THE JOURNAL OF PHYSICAL CHEMISTRY B

Article

pubs.acs.org/JPCB

Noise Control in Gene Regulatory Networks with Negative Feedback

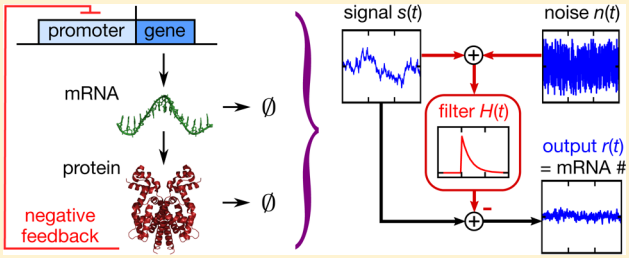
Michael Hinczewski*,† and D. Thirumalai*,‡

†Department of Physics, Case Western Reserve University, Cleveland, Ohio 44106, United States

‡Department of Chemistry, The University of Texas at Austin, Austin, Texas 78712, United States

Supporting Information

ABSTRACT: Genes and proteins regulate cellular functions through complex circuits of biochemical reactions. Fluctuations in the components of these regulatory networks result in noise that invariably corrupts the signal, possibly compromising function. Here, we create a practical formalism based on ideas introduced by Wiener and Kolmogorov (WK) for filtering noise in engineered communications systems to quantitatively assess the extent to which noise can be controlled in biological processes involving negative feedback. Application of the theory, which reproduces the previously proven scaling of the lower bound for noise suppression in terms of the number of signaling events, shows that a tetracycline repressor-based negative-regulatory gene circuit behaves as a WK filter. For the class of Hill-like nonlinear regulatory functions, this type of filter provides the optimal reduction in noise. Our theoretical approach can be readily combined with experimental measurements of response functions in a wide variety of genetic circuits, to elucidate the general principles by which biological networks minimize noise.



The genetic regulatory circuits that control all aspects of life are inherently stochastic. They depend on fluctuating populations of biomolecules interacting across the crowded, thermally agitated interior of the cell. Noise is also exacerbated by low copy numbers of particular proteins and mRNAs, as well as variability in the local environment.[1−8] Yet the robust and reproducible functioning of key systems requires mechanisms to filter out fluctuations. For example, regulating noise is relevant in stabilizing cell-fate decisions in embryonic development,[9] prevention of random switching to proliferating states in cancer-regulating miRNA networks,[10] and maximization of the efficiency of bacterial chemotaxis along attractant gradients.[11] Comprehensive analysis of yeast protein expression reveals that proteins involved in translation initiation, ribosome formation, and protein degradation have lower relative noise levels,[12] suggesting natural selection could favor noise reduction for certain essential cellular components.[13,14]

A common regulatory motif capable of suppressing noise is the negative feedback loop,[1,2,15−21] as has been explicitly demonstrated in synthetic gene circuits.[1,16,17] Feedback pathways for a given chemical species can be mediated by numerous signaling molecules, each with its own web of interactions and stochastic characteristics that determine the ultimate effectiveness of the system in damping the fluctuations of the target population and maintaining homeostasis. Thus, uncovering generic laws governing the behavior of such control networks is difficult. A major advance was made by Lestas, Vinnicombe, and Paulsson (LVP),[22] who showed that information theory can set a rigorous lower bound on the magnitude of fluctuations within an arbitrarily complicated homeostatic negative feedback network. Since the bound scales like the fourth root of the number of signaling events, noise reduction is extremely expensive. This underscores the pervasiveness of biological noise, even in cases where there may be evolutionary pressure to minimize it.

The existence of a rigorous bound raises a number of intriguing issues. Can a biochemical network actually reach this lower bound, and thus optimally suppress fluctuations? What would be the dynamic behavior of such an optimal system, and how would it depend on the noise spectrum of the system components? Here we answer these equations using a theory related to the optimal linear noise-reduction filter, developed by Wiener[23] and Kolmogorov.[24] Though the original context of Wiener−Kolmogorov (WK) filter theory was removing noise from corrupted signals in engineered communications systems, it has become a powerful tool for characterizing the constraints on signaling in biochemical networks.[25,26] Recently, we showed that the action of kinase and phosphatase enzymes on their protein substrates, the basic elements of many cellular signaling pathways, can in fact effectively be represented as an optimal WK filter.[25] The WK theory also describes how systems like *E. coli* chemotaxis can optimally anticipate future changes in concentrations of extracellular ligands.[26] Although the classic WK theory is strictly defined for linear filtering of continuous signals (a reasonable approximation for certain biochemical networks), it can also be extended to yield constraints in the more general case of nonlinear production of molecular species with discrete population values.[25]

Special Issue: William M. Gelbart Festschrift

Received: February 29, 2016
Revised: April 19, 2016
Published: April 19, 2016

DOI: 10.1021/acs.jpcb.6b02093
J. Phys. Chem. B 2016, 120, 6166−6177

Interestingly, for a broad class of systems the WK linear solution turns out to be the global optimum among all nonlinear or linear networks, allowing us to delineate where nonlinearity is potentially advantageous in biochemical noise control. Most importantly, since the WK theory is formulated in terms of experimentally accessible dynamic response functions, it also provides a design template for realizing optimality in synthetic circuits. As an illustrative example, we predict that a synthetic autoregulatory TetR loop, engineered in yeast,[27] can be fine-tuned to approximate an optimal WK filter for TetR mRNA levels. Though a simple design, similar filters could be employed in nature to cope with Poisson noise arising from small copy numbers of mRNAs, often on the order of 10 per cell.[28] Based on the application of the theory to the synthetic gene network, we propose that the extent of noise reduction in biological circuits is determined by competing factors such as functional efficiency, adaptation, and robustness.

RESULTS

To make the paper readable and as self-contained as possible, many of the details of the calculation are relegated to the Supporting Information (SI). The main text contains only the necessary details needed to follow the results without the distraction of the mathematics.

Linear response theory for a general control network. To motivate the WK approach for a general control network, we start with the simple case where two species within the network are explicitly singled out:[22] a target R with time-varying population $r(t)$ fluctuating around mean $\bar{r}$, and one of the mediators in the feedback signaling pathway P, with population $p(t)$ varying around $\bar{p}$. We assume a continuum Langevin description of the dynamics,[15,18,29,30] where the rate

$$\dot{\alpha}(t) = k_\alpha(t) + n_\alpha(t) \tag{1}$$

for $\alpha = r$ or p, can be broken down into deterministic (k_α) and stochastic (n_α) parts. The function $k_\alpha(t)$ encapsulates the entire web of biochemical reactions underlying synthesis and degradation of species α, and can be an arbitrary functional of the past history of the system up to time t. It is typically divided into two parts, $k_\alpha(t) = k_\alpha^+(t) - k_\alpha^-(t)$, corresponding to the production (+) and destruction (−) rates of the species α. The term $n_\alpha(t)$ is the additive noise contribution, which can also be divided into two parts, $n_\alpha(t) = n_\alpha^{\mathrm{int}}(t) + n_\alpha^{\mathrm{ext}}(t)$. The first is the "intrinsic" or shot noise, arising from the stochastic Poisson nature of the α generation, $n_\alpha^{\mathrm{int}}(t) = \sqrt{2\bar{k}_\alpha}\,\eta_\alpha(t)$, where $\bar{k}_\alpha$ is the mean production rate, or equivalently the mean destruction rate, $\bar{k}_\alpha = \overline{k_\alpha^+(t)} = \overline{k_\alpha^-(t)}$, and $\eta_\alpha(t)$ is a Gaussian white noise function with correlation $\overline{\eta_\alpha(t)\,\eta_{\alpha'}(t')} = \delta_{\alpha\alpha'}\delta(t-t')$. The second part, $n_\alpha^{\mathrm{ext}}(t)$, is "extrinsic" noise, which arises due to fluctuations in cellular components affecting the dynamics of R and P that are not explicitly taken into account in the two-species picture. These could include mediators in the signaling pathway, or global factors such as ribosome and RNA polymerase levels. For simplicity, our main focus will be the case of no extrinsic noise. However, we will show later how a straightforward extension of the theory reveals that the same system can behave like an optimal WK filter under a variety of extrinsic noise conditions.

For small deviations $\delta\alpha(t) = \alpha(t) - \bar{\alpha}$ from the mean populations $\bar{\alpha}$, $k_\alpha(t)$ can be linearized with respect to $\delta\alpha(t)$,

$$k_\alpha(t) = \sum_{\alpha'=r,p} \int_{-\infty}^{t} dt' G_{\alpha\alpha'}(t-t')\delta\alpha'(t') \tag{2}$$

where $G_{\alpha\alpha'}(t)$ are linear response functions, which express the dependence of $k_\alpha(t)$ on the past history of $\delta\alpha'(t)$. The functions $G_{\alpha\alpha'}(t)$ capture the essential characteristic responses of the control network to perturbations away from equilibrium (Figure 1). In the static limit, $G_{\alpha\alpha'}(t)$ have appeared in various guises as gains,[6] susceptibilities,[19] or steady-state Jacobian matrices,[30] and in the frequency domain as loop transfer functions.[15,18] Feedback between R and P is encoded in the cross-responses $G_{rp}(t)$ and $G_{pr}(t)$. If the feedback occurs through slow, intermediate steps, involving additional chemical species besides R and P, the time-dependence of the cross-response functions will reflect the time scales of those intermediate processes. In the absence of intermediates, or if those steps are very fast compared to the production/degradation of R and P, the cross-responses depend only on the instantaneous population deviations, and hence $G_{\alpha\alpha'}(t) = G_{\alpha\alpha'}\,\delta(t)$, where $G_{\alpha\alpha'}$ is a constant. In the simplest scenario, the only nonzero self-responses $G_{\alpha\alpha}(t)$ are decay terms, $G_{\alpha\alpha}(t) = -\tau_\alpha^{-1}\delta(t)$, where τ_α is the decay time scale for species α. However, the theory works generally for more complicated self-response mechanisms.

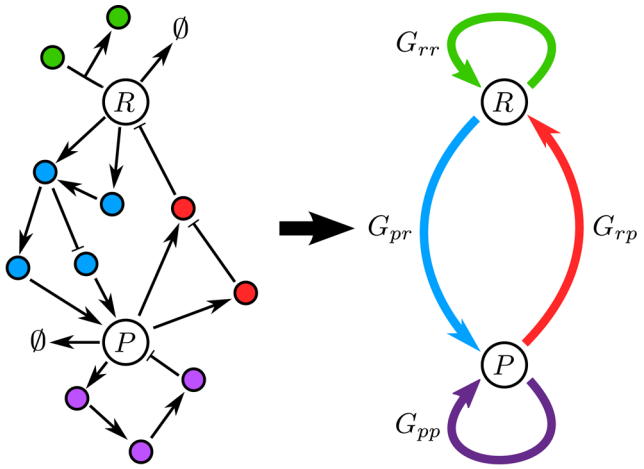


Figure 1. Schematic of a complex signaling network with the target species R and one mediator P singled out. In focusing on two species, the action of all the other components is effectively encoded in four response functions—$G_{rr}(t)$, $G_{pp}(t)$, $G_{rp}(t)$, $G_{pr}(t)$—that describe how the entire dynamical system responds to perturbations in R and P.

Control network as a noise filter. The connection between the linearized dynamical description and WK filter theory arises from comparing the original system to the case where feedback is turned off (i.e., setting $G_{rp}(\omega)$ or $G_{pr}(\omega)$ to zero). Let us define a few terms to make the noise filter analogy clear. Without feedback, the target fluctuations are $\delta r_0(t) \equiv s(t)$, where we denote $s(t)$ the *signal*. This is to distinguish it from $\delta r(t)$ in the original system, which is the *output*. The difference between the two, which reflects the impact of the feedback network, we express as $\delta r(t) = s(t) - \tilde{s}(t)$, where $\tilde{s}(t)$ is referred to as the *estimate*. In this analogy, minimizing $\delta r(t)$ requires a feedback loop where the estimate $\tilde{s}(t)$ is as close as possible to the signal $s(t)$. The only thing left to specify is the relationship between $\tilde{s}(t)$ and $s(t)$.

DOI: 10.1021/acs.jpcb.6b02093
J. Phys. Chem. B 2016, 120, 6166−6177

The dynamical system in eqs 1−2 takes a simple form in Fourier space, where the fluctuations $\delta\alpha(\omega)$ satisfy:

$$-i\omega\delta\alpha(\omega) = \sum_{\alpha'=r,p} G_{\alpha\alpha'}(\omega)\delta\alpha'(\omega) + n_\alpha(\omega) \tag{3}$$

We solve eq 3 for $\delta r(\omega)$ and break up the R fluctuation into two contributions, $\delta r(\omega) = s(\omega) - \tilde{s}(\omega)$, with the signal $s(\omega)$ and estimate $\tilde{s}(\omega)$ given by

$$s(\omega) = -\frac{n_r(\omega)}{G_{rr}(\omega) + i\omega}, \quad \tilde{s}(\omega) = H(\omega)[s(\omega) + n(\omega)] \tag{4}$$

Here we have introduced a noise function $n(\omega)$,

$$n(\omega) = \frac{n_p(\omega)}{G_{pr}(\omega)} \tag{5}$$

and a filter function $H(\omega)$:

$$H(\omega) \equiv \frac{G_{rp}(\omega)G_{pr}(\omega)}{G_{rp}(\omega)G_{pr}(\omega) - (G_{rr}(\omega) + i\omega)(G_{pp}(\omega) + i\omega)} \tag{6}$$

Thus, in the time domain the estimate $\tilde{s}(t)$ is the convolution of the filter function $H(t)$ and a noise-corrupted signal $y(t) \equiv s(t) + n(t)$,

$$\tilde{s}(t) = \int_{-\infty}^{\infty} dt' H(t - t')y(t') \tag{7}$$

Equations 4−6 constitute a one-to-one mapping between the linear response and noise filter descriptions of the system in Fourier space. They relate the four filter quantities, $s(\omega)$, $\tilde{s}(\omega)$, $n(\omega)$, and $H(\omega)$, to the four linear response functions $G_{rr}(\omega)$, $G_{rp}(\omega)$, $G_{pr}(\omega)$, and $G_{pp}(\omega)$.

The entire noise filter system is illustrated schematically in Figure 2. Note that the noise function in the filter analogy, $n(t)$, is related to $n_p(t)$ in Fourier space as $n(\omega) = n_p(\omega)/G_{pr}(\omega)$. It depends not just on the intrinsic P noise $n_p(\omega)$, but on the larger network through the cross-response $G_{pr}(\omega)$. In the optimization procedure below, we will keep the signal $s(t)$ and noise $n(t)$ properties fixed, while varying $H(t)$ to try to filter out the $n(t)$ component in $y(t)$ in order to produce $\tilde{s}(t)$ close to $s(t)$. This means fixing both G_{rr} and G_{pr}, while allowing H to vary through the remaining response functions G_{rp} and G_{pp}. Though we confine ourselves throughout this work to the case of a dynamical system with a single target and mediator species, one can easily generalize the entire approach to explicitly include many mediators, which could potentially be involved in a complex signaling pathway. The linearized dynamical system in eqs 1−2 would still have the same form (with index α running over all the species of interest), and the mapping onto the filter problem for the target species would be analogous. The only difference is that $n(\omega)$ and $H(\omega)$ would be more complicated functions of the various individual noise terms $n_\alpha(\omega)$ and the response functions $G_{\alpha\alpha'}(\omega)$ of the mediators. In our reduced, two species description, the action of all the unspecified chemical components is effectively included in the four response functions described above, with their stochastic effects contributing to the extrinsic noise. Figure 1 shows a schematic of such a reduction. The fine-grained details of the signaling pathways connecting our target R and mediator P, potentially involving many interacting species, are encoded in G_{rr}, G_{pp}, G_{rp}, and G_{pr}. As an example of how this two-species

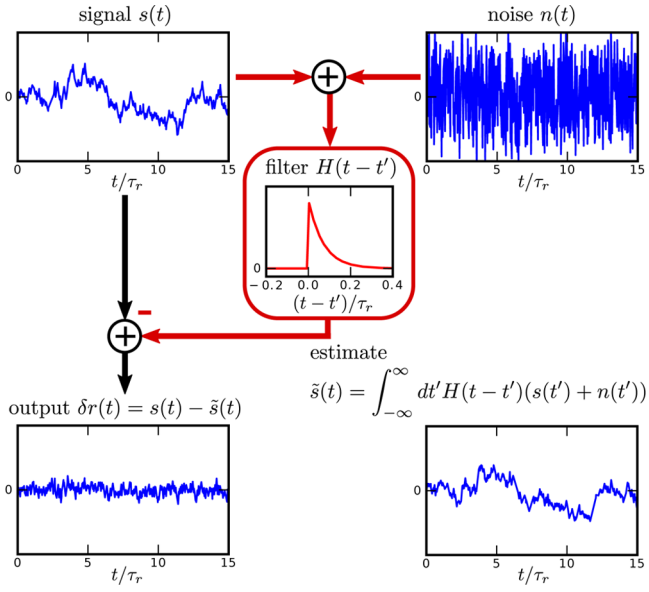


Figure 2. Signal processing diagram illustrating noise suppression in a negative feedback loop reinterpreted as a linear filter. The fluctuations in the target species $\delta r(t)$ (lower left) are expressed as $\delta r(t) = s(t) - \tilde{s}(t)$, where the raw signal $s(t)$ (upper left) equals $\delta r(t)$ in the absence of feedback control, and the estimate $\tilde{s}(t)$ (lower right) is the contribution of the feedback loop. This estimate is given by the convolution of a filter function $H(t)$ (center) and the corrupted signal $s(t) + n(t)$, where $n(t)$ is the noise (upper right). The goal of Wiener−Kolmogorov theory is to find a causal $H(t)$ such that the standard deviation of $\delta r(t)$ is minimized. All sample trajectories shown in the figure are generated from numerically solving the linearized version of the dynamical system in eq 10.

reduction would work in practice, in SI section 2 we treat an important example of a feedback loop involving multiple mediators, representing a signaling cascade in series.

Wiener−Kolmogorov theory yields the optimal filter. The WK optimization problem consists of minimizing $\sigma_r^2 = \overline{(\delta r)^2}$, the variance of target fluctuations, which are related to $H(t)$, $s(t)$, and $n(t)$ through the frequency domain integral[31] (see derivation based on the Wiener−Khinchin theorem in SI section 1):

$$\sigma_r^2 = \int_{-\infty}^{\infty} \frac{d\omega}{2\pi}[|H(\omega)|^2 P_n(\omega) + |H(\omega) - 1|^2 P_s(\omega)] \tag{8}$$

where $H(\omega)$ is the Fourier transform of $H(t)$, and $P_n(\omega)$, $P_s(\omega)$ are the power spectral densities (PSD) of $n(t)$ and $s(t)$, respectively, i.e. the Fourier transforms of their autocorrelation functions. If $P_n(\omega)$ and $P_s(\omega)$ are given, the task is to minimize σ_r^2 in eq 8 over all possible $H(\omega)$. The main constraint that makes the solution mathematically difficult is that $H(\omega)$ must correspond to a physically realizable control network, which imposes the crucial restriction that the time-domain convolution function $H(t)$ must be causal, depending only on the past history of the input, $H(t) = 0$ for $t < 0$. The great achievement of Wiener and Kolmogorov was to derive the form of the optimal causal solution $H_{opt}(\omega)$:

$$H_{opt}(\omega) = \frac{1}{P_y^c(\omega)}\left\{\frac{P_s(\omega)}{P_y^c(\omega)^*}\right\}_c \tag{9}$$

The c super/subscripts refer to two different decompositions in the frequency domain which enforce causality: (i) Any physical

DOI: 10.1021/acs.jpcb.6b02093
J. Phys. Chem. B 2016, 120, 6166−6177

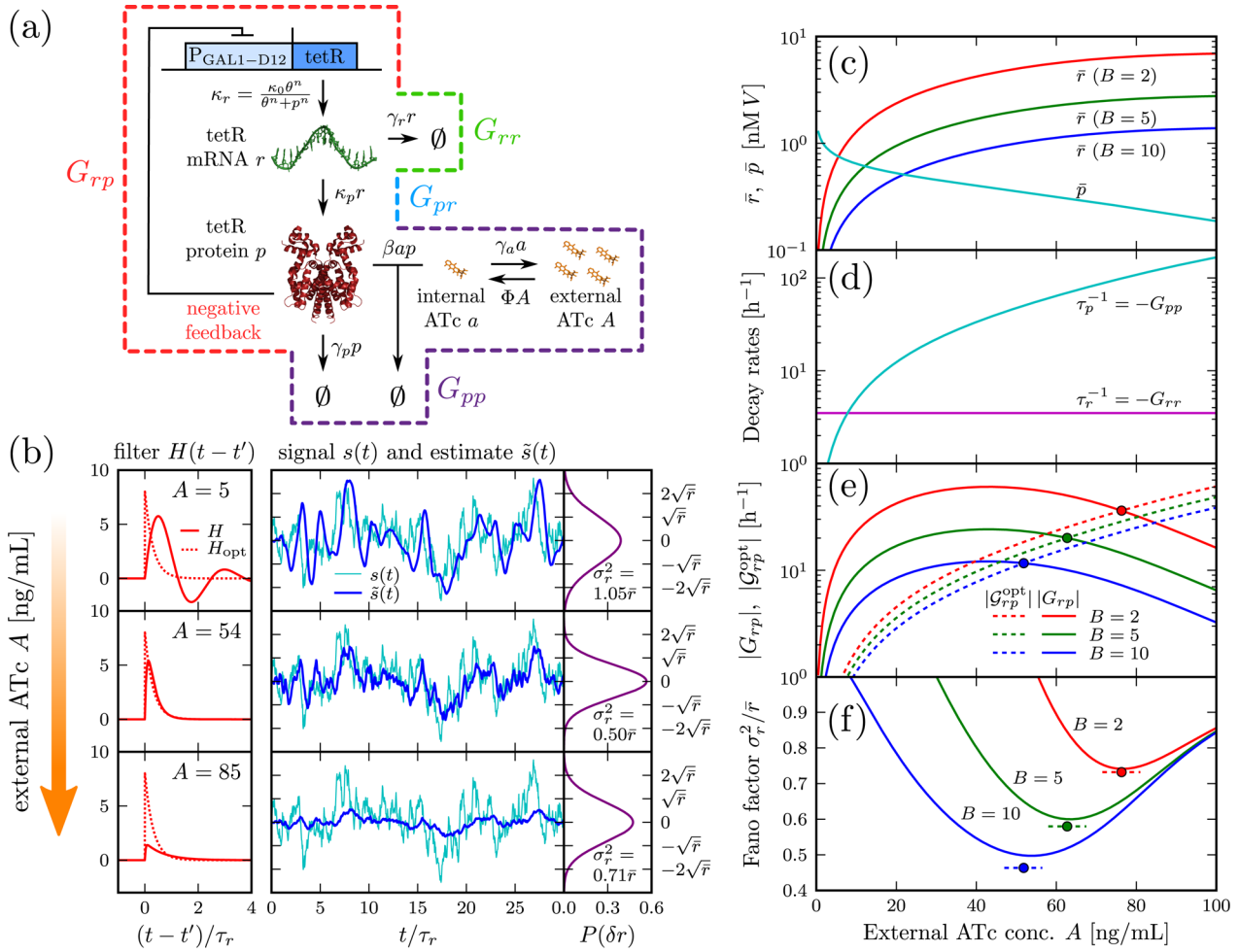


Figure 3. (a) Synthetic yeast gene circuit designed by Nevozhay et al.[27] The TetR protein negatively regulates itself by binding to its own promoter. The inducer molecule ATc associates with TetR, inhibiting its repressor activity. The subsequent panels show results for this gene circuit using the linear filter theory applied to the dynamical model of eq 10, with experimentally derived parameters (Table 1). (b) Filter functions $H(t)$ and $H_{opt}(t)$, sample signal $s(t)$, and estimate $\tilde{s}(t)$ time series for burst ratio $B = 10$ and three different values of extracellular ATc concentration A [ng/mL]. $H(t)$ is from eq 19, while $H_{opt}(t)$ is from eq 18. The sample time series trajectories are numerical solutions of the linearized eq 10. On the right are the resulting equilibrium probability distributions $P(\delta r)$, where $\delta r(t) = s(t) - \tilde{s}(t)$, which are Gaussians with variance σ_r^2. For $A \approx 54$ ng/mL, the circuit approximately functions as an optimal WK filter ($H(t)$ is close to $H_{opt}(t)$), maximally suppressing fluctuations in the population levels of TetR mRNA (minimizing $\sigma_r^2/\bar{r}$). (c) Mean populations of free intracellular TetR mRNA, $\bar{r}$, and TetR protein dimers, $\bar{p}$. (d) The decay rates of free mRNA and proteins, τ_r^{-1} and τ_p^{-1}, which are related to the network self-response functions G_{rr} and G_{pp} (both are constants in the frequency domain as shown in eq 11). (e) Magnitude of the network cross-response, $|G_{rp}|$ (solid lines), plotted together with the optimal magnitude $|\mathcal{G}_{rp}^{opt}| = \tau_p^{-1}(1 + \sqrt{1 + B})^{-1}$ (dashed lines). Filled circles mark the intersection defining $A = A_{opt}$, where the system behaves approximately like an optimal WK filter. (f) Fano factor $\sigma_r^2/\bar{r}$ (solid lines), compared to the optimal WK value $\sigma_{r,opt}^2/\bar{r} = 2/(1 + \sqrt{1 + B})$ (horizontal dashed lines). Filled circles mark the position $A = A_{opt}$.

PSD, in this case $P_y(\omega)$ corresponding to the corrupted signal $y(t) = s(t) + n(t)$, can be written as $P_y(\omega) = |P_y^c(\omega)|^2$. The factor $P_y^c(\omega)$, if treated as a function over the complex ω plane, contains no zeros and poles in the upper half-plane (Im $\omega > 0$).[32] (ii) We also define an additive decomposition denoted by $\{F(\omega)\}_c$ (see SI section 1) for any function $F(\omega)$, which consists of all terms in the partial fraction expansion of $F(\omega)$ with no poles in the upper half-plane. In SI section 1, we provide in detail a new derivation of eq 9, the heart of the WK theory.

Optimal noise control in a yeast gene circuit with feedback. To illustrate the nature of the optimal WK solution we choose as a case study the yeast negative autoregulatory gene circuit designed by Nevozhay et al.,[27] drawn schematically in Figure 3(a). The gene encoding for the TetR protein is under the control of the $P_{GAL1-D12}$ promoter, whose activity can be repressed by binding TetR dimers. The strength of the feedback can be modulated by changing the extracellular concentration A of the inducer anhydrotetracycline (ATc), which enters the cell, binds to TetR and prevents its association with the promoter, thus weakening repression.

In order to analyze the TetR negative feedback gene circuit, we start with the simple mathematical model introduced in ref 27, which provided results that are consistent with the experimental data. The simplified model, which captures the essence of the synthetic gene network, features as the main variables the population of free intracellular TetR dimer, $p(t)$, and free intracellular ATc molecules, $a(t)$. In addition to the regulatory loop, the experimental gene circuit has a parallel yEGFP reporter portion, which acts as a monitor of TetR

DOI: 10.1021/acs.jpcb.6b02093
J. Phys. Chem. B 2016, 120, 6166−6177

protein levels. Because we focus on the system as a noise filter for the TetR mRNA population, and the yEGFP part does not influence this analysis,[27] we ignore the reporter circuit.

The production of the TetR dimers occurs in a single step, with the autoregulation of the rate described by a repressory Hill function. We divide this step into two parts, introducing as an additional variable the population of TetR mRNA $r(t)$. The feedback loop (Figure 3(a)) consists of mRNA production at a rate given by the Hill function $\kappa_r(t) = \kappa_0\theta^n/(\theta^n + p^n(t))$, followed by TetR dimer generation at a rate given by $\kappa_p\, r(t)$. The degradation/dilution of the mRNA and dimers is modeled through decay terms $\gamma_r r(t)$ and $\gamma_p p(t)$. All other chemical substeps involved in this loop are comparatively fast, such as TetR dimerization, the binding of the repressor to the individual promoter sites, or the role of RNAP and ribosomes in the transcription and translation processes. We thus confine ourselves to an effective two substep description to illustrate the filter theory, though the stochastic effects of additional complexity can be approximately treated through general "extrinsic" noise terms incorporated into $n_r(t)$ and $n_p(t)$.

The main experimental variable that allows tuning of the yeast gene network behavior is the external ATc concentration A, which is assumed to be time independent. As illustrated in Figure 3(a), there is an influx ΦA of ATc molecules into the cell. Once inside, the ATc molecules associate with the TetR at a rate $\beta a(t)p(t)$. Additional loss of intracellular ATc through degradation, outflux, and dilution is modeled through an effective decay rate $\gamma_a a(t)$. We assume that the dissociation of ATc from TetR occurs on long enough time scales that it can be ignored. Since the influx/association/outflux of ATc is fast compared to the transcription and translation processes of the main feedback loop, we further assume that $a(t)$ instantaneously equilibriates at the current value of $p(t)$. Thus, the dependence of $a(t)$ on $p(t)$ is determined by equating the influx and total loss rate, which leads to $a(p(t)) = \Phi\, A/\,(\gamma_a + \beta p(t))$.

For the model described above, the dynamical equations for $r(t)$ and $p(t)$ are,

$$\dot{r}(t) = -\gamma_r r(t) + \frac{\kappa_0\theta^n}{\theta^n + p^n(t)} + n_r(t)$$

$$\dot{p}(t) = -\gamma_p p(t) + \kappa_p r(t) - \frac{\beta\Phi A p(t)}{\gamma_a + \beta p(t)} + n_p(t) \qquad (10)$$

The parameters, with values derived from experimental fitting,[27] are listed in Table 1. The only quantity that is not independently known from the fit is the rate κ_p, which we allow to vary in the range $\kappa_p/\gamma_r \equiv B = 2-10$, comparable to typical experimentally measured protein burst sizes.[33] The noise terms are $n_r(t) = \sqrt{2\gamma_r \bar{r}}\,\eta_r(t)$ and $n_p(t) = \sqrt{2\kappa_p \bar{r}}\,\eta_p(t)$, assuming only intrinsic noise contributions. Setting the right sides of eq 10 to zero, and averaging over $n_r(t)$ and $n_p(t)$, we numerically solve for the equilibrium values $\bar{r}$ and $\bar{p}$ as a function of external ATc concentration A [Figure 3(c)]. For $A = 0$, the promoter is nearly fully repressed, but with increasing A, the mean population $\bar{p}$ of free TetR dimers is reduced, weakening the repression and boosting the mean mRNA population $\bar{r}$. Changing A allows us to explore a wide range of control network behavior. Note that since $\bar{p}$ depends on B only through the product $\kappa_0\, B$, and the value of this product is fixed at a constant value from the experimental fit (Table 1), $\bar{p}$ is independent of B. On the other hand, $\bar{r}$, which is proportional to κ_0, is inversely proportional to B.

Table 1. Parameter Values for the Dynamical Model of the Yeast Synthetic Gene Circuit (Eq 10)[a]

Parameter	Value
n	4
θ	0.44 nM V
γ_r	3.5 h^{-1} [b]
γ_p	0.12 h^{-1}
γ_a	1.2 h^{-1}
β	3.6 $nM^{-1}\ h^{-1} V^{-1}$
Φ	0.6 $h^{-1}\ V$
κ_0	50 nM $h^{-1}\ V\ B^{-1}$ [c]
A	0–500 ng/mL [d]

[a]The cell volume V is assumed fixed. Unless otherwise noted, all values are taken from the experimental fit of ref 27. [b]Reference 34. [c]The burst ratio $B \equiv \kappa_p/\gamma_r$. Though not independently determined by the experimental fit, we assume that B is in the range $B = 2-10$.[33] [d]For external ATc concentration A, 1 ng/mL corresponds to 2.25 nM.

Linearizing eq 10 around $\bar{r}$ and $\bar{p}$, we extract the following frequency domain response functions:

$$G_{rr}(\omega) = -\tau_r^{-1} = -\gamma_r, \quad G_{rp}(\omega) = -\frac{\kappa_0 n\theta^n \bar{p}^{n-1}}{(\theta^n + \bar{p}^n)^2},$$

$$G_{pp}(\omega) = -\tau_p^{-1} = -\gamma_p - \frac{\beta\gamma_a\Phi A}{(\gamma_a + \beta\bar{p})^2}, \quad G_{pr}(\omega) = \kappa_p \qquad (11)$$

All the functions are constants in the frequency domain. Here τ_r and τ_p are effective decay times for the mRNA and proteins, respectively. The value of τ_r is fixed, and sets the intrinsic time scale of mRNA fluctuations, but τ_p and G_{rp} depend on $\bar{p}$, which is a function of the external ATc concentration A. In fact, association with intracellular ATc, described by the second term in the G_{pp} expression above, is the dominant form of decay for the free TetR dimers. Figure 3(d) plots the effective decay constants τ_r^{-1} and τ_p^{-1} as a function of A. Except for $A \lesssim 8$ ng/mL we are in the regime where $\tau_p^{-1} \gg \tau_r^{-1}$, which is relevant in simplifying the optimality condition for $G_{rp}(\omega)$ discussed below.

The optimal filter calculation for the TetR gene circuit depends on the linear response functions of eq 11. Using eqs 4 and 5, we obtain the following power spectra for the signal and noise in the absence of extrinsic noise:

$$P_s(\omega) = \frac{\overline{n_r(\omega)n_r(-\omega)}}{(G_{rr}(\omega) + i\omega)(G_{rr}(-\omega) - i\omega)} = \frac{2\bar{r}\tau_r}{1 + (\omega\tau_r)^2},$$

$$P_n(\omega) = \frac{\overline{n_p(\omega)n_p(-\omega)}}{G_{pr}(\omega)G_{pr}(-\omega)} = \frac{2\bar{r}\tau_r}{B} \qquad (12)$$

where the burst ratio $B \equiv \kappa_p\tau_r$ is the mean number of proteins synthesized per mRNA during the lifetime τ_r. The problem is to evaluate eq 9 for $H_{opt}(\omega)$. The sum of signal plus noise, $y(\omega) = s(\omega) + n(\omega)$, has a power spectrum $P_y(\omega) = P_s(\omega) + P_n(\omega)$, which we can rewrite as follows:

DOI: 10.1021/acs.jpcb.6b02093
J. Phys. Chem. B 2016, 120, 6166–6177

$$P_y(\omega) = 2\bar{r}\tau_r\left[\frac{1}{1+(\omega\tau_r)^2} + \frac{1}{B}\right] = \left|\left(\frac{2\bar{r}\tau_r}{B}\right)^{1/2}\frac{\sqrt{1+B} - i\omega\tau_r}{1 - i\omega\tau_r}\right|^2 \quad (13)$$

The expression within the absolute value brackets is zero only at $\omega = -i\tau_r^{-1}\sqrt{1+B}$, and has a simple pole at $\omega = -i\tau_r^{-1}$. Since all the zeros and poles are in the lower complex ω half-plane, it satisfies the criterion for the causal term in the factorization $P_y(\omega) = |P_y^c(\omega)|^2$. Thus

$$P_y^c(\omega) = \left(\frac{2\bar{r}\tau_r}{B}\right)^{1/2}\frac{\sqrt{1+B} - i\omega\tau_r}{1 - i\omega\tau_r} \quad (14)$$

The other causal term in eq 9 involves the additive decomposition $\{P_s(\omega)/P_y^c(\omega)^*\}_c$. This is calculated by looking at the partial fraction expansion of $P_s(\omega)/P_y^c(\omega)^*$:

$$\frac{P_s(\omega)}{P_y^c(\omega)^*} = \frac{(2\bar{r}\tau_r B)^{1/2}}{(1 - i\omega\tau_r)(\sqrt{1+B} + i\omega\tau_r)} = \frac{(2\bar{r}\tau_r B)^{1/2}}{(1 - i\omega\tau_r)(\sqrt{1+B} + 1)} + \frac{(2\bar{r}\tau_r B)^{1/2}}{(1 + \sqrt{1+B})(\sqrt{1+B} + i\omega\tau_r)} \quad (15)$$

Of the two terms in the partial fraction expansion, only the first has poles solely in the lower complex ω half-plane. Hence, it is the only one that contributes to $\{P_s(\omega)/P_y^c(\omega)^*\}_c$:

$$\left\{\frac{P_s(\omega)}{P_y^c(\omega)^*}\right\}_c = \frac{(2\bar{r}\tau_r B)^{1/2}}{(1 - i\omega\tau_r)(\sqrt{1+B} + 1)} \quad (16)$$

Inserting eqs 14 and 15 into eq 9, we finally find that the optimal filter is

$$H_{\mathrm{opt}}(\omega) = \frac{\sqrt{1+B} - 1}{\sqrt{1+B} - i\omega\tau_r} \quad (17)$$

Transforming $H_{\mathrm{opt}}(\omega)$ into the time domain, we find

$$H_{\mathrm{opt}}(t) = (\tau_{\mathrm{avg}}^{-1} - \tau_r^{-1})\mathrm{e}^{-t/\tau_{\mathrm{avg}}}\Theta(t) \quad (18)$$

where $\tau_{\mathrm{avg}} = \tau_r/\sqrt{1+B}$, and $\Theta(t)$ is a unit step function ensuring that the filter operates only on the past history of its input. For $B \gg 1$ the prefactor in eq 18 is $\approx \tau_{\mathrm{avg}}^{-1}$, and $H_{\mathrm{opt}}(t)$ has a straightforward interpretation: it approximately acts as a moving average of the corrupted signal $y(t) = s(t) + n(t)$ over a time scale τ_{avg}. In order to get the best estimate $\tilde{s}(t)$, the averaging interval τ_{avg} can neither be too long, since it would blur out the features of the signal $s(t)$ (which vary on the time scale τ_r), nor too short, since it would be ineffective at smoothing out the noise distortion $n(t)$. Hence, there must exist an optimum τ_{avg}, which is naturally proportional to τ_r, the main time scale for the mRNA.

In Figure 3(b), we show how the noise filter properties of the system vary with A for a burst ratio of $B = 10$. The filter function $H(t)$ (solid red curve) differs substantially from $H_{\mathrm{opt}}(t)$ (dotted red curve) for large and small A, but approaches the optimal form near $A = 54$ ng/mL. Consequently, at this value of A we get the closest correspondence between the plotted sample trajectories of signal $s(t)$ (cyan curve) and estimate $\tilde{s}(t)$ (blue curve). Similarly, the equilibrium probability distribution of the output, $P(\delta r)$, shown to the right of the trajectories, exhibits the smallest Fano factor $\sigma_r^2/\bar{r}$. The latter is a measure of noise magnitude, and has a reference value of unity if mRNA production was a pure Poisson process, as would be the case without feedback. Optimality is realized in the intermediate A regime of partial repression, where the R to P responsiveness, as measured by $|G_{rp}|$, is large. Effective noise suppression requires that R be sensitive to changes in P, so that information about R fluctuations can be transmitted through the negative feedback loop.

In order to understand the optimality condition for $H(t)$ in more detail, let us look at the explicit expression for $H(t)$ in the TetR system, given by the inverse Fourier transform of eq 6 with the response functions of eq 11:

$$H(t) = \frac{G_{rp}\kappa_p}{\omega_1 - \omega_2}(e^{-\omega_1 t} - e^{-\omega_2 t})\Theta(t) \quad (19)$$

where ω_1, ω_2 are the two ω roots of the denominator in eq 6. Assuming $\tau_p \ll \tau_r$ (which holds good except for small values $A \lesssim 8$ ng/mL, as seen in Figure 3(d)), we can directly show the approach of $H(t)$ to optimality at a specific intermediate value of G_{rp}. When G_{rp} equals $\mathcal{G}_{\mathrm{rp}}^{\mathrm{opt}}(B, \tau_p) = -1/(\tau_p(1 + \sqrt{1+B}))$, the roots $\omega_1 \approx \tau_{\mathrm{avg}}^{-1}$, $\omega_2 \approx \tau_p^{-1} + \tau_r^{-1} - \tau_{\mathrm{avg}}^{-1}$, up to corrections of order τ_p/τ_r^2. In this case, eq 19 becomes

$$H(t)|_{G_{rp}=\mathcal{G}^{\mathrm{opt}}} \approx H_{\mathrm{opt}}(t)\left[\frac{1 - e^{-(\tau_p^{-1}+\tau_r^{-1}-2\tau_{\mathrm{avg}}^{-1})t}}{1 + \tau_p(\tau_r^{-1} - 2\tau_{\mathrm{avg}}^{-1})}\right] \quad (20)$$

where the factor in the brackets on the right equals 1 in the limit $\tau_p \to 0$ for all $t > 0$. Up to this correction factor, we thus expect the system to behave optimally at $A = A_{\mathrm{opt}}$, defined by the condition $G_{\mathrm{rp}} = \mathcal{G}_{\mathrm{rp}}^{\mathrm{opt}}(B, \tau_p)$, so long as A_{opt} is large enough to satisfy $\tau_p \ll \tau_r$. Figure 3(e) shows G_{rp} and $\mathcal{G}_{\mathrm{rp}}^{\mathrm{opt}}$ curves for $B = 2$, 5, and 10, with dots marking the intersection points that define A_{opt} for each B. As explained above, $|G_{rp}|$ is small at small and large A, and reaches a maximum in between. At fixed B, $|\mathcal{G}_{\mathrm{rp}}^{\mathrm{opt}}(B, \tau_p)| \propto \tau_p^{-1}$, so it increases monotonically with A, as larger concentrations of the inducer increase the effective decay rate of free proteins. Thus, for each B there is a single intersection point A_{opt} at an intermediate concentration of the inducer.

Figure 3(f) shows the Fano factor $\sigma_r^2/\bar{r}$ versus A for various B. As the control network approximates optimality at A_{opt} for each B, the Fano factor nears its minimum, close to the theoretical limit marked by the horizontal dashed lines. This limit is the minimal possible $\sigma_r^2/\bar{r}$, calculated from eq 8 using $H_{\mathrm{opt}}(t)$ from eq 18:

$$\frac{\sigma_{r,\mathrm{opt}}^2}{\bar{r}} = \frac{2}{1 + \sqrt{1+B}} \geq \frac{2}{1 + \sqrt{1+4B}} \quad (21)$$

A few comments concerning the above equation are in order. (1) The result on the far right-hand side is the rigorous lower bound derived by LVP.[22] In their case, the feedback mechanism through the rate function $k_r(t)$ could be any causal functional of $p(t)$, linear or nonlinear. The Fano factor of the optimal linear filter differs in form only by the coefficient of B, and is always within a factor of 2 of the lower bound for any value of B. (2)

DOI: 10.1021/acs.jpcb.6b02093
J. Phys. Chem. B 2016, 120, 6166−6177

For Gaussian-distributed signal $s(t)$ and noise $n(t)$ time series, the linear filter is optimal among all possible filters.[31] If the system fluctuates around a single stable state, and the copy numbers of the species are large enough that their Poisson distributions converge to Gaussians (mean populations $\gtrsim$10), the signal and noise are usually approximately Gaussian. This is a wide class of systems where the rigorous lower bound (the last term in eq 21) can never be achieved. In other words, here the WK filter yields the most efficient feedback mechanism. Although, as pointed out by LVP, nonlinearity could lead to additional noise reduction, the benefits are likely to be restricted to those systems where the signal and/or noise are substantially non-Gaussian. However, since the form of the optimal control network has not been found in the general nonlinear case, it remains an interesting open question whether the LVP bound can actually be reached even within this category of systems. We will return to this issue in the next section. (3) The parameter B is the key determinant of noise reduction. For $B \ll 1$, there are not enough signaling events to control the mRNA fluctuations, and as $B \to 0$ we approach $\sigma^2_{r,\mathrm{opt}}/\bar{r} \to 1$, the no-feedback Poisson result. In the limit $B \gg 1$ signaling is effective, and the Fano factor decreases with B as $\sigma^2_{r,\mathrm{opt}}/\bar{r} \approx 2/\sqrt{B}$. For large enough B we approach perfect control, but at extreme expense: the standard deviation of the mRNA fluctuations $\sigma_{r,\mathrm{opt}} \propto B^{-1/4}$, the same scaling derived by LVP.

WK theory constrains the performance of a broad class of nonlinear, discrete regulatory networks. The results in Figure 3 rely on a linearized, continuum approach to the TetR dynamical system. To assess if the conclusions based on the WK optimal filter hold if these approximations are relaxed, we first performed kinetic Monte Carlo simulations of the full nonlinear system (eq 10) using the Gillespie algorithm.[35] We chose a cell volume of $V = V_0 = 60$ fL, within the observed range for yeast,[36] which corresponds to the mean populations $\bar{r}$ and $\bar{p}$ shown in Figure 4(a) as a function of A. (For example, at $A = A_{\mathrm{opt}} = 62.7$ ng/mL when $B = 5$, $\bar{r} \approx 84$ and $\bar{p} \approx 11$. In addition to the nonlinearity, the discrete nature of the populations in the simulation might play a role at these low copy numbers.) The numerical results for the Fano factor $\sigma^2_r/\bar{r}$ are plotted in Figure 4(b) at $B = 2$, 5, 10, for $V = V_0$ (circles) and also for comparison at a larger volume $V = 10V_0$ (squares). The blue curves show the linear theory results, and the dashed lines are the optimality predictions for $\sigma^2_{r,\mathrm{opt}}/\bar{r}$. Although nonlinearity and discreteness effects do change the results, the linear theory gives a reasonable approximation, and the minimum is still near A_{opt}. The feedback mechanism is nonlinear in the simulations, but it does not do better than the linear predictions for $\sigma^2_{r,\mathrm{opt}}/\bar{r}$ for the parameters used to describe the experimental results. Though the intrinsic population noise is Poisson-distributed in the simulations, the Poisson distribution is very close to Gaussian, even for copy numbers as low as $\sim O(10)$. Since the linear filter is the true optimum for a Gaussian-distributed signal and noise,[31] we do not expect improvements in noise suppression by employing a nonlinear version. In the opposite limit of large copy numbers, $V \to \infty$, the continuum approximation should be valid, and population fluctuations increasingly negligible relative to the mean. Thus, the linear theory should directly apply in this limit, and indeed we see that for $V = 10V_0$ the discrepancies between numerical and theory results are substantially reduced (Figure 4(b)). It is worth emphasizing, that even at the realistically small cell volume V_0, the linear theory retains much of its predictive power. More generally, the conditions for WK optimality do not have to be perfectly satisfied in order for the filter to perform close to maximum efficiency. There is an inherent adaptability and robustness in near-optimal networks, as reflected in the broad minima of $\sigma^2_r/\bar{r}$ as a function of A (Figure 4(b)).

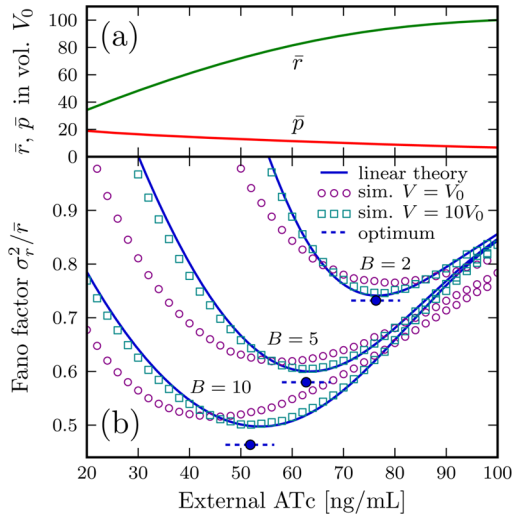


Figure 4. Results of simulation and theory for the yeast synthetic gene circuit,[27] as a function of extracellular ATc concentration A. (a) Mean populations of free TetR mRNA $\bar{r}$ and TetR dimer $\bar{p}$, assuming a cell volume $V_0 = 60$ fL. (b) The Fano factor $\sigma^2_r/\bar{r}$ for burst factor $B = 2$, 5, 10, as predicted by the linear filter theory (solid lines), versus stochastic numerical simulations at two different volumes, $V = V_0$ (circles) and $V = 10V_0$ (squares). The WK filter theory predicts the minimal Fano factor $\sigma^2_{\mathrm{opt}}/\bar{r}$ given by eq 21 (horizontal dashed lines). The system can be tuned to approach optimality near a particular A_{opt} obtained by the condition $G_{\mathrm{rp}} = \mathcal{G}^{\mathrm{opt}}_{\mathrm{rp}}$ (filled circles).

The semiquantitative agreement between the linearized theory and the simulation results displayed in Figure 4 still leaves open the possibility that some type of nonlinear, discrete filter, not described by the experimentally fitted parameters of the TetR gene network, could perform better than the WK optimum at sufficiently small volumes. Figure 5 plots both the WK value for the Fano factor (solid curve) and the rigorous lower bound of LVP (dashed curve) as a function of B (eq 21).

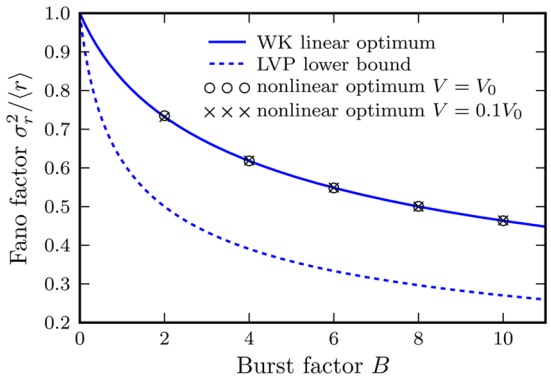


Figure 5. Fano factor $\sigma^2_r/\langle r\rangle$ as a function of burst ratio B. The solid curve is the optimal result predicted by the WK linear theory, and the dashed curve is the rigorous lower bound derived by LVP.[22] Symbols show numerical optimization for the generalized nonlinear TetR feedback system (eq 22) at two volumes, $V = V_0$ and $V = 0.1V_0$.

DOI: 10.1021/acs.jpcb.6b02093
J. Phys. Chem. B 2016, 120, 6166−6177

The above question can be posed as follows: is it possible to achieve a Fano factor that falls between the two curves by taking advantage of nonlinearity and discreteness? Ideally, one should do an optimization over all possible nonlinear regulatory functions that could describe feedback between the TetR protein and mRNA. In full generality, such an optimization appears intractable, but one can tackle a limited version of the nonlinear optimization. We will confine ourselves to Hill-like regulatory functions, which describe the experimental behavior of many cellular systems,[37] and explore whether it is possible to find any scenario where this type of nonlinear feedback outperforms the linear WK optimum. We consider the following generalized TetR feedback loop:

$$\dot{r}(t) = -\gamma_r r(t) + K_r(p(t)) + n_r(t),$$
$$\dot{p}(t) = -\gamma_p p(t) - \Gamma_p(p(t)) + \kappa_p r(t) + n_p(t) \quad (22)$$

where $n_r(t) = \sqrt{2\gamma_r \bar{r}}\,\eta_r(t)$ and $n_p(t) = \sqrt{2\kappa_p \bar{r}}\,\eta_p(t)$. This system has two Hill-like regulatory functions,

$$K_r(p) = \frac{A_1\theta_1^{n_1}}{\theta_1^{n_1} + p^{n_1}}, \quad \Gamma_p(p) = \frac{A_2 p^{n_2}}{\theta_2^{n_2} + p^{n_2}} \quad (23)$$

involving arbitrary non-negative parameters A_i, n_i, θ_i, $i = 1, 2$. The original TetR system (eq 10) is a special case of the equations above with

$$A_1 = \kappa_0, \quad n_1 = n, \quad \theta_1 = \theta, \quad A_2 = \beta\Phi A, \quad n_2 = 1,$$
$$\theta_2 = \gamma_a \quad (24)$$

The production function $K_r(p)$ is a monotonically decreasing function of p, as is expected for negative feedback, while $\Gamma_p(p)$ is monotonically increasing, a generalization of some regulatory network which effectively removes the TetR protein from the feedback loop (the role played by ATc binding in the experimental system). With these monotonicity constraints, there is always only one steady-state solution $\bar{r}$ and $\bar{p}$ to eq 22.

The optimization consists of searching for $K_r(p)$ and $\Gamma_p(p)$ that minimize the Fano factor $\sigma_r^2/\langle r\rangle$. The following quantities are fixed during the search: the degradation rates γ_r, γ_p, the P production rate κ_p (or equivalently the burst ratio $B = \kappa_p/\gamma_r$), and the steady state values $\bar{r}$, $\bar{p}$. Note that in the general nonlinear case, the steady state values do not necessarily coincide with the mean values $\langle r\rangle$, $\langle p\rangle$, since the equilibrium distributions are generally asymmetric with respect to the steady state. Fixing $\bar{r}$ and $\bar{p}$ during the optimization is one way to set an overall copy number scale, to investigate the role of discreteness. It turns out that the optimization results described below end up being independent of $\bar{r}$ and $\bar{p}$. In terms of the Hill function parameters, fixing $\bar{r}$ and $\bar{p}$ means setting A_1 and A_2 to the following values,

$$A_1 = \theta_1^{-n_1}\gamma_r\bar{r}(\theta_1^{n_1} + \bar{p}^{n_1}),$$
$$A_2 = \bar{p}^{-n_2}(\gamma_p\bar{p} - \kappa_p\bar{r})(\theta_2^{n_2} + \bar{p}^{n_2}) \quad (25)$$

Thus, the goal of optimization is to minimize $\sigma_r^2/\langle r\rangle$ over the four remaining free parameters: n_1, θ_1, n_2, θ_2.

In order to carry out this minimization, one needs an efficient procedure to calculate $\sigma_r^2/\langle r\rangle$ from eq 22, keeping both the full nonlinearity of the dynamical system, and the discreteness of the $r(t)$ and $p(t)$ populations, which means going beyond the continuum Langevin description in eq 22. The system can always be simulated through the Gillespie algorithm,[35] and accurate estimates of $\langle r\rangle$ and σ_r^2 determined from sufficiently long trajectories. However, this approach is too slow for searching over the four-dimensional parameter space, since each distinct set of parameters would require a separate long simulation run. An equivalent, faster alternative is to directly solve the system's master equation for the steady state probability distribution, which then yields $\langle r\rangle$ and σ_r^2. The joint probability distribution $P_{r,p}(t)$ of finding r mRNAs and p proteins at time t is governed by the master equation,

$$\frac{\partial}{\partial t}P_{r,p} = \gamma_r[(r+1)P_{r+1,p} - rP_{r,p}] + K_r(p)[P_{r-1,p} - P_{r,p}]$$
$$+ \gamma_p[(p+1)P_{r,p+1} - pP_{r,p}]$$
$$+ [\Gamma_p(p+1)P_{r,p+1} - \Gamma_p(p)P_{r,p}] + \kappa_p r[P_{r,p-1} - P_{r,p}] \quad (26)$$

The steady state distribution $P^s_{r,p}$ is the solution obtained by setting to zero the right-hand side of the above equation, which we denote $\mathcal{R}_{r,p}$:

$$0 = \mathcal{R}_{r,p} \equiv \gamma_r[(r+1)P^s_{r+1,p} - rP^s_{r,p}]$$
$$+ K_r(p)[P^s_{r-1,p} - P^s_{r,p}] + \gamma_p[(p+1)P^s_{r,p+1} - pP^s_{r,p}]$$
$$+ [\Gamma_p(p+1)P^s_{r,p+1} - \Gamma_p(p)P^s_{r,p}] + \kappa_p r[P^s_{r,p-1} - P^s_{r,p}] \quad (27)$$

The result is linear in the components $P^s_{r,p}$ for various r and p, and thus the set $\{\mathcal{R}_{rp} = 0\}$ for $r = 0, 1, \ldots$ and $p = 0, 1, \ldots$ constitutes a linear system of equations for $P^s_{r,p}$. The master equation can be solved by spectral methods, which are generally more efficient than brute force Gillespie simulations.[38] However, we use a different approach, described below, to solve eq 27, which is sufficiently fast for our numerical optimization purposes. Since r and p can take on any integer values between 0 and ∞, we truncate the system to focus only on the non-negligible $P^s_{r,p}$, in other words (r,p) within several standard deviations of the mean $(\langle r\rangle, \langle p\rangle)$. Specifically, we keep only those equations $\mathcal{R}_{rp} = 0$ which involve $r_{min} \le r \le r_{max}$ and $p_{min} \le p \le p_{max}$. The largest truncation range required for accurate results was $r_{max} - r_{min} = 100$ and $p_{max} - p_{min} = 50$. All $P^s_{r,p}$ outside the range which appear in the truncated system of equations are set to a positive constant $\epsilon > 0$. (The precise value of ϵ is unimportant since the distribution is subsequently normalized, and the truncation range is chosen large enough so that the boundary condition does not significantly affect the outcome.) The resulting finite linear system, which is sparse, can be efficiently solved using an unsymmetric-pattern multifrontal algorithm.[39] Knowing $P^s_{r,p}$, we then directly calculate the moments of the distribution to find $\langle r\rangle$ and σ_r^2. The numerical accuracy of the procedure is verified by comparison to Gillespie simulation results.

In order to set a starting point for each round of nonlinear optimization, we use the following initialization procedure: we take the original TetR system at a given volume V and burst ratio B (fixing the Hill function parameters according to eq 24) and find the ATc concentration A_{min} where $\sigma_r^2/\langle r\rangle$ is smallest, evaluating the Fano factor using the linear solver described above. The $\bar{r}$ and $\bar{p}$ at this concentration are then chosen to be fixed constants for the nonlinear optimization, where we vary the parameters n_1, θ_1, n_2, θ_2 from the initial values given by eq 24 to minimize $\sigma_r^2/\langle r\rangle$. The minimization is carried out using

DOI: 10.1021/acs.jpcb.6b02093
J. Phys. Chem. B 2016, 120, 6166−6177

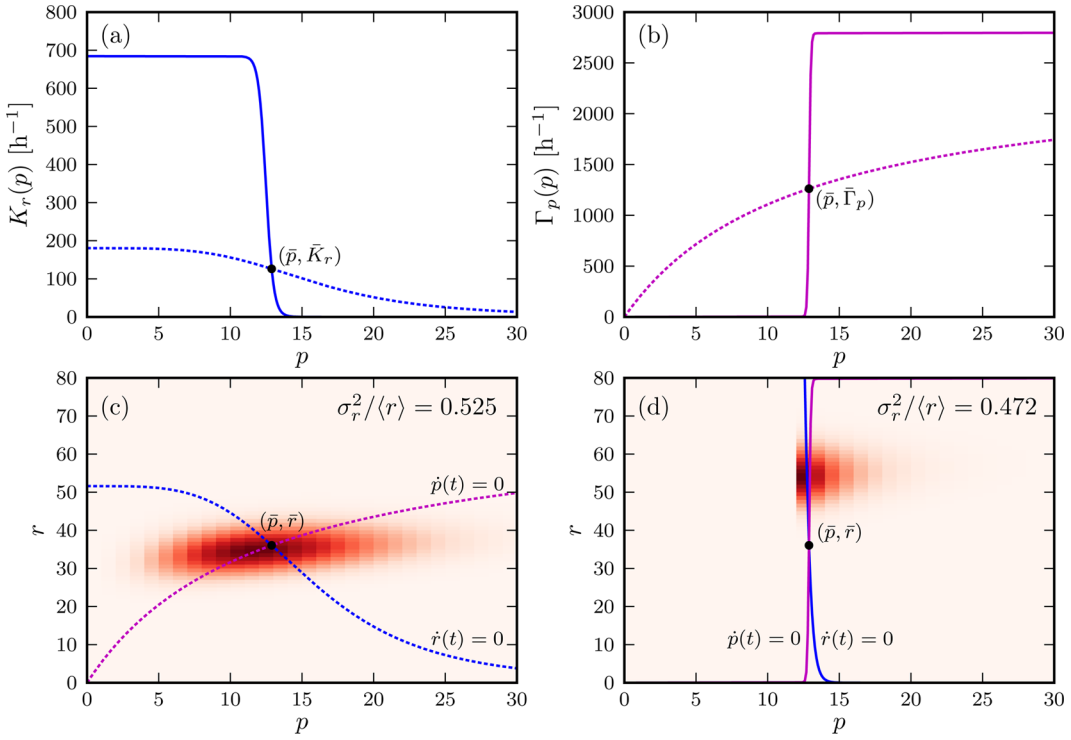


Figure 6. Results for numerical optimization of the generalized nonlinear TetR feedback system of eq 22, with starting parameters $B = 10$ and $V = V_0$. (a) The mRNA production regulation function $K_r(p)$ in its initial form before optimization (dashed curve), and after several steps of the minimization algorithm (solid curve). (b) Similar to (a), but showing the protein degradation function $\Gamma_p(p)$. (c) Heat map of the steady-state probability distribution $P^s_{r,p}$ before optimization, corresponding to regulation governed by the dashed curves in the top panels. The nullclines $\dot{r}(t) = 0$ and $\dot{p}(t) = 0$ are superimposed. (d) Similar to (c), but after several steps of the minimization algorithm, corresponding to regulation governed by the solid curves in the top panels.

Brent's principal axis method,[40] which is feasible due to the fast evaluation of $\langle r\rangle$ and σ_r^2 at each different parameter set through the linear solver.

Figure 6 shows results of a typical minimization run, where the initial system is at volume $V = V_0$ with $B = 10$, with a corresponding $A_{\min} = 50$ ng/mL. The dashed lines in Figure 6(a) and (b) show the Hill functions $K_r(p)$ and $\Gamma_p(p)$ of the original TetR system at these parameter values, and the heat map in Figure 6(c) represents the associated steady-state probability distribution $P^s_{r,p}$. The dashed lines superimposed on the heat map are the loci of solutions to $\dot{r}(t) = 0$ and $\dot{p}(t) = 0$ (the right-hand sides of eq 22 set to zero), which intersect at the steady state $(\bar{r}, \bar{p})$. The Fano factor for this distribution, which represents the best the TetR system can perform given the experimentally fitted parameters, is $\sigma_r^2/\langle r\rangle = 0.525$. This is above the linear WK optimum for $B = 10$, $2/(1 + \sqrt{1 + B}) = 0.463$, and significantly larger than the rigorous LVP lower bound of $2/(1 + \sqrt{1 + 4B}) = 0.270$. Once we relax the experimental constraints, and carry out the numerical minimization, the Fano factor decreases. The solid lines in Figure 6(a) and (b) show $K_r(p)$ and $\Gamma_p(p)$ after several steps of the minimization algorithm, and Figure 6(d) shows the corresponding $P^s_{r,p}$. The Hill functions have become very steep steps around $\bar{p}$, while the average of the distribution $\langle r\rangle$ has been pushed above $\bar{r}$. The probabilities $P^s_{r,p}$ for $p < p_0$ become negligible, where $p_0 \equiv \lfloor \bar{p} \rfloor$ is the largest integer value below $\bar{p}$. For $p > p_0$, $P^s_{r,p}$ rapidly decay to zero. The Fano factor, $\sigma_r^2/\langle r\rangle = 0.472$, approaches closer to the linear WK optimum, but is still above it. If we allow the minimization to proceed, these trends continue: at each iteration the Hill functions get steeper, $\langle r\rangle$ increases, $P^s_{r,p}$ for $p < p_0$ tends to zero, and $\sigma_r^2/\langle r\rangle$ approaches arbitrarily close to the linear WK optimum from above.

In fact, the same behavior is seen irrespective of the volume V and burst ratio B used to define the initial point of the optimization. Figure 5 shows the results of nonlinear optimization for $B = 2$–10 at two volumes, $V = V_0$ and $V = 0.1V_0$. Even for the smallest volume, the nonlinear optimization results can get arbitrarily close to the WK optimum, but never do better. No generalized nonlinear system based on Hill function regulation brings us close to the theoretically possible LVP lower bound. This overall conclusion holds even when we change the functional form for the generalized feedback. We tried two alternatives: (i) using sigmoidal (logistic) functions instead of Hill functions; (ii) expanding $K_r(p)$ and $\Gamma_p(p)$ in a Taylor series around $\bar{p}$, truncating after the third order term, and minimizing with respect to the Taylor coefficients. In both cases numerical minimization of the Fano factor led to similar step-like behavior for $K_r(p)$ and $\Gamma_p(p)$, and the Fano factor tended to WK optimum from above.

From the $P^s_{r,p}$ distribution in Figure 6(d) we see that the step-function limit leads to a system which is highly nonlinear along the p axis: in fact the gene network spends most of its time at $p = p_0$, just below the sudden change in regulation due to the steep Hill functions, and $p > p_0$ just above the sudden regulatory change. The feedback on the TetR mRNA population is mediated by p fluctuations between the two regimes, resulting in threshold-like regulatory behavior. Remarkably, despite this discrete, nonlinear character, the network can still approach the efficiency of an optimal WK linear filter. To gain a deeper understanding of how the step-like regulation can match WK optimality, we used the

DOI: 10.1021/acs.jpcb.6b02093
J. Phys. Chem. B 2016, 120, 6166–6177

numerical optimization results described above to posit a limiting form of the nonlinear gene network that can be solved analytically (details in SI Sec. 3). The analytic results explicitly show that we can asymptotically approach the WK optimum behavior from above, even in systems where the protein copy numbers are very small. Thus, at least for a two-component TetR-like system regulated by biologically realistic Hill functions, the constraint derived from the WK theory has a broader validity than one would guess from the underlying continuum, linear assumptions. It thus becomes an interesting and a nontrivial problem, left for future studies, to find an example of a gene network where the rigorous lower bound of LVP could be directly achieved.

Realizing optimality under the influence of extrinsic noise. Extrinsic noise is ubiquitous and hence must also be considered in any effective description of the control network. Inevitably, certain cellular components are not explicitly included in such a description, which in our case study could include RNA polymerase, ribosomes, and transcription factors that bind to the same promoter. Each of these components have their own stochastic characteristics and may contribute noise to a smaller or greater extent. Particularly for eukaryotes like yeast, the extrinsic noise contribution may be significantly larger than the intrinsic component.[41,42] We adopt a simple model for the extrinsic noise based on earlier approaches,[16,18] which assume that it is band-limited at a low frequency τ_e^{-1}, where τ_e is on the order of the cell growth time scale. The justification is that higher frequency contributions to the extrinsic noise are filtered out by the gene circuits associated with its sources. This idea is consistent with the experimental observation of extrinsic noise in protein production in *E. coli*, which found long autocorrelation times for the extrinsic noise on the order of the cell cycle period.[43]

For the TetR system, our theory is extended to the extrinsic noise case in SI section 4, with the results illustrated in Figure 7. The outcome is that a given TetR gene circuit, tuned appropriately such that $A = A_{opt}$, can act as a WK filter for an entire family of extrinsic noise scenarios. A single set of parameters can approximately represent the optimal solution for a variety of extrinsic inputs. This makes the WK concept a versatile design tool for noise suppression in biological systems: the same control network can act with maximum efficiency in a variety of different contexts. It is possible that the requirement of adaptability to a wide range of conditions has resulted in the evolution of control networks acting as WK filters. It remains to be seen whether nature has exploited this feature *in vivo*.

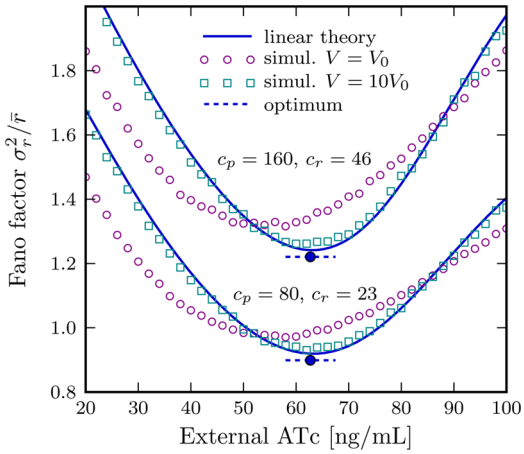


Figure 7. Comparison of simulation and theory results based on the dynamical model (eq 10) of the yeast synthetic gene circuit,[27] in the presence of extrinsic noise given by SI eq 46. All quantities are plotted as a function of extracellular ATc concentration A for the burst ratio $B = 5$. Each set of curves shows the Fano factor $\sigma_r^2/\bar{r}$, as predicted by the linear filter theory (solid lines), versus stochastic numerical simulations at two different volumes, $V = V_0 = 60$ fL (circles) and $V = 10V_0$ (squares). The two sets correspond to noise magnitudes $c_p = 80$, $c_r = 23$ and $c_p = 160$, $c_r = 46$. In both cases c_r and c_p are related through the condition in SI eq 56, and the minimal Fano factor predicted by WK filter theory (horizontal dashed lines) is modified as shown in SI eq 57. The system can be tuned to approach optimality near a particular A_{opt} obtained by the condition $G_{rp} = \mathcal{G}_{rp}^{opt}$ (filled circles).

CONCLUSION

The TetR feedback loop is a concrete example of how a WK filter can be implemented in a gene network driven by a complex set of biochemical reaction rates, but the overall approach outlined here has far reaching implications, thus highlighting the appeal of engineering paradigms in biology.[44] With the entire network complexity encoded in a handful of response functions, we can derive fundamental limits and design principles governing biological regulation. The key step is to map the linear response picture onto a signal estimation problem, whose solution is given by WK theory. This idea allows us to predict the dynamic properties of the feedback pathway required to optimally filter noise in a broad class of negative feedback circuits. As already demonstrated in earlier works,[25,26] the mapping, and the potential utility of the WK approach, is not unique to the negative feedback loop. Another important byproduct of the theory is that the behavior of gene circuits away from optimality can also be predicted. In this sense, our practical approach goes beyond just obtaining rigorous bounds, and allows us to characterize how close or far gene networks are from optimality for biologically relevant parameters.

We have derived response functions by linearizing a minimal model extracted from experimental observations, but it is also possible to directly apply small perturbations to a system, and measure the resulting time-dependent changes in populations of species. Recently, the yeast hyperosmolar signaling pathway has been probed by perturbations in the form of salt shocks.[45−47] Despite the underlying complex nonlinear network, the details of which are not completely characterized, a linear response description quantitatively captures the frequency-dependent behavior of the pathway over a wide range of inputs. *E. coli* chemotaxis signaling also exhibits a linear regime,[48] where the fluctuation−dissipation relationship between the system's unperturbed behavior and its reaction to external stimuli has been explicitly verified.

Linear response functions can thus become a fundamental tool in analyzing biochemical circuits, analogous to their established role in control engineering and signal processing. More extensive experimental measurements will be critical in this effort, in order to ascertain how varied the response relationships between regulatory components are in nature. Once we understand the essential dynamical building blocks out of which a complex biological function is realized, we can map out the hidden constraints that control the behavior of living systems.

DOI: 10.1021/acs.jpcb.6b02093
J. Phys. Chem. B 2016, 120, 6166−6177

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jpcb.6b02093.

Detailed derivations of our analytical results (PDF)

■ AUTHOR INFORMATION

Corresponding Authors

*E-mail: mxh605@case.edu.
*E-mail: dave.thirumalai@gmail.com.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

DT is grateful to Bill Gelbart for stimulating discussions on a variety of topics over the years. Bill's pioneering of studies in theory and experiments have been more than inspirational. Of utmost importance is Bill's friendship and counsel that DT has enjoyed for over twenty five years. This work was done while the authors were at the Institute for Physical Sciences and Technology in the University of Maryland, College Park. We would like to thank C. Güven, G. Reddy, Z. Zhang, and P. Zhuravlev for useful discussions. The research was supported by a grant from the National Science Foundation (CHE 13-61946).

■ REFERENCES

(1) Becskei, A.; Serrano, L. Engineering Stability in Gene Networks by Autoregulation. *Nature* **2000**, *405*, 590−593.
(2) Thattai, M.; van Oudenaarden, A. Intrinsic Noise in Gene Regulatory Networks. *Proc. Natl. Acad. Sci. U. S. A.* **2001**, *98*, 8614−8619.
(3) Swain, P. S.; Elowitz, M. B.; Siggia, E. D. Intrinsic and Extrinsic Contributions to Stochasticity in Gene Expression. *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99*, 12795−12800.
(4) Ozbudak, E. M.; Thattai, M.; Kurtser, I.; Grossman, A. D.; van Oudenaarden, A. Regulation of Noise in the Expression of a Single Gene. *Nat. Genet.* **2002**, *31*, 69−73.
(5) Elowitz, M. B.; Levine, A. J.; Siggia, E. D.; Swain, P. S. Stochastic Gene Expression in a Single Cell. *Science* **2002**, *297*, 1183−1186.
(6) Paulsson, J. Summing Up the Noise in Gene Networks. *Nature* **2004**, *427*, 415−418.
(7) Iyer-Biswas, S.; Hayot, F.; Jayaprakash, C. Stochasticity of Gene Products from Transcriptional Pulsing. *Phys. Rev. E* **2009**, *79*, 1−9.
(8) Pulkkinen, O.; Metzler, R. Variance-corrected Michaelis-Menten Equation Predicts Transient Rates of Single-enzyme Reactions and Response Times in Bacterial Gene-regulation. *Sci. Rep.* **2015**, *5*, 1−11.
(9) Arias, A. M.; Hayward, P. Filtering Transcriptional Noise During Development: Concepts and Mechanisms. *Nat. Rev. Genet.* **2006**, *7*, 34−44.
(10) Tsang, J.; Zhu, J.; van Oudenaarden, A. MicroRNA-mediated Feedback and Feedforward Loops are Recurrent Network Motifs in Mammals. *Mol. Cell* **2007**, *26*, 753−767.
(11) Andrews, B. W.; Yi, T.-M.; Iglesias, P. A. Optimal Noise Filtering in the Chemotactic Response of Escherichia coli. *PLoS Comput. Biol.* **2006**, *2*, 1407−1418.
(12) Newman, J. R. S.; Ghaemmaghami, S.; Ihmels, J.; Breslow, D. K.; Noble, M.; DeRisi, J. L.; Weissman, J. S. Single-cell Proteomic Analysis of S. cerevisiae Reveals the Architecture of Biological Noise. *Nature* **2006**, *441*, 840−846.
(13) Fraser, H. B.; Hirsh, A. E.; Giaever, G.; Kumm, J.; Eisen, M. B. Noise Minimization in Eukaryotic Gene Expression. *PLoS Biol.* **2004**, *2*, 834−838.
(14) Lehner, B. Selection to Minimise Noise in Living Systems and its Implications for the Evolution of Gene Expression. *Mol. Syst. Biol.* **2008**, *4* (170), 1−6.
(15) Simpson, M. L.; Cox, C. D.; Sayler, G. S. Frequency Domain Analysis of Noise in Autoregulated Gene Circuits. *Proc. Natl. Acad. Sci. U. S. A.* **2003**, *100*, 4551−4556.
(16) Austin, D. W.; Allen, M. S.; McCollum, J. M.; Dar, R. D.; Wilgus, J. R.; Sayler, G. S.; Samatova, N. F.; Cox, C. D.; Simpson, M. L. Gene Network Shaping of Inherent Noise Spectra. *Nature* **2006**, *439*, 608−611.
(17) Dublanche, Y.; Michalodimitrakis, K.; Kummerer, N.; Foglierini, M.; Serrano, L. Noise in Transcription Negative Feedback Loops: Simulation and Experimental Analysis. *Mol. Syst. Biol.* **2006**, *2* (41), 1−12.
(18) Cox, C. D.; McCollum, J. M.; Austin, D. W.; Allen, M. S.; Dar, R. D.; Simpson, M. L. Frequency Domain Analysis of Noise in Simple Gene Circuits. *Chaos* **2006**, *16*, 1−15.
(19) Zhang, J. J.; Yuan, Z. J.; Zhou, T. S. Physical Limits of Feedback Noise-Suppression in Biological Networks. *Phys. Biol.* **2009**, *6*, 1−10.
(20) Singh, A.; Hespanha, J. P. Optimal Feedback Strength for Noise Suppression in Autoregulatory Gene Networks. *Biophys. J.* **2009**, *96*, 4013−4023.
(21) Iyer-Biswas, S.; Jayaprakash, C. Mixed Poisson Distributions in Exact Solutions of Stochastic Autoregulation Models. *Phys. Rev. E* **2014**, *90*, 1−5.
(22) Lestas, I.; Vinnicombe, G.; Paulsson, J. Fundamental Limits on the Suppression of Molecular Fluctuations. *Nature* **2010**, *467*, 174−178.
(23) Wiener, N. *Extrapolation, Interpolation and Smoothing of Stationary Times Series*; Wiley: New York, 1949.
(24) Kolmogorov, A. N. Interpolation and Extrapolation of Stationary Random Sequences. *Izv. Akad. Nauk SSSR., Ser. Mater.* **1941**, *5*, 3−14.
(25) Hinczewski, M.; Thirumalai, D. Cellular Signaling Networks Function as Generalized Wiener-Kolmogorov Filters to Suppress Noise. *Phys. Rev. X* **2014**, *4*, 1−15.
(26) Becker, N. B.; Mugler, A.; ten Wolde, P. R. Optimal Prediction by Cellular Signaling Networks. *Phys. Rev. Lett.* **2015**, *115*, 1−5.
(27) Nevozhay, D.; Adams, R. M.; Murphy, K. F.; Josic, K.; Balazsi, G. Negative Autoregulation Linearizes the Dose-Response and Suppresses the Heterogeneity of Gene Expression. *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106*, 5123−5128.
(28) Schwanhaeusser, B.; Busse, D.; Li, N.; Dittmar, G.; Schuchhardt, J.; Wolf, J.; Chen, W.; Selbach, M. Global Quantification of Mammalian Gene Expression Control. *Nature* **2011**, *473*, 337−342.
(29) Gillespie, D. T. The Chemical Langevin Equation. *J. Chem. Phys.* **2000**, *113*, 297−306.
(30) de Ronde, W. H.; Tostevin, F.; ten Wolde, P. R. Effect of Feedback on the Fidelity of Information Transmission of Time-varying Signals. *Phys. Rev. E* **2010**, *82*, 1−21.
(31) Bode, H. W.; Shannon, C. E. A Simplified Derivation of Linear Least Square Smoothing and Prediction Theory. *Proc. IRE* **1950**, *38*, 417−425.
(32) Chaikin, P. M.; Lubensky, T. C. *Principles of Condensed Matter Physics*; Cambridge University Press, 1995.
(33) Cai, L.; Friedman, N.; Xie, X. S. Stochastic Protein Expression in Individual Cells at the Single Molecule Level. *Nature* **2006**, *440*, 358−362.
(34) Garcia-Martinez, J.; Aranda, A.; Perez-Ortin, J. E. Genomic Run-on Evaluates Transcription Rates for All Yeast Genes and Identifies Gene Regulatory Mechanisms. *Mol. Cell* **2004**, *15*, 303−313.
(35) Gillespie, D. T. Exact Stochastic Simulation of Coupled Chemical Reactions. *J. Phys. Chem.* **1977**, *81*, 2340−2361.
(36) Jorgensen, P.; Nishikawa, J. L.; Breitkreutz, B. J.; Tyers, M. Systematic Identification of Pathways that Couple Cell Growth and Division in Yeast. *Science* **2002**, *297*, 395−400.
(37) Alon, U. *An Introduction to Systems Biology: Design Principles of Biological Circuits*; Chapman and Hall/CRC, 2006.

DOI: 10.1021/acs.jpcb.6b02093
J. Phys. Chem. B 2016, 120, 6166−6177

(38) Mugler, A.; Walczak, A. M.; Wiggins, C. H. Spectral Solutions to Stochastic Models of Gene Expression with Bursts and Regulation. *Phys. Rev. E* **2009**, *80*, 1−19.
(39) Davis, T. A. Algorithm 832: UMFPACK V4.3−an Unsymmetric-pattern Multifrontal Method. *ACM Trans. Math. Software* **2004**, *30*, 196−199.
(40) Brent, R. *Algorithms for Minimization without Derivatives*; Dover, 2002.
(41) Raser, J. M.; O'Shea, E. K. Control of Stochasticity in Eukaryotic Gene Expression. *Science* **2004**, *304*, 1811−1814.
(42) Volfson, D.; Marciniak, J.; Blake, W. J.; Ostroff, N.; Tsimring, L. S.; Hasty, J. Origins of Extrinsic Variability in Eukaryotic Gene Expression. *Nature* **2006**, *439*, 861−864.
(43) Rosenfeld, N.; Young, J. W.; Alon, U.; Swain, P. S.; Elowitz, M. B. Gene Regulation at the Single-cell Level. *Science* **2005**, *307*, 1962−1965.
(44) Csete, M. E.; Doyle, J. C. Reverse Engineering of Biological Complexity. *Science* **2002**, *295*, 1664−1669.
(45) Mettetal, J. T.; Muzzey, D.; Gomez-Uribe, C.; van Oudenaarden, A. The Frequency Dependence of Osmo-adaptation in Saccharomyces cerevisiae. *Science* **2008**, *319*, 482−484.
(46) Hersen, P.; McClean, M. N.; Mahadevan, L.; Ramanathan, S. Signal Processing by the HOG MAP Kinase Pathway. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105*, 7165−7170.
(47) Muzzey, D.; Gomez-Uribe, C. A.; Mettetal, J. T.; van Oudenaarden, A. A Systems-Level Analysis of Perfect Adaptation in Yeast Osmoregulation. *Cell* **2009**, *138*, 160−171.
(48) Park, H.; Pontius, W.; Guet, C. C.; Marko, J. F.; Emonet, T.; Cluzel, P. Interdependence of Behavioural Variability and Response to Small Stimuli in Bacteria. *Nature* **2010**, *468*, 819−823.

DOI: 10.1021/acs.jpcb.6b02093
J. Phys. Chem. B 2016, 120, 6166−6177