

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27

Fitness costs of noise in biochemical reaction networks and the evolutionary limits of cellular robustness

J. David Van Dyken

Department of Biology, Institute of Theoretical and Mathematical Ecology, University of Miami, Coral Gables, FL 33146

vandyken@bio.miami.edu

Keywords: gene expression noise; drift barrier hypothesis; robustness; noise propagation; allometric scaling; coarse-grained cell model; linear noise approximation; Michaelis-Menten reaction; non-linear stochastic model

ABSTRACT

Gene expression is inherently noisy, but little is known about whether noise affects cell function or, if so, how and by how much. Here I present a theoretical framework to quantify the fitness costs of gene expression noise and identify the evolutionary and synthetic targets of noise control. I find that gene expression noise reduces fitness by slowing the average rate of nutrient uptake and protein synthesis. This is a direct consequence of the hyperbolic (Michaelis-Menten) kinetics of most biological reactions, which I show cause “hyperbolic filtering”, a process that diminishes both the average rate and noise propagation of stochastic reactions. Interestingly, I find that transcriptional noise directly slows growth by slowing the average translation rate. Perhaps surprisingly, this is the largest fitness cost of transcriptional noise since translation strongly filters mRNA noise, making protein noise largely independent of transcriptional noise, consistent with empirical data. Translation, not transcription, then, is the primary target of protein noise control. Paradoxically, selection for protein-noise control favors increased ribosome-mRNA binding affinity, even though this increases translational bursting. However, I find that the efficacy of selection to suppress noise decays faster than linearly with increasing cell size. This predicts a stark, cell-size-mediated taxonomic divide in selection pressures for noise control: small unicellular species, including most prokaryotes, face fairly strong selection to suppress gene expression noise, whereas larger unicells, including most eukaryotes, experience extremely weak selection. I suggest that this taxonomic discrepancy in selection efficacy contributed to the evolution of greater gene-regulatory complexity in eukaryotes.

ARTICLE SUMMARY

Gene expression is a probabilistic process, resulting in random variation in mRNA and protein abundance among cells called “noise”. Understanding how noise affects cell function is a major problem in biology. Here I present theory demonstrating that gene expression noise slows the average rate of cell division. Furthermore, by modeling stochastic gene expression with non-linearity, I identify novel mechanisms of cellular robustness. However, I find that the cost of noise, and therefore the strength of selection favoring robustness, decays faster than linearly with increasing cell size. This may help explain the vast differences in gene-regulatory complexity between prokaryotes and eukaryotes.

INTRODUCTION

The probabilistic nature of chemical reactions, along with the small volumes of living cells, create substantial molecular randomness called “noise” (Berg 1978; McAdams and Arkin 1997). Gene expression noise, for example, is measured as variation in mRNA or protein abundance within a single cell over time (Taniguchi et al. 2010), or, equivalently, among cells in an isogenic cell population (Elowitz et al. 2002). Because the abundance of most mRNAs and proteins are on the order of 1 to 1000 copies per cell (Taniguchi et al. 2010), a random birth or death of a single copy can instantaneously change the cellular concentration by 0.1-100%. These random fluctuations are amplified by bursting, the synthesis of multiple mRNAs or proteins in brief pulses (Raser and O’Shea 2004; Raser and O’Shea 2005; Raj et al. 2006; Pedraza and Paulsson 2008; Thattai and van Oudenaarden 2001). Using fluorescent probes and fluorescently tagged proteins,

researchers have quantified the noise statistics of mRNA and/or proteins in *E. coli*, budding yeast, and mammalian cell lines, demonstrating the ubiquity of gene expression noise and its considerable variability across the genome and among taxa (Elowitz et al. 2002; Blake et al. 2003; Golding et al. 2005; Raj et al. 2006; Zenklusen, Larson, and Singer 2008; Taniguchi et al. 2010; Newman et al. 2006; Balázsi, van Oudenaarden, and Collins 2011).

A central question in cell biology and evolution is whether or not cellular noise has a functional consequence. Answering this question is necessary for establishing rational design criteria for synthetic genetic circuits and for providing evolutionary explanations for empirical patterns of noise among genes and taxa. For example, why does *E. coli* have an average mRNA Fano factor (variance/mean) of $F = 1.6$ (Taniguchi et al. 2010), which is just above the theoretical minimum of $F = 1$, whereas mammalian cells have 40-100 times stronger mRNA noise (Raj et al. 2006; Balázsi, van Oudenaarden, and Collins 2011; Dar et al. 2012)? While the mechanistic causes of this discrepancy are becoming clearer (Suter et al. 2011), the evolutionary forces permitting such a discrepancy remain murky.

Mounting empirical evidence suggests that natural selection has acted to limit gene expression noise in many cases (Metzger et al. 2015; Newman et al. 2006; Alemu et al. 2014; Batada and Hurst 2007; Fraser et al. 2004; Lehner 2008), implying that noise is often costly to cells. Yet identifying and quantifying this cost remains a major challenge. Perhaps the most widely invoked cost of cellular noise is imprecision, the lack of fine control over cellular processes. Evolutionarily, this cost of noise is typically modeled as random phenotypic deviations from the adaptive optimum, which is opposed by

stabilizing selection (Wagner et al. 1997; Draghi and Whitlock 2015). Several molecular mechanisms linking imprecision to suboptimal fitness deviations have been proposed. The dosage-control hypothesis proposes that noise disrupts control over the stoichiometric balance of interacting proteins (Fraser et al. 2004; Lehner 2008). Consistent with this hypothesis, proteins that form complexes have lower average noise than other proteins (Fraser et al. 2004; Lehner 2008; Bar-Even et al. 2006), but models currently do not predict the magnitude of this effect so that quantitative correspondence between predictions and observations remain unknown. More commonly, researchers focus on how noise disrupts the fine control over cell states regulated by bistable (ON/OFF) switches (Balázsi, van Oudenaarden, and Collins 2011; Csete and Doyle 2002). Noise in the input parameter can cause the cell to flip stochastically from the desired to undesired state, or to respond slowly to environmental cues, creating a maladaptive mismatch between the cell state and the selecting environment. However, the least noisy genes, those presumably experiencing the strongest selection to suppress noise, are constitutively expressed and involved in core cell functions, not control over bistability (Fraser et al. 2004; Lehner 2008; Bar-Even et al. 2006). This cost is also difficult to quantify in any general way because its effects are highly context dependent, to the extent that in some contexts the cost of imprecision can change signs, becoming evolutionarily beneficial (Acar, Mettetal, and van Oudenaarden 2008; Rao, Wolf, and Arkin 2002; Kussell and Leibler 2005; Eldar and Elowitz 2010; Kaern et al. 2005; Thattai and Van Oudenaarden 2004).

As an alternative, here I demonstrate a general, readily quantifiable and consistent cost of cellular noise. I show that noise in substrate abundance slows the average rate of

product formation in Michaelis-Menten-type (i.e., non-cooperative bimolecular) chemical reactions, which constitute most reactions in the cell including those most closely tied to fitness: nutrient uptake and protein synthesis. This result is a straightforward consequence of the non-linearity of bimolecular reactions. While it is well known that noise alters the mean of nonlinear stochastic systems, this effect is typically ignored, giving rise to the widespread use of linear approximations (e.g., the “linear noise approximation” (LNA) (van Kampen 2007)), particularly in studies of gene expression noise (Thattai and van Oudenaarden 2001; Shahrezaei and Swain 2008; Paulsson 2004). The LNA is well-justified when applied to investigating steady-state concentrations in noisy chemical networks, since the difference between linear and non-linear predictions at steady-state are on the order of a single molecule. However, fitness in living cells is determined by rates—not concentrations--integrated over an enormous number of reactions over the lifetime of a cell and its genetic lineage. For example, an *E. coli* cell must carry out a minimum of 2.5×10^6 translation reactions to produce a single daughter cell. A minor slowing of reaction rates within cells, if heritable, may pose a non-negligible selective cost. From this perspective, the effects of noise on fitness cannot be ignored, and indeed, I find that they can be substantial in small cells.

RESULTS

The goal of the paper is to quantify the strength of selection favoring the attenuation of gene expression noise, and to identify the targets of noise attenuation. The paper proceeds as follows. First, I show that noise slows the average rate of biochemical reactions that have non-linear, hyperbolic kinetics. Then, I extend this result to networks of coupled

non-linear reactions in order to determine how noise effects total flux through biochemical networks, as well as how the network architecture itself affects the propagation of noise. This enables us to quantify the cost of noise, measured as the loss in average rate of end-product formation, and to quantify the noise statistics of the system. I then apply the model to gene expression, focusing on the consequences of non-linear, hyperbolic translation kinetics on the propagation of mRNA noise to the protein level, and to the average rate of biomass synthesis. Finally, I investigate how the volume of reaction compartments and whole-cells influences the cost of noise, concluding that the cost of noise, and therefore the selection pressures favoring low gene expression noise, decays rapidly with increasing cell size, leading to a testable prediction about levels of gene expression noise across unicellular species.

Noise-induced slowdown of Michaelis-Menten reactions. The purpose of this section is to begin building the theory from the simplest component of a network: a single isolated reaction. We will see that input noise slows the reaction rate of a single, isolated reaction obeying Michaelis-Menten type kinetics. In the following section, the analytical results from this section are put on a more rigorous theoretical footing and then extended to networks of coupled reactions in open systems.

The cornerstone of biochemistry is the Michaelis-Menten (MM) reaction, which describes the enzyme catalyzed conversion of substrate, S , into product, P . The rate of product formation is described by the hyperbolic MM equation (Michaelis and Menten 1913; Briggs and Haldane 1925; Ingalls 2013):

$$\frac{dp}{dt} = V = V_{max} \frac{s}{K_m + s} \quad 1$$

s , e and p are reactant, enzyme and product concentration, respectively (following convention lowercase denotes concentration, which unless otherwise noted will take units of molecules per cell), $V_{max} = k_{cat}e$ is the asymptotic reaction rate (k_{cat} is the number of substrate molecules converted to product per enzyme per unit time), and K_m , the Michaelis or half-saturation constant, is the substrate concentration at $V = V_{max}/2$. This equation is derived from a system of elementary reactions by applying the law of mass action and the quasi-steady state approximation (QSSA) (Briggs and Haldane 1925; Ingalls 2013) (Sup. Mat.). The QSSA is a separation of timescales approximation that eliminates fast variables by setting them to their steady-state. The law of mass action is valid in cases where the timescale of molecular diffusion is fast relative to the reaction rate, so that the system behaves as if it were well-mixed. The rate of enzymatic reactions in cells is far below the diffusion limit (Bar-Even et al 2006), so that mass action is a reasonable approximation in most cases. For stochastic reactions, wherein s , e , and p experience random number fluctuations, Eqn. 1 takes on a probabilistic interpretation as the “reaction propensity” (van Kampen 2007) giving the transition probability for $n_p \rightarrow n_p + 1$ and $n_s \rightarrow n_s - 1$, where n denotes particle number. In general, the QSSA is valid for stochastic systems under the same conditions for which it is valid in deterministic systems (Rao and Arkin 2003; Sanft, Gillespie, and Petzold 2011), notwithstanding finite-volume corrections (Grima 2009b; Grima 2009a; Grima 2010) (see below).

Because Eqn. 1 is hyperbolic in s , Jensen’s inequality proves that, for random s , $\langle V(s) \rangle \leq V(\langle s \rangle)$ (angle brackets denote ensemble average). That is, the average reaction rate is less than or equal to a noise-free (i.e., deterministic) reaction with $s = \langle s \rangle$ (Figs. 1, S1). This is seen by expanding Eqn. 1 to second order in Taylor series about $\langle s \rangle$ and

189 taking the expectation (see SI),

$$190 \quad \langle V(s) \rangle = \underbrace{V_{max} \frac{\langle s \rangle}{K_m + \langle s \rangle}}_{V(\langle s \rangle)} - \underbrace{\frac{\langle \epsilon_s^2 \rangle}{\langle s \rangle^2} V_{max} a(1-a)^2}_{\text{noise-induced slowdown}} + O(\epsilon_s^3) \quad 2$$

191 $a = K_m/(K_m + \langle s \rangle)$ is the “kinetic order” of the MM reaction evaluated at the average
 192 substrate concentration, and $(\eta^2 = \langle \epsilon_s^2 \rangle / \langle s \rangle^2)$ the squared coefficient of variation, which
 193 is often simply called “the noise”. The leading order correction in Eqn. 2 takes the sign of
 194 the second derivative of Eqn. 1, which is everywhere negative for hyperbolic transition
 195 rates, thus demonstrating the noise-induced slowdown of MM type reactions.

196 The kinetic order parameter, a , is a key parameter that appears repeatedly below.
 197 a is identical to the steady-state fraction of unbound enzyme. In metabolic control
 198 analysis (MCA), a is equivalent to the “elasticity coefficient” (Kacser and Burns 1973).
 199 In general, a measures the nonlinearity of the reaction rate at a given mean substrate
 200 concentration. First and zeroth order reactions ($a = 1$ and 0 , respectively) are both linear,
 201 and as a consequence there is no noise-induced slowdown, consistent with Jensen’s
 202 Inequality. The rate of a first order reaction is linearly dependent on the substrate
 203 concentration, whereas the rate of a zeroth order reaction is independent of substrate
 204 concentration. In a MM type reaction, the reaction is in its first order regime when
 205 substrate is scarce, and in the zeroth order regime when substrate is abundant (the flat
 206 part of the rate law curve in Fig. 1).

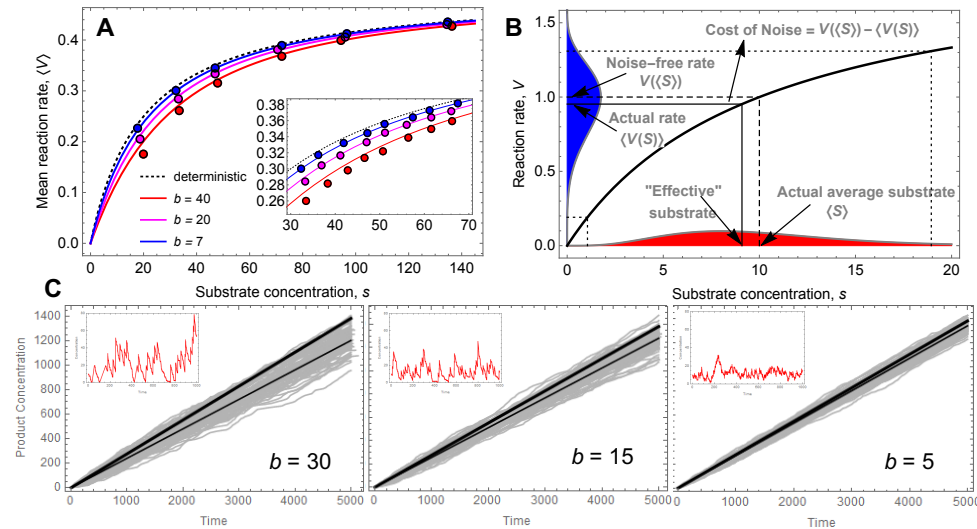


Figure 1: Noise-induced slowdown of hyperbolic reactions. Stochastic simulations employing Gillespie’s (Gillespie 1977) exact stochastic simulation algorithm (SSA) of the microscopic system of elementary reactions for the Michaelis-Menten reaction (see SI). Points represent the rate of product formation over time steps 100-800 (providing a burn-in more than 10 times the half life of P or S) ensemble averaged over 1000 realizations of the stochastic process. Results fit well with the QSSA model (Eqn. 3, solid lines in **a**, **b** and thin black line in **c**). The substrate burst input, b , is varied to alter the substrate noise level. Substrate input rate, k_{in} is divided by b so that mean substrate is not affected by bursting, and is varied to generate different substrate concentrations. All other parameters are constant ($e = 1$, $k_{ON} = 1$, $k_{OFF} = 19.5$, $k_{cat} = 0.5$, $\delta_S = 0.1$, $\delta_P = 0.1$). **A**) The rate law curve demonstrating that, for a given average substrate concentration, substrate noise slows the reaction, but only if the reaction is not in its first or zeroth order regime (low or high substrate, respectively). Note that the accuracy of the stochastic QSSA requires that $s, K_m \ll F$, which fails for the red points in **a** and **b** (where $K_m = 20$ and $F = 20.5$). Inset is a zoomed in view. **B**) A geometric illustration of noise-induced slowdown, showing that variance (noise) in a state variable reduces the average of a hyperbolic function (i.e., Jensen’s inequality). **C**) 100 realizations of the stochastic process (gray) compared to the deterministic (thick black line) and stochastic (thin black line) predictions. Red traces in inset are example outputs from a single realization of the substrate dynamics for a short time window.

A more rigorous and general expansion method is used in the following section (“Noisy Non-linear Reaction Networks in Finite Volumes”). For now, there are two main points: 1) reaction kinetics depend not only on average molecule concentrations, but on the mean and variance in molecule numbers, and 2) for a given average substrate concentration, substrate noise reduces the rate of biochemical reactions. The difference

between the average stochastic rate and deterministic rate grows monotonically with increasing noise, and equality occurs only with zero noise or zero curvature of $V(s)$. Mechanistically, this can be understood as follows: positive fluctuations in substrate oversaturate available catalysts leaving many unbound S molecules that are not actively creating product, causing inefficiency. As a consequence, the same amount of substrate would produce a faster net reaction rate if spread out evenly over the same time interval (i.e., the same mean with zero noise).

If the substrate statistics are known, then the probability density of V with random s can be solved explicitly for a number of common probability distributions. mRNA and protein counts are typically gamma distributed within a cell over time or among cells in a clonal population (Taniguchi et al. 2010). The gamma distribution has shape parameter $\alpha = CV^{-2}$ and rate parameter $\beta = F$, where $CV^2 = \sigma_x^2 / \langle x \rangle^2$ and $F = \sigma_x^2 / \langle x \rangle$ are the squared coefficient of variation (the noise) and the Fano factor (often called “the noise strength”), respectively (Taniguchi et al. 2010). F and CV^2 are common measures of noise intensity, with convenient mechanistic interpretations in gene expression: F is closely related to the burst size (e.g., number of proteins synthesized per mRNA lifetime) and CV^2 to the burst frequency (e.g., the number of mRNA’s synthesized per protein lifetime). F is independent of cell volume, whereas CV^2 decreases with cell volume (they are intensive and extensive variables, respectively). For gamma distributed s , the probability density function of the Michaelis-Menten reaction rate (for $V < V_{max}$) is exactly,

$$Pr(V = v) = \frac{V_{max}}{k_m} \frac{v^{\alpha-1} e^{-v/z\beta}}{z(z\beta)^\alpha \Gamma(\alpha)}, \quad z = \frac{V_{max} - v}{k_m} \quad 3$$

Eqn. 3 is verified by Monte Carlo simulations implemented with Gillespie’s exact stochastic simulation algorithm (SSA) (Gillespie 1977) (Fig. 1). Eqn. 3 fits data from

stochastic simulations very well provided that $F, K_m \ll s$. Equation 2 gives a very good approximation to 3 under this same restriction (Fig. S2), which provides an important benchmark for the accuracy of the more rigorous expansion approximations used below.

Figure 1 shows that the effect of noise on MM-type reactions is to increase the observed Michaelis constant. Therefore, a stochastic formulation of Eqn. 1 that accounts for noisy substrate while retaining the classical form is obtained simply by substituting into Eqn. 1 an “effective”, stochastic Michaelis constant, \tilde{K}_m , where $\tilde{K}_m \geq K_m$ with equality occurring only with zero noise, zero curvature of the rate law, and macroscopic reaction volume (S6).

Noisy Non-linear Reaction Networks in Finite Volumes. In the previous section, we showed that substrate noise slows the rate of product formation in a single hyperbolic reaction in a closed system. To apply this result, the statistics of the substrate species are required. We now take a more rigorous approach and generalize the noise-induced slowdown phenomenon to reaction networks in open systems. Further, the model allows one to predict the noise statistics of the system solely with knowledge of the network structure and reaction kinetics. Of particular interest is how network architectures may create feedbacks that buffer the system flux against noise-induced slowdown. The following theory builds on the work of van Kampen (van Kampen 2007) and Elf and Erhenberg (Elf and Ehrenberg 2003) (see also Grima (Grima 2010)).

The time evolution of the joint probability density of all N nodes in a chemical reaction network with R edges is given by the following “system-size expansion” (van Kampen 2007) of the multidimensional chemical master equation (derived in the SI):

$$\begin{aligned}
 & \text{linear Fokker-Planck equation} \\
 280 \quad \frac{\partial \Pi(\vec{\epsilon}, t)}{\partial t} = & - \sum_{i,k}^N J_{ik} \frac{\partial(\epsilon_k \Pi)}{\partial \epsilon_i} + \frac{1}{2} \sum_{i,k=1}^N D_{ik} \frac{\partial^2 \Pi}{\partial \epsilon_i \partial \epsilon_k} \\
 281 \quad & + \underbrace{\frac{\Omega^{-1/2}}{2} \left(- \sum_{i,k,l=1}^N \frac{\partial J_{ik}}{\partial \bar{x}_l} \frac{\partial(\epsilon_l \epsilon_k \Pi)}{\partial \epsilon_i} + \sum_{i,k,l=1}^N \frac{\partial D_{ik}}{\partial \bar{x}_l} \frac{\partial^2(\epsilon_l \Pi)}{\partial \epsilon_i \partial \epsilon_k} \right)}_{\text{correction for non-linearity and finite volume}} + O(\Omega^{-1}) \quad 4
 \end{aligned}$$

282 $\Pi = \Pi(\vec{\epsilon}, t)$ is the joint probability density of all reacting species, Ω is the reaction
 283 volume, ϵ_i is the random perturbation of the i^{th} species from \bar{x}_i , its macroscopic steady-
 284 state (overbars will denote the macroscopic steady-state value of a variable throughout),
 285 S_{ij} is the stoichiometric coefficient of species i in the j^{th} reaction, and
 286 $J_{ik} = \sum_{j=1}^R S_{ij} \frac{\partial f_j(\vec{x})}{\partial \bar{x}_k}$, $D_{ik} = \sum_{j=1}^R S_{ij} S_{kj} f_j(\vec{x})$
 287 are entries of the Jacobian and diffusion matrices, respectively, of the continuous
 288 deterministic system of equations. Equation 4 is a linearization of the chemical master
 289 equation obtained by applying the ansatz that concentrations fluctuate about their
 290 macroscopic steady-state value, with fluctuation size scaling inversely with the square
 291 root of the reaction volume, Ω (van Kampen 2007). This inverse square-root dependence
 292 of noise on volume is a formal statement of the Law of Large Numbers, and is formally
 293 equivalent to the relationship between sample size and the standard deviation of a
 294 sampling distribution. One thus substitutes, $x_i = \bar{x}_i + \Omega^{-1/2} \epsilon_i$, for each species i in the
 295 deterministic system of equations, and then expands each transition rate, $V_j(x)$, in powers
 296 of $\Omega^{-1/2}$ about \bar{x}_i , truncating the series at the desired order (van Kampen 2007). The terms
 297 of order $O(\Omega^0)$ in (4) give the “linear noise approximation” (LNA) (van Kampen 2007),
 298 which is equivalent to the standard linear Fokker-Planck equation, while the terms in $\Omega^{-1/2}$

299 ^{1/2} account for noise-induced deviations caused by non-linear rate laws in finite volumes,
300 which we show cannot be ignored in living cells.

301 To apply 4, one begins by writing a macroscopic description of a reaction network
302 in QSSA reduced form. Here we analyze a generalized MM scheme (S17) in an open
303 system obeying the following macroscopic system of ODEs (Sup. Mat.):

$$304 \quad \frac{ds}{dt} = b_S k_{in,S} - hV - \delta_S s, \quad \frac{de}{dt} = b_E k_{in,E} - \delta_E e, \quad \frac{dp}{dt} = V - \delta_P p \quad 5$$

305 δ_i and $k_{in,i}$ are the decay and input rates, respectively, of i , and $V = k_{cat} e s / (K_m + s)$.

306 Substrate (S) and enzyme (E) are fed into the system via a burst process with burst size b_S
307 and b_E , respectively ($n_S \rightarrow n_S + b_S$; $n_E \rightarrow n_E + b_E$). $h = \{0,1\}$ is an indicator variable
308 that toggles between reactions where S is consumed ($h = 1$), as in most metabolic
309 reactions, and reactions where S is not consumed ($h = 0$), as with template-mediated
310 reactions such as translation where E denotes ribosomes and S mRNA, which is not
311 consumed in translation.

312 From System 5 and Eqn 4, we obtain the following steady-state average
313 concentrations of each species in a generalized, open MM reaction network (Sup. Mat.):

$$314 \quad \langle s \rangle = \bar{s} + h \underbrace{\left[a \frac{\langle \epsilon_S^2 \rangle}{\bar{s}} - K_m \frac{\langle \epsilon_S \epsilon_E \rangle}{\bar{e} \bar{s}} \right]}_{\text{stochastic amplification}} A \Omega^{-1}, \quad \langle e \rangle = \bar{e},$$

$$315 \quad \langle p \rangle = \bar{p} - \frac{\delta_S}{\delta_P} \left[a \frac{\langle \epsilon_S^2 \rangle}{\bar{s}} - K_m \frac{\langle \epsilon_S \epsilon_E \rangle}{\bar{e} \bar{s}} \right] A \Omega^{-1}, \quad 6$$

$$316 \quad \langle V \rangle = \bar{V} - \delta_S \underbrace{\left[a \frac{\langle \epsilon_S^2 \rangle}{\bar{s}} - K_m \frac{\langle \epsilon_S \epsilon_E \rangle}{\bar{e} \bar{s}} \right]}_{\text{noise-induced slowdown}} A \Omega^{-1}$$

317 Where $A = (k_{cat} \bar{e} \bar{s} / (h k_{cat} \bar{e} K_m + \delta_S (K_m + \bar{s})^2))$ and $a = K_m / (K_m + \bar{s})$ is the MM
318 reaction order. Note that in the general case, $\langle s \rangle$ is increased and $\langle p \rangle$ is decreased by
319 noise compared to the macroscopic system. The non-linear mesoscopic correction terms

vanish when $s \rightarrow 0$ (first order kinetics), $s \rightarrow \text{Infinity}$ (zeroth order kinetics) or when $\Omega \rightarrow \text{Infinity}$ (macroscopic volume). Two of the second moments (Sup. Mat.) for the general MM case can be written compactly:

$$\langle \epsilon_E^2 \rangle = \frac{1}{2} \bar{e}(b_E + 1), \quad \langle \epsilon_S \epsilon_E \rangle = -h \langle \epsilon_E^2 \rangle \frac{k_{cat} \bar{s}}{h a k_{cat} \bar{e} + (\delta_E + \delta_P)(K_m + \bar{s})} \quad 7$$

While the present approach builds on that derived by Grima (Grima 2010) the results and conclusions differ: Grima (Grima 2010) found that noise had no effect on product concentration in MM reactions. There are two reasons for this discrepancy. First, Grima assumed no dilution of the substrate ($\delta_S = 0$). As can be seen from equation 6, this causes the finite volume correction term for $\langle p \rangle$ to vanish. Thus, Grima's results are in agreement with the present results for this special case, and all of the results in his papers are correct. However, following (Grima 2010) exactly for $\delta_S > 0$ also gives zero stochastic deviation of the product concentration (see SI), which is incorrect. Here is why: when the system-size expansion of the chemical master equation is applied to a system of elementary reactions, as in (Grima 2010) and most other methods, the rate of product formation, V , is proportional to the intermediate complex concentration, c . Because c is unimolecular, the method reads this rate as linear (its second derivative is zero), and so the mean cannot be affected by noise (as proven by Jensen's inequality). However, in MM reactions V is decidedly non-linear. Indeed, it is hyperbolic. Thus, the elementary reaction conceals an essential non-linearity of the system, and will give incorrect results in the general case. To rectify this, one must apply Eqn. 4 to a QSSA reduced system of biochemical equations as shown in the SI. Note that the use of elementary or QSSA reactions is not up to choice: the elementary reaction scheme will

give the wrong answer. Note also that applying the method of (Grima 2010) to a QSSA reduced system also gives incorrect results, for reasons discussed in the SI.

Noise propagation in gene expression with hyperbolic filtering. Because gene expression noise is costly, there will ostensibly be selection pressure for noise control. But how can a cell suppress gene expression noise? That is, what are the evolutionary targets of noise-control? To answer these questions, we must model the process of gene expression, from transcription through translation, to identify how noise is generated and filtered.

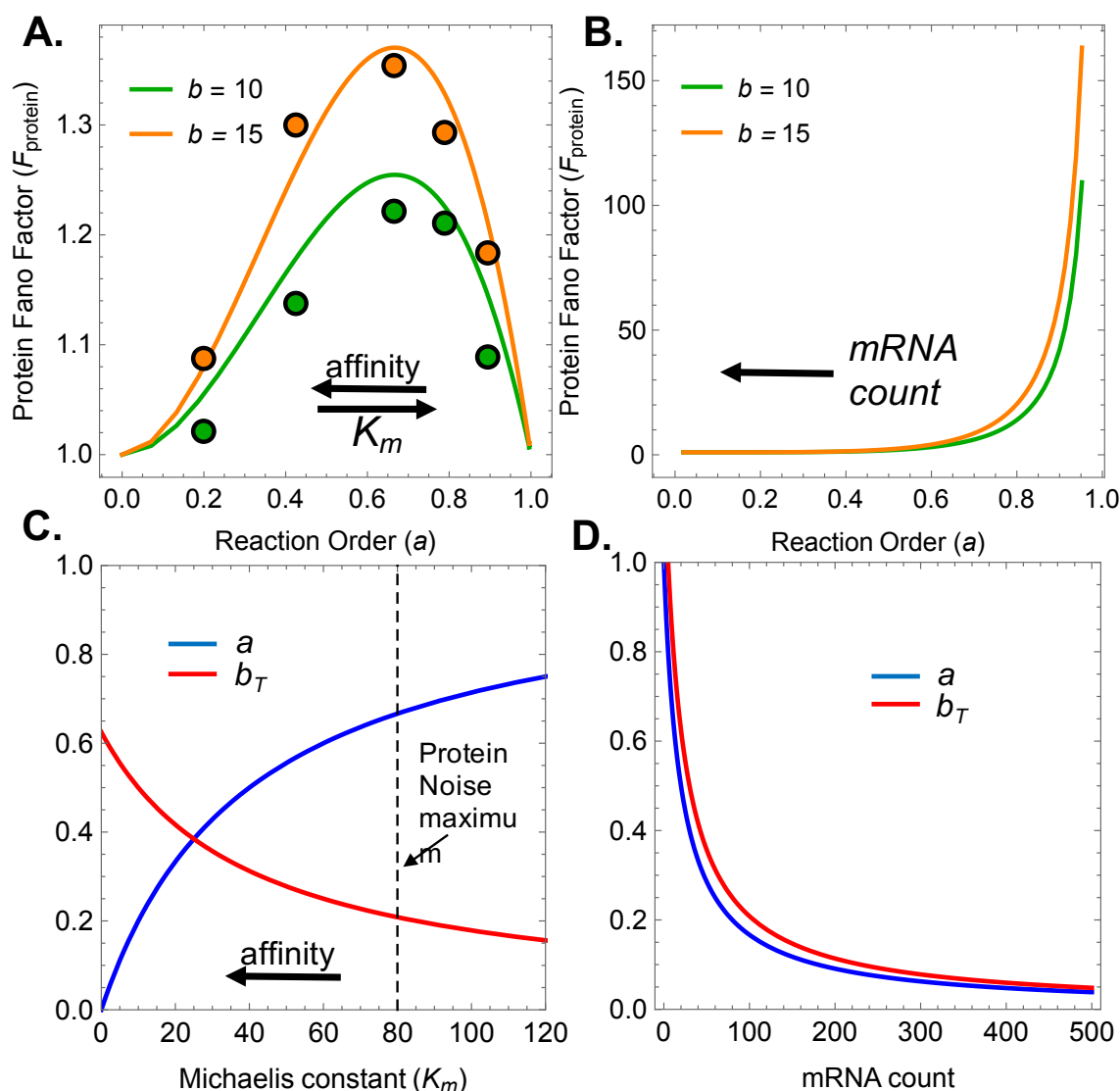
We deviate from previous models of gene expression noise by assuming that translation obeys non-linear reaction kinetics. Previous models of gene expression noise have considered translation as a first-order process (Thattai and van Oudenaarden 2001; Shahrezaei and Swain 2008; Paulsson 2004; Bar-Even et al. 2006; Pedraza and Paulsson 2008; McAdams and Arkin 1997). However, translation in living cells may actually be closer to zeroth-order than first-order, with a reaction order of $\alpha = 0.1$ - 0.2 in log-phase growth in *E. coli* and budding yeast, based on the fraction of unoccupied ribosomes (Arava et al. 2003; Zenklusen, Larson, and Singer 2008; Ingolia 2014). We now investigate the effects of hyperbolic translation kinetics on gene expression noise.

Following previous models of gene expression noise (Thattai and van Oudenaarden 2001; Shahrezaei and Swain 2008; Paulsson 2004; Bar-Even et al. 2006; Pedraza and Paulsson 2008; McAdams and Arkin 1997), we assume for now that translation is initiation limited (see “Cell fitness” section below for generalization), such that it relies only on the concentrations of ribosomes, r , and mRNA, m . If mRNA is

transcribed in bursts of size b_t , then the steady-state variance and Fano factor of mRNA
are (Sup. Mat.),

$$\sigma_{mRNA}^2 = \frac{1}{2} \bar{m}(b_t + 1), \quad F_{mRNA} = \frac{1}{2}(b_t + 1) \quad 8$$

which is a special case of (Elf and Ehrenberg 2003) and which is in excellent agreement
with simulations (Figs 2, 3).



370

371 **Figure 2: Protein noise with hyperbolic filtering.** Translation propagates mRNA noise
372 to the protein level, but hyperbolic reaction kinetics ($0 < a < 1$) can substantially ameliorate
373 this noise and cause highly non-linear relationships between noise and underlying

parameters. Points represent ensemble variance of protein number from the last of 110 time steps over 10,000 realizations of the SSA (see SI), lines represent theoretical prediction (from Eqns 9 and S21). **A, C**) Increasing ribosomal binding affinity (i.e., reducing K_m) has two opposing effects on protein noise: it reduces reaction order but increases translational bursting. Which effect dominates depends on the population's starting point relative to the protein noise peak at $a^* = 2/3$. Most species have translational $a < 2/3$, suggesting that increasing ribosomal binding affinity will actually reduce protein noise, even though it increases bursting. **B, D**) Increasing mRNA abundance, holding mRNA noise constant, monotonically decreases protein noise for all parameter values, as both the reaction order and translational burst parameters decrease with increasing mRNA (Sup. Mat. for theoretical derivation of this result). Parameters: $e = 5$, $k_{ON} = 1$, $k_{OFF} = 19.5$ (varied in **A**), $k_{cat} = 0.5$, $\delta_S = 0.1$, $\delta_P = 0.1$, $k_{in} = 4/b$ (varied in **B**). K_m was tuned by varying k_{OFF} , s was tuned by varying k_{in} , while mRNA noise was held constant with a burst input of either $b = 10$ or 15 .

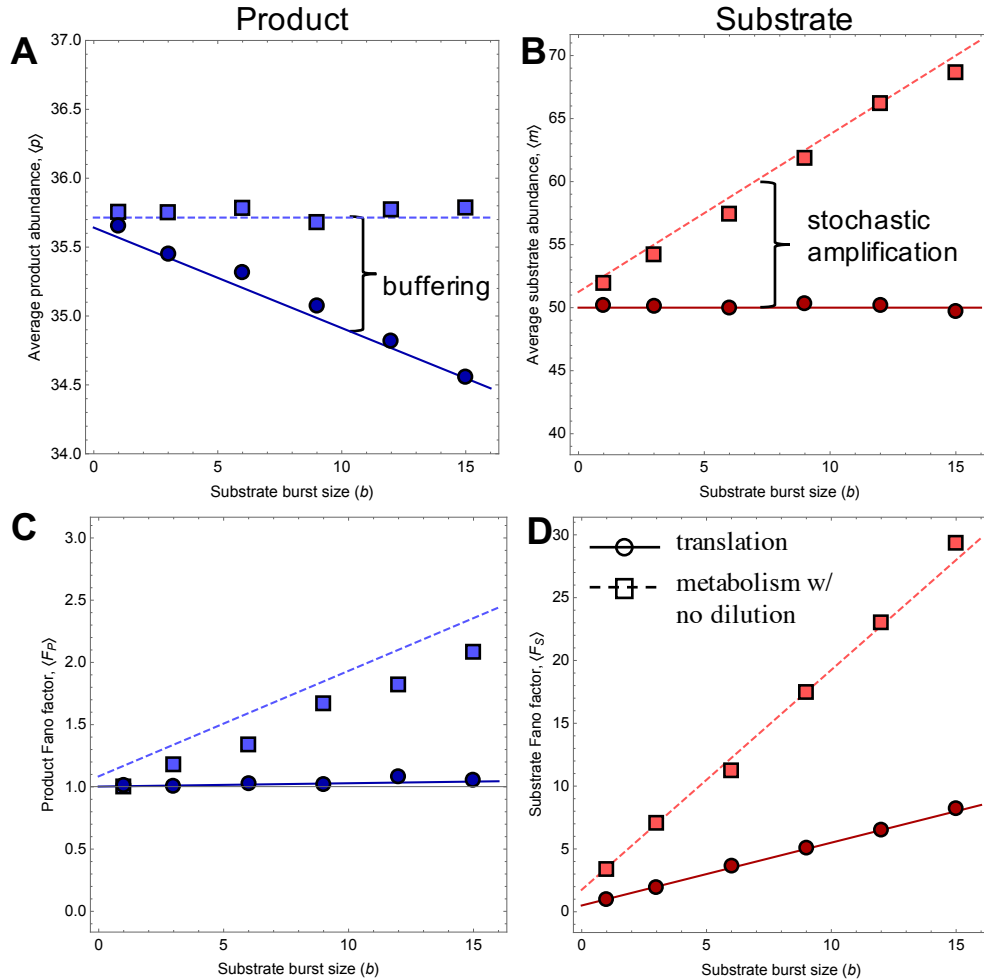


Figure 3: Effect of network architecture on the cost of noise. Theoretical predictions (Eqns. 6-8, S20, S21) for metabolic reactions (dashed lines; $h = 1$, $\delta_S = 0$) and translation (solid lines; $h = 0$, $\delta_S = 0.1$) compared to ensemble average (blue) and variance (red) of

1,000-5,000 SSA realizations (points). **A)** Noise-induced slowdown causes an accumulation of substrate (“stochastic amplification”) which completely buffers reaction flux for metabolic reactions when substrate dilution/degradation is zero ($\delta_S = 0$) **(B)**. Because mRNA is not consumed in translation ($h = 0$), its concentration is independent of reaction rate so does not experience stochastic amplification. Thus, translation is not buffered by network feedback. **C,D)** The effect of network architecture on substrate (C) and product (D) noise. $e = 1$, $k_{ON} = 1$, $k_{OFF} = 19.5$, $k_{cat} = 0.5$, $\delta_P = 0.01$, and $k_{in} = 5/b$ (solid lines) or $0.357/b$ (dashed lines).

The variance in protein numbers at steady-state is, $\langle \epsilon_{protein}^2 \rangle \cong \bar{p}(1 + a^2 b_T F_{mRNA} + CV_R^2 \bar{p}/2)$, where the parameter $a^2 = [K_m/(K_m + m)]^2$ is the squared reaction order and $CV_{ribo}^2 = \sigma_{ribo}^2/\bar{r}^2$ is the ribosomal noise (Sup. Mat.). The translational burst parameter, b_T , is the number of proteins synthesized by a single mRNA molecule over its lifetime, which with hyperbolic translation kinetics is, $b_T = k_T r / ((K_m + m)\delta_S)$. Previous work assumed first-order translation, such that $b_T = k_T r / (K_m \delta_S)$ (Thattai and van Oudenaarden 2001; Shahrezaei and Swain 2008). Introducing non-linearity causes the translational burst size to become a non-linear, decreasing function of mRNA abundance: all else equal, an mRNA molecule competing with fewer other mRNA’s will be translated more often in its lifetime, increasing its burst size. Alternatively, in the zeroth order regime ($m \gg K_m$), translational bursting goes to zero because ribosomes are so oversaturated with mRNA’s that any given mRNA is more likely to be degraded than translated. The protein noise strength and noise, then, are, respectively,

$$F_{protein} \cong 1 + a^2 b_T F_{mRNA} + CV_{ribo}^2 \bar{p}/2$$

416

$$CV_{protein}^2 \cong \underbrace{\frac{1}{\bar{p}} + \frac{a^2 b_T F_{mRNA}}{\bar{p}}}_{\text{intrinsic noise}} + \underbrace{\frac{CV_{ribo}^2}{2}}_{\text{extrinsic noise}}$$

9

417

Eqn. 9 agrees well with simulation results (Figs. 2, 3) and retrieves the result of

418

Thattai and van Oudenaarden (Thattai and van Oudenaarden 2001) ($F_{protein} \cong 1 + b_T$)

419

as a special case when there is no transcriptional bursting ($b_t = 1 \Rightarrow F_{mRNA} = 1$), no

420

ribosome fluctuations ($CV_{ribo}^2 = 0$), and translation is first order (i.e., linear) in mRNA

421

abundance ($a = 1$). The first terms in $F_{protein}$ and $CV_{protein}^2$ (1 and $1/\bar{p}$, respectively) are

422

simply Poisson terms reflecting low copy number fluctuations (Paulsson 2004; Pedraza

423

and Paulsson 2008), while the second and third terms introduce factors that increase noise

424

above Poisson levels. The second terms ($a^2 b_T F_{mRNA}$ and $a^2 b_T F_{mRNA}/\bar{p}$, respectively)

425

incorporate the consequences of translational and transcriptional bursting. These terms

426

propagate mRNA fluctuations, F_{mRNA} , to the protein level. Importantly, hyperbolic

427

translation kinetics ($0 < a < 1$) dampen the effect of mRNA fluctuations on protein noise,

428

thus attenuating the propagation of mRNA noise via a process that might be called

429

“hyperbolic filtering”. Together the first and second terms are considered “intrinsic

430

noise” (Swain, Elowitz, and Siggia 2002; Raser and O’Shea 2005; Paulsson 2004; Bar-

431

Even et al. 2006; Sanchez and Golding 2013). The last term ($CV_{ribo}^2 \bar{p}/2$ and $CV_{ribo}^2/2$)

432

accounts for fluctuations in ribosome numbers, which create positively correlated

433

fluctuations among protein species called “extrinsic noise” (Swain, Elowitz, and Siggia

434

2002; Raser and O’Shea 2005; Paulsson 2004; Bar-Even et al. 2006; Sanchez and

435

Golding 2013). This extrinsic noise term dominates for abundant proteins (large p),

436

consistent with the observation that protein noise scales inversely with protein abundance

for low copy number proteins, but is independent of abundance for high copy number proteins (Blake et al. 2003; Raser and O'Shea 2004; Bar-Even et al. 2006).

Translation initiation in *E. coli* and yeast has a reaction order of $a \sim 0.1-0.2$ (3, 49), which is comfortably to the left of the protein noise peak at $a^* = 2/3$ (see Sup. Mat.; Fig 2a). This implies that, contrary to predictions based on first-order translation initiation models (Thattai and van Oudenaarden 2001; Raser and O'Shea 2005), evolving or engineering increased binding affinity of mRNA to ribosomes (i.e., reduced translational K_m) will actually lower protein noise, not increase it. This occurs because reducing K_m has two complementary effects on protein noise: it increases translational bursting, b_T , but reduces reaction order, a , which attenuates noise via a phenomenon we term “zeroth-order insensitivity”. The latter effect is not accounted for in previous gene expression models, but has implications for the evolutionary and synthetic targets of noise-amelioration (see Discussion). Hyperbolic filtering also has practical implications when making inferences from data. For example, translational burst size, b_T , can be estimated from Eqn 9 with knowledge of the mRNA and protein statistics; however, assuming first order reaction kinetics ($a = 1$) can lead to large underestimates of this parameter from data. For example, for $a \sim 1/5$, assuming first order translation kinetics can lead to a 25-fold underestimate of translational burst size.

Correlation between mRNA and protein noise with hyperbolic translational

filtering: We now quantify the effect of hyperbolic translation kinetics on the correlation between mRNA and protein noise. Empirical work in yeast, *E. coli* and human cell lines have found that there is no correlation between mRNA and protein levels in the cell, or

between mRNA and protein noise (Chen et al. 2002; Ghaemmamghami et al. 2003; Taniguchi et al. 2010; Gygi et al. 1999). Here I show that this lack of correlation can be explained, in part, by modeling translation as a hyperbolic reaction.

Assuming as before that mRNA lifetimes are much shorter than protein lifetimes ($\delta_{mRNA} \gg \delta_{protein}$), the covariance between mRNA and protein noise, σ_{mp} , is found to be, $\sigma_{mp} \cong ab_T \sigma_{mRNA}^2$. The linear regression coefficient, $\beta_{mp} \equiv \sigma_{mp} / \sigma_{mRNA}^2$ is then, $\beta_{mp} \cong ab_T$. Both of these show that hyperbolic translation kinetics ($a < 1$) diminish the statistical association between mRNA and proteins. This is most usefully illustrated by the Pearson correlation coefficient, $\rho \equiv \sigma_{mp} / \sigma_{mRNA} \sigma_{protein}$, which is approximately equal to,

$$\rho \cong a \sqrt{\left(\frac{\delta_{protein}}{\delta_{mRNA}}\right) \left(\frac{b_T F_{mRNA}}{1 + a^2 b_T F_{mRNA} + CV_{ribo}^2 \bar{p} / 2}\right)} \quad 10$$

Consistent with the theoretical predictions of Tanaguchi et al (3), Eqn. 10 shows that the mRNA-protein correlation is increased with increasing transcriptional noise strength, F_{mRNA} , is diminished with larger differences in lifetimes of mRNA and protein ($\delta_{protein} / \delta_{mRNA} < 1$), and is diminished by extrinsic noise ($CV_{ribo}^2 \bar{p} / 2$). In addition, the covariance, linear regression coefficient, and correlation between mRNA and protein fluctuations all increase with increasing translational bursting, b_T . The most important novel result here is that the correlation between mRNA noise and protein noise is diminished by the hyperbolic filtering of translation (that is, when $a < 1$).

Network architecture, feedback, and buffering: While noise slows a reaction for a given substrate concentration, network feedbacks cause adjustment to substrate levels

that may in turn alter reaction rates. Specifically, if a reaction proceeds more slowly, then substrate levels accumulate, which then cause the reaction to speed up again. Accounting for this feedback is crucial in quantifying the costs of noise. We now ask: How exactly does this feedback influence the total flux through noisy reaction networks? To address this problem, we now analyze how network architecture creates feedbacks that alter chemical flux, potentially compensating for noise-induced slowdown. We treat three limiting cases: 1) $h = 1$, $\delta_S = 0$; 2) $h = 1$, $\delta_S > 0$; and 3) $h = 0$.

Cases 1 ($h = 1$, $\delta_S = 0$): Stochastic substrate amplification and buffering: One important feedback in noisy networks is the amplification of substrate concentration by noise (Eqn. 6), which has been studied in a process called “stochastic focusing” (Paulsson, Berg, and Ehrenberg 2000). Eqn. 6 shows that stochastic amplification will occur whenever substrate is consumed by the reaction ($h = 1$), as in metabolic reactions. Consequently, this steady-state adjustment of substrate creates intrinsic buffering of reaction flux that ameliorates noise-induced slowdown: the reaction is slowed by noise, allowing substrate to accumulate, which then speeds up the reaction. The extent of buffering depends on the substrate dilution parameter, δ_S . In the MM network studied by Grima (Grima 2010; Grima 2009b), for example, the substrate entered the reaction volume at rate k_{in} and exited only once consumed by the focal reaction at rate V . Flux balance requires that $V = k_{in}$ in such a system, so that the reaction rate in steady-state is unaffected by noise, as found by Grima (Grima 2010; Grima 2009b). Thus, under the flux constraint $V = k_{in}$, stochastic amplification completely buffers the system against noise-induced slowdown (Fig 3).

504 *Case 2 ($h = 1, \delta_S > 0$): Incomplete amplification and buffering.* However, the
505 precise balancing of V with k_{in} does not apply in living cells, where substrate exits the
506 system via multiple alternative channels ($\delta_S > 0$), primarily dilution by cell division, but
507 also through excretion, titration by other reactions and enzymatic degradation.
508 Extracellular substrate also experiences non-zero δ_S , as it is diluted by bulk flow (such as
509 in a chemostat) or by uptake from neighboring cells. Accounting for these effects
510 modifies the flux constraint to: $k_{in} = V + \delta_S s$. With this network architecture, stochastic
511 amplification compensates for some of the diminution of V by noise, but not all.
512 Consequently, stochastic amplification fails to completely buffer metabolic reactions in
513 living cells against noise-induced slowdown.

514 *Case 3 ($h = 0, \delta_S > 0$): Gene expression is especially sensitive to noise-induced*
515 *slow-down; no stochastic amplification, no buffering.* The $V = k_{in}$ flux constraint is
516 entirely abolished in template-mediated reactions ($h = 0$), such as protein synthesis via
517 translation of mRNA. Because mRNA is not consumed by translation, steady-state
518 mRNA concentration is set by the flux constraint $k_{in} = \delta_S s$, which is independent of
519 translation rate, V . Thus, stochastic amplification does not occur in translation and so
520 cannot buffer translation from mRNA noise (Fig. 3). As a consequence, gene expression
521 is unusually sensitive to noise.

522

523 **Cell fitness.** Noise slows biochemical reactions, but how does it affect fitness? We now
524 consider a reaction network representing a coarse-grained model of cell growth. We then
525 analyze this reaction network in the stochastic case using the method developed above in
526 order to quantify the effects of gene expression noise on the average rate of cell division.

During balanced exponential growth, it has been shown empirically that fitness, here defined as the division rate of a cell, is equivalent to the total rate of protein synthesis by translation in a cell (Shahrezaei and Marguerat 2015; Scott et al. 2010; Scott et al. 2014). We denote the rate of cell division by the parameter λ . In general, biomass synthesis involves the import of extracellular nutrients followed, ultimately, by their conversion into new biomass via translation. Translation initiation involves the binding of mRNA, in abundance m , to a ribosome, in abundance r , followed by elongation via the polymerization addition of amino acids (aa) to a growing protein chain. Elongation rate is limited by the intracellular nutrient, in abundance s_I , in scarcest supply (i.e., von Leibig's Law of the Minimum (Droop 1974; De Baar 1994; Tilman 1982)), which is typically either ATP or aa-charged tRNA's. In eutrophic conditions, where nutrients are saturating, translation is initiation-limited (Shah et al. 2013), whereas in oligotrophic conditions, where nutrients are scarce, translation is elongation-limited (Tuller et al. 2010).

This coarse-grained cell model follows a QSSA system of ODE's (S22) with growth rate,

$$\frac{dp}{dt} = k_T r \frac{s_I m}{K_1 + K_2 s_I + K_3 m + s_I m} - \lambda p \quad 11$$

which is equivalent to a compulsory-order two-substrate enzymatic reaction with independent (i.e., non-cooperative) binding events (Ingalls 2013). k_T is the translation elongation rate per mRNA per ribosome (which absorbs the number of s_I per protein), $K_1/ K_2/ K_3$ are affinity constants, and λ is the protein dilution rate, which during steady-state growth is equivalent to the rate of cell division. Eutrophic and oligotrophic regimes, corresponding to initiation- and elongation-limited translation, respectively, are found from this equation in the limit as s_I or m , respectively, goes to infinity. Eqn. 11 can be

rearranged to form a Monod equation, which maps growth rate to extracellular nutrient supply with a number of practical and conceptual benefits (S24).

The expected steady-state rate of cell division (cells/time) is found by setting Eqn. 11 to zero and solving for λ ,

$$\langle \lambda \rangle = \frac{\langle V \rangle}{p_{ss}} \quad 12$$

p_{ss} is defined as the species-specific total protein content of a cell (protein/cell), which is assumed constant for a given species, V is the synthesis rate of proteins (protein/time).

Fitness with noise. Following the method above, we write down a QSSA reduced system of equations for growth rate according to our coarse-grained cell model (S22), giving a system in five equations describing the time evolution of mRNA, uptake proteins, extracellular substrate, intracellular substrate and protein biomass. We assume that mRNA and extracellular nutrients undergo a linear birth-death process, supplied with burst inputs of size b_t and b_{SE} respectively, and ignore ribosome fluctuations. Intracellular nutrients are consumed by protein synthesis, but mRNA is not. Expected fitness, then, is (Sup. Mat.):

$$\langle \lambda \rangle \cong \frac{\bar{V}}{p_{ss}} \left[1 - \underbrace{\frac{\overbrace{\langle \epsilon_m^2 \rangle}^{mRNA\ noise} A_1}{\bar{m} \Omega}}_{cost\ of\ gene\ expression\ noise} - \frac{\overbrace{\langle \epsilon_{PU}^2 \rangle}^{protein\ noise} A_2}{\bar{p}_U \Omega} - \frac{\overbrace{\langle \epsilon_{SE}^2 \rangle}^{extracellular\ noise} A_3}{\bar{s}_E \Omega} \right] \quad 13$$

Where the A_i 's are positive constants given in the SI. In reality, the protein noise term will contain contributions from all proteins involved in the pathway importing and processing nutrients to their final form utilized in translation. The MCA model of Wang and Zhang (Wang and Zhang 2011) then fits within this term. The extracellular noise

term quantifies the strength of selection favoring phenotypic “bet hedging” (Starrfelt and Kokko 2012) strategies in response to environmental unpredictability.

Surprisingly, transcriptional noise directly reduces fitness. This occurs because fluctuations in mRNA levels reduce the average translation rate. This is counter to the typical view, which holds that transcriptional noise impinges on cell function only to the extent that it propagates to generate protein noise. Unlike the cost of protein or extracellular nutrient noise, which require a number of parameter estimates that are highly context-dependent, the cost of mRNA noise is easily estimated from available data if we assume that cells are growing in their eutrophic regime ($s_I \rightarrow \infty$), where growth rate is limited by translation initiation. Taking the eutrophic limit of Eqn 11 gives, $dp/dt \cong k_T rm/(K_m + m) - \lambda p$, which is simply a MM reaction with mRNA serving as the substrate and ribosomes as the catalyst. Noting that fluctuations in total ribosome number and mRNA number are uncorrelated, we find from equations 4, 7 and 9,

$$\langle \lambda \rangle \cong \frac{\bar{V}}{p_{ss}} \left[1 - \underbrace{\frac{(b_t+1)}{2\bar{m}} a(1-a)\Omega^{-1}}_{\text{cost of transcriptional bursting}} \right] \quad 14$$

Where $\bar{V} = k_T rm/(K_m + m)$. The cost of transcriptional bursting vanishes with zeroth- or first-order translation ($a = 0$ and 1 , respectively) and/or as cell volume and/or mRNA abundance become large.

During balanced exponential growth of *E. coli* in rich medium: $m \sim 1400$ molecules, $r \sim 21,000$ molecules, and $k_T \sim 21aa/s \sim 4.6$ proteins/min (given a mean protein length of $275aa$). The total protein count of a cell is $p_{ss} = 2.4 \times 10^6$ molecules. During growth in nutrient rich medium, $(1 - a) \sim 0.85$ (Arava et al. 2003; Ingolia 2014;

Zenklusen, Larson, and Singer 2008). Applying these parameter estimates to \bar{V}/p_{ss} gives a doubling time for *E. coli* of ~30 minutes. Similar calculations with parameters for the budding yeast *Saccharomyces cerevisiae* ($m \sim 15,000$ -60,000, $r \sim 243,000$, $k_T \sim 15aa/s \sim 2$ proteins/min (average protein length 450aa), $p_{ss} = 5 \times 10^7$, $a = .85$) gives a doubling time of ~120 minutes. Both values are in close agreement with experiments in nutrient rich media, providing an important benchmark.

In an *E. coli* cell, according to Eqn. 14, a mutation causing one unit increase in the total mRNA Fano factor results in an average loss of 7.5 proteins/min, 223 proteins/cell division, and 1.5×10^{-4} cells/generation, such that selection against mRNA noise is more than 10,000 times stronger than random drift in this species (the effective population size determines the strength of random allele frequency fluctuations and is $N_e = 2 \times 10^8$ in *E. coli*). In contrast, the same mutation in yeast, which has 40-fold higher m , has a cost of 2.1×10^{-6} cells/generation, which is only about 20 times stronger than random drift ($N_e = 1 \times 10^7$ in *S. cerevisiae*). This illustrates how fitness magnifies small differences in reaction rates by integrating over the whole cell and cell lifetime, and highlights the importance of cell volume (or more directly, molecule number) in influencing the efficacy of selection against noise.

Rapid decay of selection efficacy with cell volume. As we just showed, the cost of noise depends on reaction volume. Because unicellular species differ over several orders of magnitude in their cell volumes, it may be possible to use these two facts in order to make predictions about taxonomic patterns of gene expression noise.

615 The fate of a mutation that modifies cellular noise depends on the relative
616 magnitude of selection and random genetic drift. Selection efficiently overcomes random
617 drift provided that $2N_e|s| > 1$ (Kimura 1957; Kimura 1962). The boundary at $2N_e|s| = 1$
618 defines the "drift barrier" separating efficient and inefficient selection regimes (Sung et
619 al. 2012; Lynch 2007b). The drift barrier acts as a kind of evolutionary attractor
620 preventing organismal perfection: recurrent deleterious mutations with effects smaller
621 than $2N_e|s|$ push species away from their adaptive optima, and selection is too weak to
622 push them back. Thus, the balance between recurrent mutation, drift and selection
623 establishes a suboptimal phenotypic steady-state. We now derive a scaling relationship
624 between cell volume and selection efficacy demonstrating that the drift boundary for
625 noise attenuation scales superlinearly with cell volume.

626

627 **Selection strength scales inversely with cell size.** Selection strength, s , is defined as the
628 difference in absolute fitness of mutant (m) and resident (r) alleles scaled by the mean
629 fitness of the population. For a rare mutant we have: $s = (\lambda_m - \lambda_r)/\lambda_r$. This is simply
630 the difference in the absolute number of daughter cells produced by the mutant per
631 resident doubling. For simplicity, we assume a haploid population and that the mutant
632 and resident genotypes differ only in transcriptional noise. When the only source of noise
633 is intrinsic mRNA fluctuations, the selection coefficient is approximately (see SI for
634 derivation),

$$635 \quad s \cong \Delta F_{mRNA} \frac{K_m}{(K_m + m)^2} \Omega_c^{-1} = \Delta CV_{mRNA}^2 a(1 - a) \Omega_c^{-1} \quad 15$$

636 Ω_c denotes the volume of the reaction compartment. The sign of selection is determined
637 by $\Delta F = (F_r - F_m)$ or $\Delta CV^2 = (CV_r^2 - CV_m^2)$, the difference in Fano factors (noise

strength) or squared coefficients of variation (noise), respectively, between resident and mutant alleles. A mutation that increases transcriptional noise will be deleterious ($s < 0$), while a decrease in noise is advantageous ($s > 0$). More generally, $s \cong \sum_i^n (\Delta F_i A_i) \Omega_{ci}^{-1}$, where the sum is over all n sources of noise affected by the mutation, the A_i are constants, and Ω_{ci} is the reaction volume of the i^{th} noise source. Note that selection cannot act on noise levels in first or zeroth order reactions ($a = 1$ or 0 , respectively), since noise does not alter the mean in these cases. Noise will, however, alter the heritability (e.g., penetrance) of a trait (Wang and Zhang 2011), which will reduce the response to selection, but we do not consider this effect here. Importantly, the strength of selection against noise explicitly depends on reaction volume.

Drift strength scales positively with cell size. A species population size tends to scale inversely with its body size: a species must divide a limited resource pool amongst its members, resulting either in a large number of small individuals or a small number of large ones (Damuth 1981; Damuth 1987). For multicellular species the empirical scaling relationship between body mass and population size, known as “Damuth’s Law”, follows a power law with a scaling exponent of $-3/4$, likely reflecting the $3/4$ power scaling of body size with energetic demands (Damuth 1981; Damuth 1987). However, in unicellular species’ metabolic rate scales either linearly (eukaryotes) or superlinearly (prokaryotes) with cell mass (DeLong et al. 2010). Assuming that N_e scales linearly with N , that mass scales linearly with cell volume (i.e., cells have constant density), then $N_e = B\Omega_w^{-g_{drift}}$, where the subscript on Ω_w denotes whole cell volume and $g_{drift} = [3/4, 5/4]$ is the scaling exponent. The constant, B , is proportional to the population size of a species with unit

volume, and absorbs cell density as the proportionality constant between mass and volume.

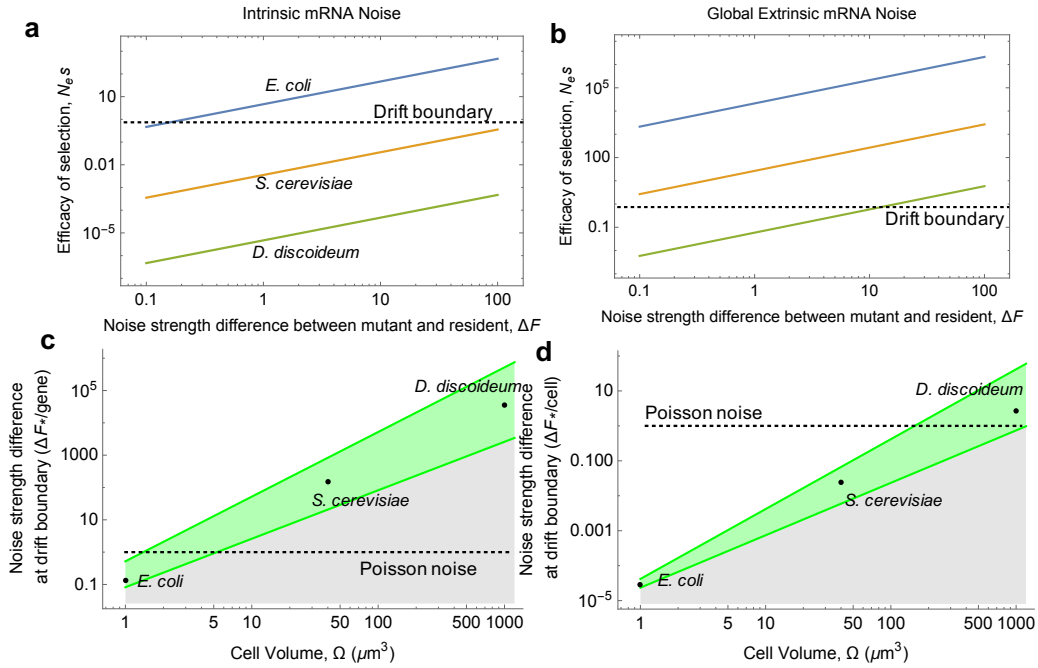
Combining forces: Selection efficacy scales superlinearly with cell size. Combining these results implies that the efficacy of selection against noise is proportional to $(\Omega_w^{-g_{drift}} \Omega_c^{-1})$. For translation, it is possible to establish a relationship between Ω_c and Ω_w , allowing us to reduce the number of independent variables.

Protein synthesis occurs in the cytoplasm/cytosol, suggesting that reaction-compartment volume may simply equal whole-cell volume ($\Omega_c \sim \Omega_w$) for translation. This relationship appears to hold between *E. coli* and yeast ($m = 1500$ and $60,000$, respectively, with a 40 fold difference in cell volume) and within mammalian cells as they change volume throughout the cell cycle (Kempe et al. 2015; Padovan-Merhar et al. 2015). However, Gillooly et al (Gillooly et al. 2005) found that RNA concentration scales as body mass to the $-1/4$ power over a broad range of eukaryotes, corresponding to RNA number scaling with mass to the $3/4$ power. Although unicells were underrepresented in their sample, this agrees well with the scaling of mRNA between *E. coli* and the amoeba *Dictyostelium discoideum* where $m = 155,000$ and cell size is approximately 500-1000 fold larger than *E. coli*, such that $\Omega_c \sim \Omega_w^{2/3}$ to $\Omega_w^{3/4}$.

Thus, for translation, we assume that reaction volume scales with whole cell volume according to, $\Omega_c = \Omega_w^{-g_{sel}}$, where $g_{sel} = [3/4, 1]$. Therefore, the efficacy of selection against mRNA noise can be written as an allometric scaling law:

$$N_e |s| \cong \Delta CV_{mRNA}^2 a(1 - a) B \Omega_w^{-g} \quad 16$$

683 The superlinear scaling exponent, $g = g_{sel} + g_{drift} = [3/2, 2^{1/4}]$, makes selection against
684 noise extremely weak in large cells. For example, all else equal, selection against a
685 mutation that increases expression noise is thirty thousand to one million times stronger
686 in *E. coli* than in a species, such as the amoeba *Dictyostelium discoideum*, which is about
687 1000 times larger (Fig. 4).



688

689 **Figure 4: Power-law scaling of selection efficacy with cell volume.** **A,B)** The predicted
690 efficacy of selection, $N_e|s|$, for *E. coli*, *S. cerevisiae* and *Dictyostelium discoideum*, as a
691 function of the jump size in noise caused by a mutation, $|\Delta F|$, for mutations acting at a
692 single locus (left-hand column) and mutations acting on the noise of the global
693 transcriptome (right-hand column). The mRNA concentrations and effective population
694 size (N_e) estimates used for the figure are found in Table S1 ($m = 1500, 60,000$ and
695 $155,000$, respectively, and ploidy-adjusted $N_e = 2 \times 10^8, 1 \times 10^7$ and 1.1×10^5 , respectively).
696 Reaction order was assumed to be $a = 0.8$ for all three species. The “drift boundary” is
697 $2N_e|s| = 1$. **C,D)** The predicted noise jump size at the drift boundary as a function of cell
698 volume. The green region represents a range of values for the scaling exponent (-2 to $-3/2$),
699 reaction order (0.75 - 0.9), and in **C)** the gene number (3250 - $12,500$). Above the
700 green region (or individual points) selection is strong, below mutations are effectively
701 neutral.
702

Solving for $|\Delta F|$ at the drift boundary ($2N_e|s| = 1$), gives the jump size in noise strength “visible” to natural selection. If the only cost of transcriptional noise is reduced translation, then the selectively impermissible jump sizes in transcriptional noise follows the scaling relationship,

$$|\Delta F_{mRNA}^*| > b\Omega_w^g \quad 17$$

where $b = (K_m + m)^2 / BK_m$. Mutations causing jumps smaller than this critical size are “effectively neutral”, and thus invisible to selection acting on translation rate. Adaptive mutations that decrease noise by an amount less than this jump size cannot efficiently spread, while deleterious mutations with effect size below this threshold cannot be efficiently purged by selection.

Parameterizing our model with literature estimates of cell volume, effective population size, genome size, ribosomal occupancy, and translation kinetics, we predict that *E. coli* should be strongly selected to minimize intrinsic transcriptional noise simply because this noise slows translation. Whereas the cost of transcriptional noise is too weak per gene (on average) in *S. cerevisiae* or *Dictyostelium* for selection to maintain intrinsic noise at the Poisson minimum if noise-induced slowdown of translation is the only cost of mRNA noise. Fig 4 shows that in cells below about 5fL (most prokaryotes), selection favors suppression of intrinsic transcriptional noise to its theoretical minimum defined by Poisson statistics, even if the only cost of transcription noise is reduced translation rate. But because of the superlinear scaling of selection efficacy, cells only 100 times larger (most eukaryotes) cannot efficiently oppose intrinsic noise jumps even as large as $\Delta F = 1000$ (again, if this is the only selection pressure for noise-minimization acting). Alternatively, jumps in global transcriptional noise (i.e., extrinsic noise) have a much

greater selective effect, meaning that traits affecting extrinsic noise, such as ribosome fluctuations (Eqn 9), should be under strong selection for noise-amelioration even in moderately sized species.

DISCUSSION

Biochemical reactions are non-linear, but are typically modeled as linear in stochastic models of gene expression due to technical limitations (van Kampen 2007). Here we have introduced a novel procedure for easily and more accurately incorporating non-linearity into models of coupled stochastic chemical reactions. We have found that hyperbolic reaction kinetics filter input noise in such a way that the average reaction rate is slowed by noise, while attenuating noise propagation. These results have a number of implications.

A drift barrier hypothesis for gene expression noise and robustness. Cell size places severe constraints on the evolution of noise suppression. The magnitude of noise-induced slowdown of Michaelis-Menten reactions is inversely proportional to the reaction volume, which follows intuitively from the fact that, for a given molecular density, greater volumes contain more molecules and thus generate less noise. Lynch (Lynch 2007b) has previously noted the “genomic perils of evolving large body size”. Large bodied species tend to be less efficient at purging deleterious mutations due to the positive scaling of genetic drift strength with body size, leading to a number of broad taxonomic patterns including greater genomic mutation rate, transposon abundance and intron number in unicellular eukaryotes and metazoans relative to prokaryotes and

viruses (Lynch 2007b; Lynch 2007a; Sung et al. 2012). Unique to gene expression noise, cell size affects not only drift but also fitness, resulting in a negative superlinear scaling of selection efficacy with cell size (Eqns. 15-17). Unicellular species span more than four orders of magnitude in cell volume ($\sim 0.1 \mu\text{m}^3$ Mycoplasma to more than $1000 \mu\text{m}^3$ in amoebazoans and other protists), with a sharp taxonomic divide separating prokaryotes and eukaryotes. “Typical” cell diameters are 0.5-5 μm for prokaryotes and 10-100 μm for unicellular eukaryotes. Taking a conservative value of a 100-fold average volume difference between prokaryotes and eukaryotes, this translates to a 1000-10,000-fold difference in the average efficacy of selection against noise for non-compartmentalized processes (Eqn. 16).

Of course, numerous other fitness effects of gene expression noise occur (Eqn. 13 and (Fraser et al. 2004; Wang and Zhang 2011; Lehner 2008; Raj and van Oudenaarden 2008; Balázsi, van Oudenaarden, and Collins 2011; Raser and O’Shea 2005)), but these must all scale with cell (or compartment) volume, such that selection for noise-amelioration will often be weaker in eukaryotes than prokaryotes. It is possible that the greater transcriptional complexity in eukaryotic cells, which employ multisubunit regulatory complexes, enhancer regions, heterochromatic gene silencing, nuclear localization into “transcription factories” and other features that greatly increase the burstiness of transcription, evolved only once greater cell size freed populations from the selective constraint of noise minimization. The cause of the size disparity between prokaryotes and eukaryotes is thought to be due to differences in cellular respiration: prokaryotes synthesize ATP via the electron transport chain embedded in the cell’s plasma membrane, which requires a high surface area to volume ratio for the whole cell

in order to balance energy production and demand. Respiration via mitochondria free eukaryotes from this constraint, allowing them to achieve large cell sizes without bankrupting the cell's energy budget. Thus, the evolution of mitochondria may have paved the way for the greater transcriptional sophistication of eukaryotes.

Translation attenuates the propagation of transcriptional noise: Gene expression noise is generated at two levels: transcription and translation. While the molecular mechanisms controlling noise at each of these levels are becoming increasingly well understood, much remains unclear about the propagation of noise between levels. I have shown here that hyperbolic translation kinetics greatly attenuate the propagation of mRNA noise to the protein level. Given parameter estimates of translation kinetics in *E. coli* and yeast, the results here imply that hyperbolic translation will reduce mRNA noise by between 25 and 100 fold, substantially attenuating the propagation of mRNA noise to the protein level. The hyperbolic filtering of mRNA noise by translation also manifests as a substantial reduction in the correlation between mRNA and protein abundance (Eqn. 10). This result is consistent with data from *E. coli* showing that there is zero correlation between a gene's mRNA noise and its protein noise (Taniguchi et al 2010). We have shown that hyperbolic translation kinetics can help explain this lack of correlation, since the correlation coefficient between mRNA and protein noise is proportional to the reaction order parameter, α , which is between 0.1 and 0.2 in *E. coli*, implying a large reduction in the correlation. Geometrically, one can visualize this diminution of correlation by observing the Michaelis-Menten rate curve: on the flat part of the curve, all

mRNA concentrations lead to the same rate of protein synthesis, thus creating a zero correlation between mRNA and protein.

A major implication of these results is that the vast majority of all intrinsic protein noise comes from translation, not transcription. Consequently, if protein noise, rather than mRNA noise, determines the functional consequences of gene expression noise, then natural selection on transcriptional noise will be extremely weak, and synthetic systems desiring control of protein noise should target translation, not transcription. We have thus come to two conclusions about transcriptional noise: 1) mRNA noise directly inhibits growth by slowing translation and 2) very little mRNA noise propagates to the protein level. These two predictions cast new light on the functional consequences of transcriptional noise in living cells.

Survival of the flattest. We have shown that hyperbolic reactions, by filtering out high-amplitude input fluctuations, cause a reduction in both signal (average output) and noise (relative output variance) compared to first-order reactions. The strength of selection favoring organismal robustness is determined by the former, while the target of selection for increased robustness is determined by the latter. Both the strength of noise and its cost are mediated by the reaction order, $a = K_m / (K_m + s)$ (Eqns. 6, 9), making this the obvious target for noise-attenuation. Noise-induced slowdown is minimal for first and zeroth order reactions ($a = 1, 0$, respectively), while noise itself is minimal only in zeroth order ($a = 0$). Indeed, zeroth order kinetics completely filter out super-Poisson intrinsic input noise (Eqn. 9; Fig. 2), making the output entirely insensitive to input noise. We might call this phenomenon “zeroth order insensitivity”. Selection for robustness, then, should push

reactions towards their zeroth order regime. Geometrically, the zeroth order regime corresponds to the flat part of the rate curve, such that adaptive robustness by this mechanism promotes “survival of flattest”, a term previously invoked to describe selection for robustness via very different mechanisms (Wilke et al. 2001; Codoñer et al. 2006).

There are two targets of modification for reaction order minimization: increasing substrate abundance, s , which moves the substrate distribution to the flatter part of the rate curve, or reducing K_m , which shifts the rate curve itself. Importantly, these strategies are not equivalent (Fig. 2). Decreasing K_m (while holding s constant) causes a unimodal change in product noise because K_m has two complementary effects: it reduces reaction order, α , which filters noise, but it also increases product burst size, b_T (Fig. 2).

Alternatively, increasing mRNA concentration (with K_m held constant) monotonically decreases protein noise because both reaction order, α , and translational bursting, b_T , decline. This strategy increases fitness monotonically, thus avoiding the mRNA/protein noise trade-off. Importantly, increasing robustness by increasing substrate concentration is different from increasing robustness by decreasing substrate noise. To highlight this, one could force noise strength to stay constant while increasing mean substrate. The result would be increased robustness without a decrease in noise strength. This strategy undoubtedly generates other costs not accounted for in the current model, such as the energetic cost of increased mRNA synthesis. Future work should explicitly examine the evolutionary feasibility of cellular robustness via this mechanism in light of its potential trade-offs.

Robustness via compartmentalization: The present theory provides new insight into the recent argument that compartmentalization of reactions in eukaryotic cells is a strategy for noise control (Stoeger, Battich, and Pelkmans 2016). If cell compartments are filled with reactants via active transport by membrane transporters, then the resulting hyperbolic filtering (described above) will attenuate copy number fluctuations of reactants potentially down to the Poisson minimum, though not lower (e.g., Eqns 9 and S21). But hyperbolic filtering also slows the average rate of transport, creating inefficiency. In addition, reducing the reaction volume increases the relative size of fluctuations (the noise), which slows the absolute average rate of reactions within the compartment. Future work is necessary to determine exactly when the conflicting costs and benefits of hyperbolic filtering favor reaction compartmentalization as a strategy for cellular robustness. In eukaryotes, this takes on another dimension, because mRNA is actively transported out of the nucleus and into the cytoplasm before translation. The hyperbolic filtering of transcriptional noise at this level will greatly ameliorate transcriptional noise even before mRNA is filtered by translation. Future work should model the consequences of this serial transcriptional noise filtering in eukaryotes.

Metabolic flux. Wang and Zhang (Wang and Zhang 2011) used metabolic control analysis (MCA) to study the effect of enzyme fluctuations on metabolic flux, concluding that enzyme noise diminishes fitness in pathways with more than one enzyme. This result arises from the hyperbolic relationship between enzyme concentration and steady state flux in multi-enzyme metabolic pathways (Kacser and Burns 1973), and so, like the present paper, is a straightforward consequence of nonlinear averaging. However, MCA

models assume first-order enzyme kinetics (Kacser and Burns 1973). In light of the present results, it is not immediately clear if this assumption makes the results of (Wang and Zhang 2011) an underestimate of the true cost of gene expression noise, because hyperbolic rate laws cause noise-induced slowdown at each pathway step, or an overestimate of cost, because hyperbolic rate laws attenuate noise at each step, thus damping noise propagation through the pathway. Future work should distinguish between these two possibilities.

The crucial assumption: All of the major results of this paper rely on the validity of hyperbolic reaction kinetics as an accurate description of *in vivo* biochemical reactions. This assumption is not guaranteed, especially for translation. Translation is a unique biochemical reaction because a single substrate molecule (mRNA) is simultaneously bound to multiple enzymes (ribosomes), and it is therefore easy to imagine that translation kinetics may not be well-described by Michaelis-Menten reaction kinetics. Therefore, the present theoretical results must remain in question until future empirical work carefully measures the *in vivo* rate law of protein translation.

METHODS

All Monte Carlo simulations were run on the elementary (microscopic) system of reactions (S17) using Gillespie's Exact Stochastic Simulation Algorithm (SSA) (Gillespie 1977) implemented in *Mathematica* v10.3.0.0 with the open source xSSALite package, freely available from the xCellerator project (<http://www.xlr8r.info/SSA/>). Each realization of the simulation is a Markov Chain with irregular (exponentially distributed) time intervals. For statistical analysis, each realization was regularized using the

TimeSeriesAggregate[] command. Mean[] and Variance[] commands of the regularized time series were used to compute the moments of the Markov Chain. Data points in figures 1,2 and 3 are ensemble averages or variances over 1000-10,000 simulated realizations of the stochastic process; each realization was run for at least 10 times the half life of S or P ($\text{half-life} = \ln(2)/\delta$), whichever was longer, and the value for that realization was s or p during the final second of the simulation run.

References

- Acar, Murat, Jerome T Mettetal, and Alexander van Oudenaarden. 2008. "Stochastic Switching as a Survival Strategy in Fluctuating Environments." *Nature Genetics* 40 (4): 471–75.
- Alemu, Elfalem Y, Joseph W Carl, Héctor Corrada Bravo, and Sridhar Hannenhalli. 2014. "Determinants of Expression Variability." *Nucleic Acids Research*. doi:10.1093/nar/gkt1364.
- Arava, Yoav, Yulei Wang, John D Storey, Chih Long Liu, Patrick O Brown, and Daniel Herschlag. 2003. "Genome-Wide Analysis of mRNA Translation Profiles in *Saccharomyces Cerevisiae*." *Proceedings of the National Academy of Sciences* 100 (7): 3889–94. doi:10.1073/pnas.0635171100.
- Balázsi, Gábor, Alexander van Oudenaarden, and James J Collins. 2011. "Cellular Decision Making and Biological Noise: From Microbes to Mammals." *Cell* 144 (6): 910–25. doi:http://dx.doi.org/10.1016/j.cell.2011.01.030.
- Bar-Even, Arren, Johan Paulsson, Narendra Maheshri, Miri Carmi, Erin O'Shea, Yitzhak Pilpel, and Naama Barkai. 2006. "Noise in Protein Expression Scales with Natural Protein Abundance." *Nat Genet* 38 (6): 636–43. doi:http://www.nature.com/ng/journal/v38/n6/supinfo/ng1807_S1.html.
- Batada, Nizar N, and Laurence D Hurst. 2007. "Evolution of Chromosome Organization Driven by Selection for Reduced Gene Expression Noise." *Nat Genet* 39 (8). Nature Publishing Group: 945–49. doi:http://www.nature.com/ng/journal/v39/n8/supinfo/ng2071_S1.html.
- Berg, Otto G. 1978. "A Model for the Statistical Fluctuations of Protein Numbers in a Microbial Population." *Journal of Theoretical Biology* 71 (4): 587–603. doi:http://dx.doi.org/10.1016/0022-5193(78)90326-0.
- Blake, William J, Mads Kaern, Charles R Cantor, and J J Collins. 2003. "Noise in Eukaryotic Gene Expression." *Nature* 422 (6932): 633–37. doi:http://www.nature.com/nature/journal/v422/n6932/supinfo/nature01546_S1.html.
- Briggs, George Edward, and John Burdon Sanderson Haldane. 1925. "A Note on the Kinetics of Enzyme Action." *Biochemical Journal* 19 (2): 338.

- Chen, Guoan, Tarek G Gharib, Chiang-Ching Huang, Jeremy M G Taylor, David E
Misek, Sharon L R Kardia, Thomas J Giordano, et al. 2002. "Discordant Protein and
mRNA Expression in Lung Adenocarcinomas." *Molecular & Cellular Proteomics* 1
(4): 304–13. doi:10.1074/mcp.M200008-MCP200 .
- Codoñer, Francisco M, José-Antonio Darós, Ricard V Solé, and Santiago F Elena. 2006.
"The Fittest versus the Flattest: Experimental Confirmation of the Quasispecies
Effect with Subviral Pathogens." *PLoS Pathog* 2 (12): e136.
- Csete, Marie E, and John C Doyle. 2002. "Reverse Engineering of Biological
Complexity." *Science* 295 (5560): 1664–69. doi:10.1126/science.1069981.
- Damuth, John. 1981. "Population Density and Body Size in Mammals." *Nature* 290
(5808): 699–700. <http://dx.doi.org/10.1038/290699a0>.
- . 1987. "Interspecific Allometry of Population Density in Mammals and Other
Animals: The Independence of Body Mass and Population Energy-use." *Biological
Journal of the Linnean Society* 31 (3): 193–246.
- Dar, Roy D, Brandon S Razooky, Abhyudai Singh, Thomas V Trimeloni, James M
McCollum, Chris D Cox, Michael L Simpson, and Leor S Weinberger. 2012.
"Transcriptional Burst Frequency and Burst Size Are Equally Modulated across the
Human Genome." *Proceedings of the National Academy of Sciences* 109 (43).
National Acad Sciences: 17454–59.
- De Baar, H J W. 1994. "Von Liebig's Law of the Minimum and Plankton Ecology
(1899–1991)." *Progress in Oceanography* 33 (4): 347–86.
- DeLong, John P, Jordan G Okie, Melanie E Moses, Richard M Sibly, and James H
Brown. 2010. "Shifts in Metabolic Scaling, Production, and Efficiency across Major
Evolutionary Transitions of Life." *Proceedings of the National Academy of Sciences*
107 (29): 12941–45. doi:10.1073/pnas.1007783107.
- Draghi, Jeremy, and Michael Whitlock. 2015. "Robustness to Noise in Gene Expression
Evolves despite Epistatic Constraints in a Model of Gene Networks." *Evolution* 69
(9). Wiley Online Library: 2345–58.
- Droop, M R. 1974. "The Nutrient Status of Algal Cells in Continuous Culture." *Journal
of the Marine Biological Association of the United Kingdom* 54 (04): 825–55.
- Eldar, Avigdor, and Michael B Elowitz. 2010. "Functional Roles for Noise in Genetic
Circuits." *Nature* 467 (7312). Nature Publishing Group, a division of Macmillan
Publishers Limited. All Rights Reserved.: 167–73.
<http://dx.doi.org/10.1038/nature09326>.
- Elf, J, and M Ehrenberg. 2003. "Fast Evaluation of Fluctuations in Biochemical
Networks with the Linear Noise Approximation." *Genome Research* 13 (11): 2475–
84. doi:10.1101/gr.1196503.
- Elowitz, Michael B, Arnold J Levine, Eric D Siggia, and Peter S Swain. 2002.
"Stochastic Gene Expression in a Single Cell." *Science* 297 (5584): 1183–86.
doi:10.1126/science.1070919.
- Fraser, Hunter B, Aaron E Hirsh, Guri Giaever, Jochen Kumm, and Michael B Eisen.
2004. "Noise Minimization in Eukaryotic Gene Expression." *PLoS Biol* 2 (6). Public
Library of Science: e137. doi:10.1371/journal.pbio.0020137.
- Ghaemmaghami, Sina, Won-Ki Huh, Kiowa Bower, Russell W Howson, Archana Belle,
Noah Dephoure, Erin K O'Shea, and Jonathan S Weissman. 2003. "Global Analysis
of Protein Expression in Yeast." *Nature* 425 (6959): 737–41.

972 <http://dx.doi.org/10.1038/nature02046>.
 973 Gillespie, Daniel T. 1977. "Exact Stochastic Simulation of Coupled Chemical
 974 Reactions." *The Journal of Physical Chemistry* 81 (25): 2340–61.
 975 Gillooly, James F, Andrew P Allen, James H Brown, James J Elser, Carlos Martinez del
 976 Rio, Van M Savage, Geoffrey B West, William H Woodruff, and H Arthur Woods.
 977 2005. "The Metabolic Basis of Whole-Organism RNA and Phosphorus Content."
 978 *Proceedings of the National Academy of Sciences of the United States of America*
 979 102 (33). National Acad Sciences: 11923–27.
 980 Golding, Ido, Johan Paulsson, Scott M Zawilski, and Edward C Cox. 2005. "Real-Time
 981 Kinetics of Gene Activity in Individual Bacteria." *Cell* 123 (6): 1025–36.
 982 doi:<http://dx.doi.org/10.1016/j.cell.2005.09.031>.
 983 Grima, R. 2009a. "Investigating the Robustness of the Classical Enzyme Kinetic
 984 Equations in Small Intracellular Compartments." *Bmc Systems Biology* 3. doi:Artn
 985 10110.1186/1752-0509-3-101.
 986 ———. 2009b. "Noise-Induced Breakdown of the Michaelis-Menten Equation in
 987 Steady-State Conditions." *Physical Review Letters* 102 (21). doi:ARTN
 988 21810310.1103/PhysRevLett.102.218103.
 989 ———. 2010. "An Effective Rate Equation Approach to Reaction Kinetics in Small
 990 Volumes: Theory and Application to Biochemical Reactions in Nonequilibrium
 991 Steady-State Conditions." *Journal of Chemical Physics* 133 (3). doi:Artn
 992 03510110.1063/1.3454685.
 993 Gygi, Steven P, Yvan Rochon, B Robert Franza, and Ruedi Aebersold. 1999.
 994 "Correlation between Protein and mRNA Abundance in Yeast." *Molecular and*
 995 *Cellular Biology* 19 (3). Am Soc Microbiol: 1720–30.
 996 Ingalls, Brian P. 2013. *Mathematical Modeling in Systems Biology: An Introduction*.
 997 Cambridge, Massachusetts: The MIT Press.
 998 Ingolia, Nicholas T. 2014. "Ribosome Profiling: New Views of Translation, from Single
 999 Codons to Genome Scale." *Nat Rev Genet* 15 (3). Nature Publishing Group, a
 1000 division of Macmillan Publishers Limited. All Rights Reserved.: 205–13.
 1001 doi:10.1038/nrg3645.
 1002 Kacser, H and, and JA34 Burns. 1973. "The Control of Flux." In *Symp. Soc. Exp. Biol.*,
 1003 27:65–104.
 1004 Kaern, Mads, Timothy C Elston, William J Blake, and James J Collins. 2005.
 1005 "Stochasticity in Gene Expression: From Theories to Phenotypes." *Nat Rev Genet* 6
 1006 (6): 451–64. <http://dx.doi.org/10.1038/nrg1615>.
 1007 Kempe, H, A Schwabe, F Crémazy, P J Verschure, and F J Bruggeman. 2015. "The
 1008 Volumes and Transcript Counts of Single Cells Reveal Concentration Homeostasis
 1009 and Capture Biological Noise." *Molecular Biology of the Cell* 26 (4): 797–804.
 1010 doi:10.1091/mbc.E14-08-1296.
 1011 Kimura, Motoo. 1957. "Some Problems of Stochastic Processes in Genetics." *The Annals*
 1012 *of Mathematical Statistics* 28 (4). Institute of Mathematical Statistics: 882–901.
 1013 <http://www.jstor.org/stable/2237051>.
 1014 ———. 1962. "On the Probability of Fixation of Mutant Genes in a Population."
 1015 *Genetics* 47 (6): 713–19. <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1210364/>.
 1016 Kussell, Edo, and Stanislas Leibler. 2005. "Phenotypic Diversity, Population Growth, and
 1017 Information in Fluctuating Environments." *Science* 309 (5743): 2075–78.

- Lehner, Ben. 2008. "Selection to Minimise Noise in Living Systems and Its Implications for the Evolution of Gene Expression." *Molecular Systems Biology* 4 (1). doi:10.1038/msb.2008.11.
- Lynch, Michael. 2007a. "The Frailty of Adaptive Hypotheses for the Origins of Organismal Complexity." *Proceedings of the National Academy of Sciences* 104 (suppl 1): 8597–8604.
- . 2007b. *The Origins of Genome Complexity*. Sunderland, MA: Sinauer Associates, Inc.
- McAdams, Harley H., and Adam Arkin. 1997. "Stochastic Mechanisms in Gene Expression." *Proceedings of the National Academy of Sciences* 94 (3): 814–19. <http://www.pnas.org/content/94/3/814.abstract>.
- Metzger, Brian P H, David C Yuan, Jonathan D Gruber, Fabien Duveau, and Patricia J Wittkopp. 2015. "Selection on Noise Constrains Variation in a Eukaryotic Promoter." *Nature* 521 (7552). Nature Publishing Group, a division of Macmillan Publishers Limited. All Rights Reserved.: 344–47. doi:10.1038/nature14244<http://www.nature.com/nature/journal/v521/n7552/abs/nature14244.html#supplementary-information>.
- Michaelis, Leonor, and Maud L Menten. 1913. "Die Kinetik Der Invertinwirkung." *Biochem. Z* 49 (333-369): 352.
- Newman, John R S, Sina Ghaemmaghami, Jan Ihmels, David K Breslow, Matthew Noble, Joseph L DeRisi, and Jonathan S Weissman. 2006. "Single-Cell Proteomic Analysis of *S. Cerevisiae* Reveals the Architecture of Biological Noise." *Nature* 441 (7095): 840–46. doi:http://www.nature.com/nature/journal/v441/n7095/supinfo/nature04785_S1.html.
- Padovan-Merhar, O, G P Nair, A G Biaesch, A Mayer, S Scarfone, S W Foley, A R Wu, L S Churchman, A Singh, and A Raj. 2015. "Single Mammalian Cells Compensate for Differences in Cellular Volume and DNA Copy Number through Independent Global Transcriptional Mechanisms." *Molecular Cell* 58 (2): 339–52. doi:10.1016/j.molcel.2015.03.005.
- Paulsson, Johan. 2004. "Summing up the Noise in Gene Networks." *Nature* 427 (6973): 415–18. doi:http://www.nature.com/nature/journal/v427/n6973/supinfo/nature02257_S1.html.
- Paulsson, Johan, Otto G Berg, and Måns Ehrenberg. 2000. "Stochastic Focusing: Fluctuation-Enhanced Sensitivity of Intracellular Regulation." *Proceedings of the National Academy of Sciences* 97 (13): 7148–53. doi:10.1073/pnas.110057697.
- Pedraza, Juan M, and Johan Paulsson. 2008. "Effects of Molecular Memory and Bursting on Fluctuations in Gene Expression." *Science* 319 (5861): 339–43. doi:10.1126/science.1144331.
- Raj, Arjun, Charles S Peskin, Daniel Tranchina, Diana Y Vargas, and Sanjay Tyagi. 2006. "Stochastic mRNA Synthesis in Mammalian Cells." *PLoS Biol* 4 (10). doi:10.1371/journal.pbio.0040309.
- Raj, Arjun, and Alexander van Oudenaarden. 2008. "Nature, Nurture, or Chance: Stochastic Gene Expression and Its Consequences." *Cell* 135 (2): 216–26. doi:<http://dx.doi.org/10.1016/j.cell.2008.09.050>.

- Rao, Christopher V, and Adam P Arkin. 2003. "Stochastic Chemical Kinetics and the Quasi-Steady-State Assumption: Application to the Gillespie Algorithm." *The Journal of Chemical Physics* 118 (11): 4999–5010.
- Rao, Christopher V, Denise M Wolf, and Adam P Arkin. 2002. "Control, Exploitation and Tolerance of Intracellular Noise." *Nature* 420 (6912): 231–37. <http://dx.doi.org/10.1038/nature01258>.
- Raser, Jonathan M, and Erin K O'Shea. 2004. "Control of Stochasticity in Eukaryotic Gene Expression." *Science* 304 (5678): 1811–14. doi:10.1126/science.1098641.
- . 2005. "Noise in Gene Expression: Origins, Consequences, and Control." *Science* 309 (5743): 2010–13. doi:10.1126/science.1105891.
- Sanchez, Alvaro, and Ido Golding. 2013. "Genetic Determinants and Cellular Constraints in Noisy Gene Expression." *Science* 342 (6163): 1188–93. doi:10.1126/science.1242975.
- Sanft, Kevin R, Daniel T Gillespie, and Linda R Petzold. 2011. "Legitimacy of the Stochastic Michaelis-Menten Approximation." *Systems Biology, IET* 5 (1): 58–69.
- Schwanhauser, Bjorn, Dorothea Busse, Na Li, Gunnar Dittmar, Johannes Schuchhardt, Jana Wolf, Wei Chen, and Matthias Selbach. 2011. "Global Quantification of Mammalian Gene Expression Control." *Nature* 473 (7347). Nature Publishing Group, a division of Macmillan Publishers Limited. All Rights Reserved.: 337–42. doi:<http://www.nature.com/nature/journal/v473/n7347/abs/10.1038-nature10098-unlocked.html#supplementary-information>.
- Scott, Matthew, Carl W Gunderson, Eduard M Mateescu, Zhongge Zhang, and Terence Hwa. 2010. "Interdependence of Cell Growth and Gene Expression: Origins and Consequences." *Science* 330 (6007): 1099–1102.
- Scott, Matthew, Stefan Klumpp, Eduard M Mateescu, and Terence Hwa. 2014. "Emergence of Robust Growth Laws from Optimal Regulation of Ribosome Synthesis." *Molecular Systems Biology* 10 (8): 747.
- Shah, Premal, Yang Ding, Malwina Niemczyk, Grzegorz Kudla, and Joshua B Plotkin. 2013. "Rate-Limiting Steps in Yeast Protein Translation." *Cell* 153 (7): 1589–1601. doi:<http://dx.doi.org/10.1016/j.cell.2013.05.049>.
- Shahrezaei, Vahid, and Samuel Marguerat. 2015. "Connecting Growth with Gene Expression: Of Noise and Numbers." *Current Opinion in Microbiology* 25: 127–35. doi:<http://dx.doi.org/10.1016/j.mib.2015.05.012>.
- Shahrezaei, Vahid, and Peter S Swain. 2008. "Analytical Distributions for Stochastic Gene Expression." *Proceedings of the National Academy of Sciences* 105 (45): 17256–61. doi:10.1073/pnas.0803850105.
- Starrfelt, Jostein, and Hanna Kokko. 2012. "Bet-Hedging—a Triple Trade-off between Means, Variances and Correlations." *Biological Reviews* 87 (3). Blackwell Publishing Ltd: 742–55. doi:10.1111/j.1469-185X.2012.00225.x.
- Stoeger, Thomas, Nico Battich, and Lucas Pelkmans. 2016. "Passive Noise Filtering by Cellular Compartmentalization." *Cell* 164 (6): 1151–61. doi:<http://dx.doi.org/10.1016/j.cell.2016.02.005>.
- Sung, Way, Matthew S Ackerman, Samuel F Miller, Thomas G Doak, and Michael Lynch. 2012. "Drift-Barrier Hypothesis and Mutation-Rate Evolution." *Proceedings of the National Academy of Sciences* 109 (45): 18488–92. doi:10.1073/pnas.1216223109.

- Suter, David M, Nacho Molina, David Gatfield, Kim Schneider, Ueli Schibler, and Felix Naef. 2011. "Mammalian Genes Are Transcribed with Widely Different Bursting Kinetics." *Science* 332 (6028): 472–74.
<http://science.sciencemag.org/content/332/6028/472.abstract>.
- Swain, Peter S, Michael B Elowitz, and Eric D Siggia. 2002. "Intrinsic and Extrinsic Contributions to Stochasticity in Gene Expression." *Proceedings of the National Academy of Sciences* 99 (20): 12795–800. doi:10.1073/pnas.162041399.
- Taniguchi, Yuichi, Paul J Choi, Gene-Wei Li, Huiyi Chen, Mohan Babu, Jeremy Hearn, Andrew Emili, and X Sunney Xie. 2010. "Quantifying E. Coli Proteome and Transcriptome with Single-Molecule Sensitivity in Single Cells." *Science* 329 (5991): 533–38. doi:10.1126/science.1188308.
- Thattai, Mukund, and Alexander van Oudenaarden. 2001. "Intrinsic Noise in Gene Regulatory Networks." *Proceedings of the National Academy of Sciences* 98 (15): 8614–19. doi:10.1073/pnas.151588598.
- Thattai, Mukund, and Alexander Van Oudenaarden. 2004. "Stochastic Gene Expression in Fluctuating Environments." *Genetics* 167 (1): 523–30.
- Tilman, D. 1982. *Resource Competition and Community Structure*. Princeton, NJ: Princeton University Press.
- Tuller, Tamir, Asaf Carmi, Kalin Vestsigian, Sivan Navon, Yuval Dorfan, John Zaborske, Tao Pan, Orna Dahan, Itay Furman, and Yitzhak Pilpel. 2010. "An Evolutionarily Conserved Mechanism for Controlling the Efficiency of Protein Translation." *Cell* 141 (2): 344–54. doi:http://dx.doi.org/10.1016/j.cell.2010.03.031.
- van Kampen, N G. 2007. *Stochastic Processes in Physics and Chemistry*. Third Edit. New York: Elsevier.
- Wagner, G, xfc, P nter, Ginger Booth, and Homayoun Bagheri-Chaichian. 1997. "A Population Genetic Theory of Canalization." *Evolution* 51 (2). [Society for the Study of Evolution, Wiley]: 329–47. doi:10.2307/2411105.
- Wang, Zhi, and Jianzhi Zhang. 2011. "Impact of Gene Expression Noise on Organismal Fitness and the Efficacy of Natural Selection." *Proceedings of the National Academy of Sciences* 108 (16): E67–76. doi:10.1073/pnas.1100059108.
- Wilke, Claus O, Jia Lan Wang, Charles Ofria, Richard E Lenski, and Christoph Adami. 2001. "Evolution of Digital Organisms at High Mutation Rates Leads to Survival of the Flattest." *Nature* 412 (6844): 331–33.
- Zenklusen, Daniel, Daniel R Larson, and Robert H Singer. 2008. "Single-RNA Counting Reveals Alternative Modes of Gene Expression in Yeast." *Nat Struct Mol Biol* 15 (12). Nature Publishing Group: 1263–71.
doi:http://www.nature.com/nsmb/journal/v15/n12/supinfo/nsmb.1514_S1.html.

Supplemental Material: “Fitness costs of noise in biochemical reaction networks and the evolutionary limits of cellular robustness”

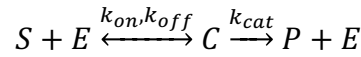
J. David Van Dyken

University of Miami

Derivations of Theoretical Results

I.a. Stochastic Michaelis-Menten quasi steady-state approximation (MM QSSA)

Consider the canonical bimolecular reaction scheme representing the catalyzed conversion of substrate into product (3, 4):



Substrate S reversibly binds catalyst E to form a reaction intermediate complex, C , that then decays irreversibly into E and product P . Irreversibility is allowed because living systems exist far from thermodynamic equilibrium. Because three or more body collisions are rare, all catalytic reactions can be decomposed into a series of bimolecular reactions of this type, which is why this reaction scheme provides the cornerstone of biochemistry.

This microscopic description is converted into a system of first-order non-linear differential equations (ODE's) by applying the law of mass action. This system cannot be solved explicitly, so we apply the quasi-steady state approximation (QSSA), which assumes that the timescale of intermediate (C) formation is fast relative to product (P) formation, such that C equilibrates to its steady-state concentration effectively instantaneously to changes in S (4, 5). This separation of timescales allows us to set the concentration of C to its steady-state value, giving the steady-state rate of product formation,

$$\frac{dp}{dt} = V = V_{max} \frac{s}{K_m + s} \quad \text{S1}$$

Here, s is reactant concentration (following convention, lowercase denotes concentration), $V_{max} = k_{cat}e$ is the maximum achievable reaction rate, and $k_m = (k_{off} + k_{cat})/k_{on}$, the Michaelis or half-saturation constant, is the substrate

concentration at $V = V_{max}/2$. This rate law function is a rectangular hyperbola, with V approaching V_{max} asymptotically as s goes to infinite (Figure 1).

Because S1 is hyperbolic, Jensen's inequality proves that, for random s , $\langle V(s) \rangle \leq V(\langle s \rangle)$. This says that the average reaction rate is less than or equal to a noise-free (i.e., deterministic) reaction with $s = \langle s \rangle$. This is shown by Taylor expanding the Michaelis-Menten reaction rate about the average value of s . First, multiply the numerator and denominator of S1 by the reaction volume, Ω , to obtain a description in terms of particle numbers, $n_S = \langle n_S \rangle + \epsilon_S$, where $\langle n_S \rangle$ is the ensemble average number of species S , and ϵ_S is a random variable with finite variance and zero mean. Expanding S1 in powers of ϵ_S and taking the expected value gives,

$$\langle V(s) \rangle = \underbrace{V_{max} \frac{\langle n_S \rangle}{k_m + \langle n_S \rangle}}_{V(\langle s \rangle)} - \underbrace{\langle \epsilon_S^2 \rangle V_{max} \frac{k_m}{(k_m + \langle n_S \rangle)^3}}_{\text{cost of noise}} + O(\epsilon^3) \quad \text{S2}$$

Where $k_m = \Omega K_m$. The variance in V is,

$$\sigma_V^2 = \langle V^2 \rangle - \langle V \rangle^2 = \langle \epsilon_S^2 \rangle V_{max}^2 \frac{k_m^2}{(k_m + \langle n_S \rangle)^4} + O(\epsilon^3) \quad \text{S3}$$

Substrate noise reduces the rate of biochemical reactions. The difference between the average stochastic rate and deterministic rate grows monotonically with increasing noise, and equality occurs only with zero noise or zero curvature of $V(s)$. For gamma distributed n_S , the probability density function of the Michaelis-Menten reaction rate (for $V < V_{max}$) is exactly,

$$P(V = v) = \frac{V_{max}}{k_m} \frac{v^{\alpha-1} e^{-v/z\beta}}{z(z\beta)^\alpha \Gamma(\alpha)}, \quad z = \frac{V_{max} - v}{k_m} \quad \text{S4}$$

which has mean,

$$\langle V \rangle = V_{max} \frac{e^{\frac{k_m}{F}}}{\eta^2} \left(\frac{F}{k_m} \right)^{-\eta^{-2}} \Gamma\left(-\eta^{-2}, \frac{k_m}{F}\right) \quad \text{S5}$$

where $\Gamma(.,.)$ is the incomplete gamma function. Equation S4 is verified by Monte Carlo simulations implemented with Gillespie's Exact Stochastic Simulation Algorithm (SSA) (1) (Fig. 1).

Figure 1 shows that the effect of noise on MM-type reactions is to increase the observed Michaelis constant. Therefore, a stochastic formulation of S1 that accounts for all stochastic effects of fluctuating s at steady-state is obtained simply by substituting an “effective”, stochastic Michaelis constant using equations or S5, respectively:

$$\langle V \rangle = V_{max} \frac{\langle s \rangle}{\tilde{K}_m + \langle s \rangle} \quad \text{S6}$$

where,

$$\tilde{K}_m = \frac{e^{-\frac{K_m}{F}} F \left(\frac{F}{K_m}\right)^{-\eta^{-2}}}{\Gamma(-\eta^{-2}, \frac{K_m}{F})} - F \eta^{-2}, \quad \tilde{K}_m \cong K_m \left[\frac{(K_m + \langle s \rangle)^2 + \langle s \rangle F}{(K_m + \langle s \rangle)^2 - K_m F} \right]$$

Note that $\tilde{K}_m \geq K_m$ and \tilde{K}_m increases with increasing noise strength (measured by the Fano factor, F). This simple substitution retains the form of the classical MM equation (Equation S1) and provides a simple, phenomenological short-hand for the effect of noise on MM reaction kinetics.

I.b. Reaction networks

I.b.i. The multivariate system-size expansion of the chemical master equation

We now solve for the time evolution of the moments (mean and variance) for each species in a chemical reaction network. We employ the chemical master equation and its expansion in orders of the inverse square root of system size, Ω . The chemical master equation describes the time evolution of the probability distribution of chemical species. Because these equations include bimolecular reactions they are non-linear, and are typically not explicitly solvable. The system size expansion is an approximation method, where the parameter Ω provides an objective criterion for truncation of the series expansion allowing for moment closure. For thorough reviews and extensions of the method see van Kampen (6), Elf and Erhenberg (7), and Grima (8). van Kampen introduced the system size expansion, Elf and Erhenberg generalized the system size expansion to systems of chemical reactions, and Grima extended this further by deriving a multivariate system size expansion that retained terms beyond the linear noise

approximation (LNA). Here we build on this work to provide a novel method that accurately models the statistics of gene expression and metabolic networks in living cells.

We define $P(\vec{n}, t)$ as the joint probability density function of particle numbers for all N species in a system of R reactions. Its time evolution follows the chemical master equation (7),

$$\frac{\partial P(\vec{n}, t)}{\partial t} = \Omega \sum_{j=1}^R \left(\prod_{i=1}^N E_i^{-S_{ij}} - 1 \right) f_j(\vec{n}, \Omega) P(\vec{n}, t) \quad S7$$

where \vec{n} is the vector of particle numbers of all species, Ω is the reaction volume, $f_j(\vec{n}, \Omega)$ is the transition rate between states for the j th reaction, S_{ij} is the stoichiometric coefficient for the net reaction, and E is the step operator which acts as follows:

$E_i^{-S_{ij}} f(n_1, \dots, n_i, \dots, n_N) = f(n_1, \dots, n_i - S_{ij}, \dots, n_N)$. The particle number is then written as a random perturbation about the macroscopic concentration (thus translating from the extensive variable n to the intensive variable x) according to the ansatz that the relative perturbation size scales as the inverse square root of the reaction volume, in accordance with the Law of Large Numbers, leading to the substitution:

$$\frac{n_i}{\Omega} = \bar{x}_i + \Omega^{-1/2} \epsilon_i \quad S8$$

where \bar{x}_i is the macroscopic steady-state concentration of species i , and ϵ_i is a continuous random variable. Substituting into Eqn. S7 and simplifying (6, 7),

$$\frac{\partial P(\vec{n}, t)}{\partial t} = \frac{\partial \Pi(\vec{\epsilon}, t)}{\partial t} - \Omega^{1/2} \sum_{i=1}^N \frac{\partial x_i}{\partial t} \frac{\partial \Pi(\vec{\epsilon}, t)}{\partial \epsilon_i} \quad S9$$

Taylor expansion of the transition rate for the j th reaction, $f_j(\vec{x})$, about the macroscopic rate gives,

$$f_j(\vec{x}) = f_j(\vec{\bar{x}}) + \Omega^{-1/2} \sum_i \frac{\partial f_j(\vec{\bar{x}})}{\partial x_i} \epsilon_i + \frac{\Omega^{-1}}{2} \sum_{i,k=1}^N \frac{\partial^2 f_j(\vec{\bar{x}})}{\partial x_i \partial x_k} \epsilon_i \epsilon_k + O(\Omega^{-3/2}) \quad S10$$

$$= f_j(\vec{\bar{x}}) + \Omega^{-1/2} \nabla f_j(\vec{\bar{x}}) \vec{\epsilon} + \frac{\Omega^{-1}}{2} \vec{\epsilon}^T \mathbf{H}(\vec{\bar{x}}) \vec{\epsilon} + O(\Omega^{-3/2})$$

where ∇ is the gradient operator and \mathbf{H} is the Hessian matrix.

Note that S10 differs from that of Grima (equation 13 in (8)), who included an additional term accounting for the special case of homodimerization. While Grima's expansion works for all elementary reactions, it gives erroneous results when applied to bimolecular reactions written in QSSA reduced form. We therefore use Eqn. S10 with the caveat that it will be inaccurate for homodimerization reactions at low copy numbers.

Taylor expansion of $\prod_{i=1}^N E_i^{-S_{ij}}$ to order $O(\Omega^{-3/2})$ gives the differential operator (6–8),

$$\begin{aligned} & \prod_{i=1}^N E_i^{-S_{ij}} \\ &= 1 - \Omega^{-1/2} \sum_i S_{ij} \frac{\partial}{\partial \epsilon_i} + \frac{\Omega^{-1}}{2!} \sum_{i,k=1}^N S_{ij} S_{kj} \frac{\partial^2}{\partial \epsilon_i \partial \epsilon_k} \\ & \quad - \frac{\Omega^{-1}}{3!} \sum_{i,k,l=1}^N S_{ij} S_{kj} S_{lj} \frac{\partial^3}{\partial \epsilon_i \partial \epsilon_k \partial \epsilon_l} + O(\Omega^{-3/2}) \end{aligned} \tag{S11}$$

Combining S8-S11, and ignoring third derivatives (which vanish for first and second moments) we obtain,

$$\begin{aligned} & \frac{\partial \Pi(\vec{\epsilon}, t)}{\partial t} - \Omega^{\frac{1}{2}} \sum_{i=1}^N \frac{\partial x_i}{\partial t} \frac{\partial \Pi(\vec{\epsilon}, t)}{\partial \epsilon_i} \\ &= \Omega \sum_{j=1}^R \left(-\Omega^{-1/2} \sum_i S_{ij} \frac{\partial}{\partial \epsilon_i} + \frac{\Omega^{-1}}{2} \sum_{i,k=1}^N S_{ij} S_{kj} \frac{\partial^2}{\partial \epsilon_i \partial \epsilon_k} \right) \left(f_j(\vec{x}) \right. \\ & \quad \left. + \Omega^{-1/2} \sum_i \frac{\partial f_j(\vec{x})}{\partial x_i} \epsilon_i + \frac{\Omega^{-1}}{2} \sum_{i,j=1}^N \frac{\partial^2 f_j(\vec{x})}{\partial x_i \partial x_j} \epsilon_i \epsilon_j \right) \Pi(\vec{\epsilon}, t) \end{aligned}$$

Noting that terms of order $O(\Omega^{1/2})$ cancel, after some algebra we arrive at the following multivariate master equation,

$$\begin{aligned}
& \frac{\partial \Pi(\vec{\epsilon}, t)}{\partial t} \\
&= - \sum_{i,k}^N J_{ik} \frac{\partial(\epsilon_k \Pi)}{\partial \epsilon_i} + \frac{1}{2} \sum_{i,k=1}^N D_{ik} \frac{\partial^2 \Pi}{\partial \epsilon_i \partial \epsilon_k} \\
&+ \frac{\Omega^{-1/2}}{2} \left(- \sum_{i,k,l=1}^N \frac{\partial J_{ik}}{\partial x_l} \frac{\partial(\epsilon_l \epsilon_k \Pi)}{\partial \epsilon_i} + \sum_{i,k,l=1}^N \frac{\partial D_{ik}}{\partial x_l} \frac{\partial^2(\epsilon_l \Pi)}{\partial \epsilon_i \partial \epsilon_k} \right) + o(\Omega^{-1})
\end{aligned}$$

1267

1268

1269 where $\Pi = \Pi(\vec{\epsilon}, t)$, and

1270

$$J_{ik} = \sum_{j=1}^R S_{ij} \frac{\partial f_j(\vec{x})}{\partial x_k}$$

1271

$$D_{ik} = \sum_{j=1}^R S_{ij} S_{kj} f_j(\vec{x})$$

1272 The first two terms of S12 (of order Ω^0) together give the standard linear Fokker-Planck
 1273 equation which is the sum of directional motion (e.g., convection, transport, drift) and
 1274 diffusion, respectively. The terms of order $\Omega^{-1/2}$ are the nonlinear, finite volume
 1275 corrections to these terms, which vanish for linear reactions and/or reactions in large
 1276 volumes (i.e., in the macroscopic limit).

1277 The stationary solution of the linear Fokker-Planck equation is a Gaussian
 1278 distribution. We expect that if the relative fluctuation size is small compared to the
 1279 curvature of the transition rate, then the stationary solution of S12 remains approximately
 1280 Gaussian even in finite volumes with non-linear transition rates. Of course, the actual
 1281 product distribution of a reaction with a hyperbolic rate law is positively skewed (and a
 1282 sigmoidal rate law is leptokurtic). Assuming relatively small fluctuations, only the first
 1283 two moments of $\Pi(\vec{\epsilon}, t)$ are needed to approximately characterize the full distribution in
 1284 small volumes.

1285 The time derivative of the first moment of the i th species is found by multiplying
 1286 both sides of S12 by ϵ_i and integrating over all ϵ :

$$\begin{aligned}
1287 \quad & \int \epsilon_i \frac{\partial \Pi(\vec{\epsilon}, t)}{\partial t} = \frac{\partial \langle \epsilon_i \rangle}{\partial t} \\
1288 \quad & = - \int \epsilon_i \sum_k^N J_{ik} \frac{\partial (\epsilon_k \Pi)}{\partial \epsilon_i} d\epsilon_i + \frac{1}{2} \int \epsilon_i \sum_{k=1}^N D_{ik} \frac{\partial^2 \Pi}{\partial \epsilon_i \partial \epsilon_k} d\epsilon_i \\
1289 \quad & - \frac{\Omega^{-1/2}}{2} \left(\int \epsilon_i \sum_{k,l=1}^N \frac{\partial J_{ik}}{\partial x_l} \frac{\partial (\epsilon_l \epsilon_k \Pi)}{\partial \epsilon_i} d\epsilon_i - \int \epsilon_i \sum_{k,l=1}^N \frac{\partial D_{ik}}{\partial x_l} \frac{\partial^2 (\epsilon_l \Pi)}{\partial \epsilon_i \partial \epsilon_k} d\epsilon_i \right)
\end{aligned}$$

1290 This is straightforwardly computed using integration by parts (iteratively for the second
1291 derivative terms) and applying the definitions of the mean and variance, leading to:

$$\frac{\partial \langle \epsilon_i \rangle}{\partial t} = - \sum_k^N J_{ik} \langle \epsilon_k \rangle + \frac{\Omega^{-1/2}}{2} \sum_{k,l=1}^N \frac{\partial J_{ik}}{\partial x_l} \langle \epsilon_l \epsilon_k \rangle + O(\Omega^{-1}) \quad S13$$

1292
1293 which has matrix form:

$$\frac{\partial \langle \vec{\epsilon} \rangle}{\partial t} = -\mathbf{J} \cdot \langle \vec{\epsilon} \rangle + \Omega^{-\frac{1}{2}} \vec{\theta} + O(\Omega^{-1}) \quad S14$$

1294
1295 where $\vec{\theta}$ is a vector with k^{th} element:

$$1296 \quad \theta_k = \frac{1}{2} \sum_{i,j}^N \frac{\partial J_{ki}}{\partial x_j} \langle \epsilon_i \epsilon_j \rangle$$

1297 Substituting $\vec{\epsilon} = \Omega^{1/2}(\vec{x} - \vec{\bar{x}})$, setting S14 equal to zero and solving for the average
1298 concentration vector, $\langle \vec{x} \rangle$, finally gives the matrix equation for first moments of the
1299 steady-state system:

$$\langle \vec{x} \rangle = \vec{\bar{x}} + \Omega^{-1} \mathbf{J}^{-1} \cdot \vec{\theta} \quad S15$$

1300 Thus, one can solve for the average of all species in a discrete, stochastic system with
1301 finite volume directly from the continuous, deterministic system of equations for the
1302 system.

1303 One solves the second moments of the system using the same approach. One
1304 multiplies S12 by ϵ_i^2 integrates (again, using integration by parts), then solves for
1305 $\langle \epsilon_i^2 \rangle$ when the distribution is stationary ($\frac{d\langle \epsilon_i^2 \rangle}{dt} = 0$). The full solution is a function of

1306 the first three moments of the joint distribution, such that the moment hierarchy does not
 1307 close. Further, solving the second moments of the linear multivariate Fokker-Planck
 1308 equation is already computationally intensive, without consider higher order corrections.
 1309 Thus, we will ignore the non-linear finite volume correction to the second moments for
 1310 the purposes of the present paper. Simulation results suggest that this leads to small error
 1311 (Fig. 2, 3 main text). Therefore, we can solve for the second moments by solving the
 1312 Lyapunov Equation (7):

$$0 = \mathbf{J} \cdot \mathbf{C} + \mathbf{C} \cdot \mathbf{J}^T + \mathbf{D} + O(\Omega^{-\frac{1}{2}}) \quad \text{S16}$$

1313
 1314 for the covariance matrix \mathbf{C} which has elements,
 1315 $C_{ij} = \langle \epsilon_i \epsilon_j \rangle$
 1316 This can be solved symbolically using the *Mathematica* (Wolfram Research) command:
 1317 `LyapunovSolve[\mathbf{J} , $-\mathbf{D}$]`. \mathbf{D} is the noise-generating diffusion matrix given by,

$$1318 \quad \mathbf{D} = \mathbf{S} \cdot \mathbf{R} \cdot \mathbf{S}^T$$

1319 \mathbf{S} is the stoichiometric matrix of the network and \mathbf{R} is the matrix of the macroscopic rate
 1320 function.

1321

1322 **Comparison to Grima (2010):**

1323 Grima (8) previously derived results for a multivariate system size expansion to order
 1324 $O(\Omega^{-1})$, as we have just done. However, our final equations differ in two ways, with both
 1325 quantitative and qualitative disagreement. Grima's master equation does not include
 1326 finite volume non-linear corrections to the diffusion term, which we include here only for
 1327 the sake of generality and completeness, although it disappears for first moments and we
 1328 neglect it for our computation of second moments. However, the following difference
 1329 prevents Grima's method from being applied accurately to a QSSA reduced system:

1330 Grima's (8) first moments are solved using:

$$1331 \quad \langle \vec{x} \rangle = \vec{\bar{x}} + \Omega^{-1} \mathbf{J}^{-1} \cdot \vec{\bar{\Delta}}$$

1332 where,

1333

$$\Delta_k = \frac{1}{2} \left(\sum_{i,j}^N \frac{\partial J_{ki}}{\partial \phi_j} \langle \epsilon_i \epsilon_j \rangle - \sum_i^N \phi_i \frac{\partial J_{ki}}{\partial \phi_i} \right)$$

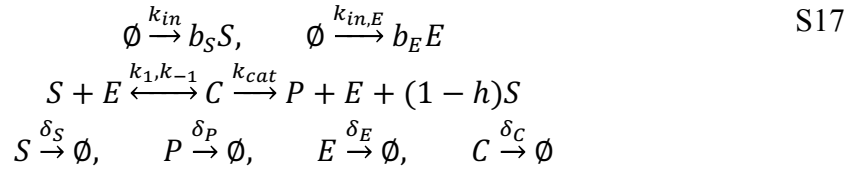
1334

Compare this Δ to our θ above. Grima's expression is valid when applied to all systems of elementary reactions, where the second term in Δ_k accounts for homodimerization reactions (since homodimerization rates are proportional to $n(n-1)$, not n^2). However, Grima's expression does not work when applied to reduced (e.g., "elementary complex", pseudo-first order, biochemical) reactions. As a trade-off, the present approach is inaccurate when applied to homodimerization reactions in small volumes.

1340

1341 **I.b.ii. Application to a general enzyme-catalyzed reaction in an open system**

1342 We consider the following generalized MM reaction scheme:



1343

1344 This scheme generalizes the three cases of interest. The parameter h is an indicator
1345 variable: $h = 1$ in reactions where S is consumed (as with most metabolic reactions) and h
1346 $= 0$ for template-mediated reactions (such as when S is mRNA (i.e., translation) or a gene
1347 (i.e., transcription)). When $\delta_S = 0$ the only output of S is via the reaction catalyzed by E ;
1348 this corresponds to the case modeled by Grima (8) who found that noise does not alter
1349 pathway flux. This scheme corresponds to the following system of ordinary differential
1350 equations (ODE's):

$$\begin{aligned} \frac{ds}{dt} &= b_S k_{in} - k_1 s e_0 + ((1-h)k_{cat} + k_{-1})c - \delta_S s & S18 \\ \frac{de}{dt} &= b_E k_{in,E} - k_1 s e_0 + (k_{cat} + k_{-1})c - \delta_E e \\ \frac{dc}{dt} &= k_1 s e_0 - (k_{cat} + k_{-1})c \\ \frac{dp}{dt} &= k_{cat} c - \delta_P p \end{aligned}$$

1351 Following convention, lower case s , e , and p denote the concentration of species S , E and
 1352 P , respectively. The subscript on e_0 is used to make clear that this refers to “free” (unbound)
 1353 enzyme concentration. Assuming that c equilibrates fast relative to the rest of the reaction,
 1354 we can replace c with its steady-state solution which leads to the following QSSA reduction
 1355 of the system,

$$\begin{aligned} \frac{ds}{dt} &= b_S k_{in,S} - hV - \delta_S s & \text{S19} \\ \frac{de}{dt} &= b_E k_{in,E} - \delta_E e \\ \frac{dp}{dt} &= V - \delta_P p \end{aligned}$$

1356 where,

$$1357 \quad V = k_{cat} e \frac{s}{K_m + s}$$

1358 and where $e = e_0 + c$ is the total concentration of enzyme.

1359 This system has corresponding Jacobian matrix,

$$1360 \quad \mathbf{J} = \begin{pmatrix} -h \left(\frac{k_{cat} e}{K_m + s} \right) \left(1 - \frac{s}{K_m + s} \right) - \delta_S & -h \left(\frac{k_{cat} s}{K_m + s} \right) & 0 \\ 0 & -\delta_E & 0 \\ \left(\frac{k_{cat} e}{K_m + s} \right) \left(1 - \frac{s}{K_m + s} \right) & \frac{k_{cat} s}{K_m + s} & -\delta_P \end{pmatrix}$$

1361

1362 rate matrix \mathbf{R} ,

$$1363 \quad \mathbf{R} = \begin{pmatrix} k_{in,S} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & k_{in,E} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & V & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & V & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \delta_S s & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \delta_E e & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \delta_P p \end{pmatrix}$$

1364 and stoichiometric matrix,

$$1365 \quad \mathbf{S} = \begin{pmatrix} b_S & 0 & -h & 0 & -1 & 0 & 0 \\ 0 & b_E & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & -1 \end{pmatrix}$$

1366

1367 The noise-generating diffusion matrix is then,

$$1368 \quad \mathbf{D} = \mathbf{S} \cdot \mathbf{R} \cdot \mathbf{S}^T = \begin{pmatrix} b_S^2 k_{in,S} + hV + \delta_S s & 0 & 0 \\ 0 & b_E^2 k_{in,E} + \delta_E e & 0 \\ 0 & 0 & V + \delta_P \end{pmatrix}$$

1369 Finally, the θ vector equals,

$$1370 \quad \vec{\theta}$$

$$1371 \quad = \left(h \left(\frac{k_{cat} K_m}{(K_m + \bar{s})^2} \right) \left(\langle e \rangle \frac{\langle \epsilon_S^2 \rangle}{K_m + \bar{s}} - \langle \epsilon_S \epsilon_E \rangle \right) \quad 0 \quad - \left(\frac{k_{cat} K_m}{(K_m + \bar{s})^2} \right) \left(\langle e \rangle \frac{\langle \epsilon_S^2 \rangle}{K_m + \bar{s}} - \langle \epsilon_S \epsilon_E \rangle \right) \right)^T$$

1372

1373 Substituting \mathbf{J} and $\vec{\theta}$ into Eqn S15 gives the noise-induced deviation of the steady-
1374 state concentrations from the macroscopic prediction:

$$1375 \quad \langle \bar{x} \rangle - \bar{x}$$

$$1376 \quad = \left(h \left(\frac{k_{cat} K_m}{\Omega(h\bar{e}k_{cat}K_m + \delta_S(K_m + \bar{s})^2)} \right) \left(\bar{e} \frac{\langle \epsilon_S^2 \rangle}{K_m + \bar{s}} - \langle \epsilon_S \epsilon_E \rangle \right), \quad 0, \quad -\delta_S \left(\frac{k_{cat} K_m}{\Omega\delta_P(h\bar{e}k_{cat}K_m + \delta_S(K_m + \bar{s})^2)} \right) \right)^T$$

1377 For translation ($h = 0$) the vector of steady-state concentrations simplifies to:

$$1378 \quad \langle \bar{x} \rangle = \begin{pmatrix} \bar{s} & \bar{e} & \bar{p} - \underbrace{\frac{\langle \epsilon_S^2 \rangle}{\delta_P(K_m + \bar{s})^3} k_{cat} \bar{e} K_m}_{\text{cost of noise}} \Omega^{-1} \end{pmatrix}^T$$

$$1379 \quad = \begin{pmatrix} \bar{s} & \bar{e} & \bar{p} \left(1 - \frac{F_S K_m}{(K_m + \bar{s})^2} \Omega^{-1} \right) \end{pmatrix}^T$$

$$1380 \quad = (\bar{s} \quad \bar{e} \quad \bar{p}(1 - \eta_S^2 a(1 - a)\Omega^{-1}))^T$$

1381

1382 where $F_S = \langle \epsilon_S^2 \rangle / \bar{s}$ is the Fano factor of the substrate (i.e., mRNA) concentration, which
1383 equals $\frac{1}{2}(1+b)$, where b is the burst size of mRNA.

1384 Four things to note from these results. First, the steady-state enzyme
1385 concentration (the second term in each vector) is not affected by noise (it has a deviation
1386 equal to 0). This is expected since de/dt has no non-linear terms. Second, if the substrate
1387 is not consumed by the reaction ($h = 0$), then the steady-state substrate concentration is
1388 unaffected by noise (the first term in the deviation vector equals 0 when $h = 0$). Again,
1389 this is as expected since the substrate concentration when not consumed follows a simple
1390 linear birth-death process, just like the enzyme. Third, noise alters the steady-state

concentration of product (i.e. the reaction flux) only if the substrate is diluted (since the third entry in the deviation vector is proportional to δ_S) or if the reaction is template-mediated ($h = 0$). This is also expected: as described in the Main Text, when $\delta_S = 0$, stochastic focusing, the increase in substrate concentration due to noise, completely buffers the system from noise-induced loss of reaction flux. Finally, the cost of noise is proportional to either the Fano factor (a.k.a. the noise strength) or the squared coefficient of variation (a.k.a. the noise) of the substrate.

I.c. Gene expression noise

The steady-state second moments are given by the covariance matrix, **C**. Many of these moments are too unwieldy to reproduce here. A *Mathematica* (Wolfram Research) notebook with the solution is provided as Supplementary Material. Three of the more compact moments are given here:

$$\begin{aligned}\langle \epsilon_E^2 \rangle &= \frac{1}{2} \bar{e}(1 + b_E) \quad \text{S20} \\ \langle \epsilon_S^2 \rangle &= \frac{(b_S^2 k_{in} + \delta_S \bar{s})(\delta_E + \delta_S)(K_m + \bar{s})^4}{2(hK_m k_{cat} \bar{e} + (\delta_E + \delta_S)(K_m + \bar{s})^2)(hK_m k_{cat} \bar{e} + \delta_S(K_m + \bar{s})^2)} \\ &\quad + h \frac{k_{cat}(K_m + \bar{s})G}{2(hK_m k_{cat} \bar{e} + (\delta_E + \delta_S)(K_m + \bar{s})^2)(hK_m k_{cat} \bar{e} + \delta_S(K_m + \bar{s})^2)}, \\ \langle \epsilon_S \epsilon_E \rangle &= -h \langle \epsilon_E^2 \rangle \frac{k_{cat} \bar{s}}{h a k_{cat} \bar{e} + (\delta_E + \delta_P)(K_m + \bar{s})}, \\ &\quad G \\ &= [2\langle \epsilon_E^2 \rangle k_{cat}(K_m + \bar{s}) + (b_S^2 k_{in} + \delta_S \bar{s})K_m \bar{e}(K_m + \bar{s}) \\ &\quad + \bar{e} \bar{s}(K_m k_{cat} \bar{e} + (\delta_E + \delta_S)(K_m + \bar{s})^2)]\end{aligned}$$

For the special case of template-mediated reactions ($h = 0$), such as translation, we find:

$$\begin{aligned}\langle \epsilon_E^2 \rangle &= \frac{1}{2} \bar{e}(1 + b_E) \\ \langle \epsilon_S^2 \rangle &= \frac{1}{2} \bar{s}(1 + b_S) \\ \langle \epsilon_S \epsilon_E \rangle &= 0\end{aligned}$$

1410 For the case when $h = 0$, such as translational, the variance of product (e.g., protein)
 1411 number is found to be:

$$\begin{aligned} \langle \epsilon_P^2 \rangle &= \bar{p} + \frac{k_{cat} \bar{e} \bar{s}}{2\delta_P(K_m + \bar{s})} + \frac{b_S^2 k_{cat}^2 \bar{e}^2 k_{in,S} K_m^2}{2\delta_S \delta_P (\delta_S + \delta_P)(K_m + \bar{s})^4} \\ &+ \frac{k_{cat}^2 \bar{e}^2 \bar{s} K_m^2}{2\delta_P (\delta_S + \delta_P)(K_m + \bar{s})^4} + \frac{k_{cat}^2 \bar{s}^2 (\delta_E \bar{e} + k_{in,E} b_E^2)}{2\delta_E \delta_P (\delta_E + \delta_P)(K_m + \bar{s})^2} \end{aligned} \quad S21$$

1412
 1413 This is further simplified by noting the following:

$$\bar{p} = \frac{k_{cat} \bar{e} \bar{s}}{\delta_P (K_m + \bar{s})}$$

$$\bar{s} = \frac{b_S k_{in,S}}{\delta_S}$$

$$\bar{r} = \frac{b_E k_{in,E}}{\delta_E}$$

$$\langle \epsilon_E^2 \rangle = \frac{(\delta_R \bar{e} + k_{in,E} b_E^2)}{2\delta_E}$$

1418 Making these substitutions and simplifying gives,

$$\langle \epsilon_P^2 \rangle = \bar{p} + \bar{p} \frac{k_{cat} \bar{e} K_m^2}{2(\delta_S + \delta_P)(K_m + \bar{s})^3} (b + 1) + \langle \epsilon_E^2 \rangle \frac{k_{cat}^2 \bar{s}^2}{\delta_P (\delta_E + \delta_P)(K_m + \bar{s})^2}$$

1420

1421 Assuming that $\delta_S \gg \delta_P$ and that $\delta_E = \delta_P$,

$$\langle \epsilon_P^2 \rangle \cong \bar{p} + \bar{p} \frac{k_{cat} \bar{e} K_m^2}{2\delta_S (K_m + \bar{s})^3} (b + 1) + \langle \epsilon_E^2 \rangle \frac{k_{cat}^2 \bar{s}^2}{2\delta_P^2 (K_m + \bar{s})^2}$$

1423

1424 Noting that,

$$\frac{k_{cat}^2 \bar{s}^2}{\delta_P^2 (K_m + \bar{s})^2} = \left(\frac{\bar{p}}{\bar{e}} \right)^2$$

1426 and defining the burst parameter as the number of product molecules produced in the
 1427 lifetime of a single substrate molecule,

$$b_T \equiv \frac{k_{cat}\bar{e}}{\delta_S(K_m + \bar{s})}$$

and using the definition of a and η_R^2 from the main text and $\langle\epsilon_S^2\rangle$ above, finally leads to,

$$\langle\epsilon_P^2\rangle \cong \bar{p}(1 + a^2 b_T F_S + \eta_E^2 \bar{p}/2)$$

where F_S is the substrate Fano factor, $\langle\epsilon_S^2\rangle/\bar{s}$.

Note that for translation with non-limiting amino acid abundance, fluctuations in ribosome concentration, $\langle\epsilon_E^2\rangle$, do not influence the the statistics mRNA and do not influence the mean protein abundance. Ribosome fluctuations do, however, alter the variance of protein abundance, $\langle\epsilon_P^2\rangle$. This latter effect constitutes one source of “extrinsic noise” (9, 10) in gene expression.

I.d. The cost of transcriptional bursting

Substituting these statistics into the steady-state product concentration gives an expression for the effect of transcriptional bursting, b_t , on protein abundance,

$$\begin{aligned} \langle p \rangle &= \bar{p} - (b_t + 1) \frac{k_{cat}\bar{e}\bar{s}K_m}{2\Omega\delta_P(K_m + \bar{s})^3} \\ &= \bar{p} \left(1 - \frac{(b_t + 1)K_m}{2\Omega(K_m + \bar{s})^2} \right) \end{aligned}$$

and the steady-state rate of protein synthesis ($\langle V \rangle = \langle p \rangle / \delta_P$)

$$\begin{aligned} \langle V \rangle &= \bar{V} - \underbrace{(b_t + 1) \frac{k_{cat}\bar{e}\bar{s}K_m}{2\Omega(K_m + \bar{s})^3}}_{\text{cost of transcriptional bursting}} \\ &= \bar{V} \left(1 - \frac{(b_t + 1)K_m}{2\Omega(K_m + \bar{s})^2} \right) \end{aligned}$$

II.a. Response of protein noise to evolutionary changes in K_m and s .

Protein noise strength ($F_{protein}$) can be adjusted in two ways: modifying K_m or s . We now show that these strategies are not equivalent. Specifically, modifying K_m results in a unimodal change in $F_{protein}$, whereas modifying s has a monotonic effect. Recall from above,

$$F_{protein} \cong 1 + a^2 b_T F_{mRNA}$$

where the kinetic order parameter, a , is defined as (e.g., (11, 12)),

$$a_{ij} \equiv \left(\frac{\partial f_i}{\partial x_j} \right) \left(\frac{x_j}{f_i} \right) = \frac{\partial \ln(f_i)}{\partial \ln(x_j)}$$

for the reaction rate f and concentration x . For Michaelis-Menten kinetics,

$$a = \frac{K_m}{K_m + \bar{s}}$$

The reaction order, a , is equivalent to the fraction of unbound catalyst.

Evolutionary or synthetic modification of K_m will alter protein noise according to,

$$\frac{\partial F_{protein}}{\partial K_m} = (2a - 3a^2) \frac{k_{cat} \bar{e} F_{mRNA}}{\delta_s (K_m + \bar{s})^2}$$

which has an interior root at,

$$a^* = 2/3$$

for all parameter values, which is also the maximum. Alternatively,

$$\frac{\partial F_{protein}}{\partial \bar{s}} = -3a^2 \frac{k_{cat} \bar{e} F_{mRNA}}{\delta_s (K_m + \bar{s})^2}$$

which is a monotonically decreasing function of a . Thus, ameliorating noise by increasing s is always a good strategy, all else equal, whereas the direction of evolution of binding affinity depends on initial conditions. Note that $a=2/3$ corresponds with weak binding affinity of an enzyme. Indeed, ribosome binding affinity ranges from $a = 0.1-0.3$, such that enhancing binding affinity will reduce protein noise.

II.b. Absolute fitness cost of noise

We analyze a coarse-grained cell model where cell growth relies on import of extracellular nutrient, s_E , by nutrient import proteins, p_U , for use in translation of mRNA, m , by ribosomes, r . Translation initiation involves the reversible binding of mRNA to ribosomes. Then elongation proceeds with the sequential addition of n intracellular nutrients, s_I , which we define as the intracellular nutrient most limiting to translational elongation (either ATP or charged amino acids). posit that fitness, the rate of new biomass synthesis, V , obeys the following QSSA reduced system of biochemical equations:

$$\frac{ds_E}{dt} = b_{sE} k_{in,SE} - \delta_{sE} s_E \quad S22$$

$$\begin{aligned}\frac{ds_I}{dt} &= V_U - \delta_{s_I}s_I \\ \frac{dm}{dt} &= b_t k_t - \delta_M m \\ \frac{dp_U}{dt} &= \phi V_T - \delta_{p_U} p_U \\ \frac{dp}{dt} &= V_T - \delta_P p\end{aligned}$$

1478

1479 where,

$$V_T = k_T r \frac{m s_I}{K_1 + K_2 s_I + K_3 m + s_I m} \quad \text{S23}$$

1480

1481

1482 V_T is the rate law for a compulsory-order (mRNA first, then s_I second) , two substrate
 1483 (mRNA and s_I) enzymatic reaction, where ribosomes are the enzyme. The parameter ϕ is
 1484 the fraction of total mRNA, m , consisting of nutrient importer protein p_U transcripts.
 1485 In practice, the extracellular nutrient supply, s_E , is much easier to measure and
 1486 experimentally manipulate than the intracellular nutrient supply, s_I . It is therefore
 1487 desirable to rewrite Eqn. S23 as a Monod Growth Equation mapping extracellular
 1488 nutrient supply onto growth rate (13–17). Assuming that nutrient uptake obeys MM
 1489 kinetics, such that in steady-state, $s_I = k_U p_U s_E / (K_S + s_E) \delta_S$, where k_U is the catalytic
 1490 constant for uptake, K_S is the uptake half-saturation constant, p_U is the concentration of
 1491 uptake proteins, and δ_{s_I} is the dilution rate of s_I . We further assume the simplest case
 1492 where extracellular nutrient is used directly for biosynthesis without modification.
 1493 Substituting into Eqn. S23 and rearranging gives the following Monod Growth equation:

$$\frac{dp}{dt} = \frac{V_{max} s_E}{K_M + s_E} - \lambda p \quad \text{S24}$$

1494 where,

$$V_{max} = \frac{k_T r k_U p_U}{k_U p_U + \delta_{s_I} K_3} \left(\frac{m}{K_A + m} \right)$$

and where $K_A = (\delta_{s_I}K_1 + k_U p_U K_2)/(k_U p_U + \delta_{s_I}K_3)$ and $K_M = K_s \delta_{s_I} (K_1 + K_3 m)/[\delta_{s_I} (K_1 + K_3 m) + k_U p_U (K_2 + m)]$. The practical and conceptual utility of the Monod equation is that it neatly disentangles cellular and environmental determinants of growth: V_{max} and K_M are intrinsic to the cell, under genetic control, and evolvable, whereas s_E is extrinsic to the cell, not under genetic control and not evolvable. The version derived here unpacks the two fitting parameters, V_{max} and K_M , into their mechanistic components, revealing them both to be jointly determined by translation and uptake kinetics. Faster uptake (higher $k_U p_U$) increases the maximum growth rate (V_{max}) and lowers the half-saturation constants, K_M and K_A , making growth faster at both high and low nutrient densities. Interestingly, this predicts an uptake-mediated negative correlation between V_{max} and K_M , which is contrary to the common view that V_{max} and K_M are positively correlated and thus in a trade-off conflict.

II.c. Selection strength

Absolute fitness has the following generic form if we ignore extrinsic noise:

$$\langle \lambda \rangle \cong \frac{\bar{V}}{p_{ss}} \left(1 - \sum_i^n A_i F_i \Omega^{-1} \right)$$

where F_i is the Fano factor for substrate i , A_i is a constant, and n is the number of substrates affecting fitness. The selection gradient on a mutation that alters fitness is,

$$s = \frac{\langle \lambda \rangle_m - \langle \lambda \rangle_r}{\langle \lambda \rangle} \cong \frac{\sum_i^n (A_{r,i} F_{r,i} - F_{m,i} A_{m,i})}{1 - \sum_i^n (f A_{r,i} F_{r,i} + (1 - f) F_{m,i} A_{m,i}) \Omega^{-1}} \Omega^{-1}$$

where f is the population relative frequency of resident alleles. Because A is typically small (in the standard MM reaction $A = K_m / (K_m + \bar{s})^2$), the denominator approximately equals 1 provided that F is not too large. For example, for $K_m = 350$ and $s = 1400$ (values corresponding to translation of mRNA's in *E. coli*), this assumption leads to only a 1% error for F as large as 100 and a 10% error for $F = 1000$. These F values are very high, and beyond any value considered in the paper. In addition, as Ω increases this error becomes even smaller. We therefore take this assumption as quite reasonable for the situations considered in our paper. If we further assume that genotypes differ only in

noise (i.e., that there is no pleiotropic effect of the mutation), and define ΔF_i as the difference in Fano factors between resident (r) and mutant (m) genotypes, then we have,

$$s \cong \sum_i^n (\Delta F_i A_i) \Omega^{-1}$$

Supporting References

1. Gillespie DT (1977) Exact stochastic simulation of coupled chemical reactions. *J Phys Chem* 81(25):2340–2361.
2. Shapiro BE, Levchenko A, Meyerowitz EM, Wold BJ, Mjolsness ED (2003) Cellerator: extending a computer algebra system to include biochemical arrows for signal transduction simulations. *Bioinformatics* 19(5):677–678.
3. Michaelis L, Menten ML (1913) Die kinetik der invertinwirkung. *Biochem z* 49(333-369):352.
4. Briggs GE, Haldane JBS (1925) A note on the kinetics of enzyme action. *Biochem J* 19(2):338.
5. Ingalls BP (2013) *Mathematical Modeling in Systems Biology: An Introduction* (The MIT Press, Cambridge, Massachusetts).
6. van Kampen NG (2007) *Stochastic Processes in Physics and Chemistry* (Elsevier, New York). Third Edit.
7. Elf J, Ehrenberg M (2003) Fast evaluation of fluctuations in biochemical networks with the linear noise approximation. *Genome Res* 13(11):2475–2484.
8. Grima R (2010) Intrinsic biochemical noise in crowded intracellular conditions. *J Chem Phys* 132(18). doi:Artn 18510210.1063/1.3427244.
9. Swain PS, Elowitz MB, Siggia ED (2002) Intrinsic and extrinsic contributions to stochasticity in gene expression. *Proc Natl Acad Sci* 99(20):12795–12800.
10. Raser JM, O’Shea EK (2005) Noise in Gene Expression: Origins, Consequences, and Control. *Science* (80-) 309(5743):2010–2013.

- 1550 11. Savageau MA, Voit EO, Irvine DH (1987) Biochemical systems theory and
1551 metabolic control theory: 1. Fundamental similarities and differences. *Math Biosci*
1552 86(2):127–145.
- 1553 12. Kacser H and, Burns J (1973) The control of flux. *Symp. Soc. Exp. Biol.*, pp 65–
1554 104.
- 1555 13. Tilman D (1982) *Resource Competition and Community Structure* (Princeton
1556 University Press, Princeton, NJ).
- 1557 14. Scott M, Hwa T (2011) Bacterial growth laws and their applications. *Curr Opin*
1558 *Biotechnol* 22(4):559–565.
- 1559 15. Scott M, Gunderson CW, Mateescu EM, Zhang Z, Hwa T (2010) Interdependence
1560 of cell growth and gene expression: origins and consequences. *Science* (80-)
1561 330(6007):1099–1102.
- 1562 16. Scott M, Klumpp S, Mateescu EM, Hwa T (2014) Emergence of robust growth
1563 laws from optimal regulation of ribosome synthesis. *Mol Syst Biol* 10(8):747.
- 1564 17. Monod J (1949) The growth of bacterial cultures. *Annu Rev Microbiol* 3:371–394.

1565

1566

1567

1568

1569

1570

1571

1572

1573 **Supplementary Figures**

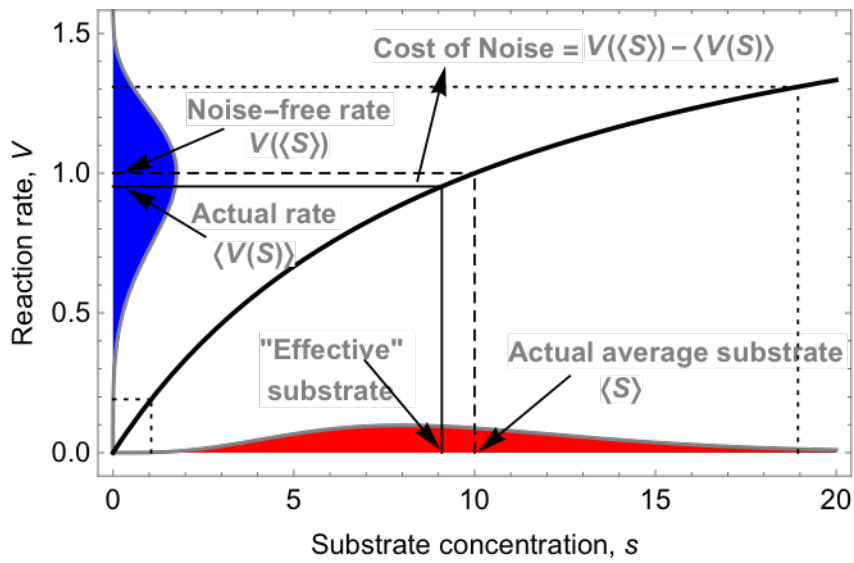


Figure S1. Schematic demonstrating the geometry of noise-induced slowdown for hyperbolic reaction.

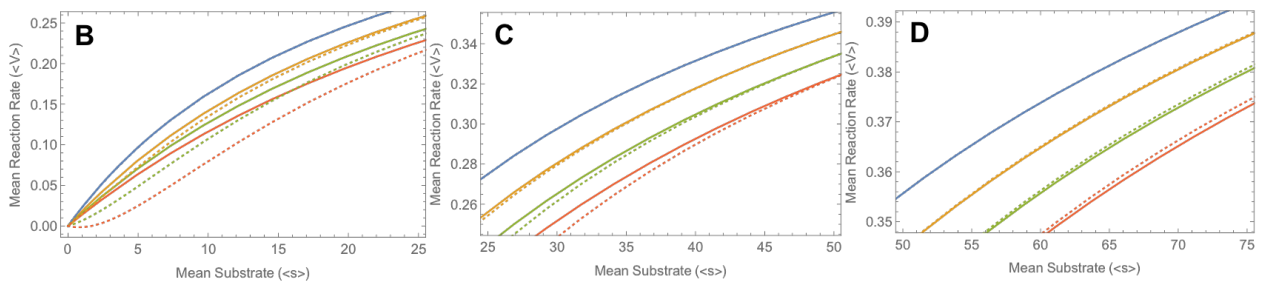
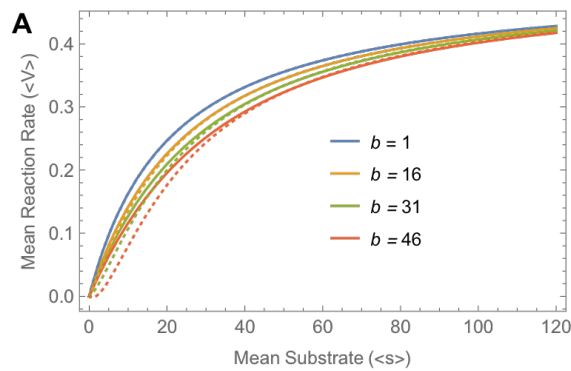


Figure S2. Comparison of exact and approximate predictions for noise-induced slowdown.