encourages us to study the sensitivity of the system. As shown in the noise profile Eq. 29, for each signal, there is a positive correlation between the level of a output noise component due to signaling and the sensitivity of the corresponding signal pathway. Interestingly, this relation suggests a tradeoff between the economy and the fine-tune ability of the generator. A high sensitivity is needed to save its metabolic cost by using only a small amount of a signal’s change for a large noise variability shift. However, it also simultaneously incurs large signal and device noise which makes fine-tuning difficult. Conversely if we desire a low level of signal and device noise for accurate modulation, a low sensitivity and a correspondingly high metabolic burden is required. Given the common rule that a higher Hill coefficient indicates a higher sensitivity and relating these conceptual ideas to the Hill functions underlying the signal transduction in the devices, we need to choose appropriate Hill coefficients to reach a compromise among metabolic burden, sensitivity, and signaling noise.
A bistable switch coupled with the dual-signal generator: Phase diagram and parameter choices

The noise generator is coupled to the bistable switch through their mutually expressed transcription factor $P$. With the tunability of the generator, we are therefore able to modulate the noise level in the system, which consequently affects the kinetics of the switch. Specifically, we have studied how different noise levels in the system can affect the flipping frequency of the switch. We thus change the generator’s noise level while keeping the bistability of the switch intact. To accomplish this we must first find a regime where the system has two stable states by analyzing its coarse grained kinetics described by

$$
\dot{p} = \alpha_0 + \alpha \frac{Wp^2}{1 + Wp^2} + \alpha_g - \beta p
$$

where $p$ represents the amount of the transcription factor molecule $P$, $\alpha_0$ and $\alpha$ are the basal and activated production rates from the switch, $\alpha_g$ is the protein production rate of the generator, and $\beta$ is the degradation rate. The Hill function $\frac{Wp^2}{1 + Wp^2}$ comes from the cooperative binding of the activator. By introducing dimensionless quantities $\tilde{t} = ((\alpha_0 + \alpha_g)\sqrt{W})t$, $\alpha' = (\alpha_0 + \alpha_g)^{-1} \alpha$, $\beta' = ((\alpha_0 + \alpha_g)\sqrt{W})^{-1} \beta$ and $\tilde{p} = \sqrt{W}p$, the deterministic equation Eq.(30) can be recast as

$$
\dot{\tilde{p}} = 1 + \alpha' \frac{\tilde{p}^2}{1 + \tilde{p}^2} - \beta' \tilde{p},
$$

which contains only two kinetic parameters $\alpha'$ and $\beta'$. We are then able to draw a phase diagram (in the inset of Fig. 5B) and find the parameter space for bistability using the nonlinear dynamics tool.

In our stochastic simulation we chose a parameter set (listed in Tab. 1) leading to the system having two stable states. This set of parameters corresponds to $\alpha' = 20$ and $\beta' = 10$ for the deterministic model which is marked with a cross in the phase diagram.
Table 1: Chemical reactions for the stochastic simulation of a bistable switch coupled with a dual-signal generator. Here $M$ and $P$ are the mRNA and the protein produced by the switch, $M_g$ and $P_g$ are the mRNA and the protein made from the generator, $G$ is the reporter protein, $D$ is the unbound promoter site of the switch, $DP_2$, $DP_{g2}$ and $DPP_g$ are promoter-protein complexes. The rate coefficients for $D_g \rightarrow D_g + M_g$ and $M_g \rightarrow M_g + P_g$ are mutually changed to tune the noise level with the mean protein level kept as a constant.