Establishing Baseline Performance Scores for the CAGEN Robust Gene Response Challenge

Shaunak Sen
Caltech CDS

February 29, 2012

Summary. The specifications of the Robust Gene Response Challenge are to design a genetic circuit that ensures fast, robust expression of a fluorescent protein upon induction. Here, we work through the specifications for a reference design and establish a baseline performance score using both computational and experimental methods. This work has helped to refine the specifications and can provide tools and protocols to help participating teams in their designs.

1 Introduction

Significant progress has occurred in the design of genetic circuits with small number of components. However, a major challenge in genetic circuit design is to ensure robust operation of these small circuits so that they can be combined to build larger circuits [8]. The Critical Assessment of Genetically Engineered Networks (CAGEN) is a competition that aims to engender the design of large interconnected genetic circuits by sequentially issuing design challenges of increasing complexity. As a first step, the Robust Gene Response Challenge specifies the design of a circuit that can quickly generate at least a tenfold change in the expression of a fluorescent protein upon induction, with minimal variation between cells and across temperatures.

Previous investigations have studied circuit properties relevant to this Challenge, such as amplitude and noise at equilibrium and speed of response (for recent reviews on noise, see [6, 9]). These investigations have combined mathematical models with experimental measurements in both *E. coli* and yeast using a variety of techniques, including single-cell measurements using fluorescence microscopy, both static and dynamic, and flow cytometry as well as population-level measurements using platereaders.

These results highlight three features that are relevant for designing Robust Gene Responses. One, noise reduces as the amplitude increases [3]. Two, noise arises at different steps of gene expression and can be quantitatively tuned [7]. Three, negative feedback can reduce noise [4] and increase the speed of response [4, 10]. However, a characterization of a circuit in a manner that establishes a baseline performance metric for the Robust Gene Response Challenge has not been performed.

Here, we propose to work through the Challenge specifications for a simple reference design. First, we state a metric that can be used to score the designs. Second, we estimate this metric using a computational model of a protein production-degradation circuit. This computation shows that the metric behavior adequately captures relevant circuit properties such as the amplitude, noise, and speed of response. Third, we estimate a baseline score from experimental single-cell time-lapse fluorescence microscopy data acquired by inducing protein expression in *E. coli*. We have also used a microfluidic platform, that offers precise control over the induction time, to establish another baseline. Finally, we use the model to show that a slight modification of the metric can be used to estimate the score from

flow cytometric measurements. The work presented here as well as the associated scripts and data can help participating teams in their designs.

2 Defining the Performance Metric

Each participating entry is required to submit time traces at a nominal temperature and at temperatures 5% above and below this nominal value, with measurements from at least five individual cells from separate colonies at each temperature. This represents a total of 15 time traces of data (denoted $y_j(t), j = 1, 2, 3 \ldots 15$). One of these traces must be designated as a reference trace, and its equilibrium value after induction (M) should be at least tenfold larger than its uninduced value (M_0). The score for each time trace (S_j), is the integrated square error between the time trace and an idealized step response with amplitude M,

$$S_j = \int_{T_1}^{T_2} (y_j(t) - M)^2 dt, j = 1, 2, 3 \ldots 15.$$

Here, T_1 is the time where the reference trace reaches 10% of M, and T_2 is the time after which the variation in the reference trace is within 10% of M. Additionally, the reference trace is required to stay within 10% of M for a time duration $T_2 - T_1$ after T_2.

The score of the design is based on worst-case analysis and will be the highest among the scores of the individual traces. Additionally, we normalize this metric by the equilibrium amplitude of the reference trace. With this consideration, the final score is,

$$S = \frac{\max_j S_j}{M^2}.$$

This generates a single number that can be used to order the designs.

3 Computing the Performance Score for a Protein Production-Degradation Circuit

Next, we estimated the performance score using a model of a simple protein production-degradation circuit. In this model, the total level of a protein X depends on the balance between its production, modeled as a zero-order process that changes from a basal rate of α_0 to α upon induction, and its dilution due to cell growth, modeled as first-order processes with rate constant γ. Analytical expressions of the fold-change upon induction, amplitude and noise at equilibrium, and the response speed for this model are well-known. The fold-change upon induction is $\frac{\alpha}{\alpha_0}$, and the Challenge specifications require that this exceed a factor of ten. The equilibrium amplitude (M) and noise (η) depend on the ratio α/γ,

$$M = \frac{\alpha}{\gamma}, \eta = \frac{1}{\sqrt{M}},$$

and the response speed (τ) depends only on γ, $\tau = 1/\gamma$. In general, these properties will change with temperature, depending on how the process parameters depend on temperature. In this estimation of the performance score, we ignore the dependence on temperature.

Calculating performance score for this model requires trajectories of the concentration as a function of time. These are generated from stochastic simulations of the model, performed using the software package BioNets [2] (Fig. 1). The performance score generated using these traces is 0.66 ($N = 15$). Recalculating the performance score for another set of stochastic traces ($= 0.62, N = 15$) or for a larger set of traces ($= 0.64, N = 100$) yields a similar value.

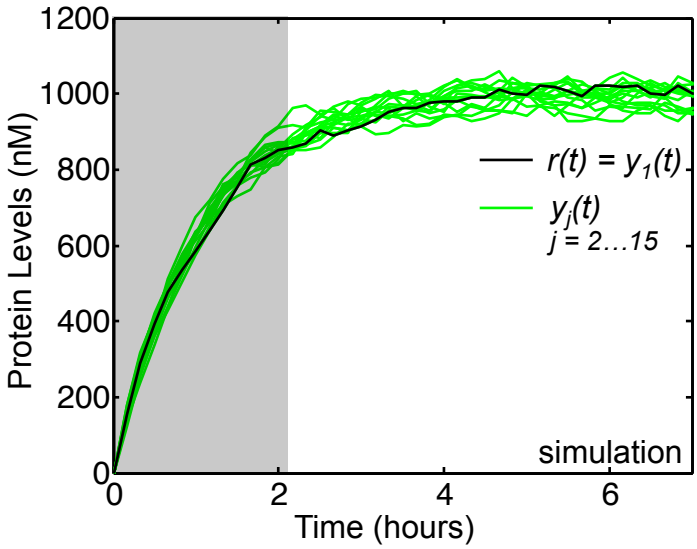

Figure 1: Trajectories of the protein production-degradation process generated using stochastic simulations. Black trace is the reference trace, arbitrarily picked as the first trace, and green traces are the other 14 traces. All traces are sampled at a 10 minute resolution. Shaded area indicates the time duration ($T_1 = 0$, $T_2 \approx 2$hr) over which the performance score integral is evaluated. The model parameters for this simulation are $\alpha = 1000$ nM/hr, $\alpha_0 = 1$nM/hr, $\gamma = 1$/hr.

To compare how the performance metric correlates with the circuit properties of noise and speed, we performed two additional set of simulations: First, the noise was increased for a fixed speed, by varying only α and not γ. Second, the speed was increased at a fixed noise, by varying γ and changing α so that α/γ is fixed. For each value of noise and speed in these cases, two sets of 15 trajectories were simulated. The performance score was calculated as outlined above. In addition, for comparison with a larger number of trajectories, the performance score was also calculated for an additional set of 100 trajectories. For all sets of trajectories, we find that the performance score worsens if noise increases or if response speed reduces (Fig. 2). The relation between the performance score and the response speed is especially strong. This is in accordance with intuitive expectations of how the score should vary relative to properties like speed and noise. Further, an implication of the broad correlation between the performance score and circuit properties like noise and speed is that population-level measurements characterizing amplitude and speed can also be used in early design iterations.

4 Measuring the Performance Score for a Inducible Protein Expression Circuit

As the next step in working through the Challenge specifications, we estimated the performance score from experimental measurements. We choose an inducible protein expression circuit in *E. coli* as a simple experimental system for this purpose. There are three parts to this task - constructing desired strain using molecular biology procedures, acquiring the data using time-lapse fluorescent microscopy, and analyzing the data using image processing tools.

We used two previously constructed strains [1, 5] to measure the single-cell protein induction dynamics (Fig. 3). In these strains, the addition of the inducer IPTG increases the expression of fluorescent proteins, CI-YFP in [5] and YFP in [1] (Fig. 2). For the measurement, cells were grown overnight in LB at 37° C. The overnight culture was diluted 1 : 100 in MGC (M9 minimal media containing 0.2%

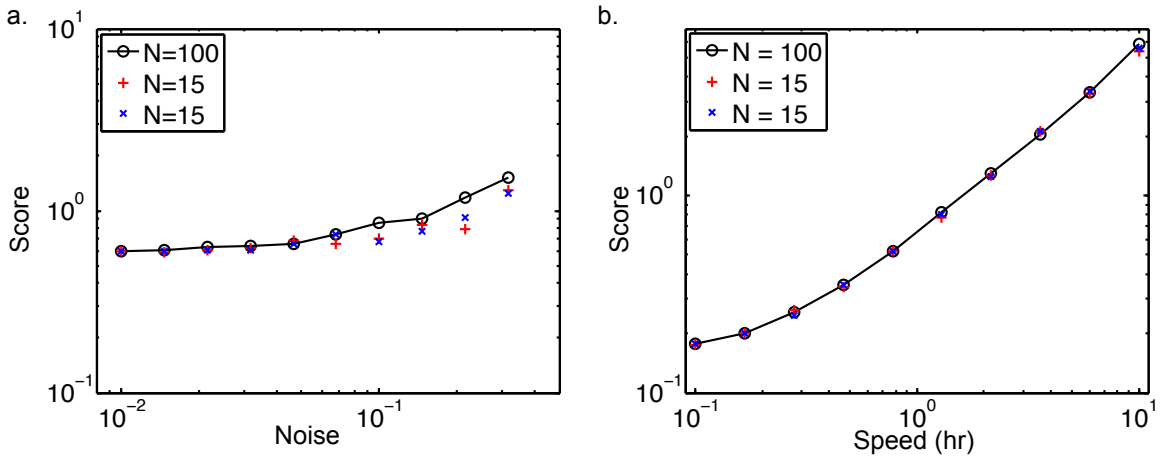

Figure 2: Performance worsens as noise increases or speed reduces. Performance scores are calculated as (a) noise is varied for fixed speed and, (b) speed is varied for fixed noise. Red crosses and blue circles are the performance scores for two sets of 15 trajectories. Black line connects points representing performance scores for a set of 100 trajectories, and are similar to the red crosses and blue circles. To change noise levels for fixed speed in (a), the values of α are varied logarithmically from 10nM/hr to 10^4nM/hr, for fixed γ. To change speed values for fixed noise in (b), the values of γ are varied logarithmically from 0.1/hr to 10/hr and α values are adjusted so that the ratio $\alpha/\gamma = 1000$ nM stays fixed.

Glycerol, 0.01% Casamino acids, 0.15 μ g/ml Biotin, and 1.5μ M Thiamine), and grown at 32°C for 3 hours. For time-lapse imaging, cells were placed on agarose pads (1.5% Low melting point Agarose in MGC) with and without IPTG and then placed under the microscope. The images were acquired every 10 minutes, at a nominal temperature, chosen to be 32°C as in [5], and at temperatures 5% (≈ 2°C) above and below this nominal value. This experimental data is in the form of a time-sequence of images. Processing these images to estimate fluorescence in each cell and how this changes over time was done through a custom software[1].

For the first strain, the reference trace reaches its equilibrium value of ≈ 60 a.u. at $T = 160$ minutes (Fig. 3a., black line). In the absence of induction, the equilibrium value for a trace is ≈ 2 a.u. Therefore, the reference trace exhibits a greater than ten-fold induction. Further, the time after which it stays within 10% of this equilibrium value is $T_2 = 150$ minutes, and so it also maintains this equilibrium value for the required time. The performance score for this dataset is 1.45. Similarly, the reference trace for the second strain reaches its equilibrium value of ≈ 45 a.u. at $T = 100$ minutes (Fig. 3b., black line). In the absence of induction, the equilibrium value of a trace is ≈ 1 a.u. Therefore, this reference trace exhibits a greater than ten-fold induction. As the time after which it stays within 10% of this equilibrium value is $T_2 = 80$ minutes, the reference trace holds the equilibrium value for the required duration of time. The performance score for this dataset is $S = 1.22$. Both performance scores are largely similar. This is to be expected as even though one of the datasets appears more variable than the other, it is also faster. Therefore, these balance each other and give rise to similar score. Interestingly, these are quite similar to the simulated ones.

In these measurements, the cells are induced just before imaging starts. An alternative to this is to use the microfluidic CellASIC ONIX platform[2] that allows cells to be induced at a defined time point during imaging. Therefore, we also repeat the experimental measurements in this fashion. For this,

[1]Software courtesy of Prof. M. B. Elowitz. Contact shaunak@cds.caltech.edu for more information.

[2]We thank Prof. E. Klavins and members of his laboratory at the University of Washington for their help with using this platform in acquiring data.

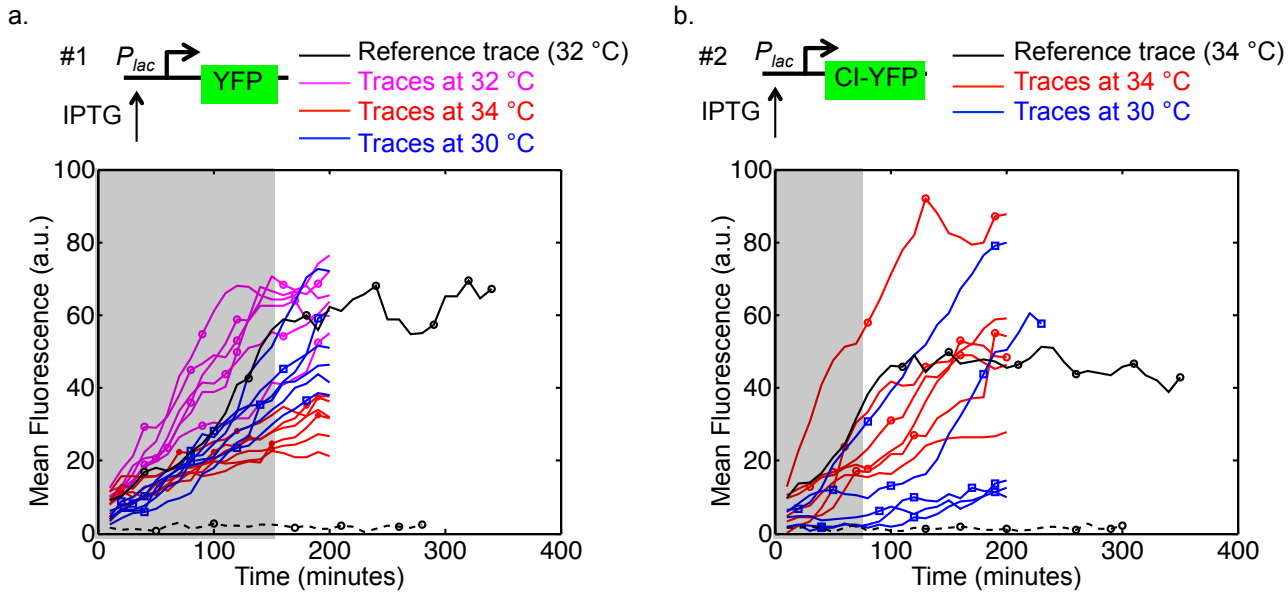

Figure 3: Single-Cell Dynamics of Inducible Protein Expression Circuits. Solid black line represents the reference trajectory. Dashed black line is a trajectory in the absence of induction. Red, magenta and blue colors represent the temperatures 34°C, 32°C, and 30°C, respectively. Symbols like circles, squares, and dots on each trace denote the time of cell division. Shaded regions indicate the time duration over which the integral is evaluated. All traces show mean single-cell fluorescence values above camera background. The time of induction is ≈ -10 min. Additional information: (a) Schematic represents induction of fluorescent protein YFP. Reference trace is at 32°C. The IPTG concentration for induction is 1mM. (b) Schematic represents induction of fluorescent protein CI-YFP. Reference trace is at 34°C. The IPTG concentration for induction is 10μM.

we use another inducible circuit, where the addition of the inducer IPTG increases the expression of GFP (Fig. 4). For this measurement, cells were grown overnight in LB at 37° C. The overnight culture was diluted 1 : 500 in minimal M9CA media (Teknova) and grown for 2 hours at either 32°C or at temperatures 5% (≈ 2°C) above and below this value. For time-lapse imaging, cells were loaded in specialized bacterial flow chambers. These chambers trap bacterial cells and can be used to image a growing colony of cells. Further they allow for media, with and without inducer, to be flowed across the growing colony. Using this technology, cells can be induced with IPTG at a well-defined time. Data was acquired at the nominal temperature of 32°C, and at temperatures 5% (≈ 2°C) above and below this value. As before, the time-sequence of data from these measurements were analyzed and found to demonstrate very precise control over the time when the inducer is applied as well as the time after which the inducer is washed off (Fig. 4). For the current purpose, we analyze the data related to induction of IPTG. For this strain, the reference trace has its equilibrium value of ≈ 1400 a.u.. Before induction, the equilibrium value for this trace is less than 50 a.u. Therefore, the reference trace exhibits a greater than ten-fold induction. The time after which it stays within 10% of this equilibrium value is $T_2 = 230$ minutes. Given that it reaches 10% of this equilibrium value at $T_1 = 80$ minutes, it also maintains this equilibrium value for the required duration $T_2 - T_1 = 150$ minutes after T_2. The performance score for this dataset is 2.29.

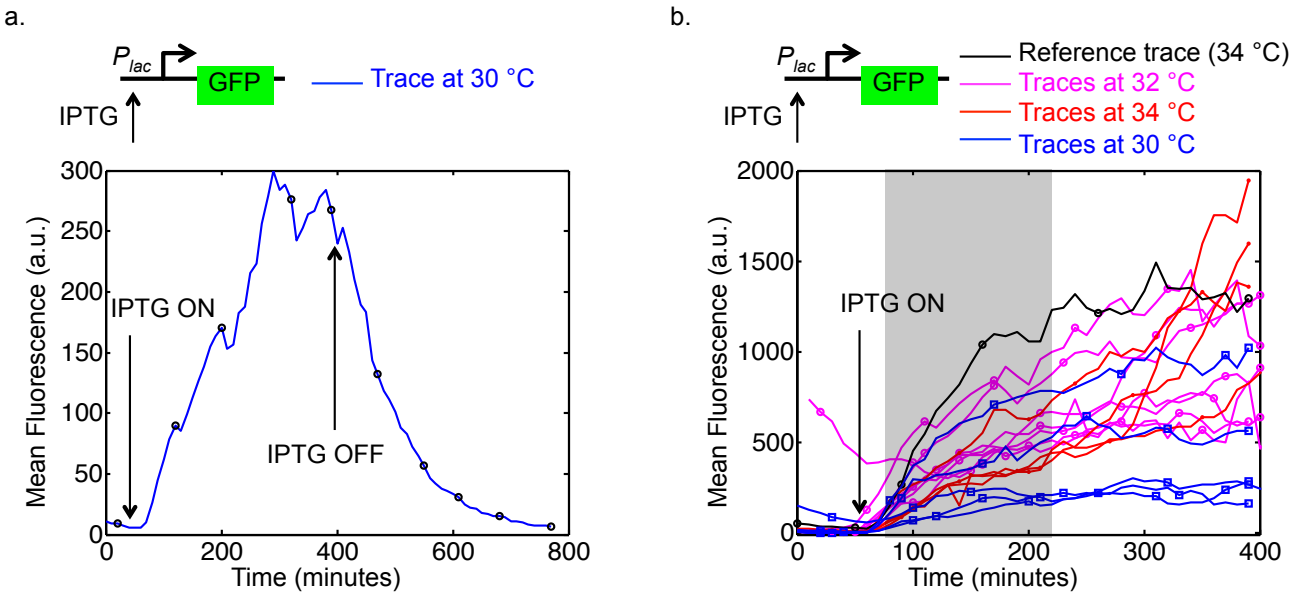

Figure 4: Single-Cell Dynamics with Controlled Induction Time. (a) Trace illustrating control over the turn-on and turn-off of induction. Schematic represents induction of fluorescent protein GFP with inducer IPTG. IPTG concentration for induction is 100μM. Circles on the trace denote the time of cell division. (b) Solid black line represents the reference trace. Reference trace is at 34°C. Red, magenta and blue colors represent the temperatures 34°C, 32°C, and 30°C, respectively. Shaded region indicates the duration over which the integral is evaluated. All traces show mean single-cell fluorescence values above camera background. The time of induction is 60 min. Symbols such as circles, squares, and dots on each trace denote the time of cell division.

Links to the scripts and data used to estimate performance score from experimental measurements will be on the CAGEN website. These will help to establish a set of reference protocols for experiments.

5 Estimating performance score using population-level measurements

Performance scores estimated so far, both computationally and experimentally, rely on single-cell time courses. Next, we explored whether similar performance estimates can be obtained from population-level measurements, which yield a a distribution of fluorescent intensities at each time point. The motivation for this is to use a simpler measurement technology in early design iterations. For this purpose, we redefine the metric based on the time course of the mean of the distribution ($\mu(t)$) and its equilibrium amplitude M,

$$S_p = \frac{1}{M^2} \int_{T_1}^{T_2} (\mu(t) - M)^2 dt.$$

Similar to the previous metric definition, the integral is evaluated over the time in which the $\mu(t)$ rises from 10% of M to a time after which it stays within 10% of its final value. In contrast to the metric S, which is a worst-case measure, the modified metric S_p is an average measure.

We have used the computational model to estimate the worst-case metric S. By ignoring the time connectivities in the simulated data (Fig. 5a), it can also be used to compute the average metric S_p. We find that the average score is close to the worst-case score when population level measurements are taken at the same time resolution. However, at a coarser time resolution, the average score significantly deviates from the worst-case score (Fig. 5b). This is to be expected as deviation from the idealized step response increases when the sampled measurements are far apart. Finally, we compared how the average metric changes relative to circuit properties like noise and speed (Fig. 5c, d). As before, we computed the average score for the case when the noise is varied for a fixed speed as well as for the case when the speed was varied for a fixed noise level. In both cases, the average score showed the expected dependence, which was also similar to the dependence of the worst case metric. Together, this analysis shows that the average score computed using population-level measurements can also be a good indicator of performance.

6 Discussion

The goal of the Robust Gene Response Challenge is to design a circuit that can quickly increase the expression of a fluorescent protein by tenfold upon induction and exhibit minimal variation in expression between cells and across temperatures. To this end, we have worked through the Challenge specifications: First, we have stated the performance metric that will be used to compare designs. Second, we have used a simple computational model to estimate this metric, and verified that it adequately captures circuit properties like noise and speed. Third, we have estimated performance scores from inducible protein expression circuits in *E. coli* using experimental single-cell time-lapse fluorescence microscopy measurements. These establish a baseline performance measure. Finally, we present a slight modification to the metric so that the performance score can be estimated using population-level dynamic measurements. This analysis establishes baseline performance criteria and provide tools and protocols that can help participating teams.

An interesting aspect of the Robust Gene Response Challenge specifications is the need for circuit performance to be maintained across temperatures. Indeed, our data suggest that temperature may have a significant effect on circuit properties. Understanding how temperature affects the operation of genetic circuits is fundamental problem in cell biology, with significant efforts devoted to understanding temperature-compensatory mechanisms in circadian rhythms. It is also interesting from a design point of view as robustness to temperature variations is likely to be an important property in the use of

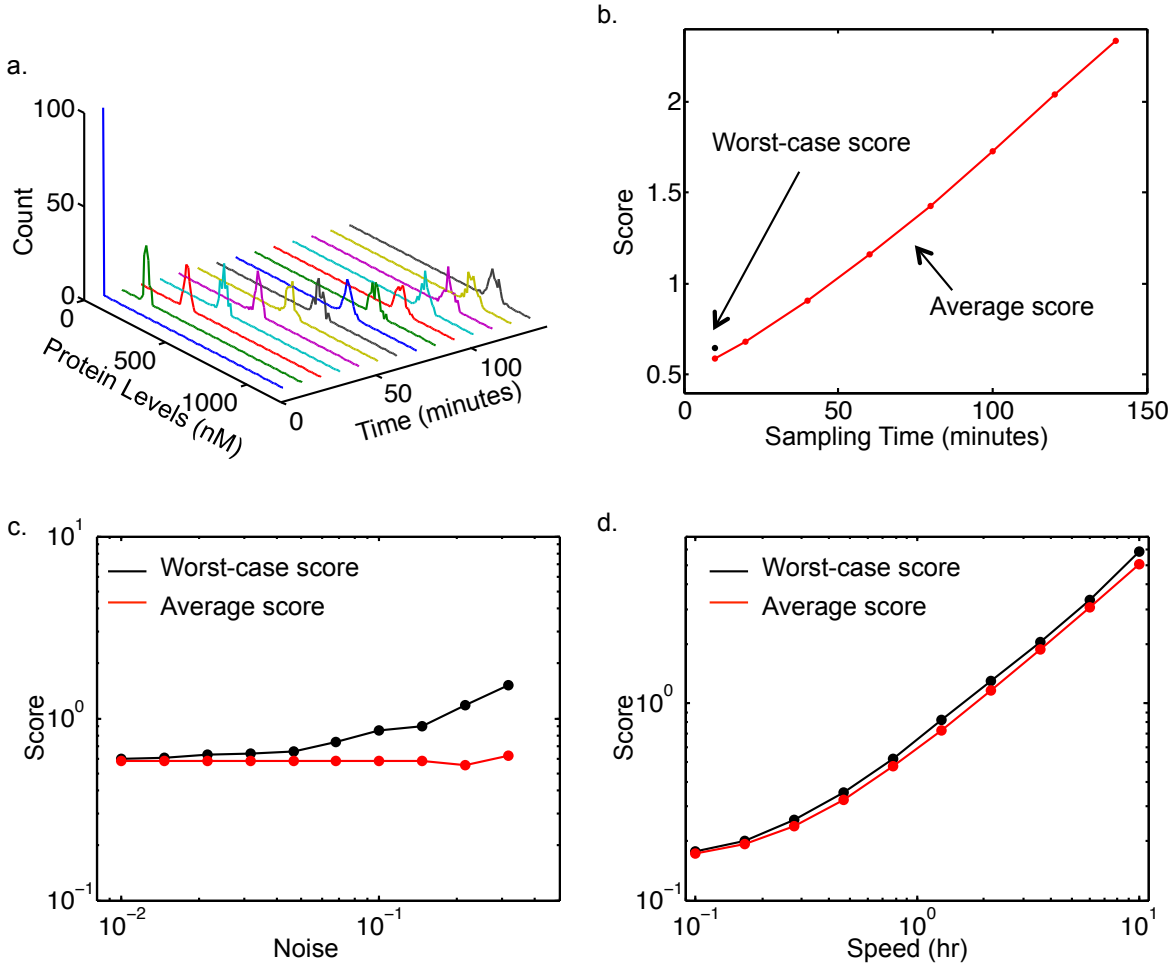

Figure 5: Population-level Metric Scores for the Protein Production-Degradation Circuit. (a) Three-dimensional plot shows evolution of fluorescence distributions over time. (b) Solid red line connects average scores computed at sampling rates 10, 20, 30, 60, 120 minutes. Black dot represents the worst-case metric score at a 10 minute sampling rate. Performance scores are calculated as (c) noise is varied for fixed speed and, (d) speed is varied for fixed noise. In each plot, solid red line connects the average score measurements ($N = 100$). Solid black line connects the worst-case score measurements ($N = 100$). To change noise levels for fixed speed in (c), the values of α are varied logarithmically from 10nM/hr to 10^4nM/hr, for fixed γ. To change speed values for fixed noise in (d), the values of γ are varied logarithmically from 0.1/hr to 10/hr and α values are adjusted so that the ratio $\alpha/\gamma = 1000$ nM stays fixed.

synthetic molecular circuits. The current Challenge is likely to result in data that will be important for both these contexts.

The current challenge is an important part of a progression of challenges leading unto the design of a linear amplifier with a fast transient response. For this, consider the present challenge followed by another challenge that specifies the design of a circuit that follows a pulse in inducer as closely as possible. In other words, this second challenge specifies that the input be a step increase in inducer followed by a step decrease. The third and final part in this series of challenges is the requirement that the magnitude of the output linearly follow the input magnitude over a large range.Together, this series of challenges will culminate in the design of a fast, linear amplifier.

Efforts to design circuits for this CAGEN Challenge will improve on the current techniques for the design of robust biological circuits. As they will be accessible to a wider research community, they will provide foundations to develop new design protocols as well as mathematical tools to address them. Most importantly, they will help the overall goal of engineering biological processes with applications in agriculture, medicine, and environmental sciences.

References

[1] R. S. Cox 3rd, M. J. Dunlop, and M. B. Elowitz. A synthetic three-color scaffold for monitoring genetic regulation and noise. *J. Biol. Eng.*, 4(10), 2010.

[2] D. Adalsteinsson, D. McMillen, and T. C. Elston. Biochemical Network Stochastic Simulator (BioNetS): software for stochastic modeling of biochemical networks. *BMC Bioinformatics*, 5(24), 2004.

[3] A. Bar-Even, J. Paulsson, N. Maheshri, M. Carmi, E. O'Shea, Y. Pilpel, and N. Barkai. Noise in protein expression scales with natural protein abundance. *Nat Genet*, 38(6):636–43, 2006.

[4] A. Becskei and L. Serrano. Engineering stability in gene networks by autoregulation. *Nature*, 405(6786):590–3, 2000.

[5] M. J. Dunlop, R. S. Cox 3rd, J. H. Levine, R. M. Murray, and M. B. Elowitz. Regulatory activity revealed by dynamic correlations in gene expression noise. *Nat Genet*, 40(12):1493–8, 2008.

[6] A. Eldar and M. B. Elowitz. Functional roles for noise in genetic circuits. *Nature*, 467(7312):167–73, 2010.

[7] E. M. Ozbudak, M. Thattai, I. Kurtser, A. D. Grossman, and A. van Oudenaarden. Regulation of noise in the expression of a single gene. *Nat Genet*, 31(1):69–73, 2002.

[8] P. E. Purnick and R. Weiss. The second wave of synthetic biology: from modules to systems. *Nat Rev Mol Cell Biol*, 10(6):410–22, 2009.

[9] A. Raj and A. van Oudenaarden. Nature, nurture, or chance: Stochastic gene expression and its consequences. *Cell*, 135(2):216–26, 2008.

[10] N. Rosenfeld, M. B. Elowitz, and U. Alon. Negative autoregulation speeds the response times of transcription networks. *J Mol Biol*, 323(5):785–93, 2002.