

University of Groningen

## Dynamic disorder in simple enzymatic reactions induces stochastic amplification of substrate

Gupta, Ankit; Miliias-Argeitis, Andreas; Khammash, Mustafa

*Published in:*  
Journal of the Royal Society Interface

*DOI:*  
[10.1098/rsif.2017.0311](https://doi.org/10.1098/rsif.2017.0311)

**IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.**

*Document Version*  
Publisher's PDF, also known as Version of record

*Publication date:*  
2017

[Link to publication in University of Groningen/UMCG research database](#)

*Citation for published version (APA):*

Gupta, A., Miliias-Argeitis, A., & Khammash, M. (2017). Dynamic disorder in simple enzymatic reactions induces stochastic amplification of substrate. *Journal of the Royal Society Interface*, 14(132), [20170311]. <https://doi.org/10.1098/rsif.2017.0311>

**Copyright**

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

**Take-down policy**

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

*Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.*

## Research



**Cite this article:** Gupta A, Miliias-Argeitis A, Khammash M. 2017 Dynamic disorder in simple enzymatic reactions induces stochastic amplification of substrate. *J. R. Soc. Interface* **14**: 20170311.

<http://dx.doi.org/10.1098/rsif.2017.0311>

Received: 27 April 2017

Accepted: 5 July 2017

**Subject Category:**

Life Sciences – Mathematics interface

**Subject Areas:**

systems biology

**Keywords:**

stochastic chemical kinetics, dynamic disorder, continuous-time Markov chains

**Author for correspondence:**

Mustafa Khammash

e-mail: [mustafa.khammash@bsse.ethz.ch](mailto:mustafa.khammash@bsse.ethz.ch)

Electronic supplementary material is available online at <https://dx.doi.org/10.6084/m9.figshare.c.3827920>.

# Dynamic disorder in simple enzymatic reactions induces stochastic amplification of substrate

Ankit Gupta<sup>1</sup>, Andreas Miliias-Argeitis<sup>1,2</sup> and Mustafa Khammash<sup>1</sup>

<sup>1</sup>Department of Biosystems Science and Engineering, ETH Zurich, Mattenstrasse 26, 4058 Basel, Switzerland

<sup>2</sup>Groningen Biomolecular Sciences and Biotechnology Institute, University of Groningen, Nijenborgh 4, 9747 AG Groningen, The Netherlands

AG, 0000-0002-1054-6133; AM-A, 0000-0002-4291-4228; MK, 0000-0002-4855-9220

A growing amount of evidence over the last two decades points to the fact that many enzymes exhibit fluctuations in their catalytic activity, which are associated with conformational changes on a broad range of timescales. The experimental study of this phenomenon, termed dynamic disorder, has become possible thanks to advances in single-molecule enzymology measurement techniques, through which the catalytic activity of individual enzyme molecules can be tracked in time. The biological role and importance of these fluctuations in a system with a small number of enzymes, such as a living cell, have only recently started being explored. In this work, we examine a simple stochastic reaction system consisting of an inflowing substrate and an enzyme with a randomly fluctuating catalytic reaction rate that converts the substrate into an outflowing product. To describe analytically the effect of rate fluctuations on the average substrate abundance at steady state, we derive an explicit formula that connects the relative speed of enzymatic fluctuations with the mean substrate level. Under fairly general modelling assumptions, we demonstrate that the relative speed of rate fluctuations can have a dramatic effect on the mean substrate, and lead to large positive deviations from predictions based on the assumption of deterministic enzyme activity. Our results also establish an interesting connection between the amplification effect and the mixing properties of the Markov process describing the enzymatic activity fluctuations, which can be used to easily predict the fluctuation speed above which such deviations become negligible. As the techniques of single-molecule enzymology continuously evolve, it may soon be possible to study the stochastic phenomena due to enzymatic activity fluctuations within living cells. Our work can be used to formulate experimentally testable hypotheses regarding the nature and magnitude of these fluctuations, as well as their phenotypic consequences.

## 1. Introduction

First made almost two decades ago, observations of enzymatic turnovers for single enzyme molecules have allowed scientists to probe enzyme behaviour beyond the regime of high copy numbers and ensemble averages [1]. Thanks to the advances brought by experimental techniques such as single-molecule fluorescence spectroscopy [2,3], the field of single-molecule enzymology developed rapidly in the subsequent years. The key observation made possible by single-molecule assays is that the catalytic rates of single enzyme molecules often display very large dynamic fluctuations over timescales much longer than the typical reaction cycle times, most probably driven by slow (spontaneous or induced) transitions in conformation [1,4–6].

Around the time when the first single enzyme molecules were observed in action, the mathematical theory of *dynamic disorder* was introduced by Zwanzig [7], motivated by several observations of different physico-chemical processes

with a seemingly common underlying cause that boiled down to random fluctuations of key process properties. The phenomenon of dynamic disorder refers to fluctuations in enzymatic reaction rates that occur at a timescale that is either slower than or comparable with the reaction timescale [8]. These fluctuations are often caused by *slow* transitions in the conformational state of enzymes. The most simple example of dynamic disorder first considered by Zwanzig [7] involved a so-called ‘rate process controlled by passage through a fluctuating bottleneck’ [9]. In the language of chemical kinetics, it describes the removal of a substrate,  $S$ , from a system at rate  $\gamma(t)S(t)$ , where the time-varying rate  $\gamma(t)$  is a (typically Markovian) stochastic process (figure 1). As the speed of  $\gamma(t)$  fluctuations tends to zero, the reaction rate  $\gamma(t)$  becomes a random variable which does not change with time, and we transition to the regime of *static* disorder [7]. On the other hand, as the speed of  $\gamma(t)$  fluctuations tends to infinity, the dynamic disorder *vanishes* on the timescale of substrate kinetics, and we recover the classical case where the reaction rate  $\gamma(t)$  becomes a deterministic constant. Our goal in this paper is to investigate the effects of  $\gamma(t)$  fluctuations between these two extremities, where most realistic systems are likely to lie.

Thanks to the mathematical theory of dynamic disorder, stochastically fluctuating enzyme activities can be understood and studied within a consistent mathematical framework that can also generate testable experimental predictions [1,5,8,10–12]. Already in [7], it was observed that dynamically disordered systems can give rise to macroscopic observations that differ from those expected in the absence of disorder. In subsequent years, a large body of theoretical and computational work has examined various alternative enzymatic reaction schemes, mostly focusing on the enzyme dynamics itself, e.g. on the autocorrelation of fluctuations and the distribution of waiting times between turnover events [13]. On the other hand, dynamic disorder has been observed for several biologically relevant enzymes [14], suggesting that it is ubiquitous in the cellular context. Early work [5] had already noted that enzymatic fluctuations could play an important biological role in a system containing only a small number of enzyme molecules, as often happens within a living cell, and the recent *in vivo* observation of fluctuating enzymatic activity confirms this claim [15].

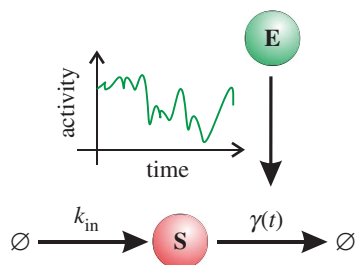
Besides studying the intrinsic mathematical properties of dynamically disordered enzymes, it would be also highly instructive and relevant for biology to examine the *consequences* of dynamic disorder on substrate statistics (a first example of such a study is given in [15]). Experimental work in this area is still done *in vitro* using constant and large substrate pools. Here on the contrary, we provide a mathematical treatment of how dynamic disorder alters the substrate mean abundance in the presence of substrate *inflow*, a condition closer to biological reality that has also been considered in [16,17]. To this end, we analytically examine a highly simplified stochastic system with a randomly fluctuating catalytic reaction rate and describe the effect of rate fluctuations on the average substrate abundance. Under fairly general conditions, we demonstrate that the relative speed of rate fluctuations can have a dramatic effect on the mean substrate, and lead to large positive deviations from predictions based on the assumption of deterministic enzyme activity. Using a Markovian model for enzyme kinetics, we mathematically characterize this effect by deriving

an explicit formula for the steady-state substrate-mean as a function of the relative speed of enzymatic fluctuations. From this formula, we show that for any finite speed value, the steady-state substrate-mean is *sandwiched* between the two values obtained in the static and the deterministic regimes. Furthermore, we demonstrate that the mapping between the relative speed of enzyme kinetics and the substrate-mean at steady state can be well approximated by a convex, monotonically decreasing function whose key shape parameter depends on the ‘mixing strength’ of the Markov process describing enzyme kinetics. This mixing strength can be measured by computing an appropriate *Dirichlet form* [18] of the Markov process. Even though we consider a highly simplified situation, our analysis can serve as a guide in the case of more realistic, but analytically intractable enzymatic reaction schemes. Our results depend *only* on the enzymatic fluctuations, but they do not depend on the fluctuations caused by the low abundance of substrate molecules [19], although we account for these fluctuations by modelling the substrate kinetics as a jump Markov chain. Indeed, the results we present remain unchanged even if we discard these fluctuations and describe the substrate kinetics as an ordinary differential equation (ODE) with a fluctuating rate constant  $\gamma(t)$ .

Enzymatic fluctuations can also arise from sources other than dynamic disorder. The abundance or the availability of enzyme molecules may also fluctuate due to gene expression noise [19,20], and their chances of finding substrate molecules can be diffusion-limited [21]. In this work, we do not distinguish between the various sources of fluctuations and model the aggregate enzymatic activity by a Markovian stochastic process.

The biological significance of our findings is manifold because enzymatic interactions are ubiquitous in cell biology and the effects of enzymatic noise in metabolic networks have only recently started being explored [20]. Using the relative speed of enzymatic fluctuations as parameter, our results provide a clear way to determine if the deterministic approximation is a faithful representation of reality. Our results can shed light on the timescale disparities that exist between enzyme and substrate kinetics. In particular, we see that enzyme kinetics needs to be ‘fast’ in order to avoid any undesirable amplification of the mean substrate abundance due to inevitable variations in the enzymatic states. On the other hand, one can envisage situations where it would be beneficial for enzymes to be ‘slow’ so that their fluctuations amplify a weak signal and enable its detection by the intracellular machinery (see §3.1). Such a signal-detection mechanism was the main motivation behind *stochastic focusing*, a sensitivity amplification phenomenon introduced in [22]. We illustrate our results on the reaction scheme of [22] in §3.2, where we characterize how the substrate-mean changes with the speed of the enzyme abundance dynamics. Note that in situations where enzyme kinetics is ‘slow’, undesirable amplification effects can be eliminated by feedback mechanisms [23]. In this context, our results can help in postulating the presence of feedback loops using experimental data. We discuss the biological importance of our results in greater detail in §4.

It is interesting to note that some of the expressions we derive are related to those obtained in the analysis of various physico-chemical quantum dynamical systems coupled to a randomly fluctuating environment. This theory dates back



**Figure 1.** Schematic of the modelled system. (Online version in colour.)

to the original work of Kubo & Anderson [24,25] and, more recently, has been generalized to arbitrary quantum systems described by the Liouville–von Neumann equation with Markovian and non-Markovian parametric noise [26–28]. The example system of [28] can be interpreted as a substrate decaying with a stochastically fluctuating rate. While similar to it, the system we consider here includes the inflow of substrate, which results in a non-zero steady state and requires a different mathematical treatment.

## 2. Results

### 2.1. The model

We consider a system into which a substrate, **S**, enters at a constant rate  $k_{\text{in}}$  and is degraded or (equivalently) converted into a product **P** that in turn leaves the system. The rate of substrate outflow depends on the activity state or abundance level of an enzyme **E**. In turn, the catalytic activity of **E**, denoted by  $(\gamma(t))_{t \geq 0}$ , is assumed to fluctuate in time  $t$  according to a continuous-time Markov chain (CTMC) over a finite state-space  $\Gamma = \{\gamma_1, \dots, \gamma_n\}$ . Here, each  $\gamma_i$  is a positive constant denoting the degradation rate constant at the  $i$ th enzymatic state or abundance level. Owing to fluctuations in the catalytic activity of **E**, the degradation rate of substrate **S** will also fluctuate in time according to a stochastic process  $(k_{\text{d},S}(t))_{t \geq 0}$  whose value at time  $t$  is given by  $k_{\text{d},S}(t) = \gamma(t)S(t)$ , where  $S(t)$  is the molecular count or concentration of the substrate. This model is summarized in figure 1.

As mentioned before, the degradation reaction  $\text{S} \xrightarrow{\gamma(t)} \emptyset$  can also be viewed as a conversion reaction  $\text{S} \xrightarrow{\gamma(t)} \text{P}$  which are catalyzed by the enzyme. Generally, this catalytic step proceeds through the reversible formation of an intermediate complex  $\text{S} \cdot \text{E}$  which is formed when an enzyme molecule binds to a substrate molecule. In other words, the single reaction  $\text{S} \xrightarrow{\gamma(t)} \text{P}$  is an abstraction for the following three reactions:



If the binding/unbinding rates of **S** and **E** molecules are much higher than the rate of the conversion reaction, then we can apply the quasi-stationary assumption to conclude that the model in figure 1 is a good approximation to the catalytic conversion dynamics (for more details, see the electronic supplementary material in [23]).

To describe the CTMC  $(\gamma(t))_{t \geq 0}$ , we need to specify its  $n \times n$  transition rate matrix  $Q = [q_{ij}]$  [29]. For any distinct  $i, j \in \{1, 2, \dots, n\}$ ,  $q_{ij} \geq 0$  denotes the rate at which the process leaves state  $\gamma_i$  and enters state  $\gamma_j$ . The diagonal entries of  $Q$  are given by  $q_{ii} = -\sum_{j \neq i} q_{ij}$ . From now on, we assume that the rate matrix  $Q$  is irreducible<sup>1</sup> which implies that there exists a unique stationary distribution  $\pi = (\pi_1, \dots, \pi_n) \in \mathbb{R}_+^n$

satisfying

$$Q\mathbf{1} = \mathbf{0}, \quad \pi^T Q = \mathbf{0}^T \quad \text{and} \quad \pi^T \mathbf{1} = 1,$$

where  $\mathbf{0}$  and  $\mathbf{1}$  denote the  $n \times 1$  vectors of all zeroes and ones, respectively. As the state-space is finite and the transition rate matrix  $Q$  is irreducible, the CTMC  $(\gamma(t))_{t \geq 0}$  is ergodic which means that the probability distribution of  $\gamma(t)$  converges to the stationary distribution  $\pi$  as  $t \rightarrow \infty$ . As we are interested in the steady-state limit, without loss of generality we can assume that the initial state  $\gamma(0)$  is distributed according to  $\pi$ , i.e.  $\mathbb{P}(\gamma(0) = \gamma_i) = \pi_i$  for each  $i = 1, \dots, n$ . This ensures that the process  $(\gamma(t))_{t \geq 0}$  is a stationary stochastic process whose various statistical properties do not depend on time.<sup>2</sup> In particular, its mean  $\mathbb{E}(\gamma(t))$  is equal to

$$\mathbb{E}(\gamma(t)) = \mathbb{E}_\pi(\gamma) = \sum_{i=1}^n \gamma_i \pi_i \quad \text{for all } t \geq 0, \quad (2.1)$$

where  $\gamma$  is a  $\Gamma$ -valued random variable with probability distribution  $\pi$  and  $\mathbb{E}_\pi(\cdot)$  denotes the expectation w.r.t. this distribution.

From now on, we regard  $(\gamma(t))_{t \geq 0}$  as the baseline process which corresponds to enzymatic dynamics at the natural timescale. In order to study the substrate behaviour, we need to view enzymatic dynamics at the timescale of substrate kinetics. For this, we define a family of processes  $(\gamma_c(t))_{t \geq 0}$  parametrized by the ‘relative speed’ parameter  $c$  as follows:

$$\gamma_c(t) = \gamma(ct) \quad \text{for all } t \geq 0. \quad (2.2)$$

Note that one time-unit of process  $(\gamma_c(t))_{t \geq 0}$  corresponds to  $c$  time-units of process  $(\gamma(t))_{t \geq 0}$ . In this sense, the parameter  $c$  sets the speed of the fluctuation dynamics for the enzyme relative to the speed of the substrate kinetics. Like  $(\gamma(t))_{t \geq 0}$ , the process  $(\gamma_c(t))_{t \geq 0}$  is also a CTMC over state-space  $\Gamma = \{\gamma_1, \dots, \gamma_n\}$  with transition rate matrix  $Q_c = cQ$  and initial distribution  $\pi$ . Since  $(\gamma(t))_{t \geq 0}$  is stationary, this process is also stationary with the same mean given by  $\mathbb{E}_\pi(\gamma) = \mathbb{E}(\gamma_c(t))$  for all times  $t \geq 0$ . Replacing  $(\gamma(t))_{t \geq 0}$  by  $(\gamma_c(t))_{t \geq 0}$  in the model depicted in figure 1, we will study how the steady-state mean of substrate abundance depends on the fluctuation speed  $c$ .

Given a sample path of the enzyme dynamics  $(\gamma_c(t))_{t \geq 0}$  with relative speed  $c$ , we regard the dynamics of substrate molecular counts as a jump Markov chain  $(S_c(t))_{t \geq 0}$  over the set of non-negative integers  $\mathbb{N}_0 = \{0, 1, 2, \dots\}$ . This Markov chain can be written in the random time change representation [30] as

$$S_c(t) = S_c(0) + Y_1(k_{\text{in}}t) - Y_2\left(\int_0^t \gamma_c(u)S_c(u) du\right), \quad (2.3)$$

where  $Y_1$  and  $Y_2$  are independent, unit rate Poisson processes. From this representation, it is immediate that the substrate-production rate is constant ( $k_{\text{in}}$ ) in time, but the substrate-degradation rate is time-varying and it is equal to  $\gamma_c(t)S_c(t)$  at time  $t$ . Here, the Poisson processes  $Y_1$  and  $Y_2$  capture the intermittency in the firing of production and degradation reactions. This intermittency becomes unimportant if the substrate is present in high copy numbers [31] and in this case one can regard  $(S_c(t))_{t \geq 0}$  as the dynamics of substrate concentration,<sup>3</sup> specified by the following ODE:

$$\frac{dS_c(t)}{dt} = k_{\text{in}} - \gamma_c(t)S_c(t). \quad (2.4)$$

Note that even if the intermittency in production/ degradation reactions is ignored and  $(S_c(t))_{t \geq 0}$  is described by the ODE (2.4), the process  $(S_c(t))_{t \geq 0}$  is still *stochastic* because it is driven by the stochastic process  $(\gamma_c(t))_{t \geq 0}$  that represents enzymatic fluctuations.

Let  $m_c(t) = \mathbb{E}(S_c(t))$  for each  $t \geq 0$ . We shall soon see that  $m_c(t)$  does not depend on whether we use representation (2.3) or (2.4) for the substrate dynamics  $(S_c(t))_{t \geq 0}$ . Our goal in this paper is to understand the role of fluctuations in the catalytic activity of enzyme E in determining the steady-state value of the mean

$$m_{\text{eq}}(c) = \lim_{t \rightarrow \infty} m_c(t). \tag{2.5}$$

In particular, we study how this steady-state mean  $m_{\text{eq}}(c)$  depends on the relative fluctuation speed  $c$  and the variability in degradation rates  $\gamma_1, \dots, \gamma_n$  at various enzymatic activity levels.

## 2.2. Expressions for $m_{\text{eq}}(c)$ : the general case

We can approximately find  $m_{\text{eq}}(c)$  by estimating  $m_c(t)$  for a very large  $t$ , using simulations of the whole system. However, this naive approach is highly unsatisfactory because these simulations can be computationally expensive and the approximation error incurred by replacing the steady-state mean by a finite-time mean is generally difficult to quantify. Moreover, this approach does not provide us with an explicit formula for  $m_{\text{eq}}(c)$  that can enable us to study its dependence on the relative speed parameter  $c$ . In the light of these difficulties, we look for alternative ways to compute  $m_{\text{eq}}(c)$ . In this section, we assume that enzymatic kinetics is given by a general stationary stochastic process with an arbitrary state-space  $\Gamma \subset (0, \infty)$ , and so we do not rely on the CTMC structure mentioned in §2.1. We specialize the results of this section to the CTMC case in §2.3.

Using representation (2.3) or (2.4), we can show that  $m_c(t) = \mathbb{E}(S_c(t))$  is given by the following formula:

$$m_c(t) = \mathbb{E}(S_c(0)) e^{-\int_0^t \gamma_c(s) ds} + k_{\text{in}} \int_0^t \mathbb{E}(e^{-\int_s^t \gamma_c(u) du}) ds. \tag{2.6}$$

From the stationarity of the process  $(\gamma_c(t))_{t \geq 0}$ , we can conclude that

$$\begin{aligned} \int_0^t \mathbb{E}(e^{-\int_s^t \gamma_c(u) du}) ds &= \int_0^t \mathbb{E}(e^{-\int_0^{t-s} \gamma_c(u) du}) ds \\ &= \int_0^t \mathbb{E}(e^{-\int_0^s \gamma_c(u) du}) ds. \end{aligned}$$

Substituting this in (2.6) and letting  $t \rightarrow \infty$ , we obtain our first formula for  $m_{\text{eq}}(c)$ , which is

$$m_{\text{eq}}(c) = \lim_{t \rightarrow \infty} m_c(t) = k_{\text{in}} \int_0^\infty \mathbb{E}(e^{-\int_0^s \gamma_c(u) du}) ds. \tag{2.7}$$

From (2.2), we obtain

$$\int_0^s \gamma_c(u) du = s \left( \frac{1}{cS} \int_0^{cS} \gamma(u) du \right). \tag{2.8}$$

As the process  $(\gamma(t))_{t \geq 0}$  is stationary, from theorem 10.6 in [32] we know that as  $c \rightarrow \infty$ , the quantity (2.8) converges a.s. to  $s\mathbb{E}_\pi(\gamma)$  (recall (2.1)). As a consequence,  $\mathbb{E}(e^{-\int_0^s \gamma_c(u) du}) \rightarrow e^{-s\mathbb{E}_\pi(\gamma)}$  and hence we get

$$\lim_{c \rightarrow \infty} m_{\text{eq}}(c) = k_{\text{in}} \int_0^\infty e^{-s\mathbb{E}_\pi(\gamma)} ds = \frac{k_{\text{in}}}{\mathbb{E}_\pi(\gamma)} := m_{\text{eq}}^{(\text{det})}. \tag{2.9}$$

This shows that as the relative speed  $c$  of enzymatic fluctuations approaches  $\infty$ , these fluctuations become *equilibrated* at the timescale of substrate kinetics, and so they do not affect the mean substrate level. In other words, from the point of view of the substrate, the enzyme kinetics is so fast that it is as if the enzyme state is constant at the equilibrium level  $\mathbb{E}_\pi(\gamma)$ . This corresponds to the classical case where there is no dynamic disorder in the enzyme activity and so this activity is well approximated by a deterministic rate constant for the substrate degradation reaction. As the mapping  $x \mapsto e^{-x}$  is convex, Jensen's inequality tells us that

$$\mathbb{E}(e^{-\int_0^s \gamma_c(u) du}) \geq e^{-\int_0^s \mathbb{E}(\gamma_c(u)) du} = e^{-s\mathbb{E}_\pi(\gamma)},$$

where the last relation follows from the fact that  $(\gamma_c(t))_{t \geq 0}$  is a stationary process with mean  $\mathbb{E}(\gamma_c(t)) = \mathbb{E}_\pi(\gamma)$  for all times  $t \geq 0$ . Substituting this in (2.7), we see that for any  $c \geq 0$

$$m_{\text{eq}}(c) \geq k_{\text{in}} \int_0^\infty e^{-s\mathbb{E}_\pi(\gamma)} ds = \frac{k_{\text{in}}}{\mathbb{E}_\pi(\gamma)} = m_{\text{eq}}^{(\text{det})}. \tag{2.10}$$

Therefore for a finite relative speed  $c$ , enzymatic fluctuations *always amplify the mean substrate abundance*, in comparison to the classical deterministic case. The natural question that now arises is—how large should speed  $c$  be in order for the deterministic approximation to be acceptable within a certain tolerance level  $\epsilon$ ? We address this question in §2.4.

Let us now consider the situation where the relative speed parameter  $c \rightarrow 0$  and so at the timescale of substrate kinetics, the enzyme dynamics  $(\gamma_c(t))_{t \geq 0}$  approaches a *static* process, i.e.  $\gamma_c(t) = \gamma(0)$  for all  $t \geq 0$ . This case corresponds to the situation where the enzyme kinetics is *very slow* in comparison to the substrate kinetics. Hence from the point of view of the substrate, the kinetics of the enzyme is almost *fixed*. In this regime, we can replace  $\gamma_c(u)$  by  $\gamma(0)$  in (2.7) to obtain

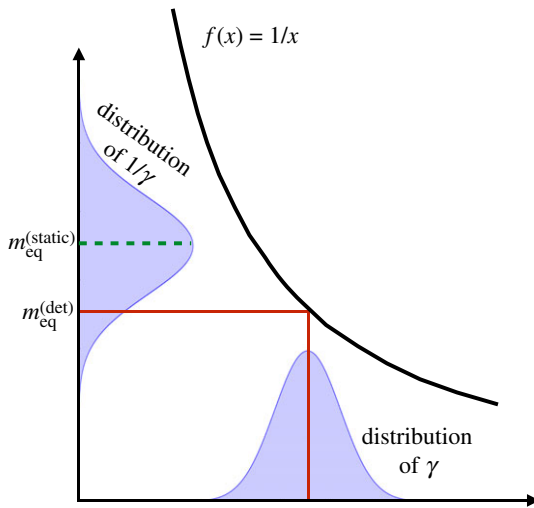
$$\begin{aligned} \lim_{c \rightarrow 0} m_{\text{eq}}(c) &= k_{\text{in}} \mathbb{E} \left( \int_0^\infty e^{-s\gamma(0)} ds \right) = k_{\text{in}} \mathbb{E} \left( \frac{1}{\gamma(0)} \right) \\ &= k_{\text{in}} \mathbb{E}_\pi \left( \frac{1}{\gamma} \right) := m_{\text{eq}}^{(\text{static})}, \end{aligned} \tag{2.11}$$

where we have used the fact that  $\gamma(0)$  has probability distribution  $\pi$  to write  $\mathbb{E}(1/\gamma(0))$  as  $\mathbb{E}_\pi(1/\gamma)$ . Observe that  $m_{\text{eq}}^{(\text{static})} \geq m_{\text{eq}}^{(\text{det})}$ , which can be readily seen by letting  $c \rightarrow 0$  in (2.10) or by directly using Jensen's inequality on the convex map  $f(x) = 1/x$  (figure 2). The two extremal cases  $c \rightarrow 0$  and  $c \rightarrow \infty$  serve as a guide to the behaviour of realistic systems with an intermediate value of  $c$ . In particular, we can expect that for such intermediate  $c$ -values, the steady-state substrate mean will lie somewhere between  $m_{\text{eq}}^{(\text{det})}$  and  $m_{\text{eq}}^{(\text{static})}$ . This is precisely what happens as we shall soon see. We will also discuss how the precise value of  $m_{\text{eq}}(c)$  can be computed or estimated from any Markovian model of enzymatic fluctuations.

Until now, the conclusions we have drawn regarding  $m_{\text{eq}}(c)$  rely on the formula (2.7) that holds for any real-valued stationary stochastic process  $(\gamma_c(t))_{t \geq 0}$  as long as its states are positive and bounded away from 0, i.e. the state-space  $\Gamma$  satisfies

$$\inf \{x: x \in \Gamma\} \geq \epsilon \tag{2.12}$$

for some  $\epsilon > 0$ . This makes this formula very general but it is difficult to work with, because it involves an indefinite integral which is generally analytically intractable as the mapping  $s \mapsto \mathbb{E}(e^{-\int_0^s \gamma_c(u) du})$  does not have an explicit form. We remedy this problem in the next section by specializing



**Figure 2.** This figure illustrates how the convex map  $f(x) = 1/x$  distorts the distribution of  $\gamma$ . Assuming  $k_{in} = 1$  here, the static mean  $m_{eq}^{(static)} = \mathbb{E}_{\pi}(f(\gamma)) = \mathbb{E}_{\pi}(1/\gamma)$  is always greater than the deterministic mean  $m_{eq}^{(det)} = f(\mathbb{E}_{\pi}(\gamma)) = 1/\mathbb{E}_{\pi}(\gamma)$ . (Online version in colour.)

this formula to the case where  $(\gamma_c(t))_{t \geq 0}$  is a finite state-space CTMC as described in §2.1. Before we come to that, we provide a numerical recipe for statistically estimating  $m_{eq}(c)$  without the need for evaluating the indefinite integral. This scheme is based on the assumption that we can efficiently generate sample paths of the stationary process  $(\gamma_c(t))_{t \geq 0}$  [33,34].

Define a random variable  $\tau_c$  by

$$\tau_c = \inf \left\{ t \geq 0: \int_0^t \gamma_c(s) ds = -\ln u \right\}, \quad (2.13)$$

where  $u$  is an independent random variable with the uniform distribution on  $[0, 1]$ . To sample  $\tau_c$ , we can adopt the following strategy. We first sample  $u$  from the uniform distribution on  $[0, 1]$ , draw an initial condition  $\gamma_c(0)$  from  $\pi$ , and then simulate the sample path  $(\gamma_c(t))_{t \geq 0}$ , keeping track of the integral  $\int_0^t \gamma_c(s) ds$ . We take  $\tau_c$  to be first time  $t$  when this integral hits the value  $(-\ln u)$ . From the samples of the random variable  $\tau_c$ , we can estimate its expectation  $\mathbb{E}(\tau_c)$  which gives us an estimate for  $m_{eq}(c)$  because it can be shown that

$$m_{eq}(c) = k_{in} \mathbb{E}(\tau_c). \quad (2.14)$$

Note that the estimator for  $m_{eq}(c)$  based on formula (2.14) will be unbiased, but it will suffer from statistical error due to a finite sample size. However, this error can be estimated and managed far more easily than the error one would incur by approximating the steady-state mean  $m_{eq}(c)$  by a finite-time mean  $m_c(t)$  (recall (2.5)). This makes this formula (2.14) useful in practice (e.g. §3.2).

The results from this section are collected in our next proposition which is proved in the electronic supplementary material.

**Proposition 2.1.** Suppose  $(\gamma(t))_{t \geq 0}$  is a real-valued stationary stochastic process with stationary distribution  $\pi$  and state-space  $\Gamma$  satisfying (2.12). Let  $(\gamma_c(t))_{t \geq 0}$  be the speed  $c$  version of this process given by (2.2) and define the substrate dynamics  $(S_c(t))_{t \geq 0}$  either by (2.3) or by (2.4). Let  $m_c(t) = \mathbb{E}(S_c(t))$  and let the steady-state limit  $m_{eq}(c)$  be given by (2.5). Then we have the following:

- (A) The value  $m_{eq}(c)$  is well defined (i.e. the limit in (2.5) exists) and it is given by (2.7).
- (B) If  $\tau_c$  is the random variable defined by (2.13), then (2.14) holds.
- (C) The limits (2.9) and (2.11) are satisfied as  $c \rightarrow \infty$  and  $c \rightarrow 0$ , respectively.

### 2.3. Expressions for $m_{eq}(c)$ : the finite continuous-time Markov chain case

In this section, we specialize expression (2.7) to the case where  $(\gamma(t))_{t \geq 0}$  is a stationary CTMC with a finite state-space  $\Gamma = \{\gamma_1, \dots, \gamma_n\}$ , as described in §2.1. Define the CTMC  $(\gamma_c(t))_{t \geq 0}$  by (2.2) and recall that its  $n \times n$  transition-rate matrix is given by  $Q_c = cQ$ . Let  $D$  be the  $n \times n$  diagonal matrix

$$D = \text{Diag}(\gamma_1, \dots, \gamma_n) \quad (2.15)$$

whose entries are the degradation rates at different enzymatic states or abundance levels. One of the main results in our paper is to show that  $m_{eq}(c)$  can be expressed as

$$m_{eq}(c) = k_{in} [\pi^T (D - cQ)^{-1} \mathbf{1}]. \quad (2.16)$$

Two alternative proofs of this result are given in the electronic supplementary material. The first proof exploits some ideas from the theory of occupation measures for Markov chains [35] while the second proof is based on the *methods of conditional moments* approach recently developed by Hasenauer *et al.* [36]. Note that this formula assumes matrix  $(D - cQ)$  is invertible for any  $c \geq 0$ , but this can be easily verified from the properties of matrix  $Q$ . Using formula (2.16), we can prove that for any  $c \geq 0$

$$m_{eq}^{(det)} \leq m_{eq}(c) \leq m_{eq}^{(static)}. \quad (2.17)$$

Therefore for any relative speed  $c$  of enzymatic fluctuations, the steady-state mean of the substrate is always sandwiched between the values obtained for the *deterministic* and the *static* cases. Moreover, since  $m_{eq}(c)$  depends continuously on  $c$ , the limits (2.11) and (2.9) imply that for any value  $m^*$  in the open interval  $(m_{eq}^{(det)}, m_{eq}^{(static)})$  there exists a relative speed value  $c^* > 0$  such that  $m_{eq}(c^*) = m^*$ . Hence, the positive deviations caused by enzymatic fluctuations (in the mean substrate abundance) range from 0 to *exactly*  $(m_{eq}^{(static)} - m_{eq}^{(det)})$ .

To determine the map  $c \mapsto m_{eq}(c)$ , we need to evaluate  $m_{eq}(c)$  at several values of  $c$ . This can be difficult with formula (2.16) because each evaluation requires inversion of a potentially large matrix. Fortunately, we can resolve this issue using simple ideas from the *theory of resolvents* for linear operators [37], as we now describe. Let  $\tilde{Q} = D^{-1}Q$  be the transition rate matrix of another CTMC over state-space  $\Gamma = \{\gamma_1, \dots, \gamma_n\}$ . The difference between this new CTMC and the original CTMC is that the rates of outflow from state  $\gamma_i$  to each state  $\gamma_j$  (for  $j \neq i$ ) are divided by the state value  $\gamma_i$ . Let  $\mathbb{C}$  denote the field of complex numbers. The resolvent for the Markov semigroup corresponding to this CTMC is the matrix-valued function over  $\mathbb{C}$  defined by

$$R(z) = (zI - \tilde{Q})^{-1}, \quad (2.18)$$

where  $I$  is the  $n \times n$  Identity matrix. This function is well defined for any  $z$  which is not an eigenvalue of

matrix  $\tilde{Q}$ . Let  $\lambda_1, \dots, \lambda_n$  be the  $n$  eigenvalues of matrix  $\tilde{Q}$ , repeated according to their algebraic multiplicity. Since  $\tilde{Q}$  is the transition rate matrix of a CTMC, it has a simple<sup>4</sup> eigenvalue (say  $\lambda_1$ ) equal to 0, while its other eigenvalues have negative real parts. This implies that the resolvent function  $R$  is well defined on the positive real line  $(0, \infty)$ .

From now on, we assume that matrix  $\tilde{Q}$  is diagonalizable<sup>5</sup> over the field  $\mathbb{C}$  of complex numbers. This assumption is not very restrictive because almost every matrix is diagonalizable [38] and so if  $\tilde{Q}$  is not diagonalizable, we can perturb matrix  $Q$  slightly to make  $\tilde{Q}$  diagonalizable and not affect the enzyme dynamics significantly. The diagonalizability of  $\tilde{Q}$  allows us to write matrix  $\tilde{Q}$  as  $\tilde{Q} = U\Lambda U^{-1}$ , where  $\Lambda = \text{Diag}(\lambda_1, \dots, \lambda_n)$  and  $U$  is an invertible matrix whose columns contain the right eigenvectors for matrix  $\tilde{Q}$  corresponding to the eigenvalues  $\lambda_1, \dots, \lambda_n$ . Similarly the rows of  $U^{-1}$  contain the left eigenvectors for matrix  $\tilde{Q}$  corresponding to the eigenvalues  $\lambda_1, \dots, \lambda_n$ . Let  $u_i$  and  $w_i$  be  $n \times 1$  vectors denoting the  $i$ th column and  $i$ th row of matrices  $U$  and  $U^{-1}$ , respectively. Therefore,

$$\tilde{Q} = \sum_{i=1}^n \lambda_i u_i w_i^T \quad (2.19)$$

and we can express the resolvent function  $R$  (see ch. 5 in [37]) as

$$R(z) = \sum_{i=1}^n \left( \frac{1}{z - \lambda_i} \right) u_i w_i^T. \quad (2.20)$$

Let  $\langle \cdot, \cdot \rangle$  denote the standard inner product on  $\mathbb{R}^n$ . Note that formula (2.16) can be expressed as

$$m_{\text{eq}}(c) = k_{\text{in}} [\pi^T (I - c\tilde{Q})^{-1} D^{-1} \mathbf{1}] = c^{-1} k_{\text{in}} [\pi^T R(c^{-1}) D^{-1} \mathbf{1}]$$

for any  $c > 0$ . Plugging  $R(c^{-1})$  from (2.20) and defining

$$\alpha_i = \langle \pi, u_i \rangle \langle w_i, D^{-1} \mathbf{1} \rangle = (\pi^T u_i) (w_i^T D^{-1} \mathbf{1}) \quad (2.21)$$

for each  $i = 1, \dots, n$ ,

we obtain the following formula for  $m_{\text{eq}}(c)$ :

$$m_{\text{eq}}(c) = k_{\text{in}} \sum_{i=1}^n \left( \frac{\alpha_i}{1 - c\lambda_i} \right). \quad (2.22)$$

Observe that since  $\alpha_i$ -s and  $\lambda_i$ -s are independent of  $c$ , they only need to be computed once to construct this expression and then we can easily compute  $m_{\text{eq}}(c)$  for several values of  $c$  without the need of evaluating the matrix inverses in (2.16). Moreover if  $n$  is large, then using the values of  $\alpha_i$  and  $\lambda_i$  as a guide, one can derive suitable approximations of formula (2.22) for  $m_{\text{eq}}(c)$ . We derive one such approximation in the next section and use it as a tool to further understand the phenomenon of stochastic amplification induced by dynamic disorder in enzymatic activity.

The results from this section are collected in our next theorem which is proved in the electronic supplementary material.

**Theorem 2.2.** Suppose  $(\gamma(t))_{t \geq 0}$  is a stationary CTMC with transition rate matrix  $Q$ , stationary distribution  $\pi$  and state-space  $\Gamma = \{\gamma_1, \dots, \gamma_n\}$  (see §2.1). Let  $(\gamma_c(t))_{t \geq 0}$  be the speed  $c$  version of this process given by (2.2) and define the substrate dynamics  $(S_c(t))_{t \geq 0}$  either by (2.3) or by (2.4). Let the steady-state substrate mean

$m_{\text{eq}}(c)$  be given by (2.5) and the diagonal matrix  $D$  be defined by (2.15). Then we have the following:

- (A) The matrix  $(D - cQ)$  is invertible and  $m_{\text{eq}}(c)$  can be expressed as (2.16).
- (B) Suppose the matrix  $\tilde{Q} = D^{-1}Q$  is diagonalizable and let  $\lambda_1, \dots, \lambda_n$  be its eigenvalues. For each  $i = 1, \dots, n$  define  $\alpha_i$  by (2.21). Then  $m_{\text{eq}}(c)$  can be expressed as (2.22).
- (C) The relation (2.17) is satisfied for any  $c \geq 0$ .

## 2.4. Approximate formula for $m_{\text{eq}}(c)$

The goal of this section is to derive an approximate formula for  $m_{\text{eq}}(c)$  using (2.22) and then use it to obtain some interesting insights. Recall from the previous section that  $\lambda_1, \dots, \lambda_n$  are the eigenvalues of matrix  $\tilde{Q}$ . Among these  $\lambda_1 = 0$ , while the eigenvalues  $\lambda_2, \dots, \lambda_n$  have negative real parts. Define a positive constant  $\epsilon_{\text{max}}$  by

$$\epsilon_{\text{max}} = -\max\{\text{Re}(\lambda_i) : i = 2, \dots, n\},$$

where  $\text{Re}(z)$  denotes the real part of a complex number  $z$ . Setting  $\lambda_1 = 0$  in (2.22) we obtain

$$m_{\text{eq}}(c) = k_{\text{in}} \left[ \alpha_1 + \sum_{i=2}^n \left( \frac{\alpha_i}{1 - c\lambda_i} \right) \right]. \quad (2.23)$$

This formula is valid for any  $c$  in the interval  $(-\epsilon_{\text{max}}, \infty)$  and its form shows that the function  $m_{\text{eq}}(c)$  is *real-analytic*<sup>6</sup> at  $c = 0$ . Therefore, all the information about function  $m_{\text{eq}}(c)$  is contained in the value of this function and its derivatives at  $c = 0$ .

Using limits (2.9) and (2.11), we can conclude that

$$\alpha_1 = \frac{m_{\text{eq}}^{(\text{det})}}{k_{\text{in}}} \quad \text{and} \quad \sum_{i=2}^n \alpha_i = \left( \frac{m_{\text{eq}}^{(\text{static})} - m_{\text{eq}}^{(\text{det})}}{k_{\text{in}}} \right). \quad (2.24)$$

Let  $\theta$  denote the following weighted combination of eigenvalues  $\lambda_2, \dots, \lambda_n$ :

$$\theta = -\frac{\sum_{i=2}^n \lambda_i \alpha_i}{\sum_{i=2}^n \alpha_i}. \quad (2.25)$$

We now propose an approximate formula for  $m_{\text{eq}}(c)$

$$\hat{m}_{\text{eq}}(c) = m_{\text{eq}}^{(\text{det})} + \left( \frac{m_{\text{eq}}^{(\text{static})} - m_{\text{eq}}^{(\text{det})}}{1 + c\theta} \right). \quad (2.26)$$

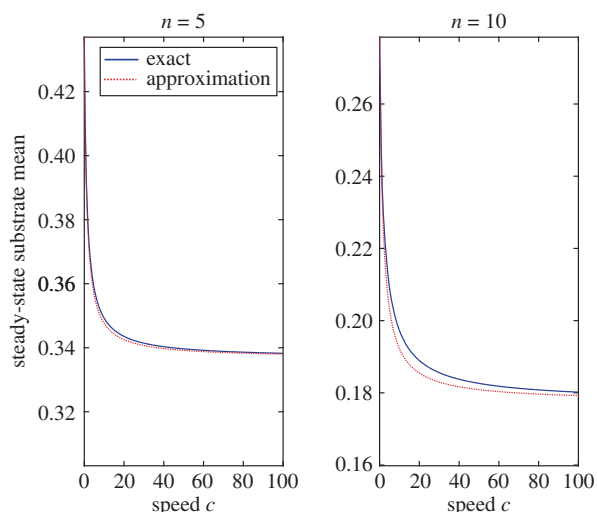
Note this formula is much easier to use than (2.23) because it contains only one rational term. From (2.24), it is immediate that  $\hat{m}_{\text{eq}}(c)$  also obeys the limits (2.9) and (2.11). Moreover, it is straightforward to check that the first derivatives of  $\hat{m}_{\text{eq}}(c)$  and  $m_{\text{eq}}(c)$  match at  $c = 0$ . Hence, the approximation error is given by a difference of second-order derivatives and we explain in the electronic supplementary material why this error is likely to be small. We can also view the approximation  $\hat{m}_{\text{eq}}(c)$  of  $m_{\text{eq}}(c)$  as replacing a weighted *arithmetic mean* of several quantities with the corresponding *harmonic mean*. To see this note that from (2.23) and (2.24) we can express  $m_{\text{eq}}(c)$  as

$$m_{\text{eq}}(c) = m_{\text{eq}}^{(\text{det})} + (m_{\text{eq}}^{(\text{static})} - m_{\text{eq}}^{(\text{det})}) \bar{x}, \quad (2.27)$$

where

$$\bar{x} = \frac{\sum_{i=2}^n \alpha_i (1 - c\lambda_i)^{-1}}{\sum_{i=2}^n \alpha_i}$$

is the weighted arithmetic mean of quantities  $(1 - c\lambda_2)^{-1}, \dots, (1 - c\lambda_n)^{-1}$  with weights  $\alpha_2, \dots, \alpha_n$ .<sup>7</sup> The corresponding



**Figure 3.** Comparison of the exact steady-state mean substrate value  $m_{\text{eq}}(c)$  with its approximation  $\hat{m}_{\text{eq}}(c)$ . (Online version in colour.)

weighted harmonic mean of these quantities is given by

$$\hat{x} = \frac{\sum_{i=2}^n \alpha_i}{\sum_{i=2}^n \alpha_i (1 - c\lambda_i)} = \frac{1}{1 + c\theta}$$

and observe that  $\hat{m}_{\text{eq}}(c)$  can be expressed as the r.h.s. of (2.27) with arithmetic mean  $\bar{x}$  replaced by the harmonic mean  $\hat{x}$ .

We now illustrate the accuracy of this approximation using a couple of randomly generated  $n \times n$ , transition rate matrices  $Q$  with  $n = 5$  and  $n = 10$ , respectively. In both cases, we choose the input rate to be  $k_{\text{in}} = 1$  and the enzymatic state values to be  $\gamma_i = i$  for  $1, 2, \dots, n$ . The exact function  $\hat{m}_{\text{eq}}(c)$  along with its approximation  $m_{\text{eq}}(c)$  are plotted in figure 3. The accuracy of this approximation can be easily seen. Note that the exact function is *slightly above* its approximation. Assuming the significant weights ( $\alpha_i$ -s) are positive reals, this can be explained by the fact that arithmetic mean is always higher than the corresponding harmonic mean.

From (2.26), it is immediate that the shape of the function  $\hat{m}_{\text{eq}}(c)$  depends crucially on the parameter  $\theta$  computed according to (2.25). We now examine  $\theta$  more closely and see how it is connected to an existing notion from the theory of Markov processes. Let us denote the numerator of (2.25) by

$$\Theta = - \sum_{i=2}^n \lambda_i \alpha_i. \quad (2.28)$$

Since  $\alpha_i$ -s are given by (2.21), using (2.19),  $\tilde{Q} = D^{-1}Q$  and  $\lambda_1 = 0$ , we can express  $\Theta$  as

$$\begin{aligned} \Theta &= - \sum_{i=1}^n \lambda_i \alpha_i = -\pi^T \left( \sum_{i=1}^n \lambda_i u_i w_i^T \right) D^{-1} \mathbf{1} \\ &= -\pi^T \tilde{Q} D^{-1} \mathbf{1} = -\pi^T D^{-1} Q D^{-1} \mathbf{1}. \end{aligned} \quad (2.29)$$

This relation shows that  $\Theta$  (and hence  $\theta$ ) is always real-valued even though some  $\lambda_i$ -s or  $\alpha_i$ -s may have imaginary parts. Moreover to compute  $\Theta$ , we do not need to compute the eigenvalues  $\lambda_1, \dots, \lambda_n$  of a potentially large matrix  $\tilde{Q}$ . Instead, we only need to evaluate the expression  $\pi^T D^{-1} Q D^{-1} \mathbf{1}$  which is computationally much easier. Interestingly, the definition of  $\Theta$  coincides with the well-known notion of *Dirichlet forms* that is extensively used in the study of mixing properties

of Markov processes [18,39]. We now discuss this connection in more detail.

Consider the CTMC  $(\gamma(t))_{t \geq 0}$  with state-space  $\Gamma = \{\gamma_1, \dots, \gamma_n\}$  and transition rate matrix  $Q = [q_{ij}]$ . The generator  $\mathbb{Q}$  of this CTMC maps any real-valued function  $f$  on  $\Gamma$  to another such real-valued function  $\mathbb{Q}f$  given by

$$\mathbb{Q}f(\gamma_i) = \sum_{j \neq i} q_{ij} (f(\gamma_j) - f(\gamma_i)).$$

Define a function  $f: \Gamma \rightarrow (0, \infty)$  by  $f(\gamma) = 1/\gamma$ . Then one can see that  $\Theta$  (2.28) can be expressed as

$$\Theta = -\mathbb{E}_\pi(f(\gamma)\mathbb{Q}f(\gamma)). \quad (2.30)$$

In other words, if a  $\Gamma$ -valued random variable  $\gamma$  has distribution  $\pi$ , then  $\Theta$  is the expectation of the random variable  $(-f(\gamma)\mathbb{Q}f(\gamma))$ . Relation (2.30) shows that  $\Theta$  is a Dirichlet form associated with the Markovian semigroup generated by  $\mathbb{Q}$  [18]. An important consequence of this connection is that  $\Theta$  is always positive (see lemma 2.1.2 in [18]) irrespective of the entries of the rate matrix  $Q$  or the state values  $\gamma_1, \dots, \gamma_n$ . The positivity of  $\Theta$  implies that  $\theta$  is also positive and hence the mapping  $c \mapsto \hat{m}_{\text{eq}}(c)$  is convex and monotonically decreasing from  $m_{\text{eq}}^{\text{(static)}}$  at  $c = 0$  to  $m_{\text{eq}}^{\text{(det)}}$  as  $c \rightarrow \infty$ . Intuitively, the magnitude of Dirichlet form  $\Theta$  corresponds to the *mixing strength* of the underlying Markov process. Therefore as  $\Theta$  increases,  $\theta$  also increases and the mapping  $c \mapsto \hat{m}_{\text{eq}}(c)$  has a sharper ‘drop’ to the deterministic value  $m_{\text{eq}}^{\text{(det)}}$ . Our next goal is to make this mathematically precise and quantitatively estimate the relative speed values  $c$  beyond which the deterministic assumption is acceptable.

In the rest of this section, our object of interest will be the *relative stochastic amplification factor* defined by

$$\rho(c) = \frac{m_{\text{eq}}(c) - m_{\text{eq}}^{\text{(det)}}}{m_{\text{eq}}^{\text{(det)}}}, \quad (2.31)$$

which measures the difference of steady-state substrate means in the presence and absence of enzymatic fluctuations, *normalized* by the steady-state substrate mean in the deterministic case. Note that  $\rho(c)$  does not depend on the input rate  $k_{\text{in}}$  and using (2.17) we can see that  $\rho(c)$  satisfies

$$0 \leq \rho(c) \leq \rho_{\text{max}} := \frac{m_{\text{eq}}^{\text{(static)}}}{m_{\text{eq}}^{\text{(det)}}} - 1 = \mathbb{E}_\pi \left( \frac{1}{\gamma} \right) \mathbb{E}_\pi(\gamma) - 1 \quad (2.32)$$

for any  $c \geq 0$ .

In order to study the dependence of  $\rho(c)$  on  $c$ , we now look at its approximation  $\hat{\rho}(c)$  which is defined analogously to (2.31), with  $m_{\text{eq}}(c)$  replaced by  $\hat{m}_{\text{eq}}(c)$ . Using (2.25), (2.28) and (2.24), we see that  $\theta$  is the same as the *normalized Dirichlet form* defined by

$$\theta = \frac{\Theta \mathbb{E}_\pi(\gamma)}{\rho_{\text{max}}}. \quad (2.33)$$

Dividing (2.26) by  $m_{\text{eq}}^{\text{(det)}}$ , we obtain the following formula after some simple algebraic manipulations:

$$\hat{\rho}(c) = \frac{\rho_{\text{max}}}{1 + c\theta}. \quad (2.34)$$

This formula clearly indicates that as  $\theta$  gets larger, the amplification factor decreases more sharply to 1 with the relative speed parameter  $c$ . One can regard  $\rho(c) \approx \hat{\rho}(c)$  as the ‘relative error’ between the actual substrate mean and the mean computed with deterministic assumption on the enzymatic kinetics. From relation (2.34), it is immediate that in order



to test if this error will exceed some tolerance level  $\epsilon > 0$ , we just need to check if the relative enzyme speed  $c$  is smaller than the threshold  $c_\epsilon$  defined by

$$c_\epsilon := \left( \frac{\rho_{\max} - \epsilon}{\theta \epsilon} \right). \quad (2.35)$$

We can expect this test to be rather conservative because as we have argued before, the exact values  $m_{\text{eq}}(c)$  will usually lie above their approximation  $\hat{m}_{\text{eq}}(c)$ .

Note that  $c_\epsilon$  is inversely proportional to  $\theta$  but directly proportional to  $\rho_{\max}$ . The first parameter  $\theta$  is the normalized Dirichlet form and it captures the ‘mixing strength’ of the underlying enzymatic dynamics (e.g. §3.1), while the second parameter  $\rho_{\max}$  can be viewed as a proxy for the variance of the stationary distribution  $\pi^9$  (e.g. §3.1). Generally, both these parameters will increase with higher levels of noise in the enzymatic dynamics. However, since they affect  $c_\epsilon$  in opposing ways, it is difficult to ascertain the overall effect of dynamical noise in setting the threshold value  $c_\epsilon$ . We explore this issue in greater detail in §3.2 and numerically show that increasing levels of dynamical noise in the enzymatic kinetics of that reaction network gives rise to decreasing values of  $c_\epsilon$ . This is surprising and counterintuitive because it suggests that this dynamical noise is actually beneficial in improving the accuracy of the deterministic assumption for the enzyme activity.

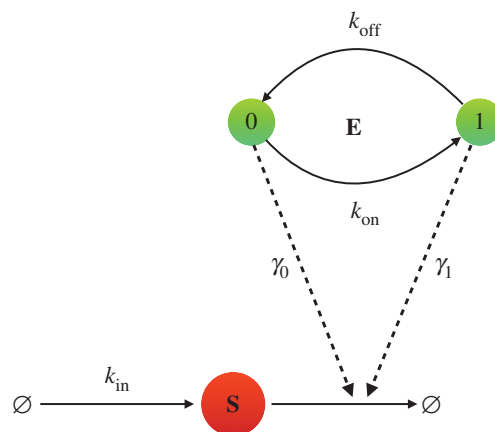
Finally, we remark that even though most of the analysis in this paper assumes that enzymatic kinetics is described by a finite Markov chain, the formulae we derive can provide insights for a more general class of stationary stochastic processes. This is because finite Markov chains can serve as good approximations of such processes [41]. Moreover if the process is Markov, even with an arbitrary state-space, we can compute expression (2.34) for  $\hat{\rho}(c)$  by sampling its stationary distribution and using this sample to estimate  $\rho_{\max}$  and the normalized Dirichlet form  $\theta$ . We illustrate this for the example network in §3.2 where the enzyme dynamics follows a Markov process over a countable state-space.

### 3. Examples

In this section, we present a couple of examples to illustrate our results. Our first example of a two-state switching enzyme is such that all calculations can be easily done analytically allowing us to clearly understand the stochastic amplification effect. We also see how enzymes can utilize their fluctuations to serve as high-gain amplifiers. Our second example is the reaction network of Paulsson *et al.* [22] which displays stochastic focusing. We apply our results to this network and demonstrate that in some cases dynamical fluctuations can actually be beneficial in reducing the unwanted stochastic amplification effects.

#### 3.1. Stochastic amplification induced by a two-state switching enzyme

Consider a simple instance of the system in figure 1, in which a single enzyme molecule is present and can fluctuate between two states of activity: with enzyme **E** in the low-activity (0) state, the degradation rate of substrate **S** is assumed to be  $\gamma_0$ , while it is equal to  $\gamma_1$  when **E** is highly active (1). The



**Figure 4.** Schematic of the two-state switching enzyme model. (Online version in colour.)

0-to-1 and 1-to-0 rates are given by  $k_{\text{on}}$  and  $k_{\text{off}}$ , respectively. A schematic of this model is presented in figure 4.

The time-varying degradation rate  $(\gamma(t))_{t \geq 0}$  induced by this fluctuating enzyme **E** is a CTMC with state-space  $\Gamma = \{\gamma_0, \gamma_1\}$  and transition-rate matrix

$$Q = \begin{bmatrix} -k_{\text{on}} & k_{\text{on}} \\ k_{\text{off}} & -k_{\text{off}} \end{bmatrix}.$$

One can check that the unique stationary distribution  $\pi = (\pi_0, \pi_1)$  for this CTMC is simply given by

$$\pi_0 = \frac{k_{\text{off}}}{k_{\text{on}} + k_{\text{off}}} \quad \text{and} \quad \pi_1 = \frac{k_{\text{on}}}{k_{\text{on}} + k_{\text{off}}}. \quad (3.1)$$

For each  $i = 0, 1$ , we can regard  $\pi_i$  as the steady-state probability of the enzyme being in state  $i$ . Owing to the ergodic theorem (see theorem 10.6 in [32]), we can also view  $\pi_i$  as the proportion of time that the enzyme spends in state  $i$  in the long run. Let  $\gamma$  be a  $\Gamma$ -valued random variable with probability distribution  $\pi$ . Then its mean and variance can be computed as

$$\left. \begin{aligned} \mathbb{E}_\pi(\gamma) &= \left( \frac{k_{\text{off}}\gamma_0 + k_{\text{on}}\gamma_1}{k_{\text{on}} + k_{\text{off}}} \right) \quad \text{and} \\ \text{Var}_\pi(\gamma) &= \frac{k_{\text{on}}k_{\text{off}}}{(k_{\text{on}} + k_{\text{off}})^2} (\gamma_0 - \gamma_1)^2. \end{aligned} \right\} \quad (3.2)$$

Suppose that the speed of enzymatic kinetics relative to the substrate is  $c$  and so the degradation rate is given by the process  $(\gamma_c(t))_{t \geq 0}$  defined by (2.2). Let  $m_{\text{eq}}(c)$  (2.5) be the steady-state substrate mean in this case. Using part (A) of theorem 2.2, we obtain

$$m_{\text{eq}}(c) = k_{\text{in}} [\pi_0 \quad \pi_1] \begin{bmatrix} \gamma_0 + ck_{\text{on}} & -ck_{\text{on}} \\ -ck_{\text{off}} & \gamma_1 + ck_{\text{off}} \end{bmatrix}^{-1} \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

This formula involves the inverse of a  $2 \times 2$  matrix, which can be easily computed explicitly. Substituting this inverse along with the expressions for  $\pi_0$  and  $\pi_1$  (see (3.1)) we get

$$\begin{aligned} m_{\text{eq}}(c) &= \frac{k_{\text{in}}}{\gamma_0\gamma_1 + c\gamma_0k_{\text{off}} + c\gamma_1k_{\text{on}}} [\pi_0 \quad \pi_1] \begin{bmatrix} \gamma_1 + ck_{\text{off}} & ck_{\text{on}} \\ ck_{\text{off}} & \gamma_0 + ck_{\text{on}} \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} \\ &= \frac{k_{\text{in}}}{(\gamma_0\gamma_1 + c\gamma_0k_{\text{off}} + c\gamma_1k_{\text{on}})(k_{\text{on}} + k_{\text{off}})} [k_{\text{off}} \quad k_{\text{on}}] \begin{bmatrix} \gamma_1 + c(k_{\text{on}} + k_{\text{off}}) \\ \gamma_0 + c(k_{\text{on}} + k_{\text{off}}) \end{bmatrix} \\ &= k_{\text{in}} \left[ \frac{\gamma_1k_{\text{off}} + \gamma_0k_{\text{on}} + c(k_{\text{on}} + k_{\text{off}})^2}{(\gamma_0\gamma_1 + c\gamma_0k_{\text{off}} + c\gamma_1k_{\text{on}})(k_{\text{on}} + k_{\text{off}})} \right]. \end{aligned} \quad (3.3)$$

It is interesting to point out the formal similarity of (3.3) with the formula for the mean transfer time of a relaxation process whose rate is modelled by a two-state CTMC [28]. From (3.3), it can be readily seen that the steady-state substrate means in the static and deterministic cases are given by

$$m_{\text{eq}}^{(\text{det})} = \lim_{c \rightarrow \infty} m_{\text{eq}}(c) = k_{\text{in}} \left[ \frac{k_{\text{on}} + k_{\text{off}}}{\gamma_0 k_{\text{off}} + \gamma_1 k_{\text{on}}} \right] \quad \text{and}$$

$$m_{\text{eq}}^{(\text{static})} = \lim_{c \rightarrow 0} m_{\text{eq}}(c) = k_{\text{in}} \left[ \frac{\gamma_1 k_{\text{off}} + \gamma_0 k_{\text{on}}}{\gamma_0 \gamma_1 (k_{\text{on}} + k_{\text{off}})} \right].$$

Hence, we can compute the maximum relative amplification factor  $\rho_{\text{max}}$  (2.32) as

$$\rho_{\text{max}} = \frac{m_{\text{eq}}^{(\text{static})}}{m_{\text{eq}}^{(\text{det})}} - 1 = \frac{(\gamma_0 k_{\text{off}} + \gamma_1 k_{\text{on}})(\gamma_1 k_{\text{off}} + \gamma_0 k_{\text{on}})}{(k_{\text{on}} + k_{\text{off}})^2 \gamma_0 \gamma_1} - 1$$

$$= \frac{k_{\text{on}} k_{\text{off}} (\gamma_0 - \gamma_1)^2}{(k_{\text{on}} + k_{\text{off}})^2 \gamma_0 \gamma_1}. \quad (3.4)$$

Recall the formula for  $\text{Var}_{\pi}(\gamma)$  from (3.2) and observe that  $\rho_{\text{max}}$  can be expressed as

$$\rho_{\text{max}} = \frac{\text{Var}_{\pi}(\gamma)}{\gamma_0 \gamma_1},$$

which reinforces the point we made in §2.4 that  $\rho_{\text{max}}$  serves as a proxy for the variance of the stationary distribution.

The Dirichlet form  $\theta$  (2.30) for this CTMC is given by

$$\theta = -[\pi_0 \quad \pi_1] \begin{bmatrix} \frac{1}{\gamma_0} & 0 \\ 0 & \frac{1}{\gamma_1} \end{bmatrix} Q \begin{bmatrix} \frac{1}{\gamma_0} & 0 \\ 0 & \frac{1}{\gamma_1} \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \left( \frac{\gamma_0 - \gamma_1}{\gamma_0 \gamma_1} \right)^2 \frac{k_{\text{on}} k_{\text{off}}}{k_{\text{on}} + k_{\text{off}}}.$$

This yields the following formula for the normalized Dirichlet form  $\theta$  (2.33):

$$\theta = \frac{\theta \mathbb{E}_{\pi}(\gamma)}{\rho_{\text{max}}} = \frac{k_{\text{off}} \gamma_0 + k_{\text{on}} \gamma_1}{\gamma_0 \gamma_1}$$

which determines the shape of our approximate formula  $\hat{m}_{\text{eq}}(c)$  (2.26) for the steady-state substrate mean. It is straightforward to check that for this example, this approximate formula is *exact* because the expression (3.3) for  $m_{\text{eq}}(c)$  can be written as

$$m_{\text{eq}}(c) = m_{\text{eq}}^{(\text{det})} + \left( \frac{m_{\text{eq}}^{(\text{static})} - m_{\text{eq}}^{(\text{det})}}{1 + \theta c} \right).$$

Note that  $\theta$  is a measure of the mixing strength of the enzymatic kinetics and so it is not surprising that it increases linearly with the transition rates  $k_{\text{on}}$  and  $k_{\text{off}}$ .

Define the relative amplification factor by (2.31). It can be exactly expressed as

$$\rho(c) = \frac{\rho_{\text{max}}}{1 + \theta c} = \frac{k_{\text{on}} k_{\text{off}} (\gamma_0 - \gamma_1)^2}{(k_{\text{on}} + k_{\text{off}})^2 (\gamma_0 \gamma_1 + c \gamma_0 k_{\text{off}} + c \gamma_1 k_{\text{on}})}.$$

We now consider the situation when the degradation rate induced by the enzyme E in the low-activity state ('0') is negligible. In this case,  $\gamma_0 \approx 0$  and  $\rho(c)$  simplifies to

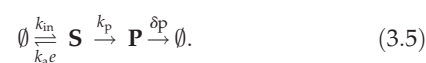
$$\rho(c) \approx \left( \frac{k_{\text{off}} \gamma_1}{(k_{\text{on}} + k_{\text{off}})^2} \right) \frac{1}{c} = \left( \frac{\pi_0 \gamma_1}{k_{\text{on}} + k_{\text{off}}} \right) \frac{1}{c},$$

which shows that the relative amplification factor is proportional to  $1/c$  and the proportionality constant is simply the product of the proportion of time ( $\pi_0$ ) the enzyme spends in the low-activity state, the degradation rate ( $\gamma_1$ ) at

the high-activity state and the reciprocal of the sum of transition rates  $k_{\text{on}}$  and  $k_{\text{off}}$ . In particular, as  $c$  approaches 0, the relative amplification factor  $\rho(c)$  can be enormous, thereby indicating that such a switching enzyme E can exploit its fluctuations to function as a biological amplifier with a very high gain.

## 3.2. Stochastic focusing network

In this section, we apply our results to the famous stochastic focusing network given in [22]. This network involves three species: substrate S, product P and enzyme E.<sup>10</sup> The molecules of substrate S are produced constitutively at rate  $k_{\text{in}}$  and converted into product P through a first-order reaction with rate constant  $k_{\text{p}}$ . Both substrate and product molecules degrade spontaneously at rates  $k_{\text{a}e}$  and  $\delta_{\text{p}}$ , respectively, where  $e$  denotes the current state or abundance level of enzyme E. The schematic of these reactions is as follows:



The enzymatic dynamics in this example is given by the Markovian birth–death process with birth-rate  $k_{\text{s}}$  and death-rate  $k_{\text{d}}$ :



This process evolves on state-space  $\mathbb{N}_0$ , which is the set of all non-negative integers and its unique stationary distribution is Poisson with mean  $k_{\text{s}}/k_{\text{d}}$ . We assume that the initial enzymatic state is a random variable with this stationary distribution.

Multiplying the rate constants  $k_{\text{s}}$  and  $k_{\text{d}}$  by  $c$ , we obtain enzymatic kinetics whose speed relative to the substrate is  $c$ . Let  $m_{\text{eq}}^{(\text{S})}(c)$  and  $m_{\text{eq}}^{(\text{P})}(c)$  denote the steady-state means of substrate and product, respectively, when the relative enzyme speed is  $c$ . From the *first-order moment equations* for the network (3.5), one can easily show (see the electronic supplementary material) that for any  $c \geq 0$

$$m_{\text{eq}}^{(\text{P})}(c) = \frac{k_{\text{p}}}{\delta_{\text{p}}} m_{\text{eq}}^{(\text{S})}(c). \quad (3.7)$$

To study the amplification of steady-state means due to enzymatic fluctuations, we use the relative stochastic amplification factor defined by (2.31). Owing to the linear relationship (3.7) between  $m_{\text{eq}}^{(\text{S})}(c)$  and  $m_{\text{eq}}^{(\text{P})}(c)$ , these amplification factors are the same for both product and substrate. Therefore, we can understand the amplification phenomenon by replacing network (3.5) with our simplified scheme (figure 1), where the degradation rate at time  $t$  is given by

$$\gamma_c(t) = k_{\text{p}} + k_{\text{a}} E_c(t),$$

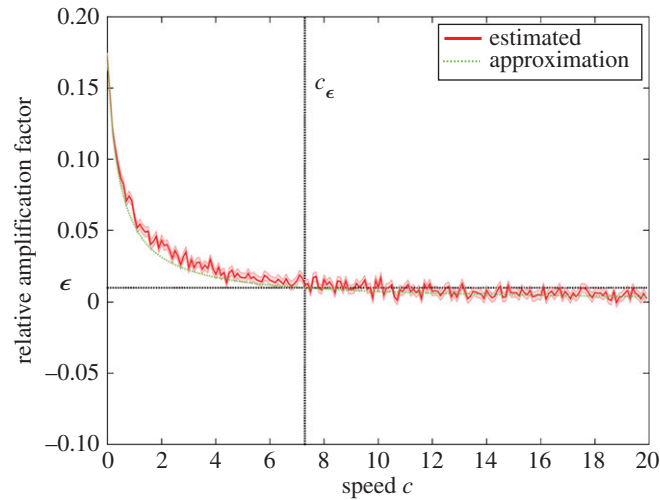
and  $E_c(t)$  denotes the state at time  $t$  of enzymatic kinetics with relative speed  $c$ . Let  $\gamma_i = k_{\text{p}} + k_{\text{a}} i$  for each  $i = 0, 1, \dots$ . Note that  $(\gamma_c(t))_{t \geq 0}$  is a CTMC with state space

$$\Gamma = k_{\text{p}} + k_{\text{a}} \mathbb{N}_0 = \{\gamma_1, \gamma_2, \dots\}$$

and stationary distribution<sup>11</sup>

$$\pi(\gamma_i) = \frac{e^{-k_{\text{s}}/k_{\text{d}}}}{i!} \left( \frac{k_{\text{s}}}{k_{\text{d}}} \right)^i \quad \text{for any } \gamma_i = k_{\text{p}} + k_{\text{a}} i \in \Gamma. \quad (3.8)$$

This CTMC transitions from state  $\gamma_i$  to state  $\gamma_{i+1}$  at rate  $ck_{\text{s}}$  and from state  $\gamma_i$  to state  $\gamma_{i-1}$  at rate  $ck_{\text{d}}$ . In other words,



**Figure 5.** Stochastic focusing network. Comparison of the estimated values of the exact relative amplification factor  $\rho(c)$  (2.31) with the approximate factor  $\hat{\rho}(c)$  obtained from formula (2.34). The threshold speed  $c_\epsilon$  for the 1% tolerance level ( $\epsilon = 0.01$ ) is also marked. The estimated values are obtained using the estimator based on formula (2.14) with  $10^5$  samples. The shaded region represents the symmetric 1 s.d. interval around the mean. (Online version in colour.)

the generator for this CTMC is given by

$$\mathbb{Q}_c f(\gamma_i) = ck_s(f(\gamma_{i+1}) - f(\gamma_i)) + cik_d(f(\gamma_{i-1}) - f(\gamma_i)) \quad (3.9)$$

for any bounded function  $f: \Gamma \rightarrow \mathbb{R}$ .

In the rest of this section, we denote the steady-state substrate mean by  $m_{\text{eq}}(c)$  instead of  $m_{\text{eq}}^{(S)}(c)$ . As the state-space  $\Gamma$  is not finite, we cannot use the results from §2.3 to compute  $m_{\text{eq}}(c)$ . However, we can easily simulate the paths of process  $(\gamma_c(t))_{t \geq 0}$  with Gillespie's algorithm [33], and obtain samples of the random variable  $\tau_c$  defined by (2.13). The corresponding sample mean then serves as an estimator for  $m_{\text{eq}}(c)$  (see part (C) of proposition 2.1). Note that the steady-state substrate mean in the absence of enzymatic fluctuations is simply given by

$$m_{\text{eq}}^{(\text{det})} = \frac{k_{\text{in}}}{k_p + k_a(k_s/k_d)}.$$

Dividing  $m_{\text{eq}}(c)$  by  $m_{\text{eq}}^{(\text{det})}$  and subtracting 1, we obtain an estimate for the relative stochastic amplification factor  $\rho(c)$  (see (2.31)). This factor only depends on four rate constants  $k_p$ ,  $k_a$ ,  $k_s$  and  $k_d$  which we now set as

$$k_p = 0.35, \quad k_a = 0.25, \quad k_d = 1 \quad \text{and} \quad k_s = 1. \quad (3.10)$$

We estimate  $\rho(c)$  for several values of  $c$  in the interval  $(0, 20)$  and plot these estimates in figure 5. For each value of  $c$ ,  $\rho(c)$  was estimated using  $10^5$  samples of  $\tau_c$  and the resulting standard error<sup>12</sup> is also displayed in figure 5. In §2.4, we develop an approximate expression  $\hat{\rho}(c)$  (2.34) for the relative amplification factor which is likely to hold even though the state-space  $\Gamma$  is not finite. Using the stationary distribution  $\pi$  (3.8) and the generator  $\mathbb{Q}_c$  (with  $c = 1$ ), we estimate the maximum amplification factor  $\rho_{\text{max}}$  (2.32) and the normalized Dirichlet form  $\theta$  (2.33) as  $\rho_{\text{max}} = 0.1703$  and  $\theta = 2.1988$ , respectively. With these values we evaluate the map  $c \mapsto \hat{\rho}(c)$  and plot it in the interval  $(0, 20)$  in figure 5. The close agreement between the estimated and the approximate values of the relative amplification factor can be easily seen. In figure 5, we also indicate the threshold speed  $c_\epsilon$  (2.35) for the 1% threshold level (i.e.  $\epsilon = 0.01$ ). This threshold speed is  $c_\epsilon = 7.2903$  which indicates that if  $c < 7.2903$ , then enzymatic fluctuations will amplify the steady-state substrate mean by more than 1% in comparison to the deterministic

case. In other words, the relative error in assuming that the enzyme activity is deterministic exceeds 1% if  $c < 7.2903$ .

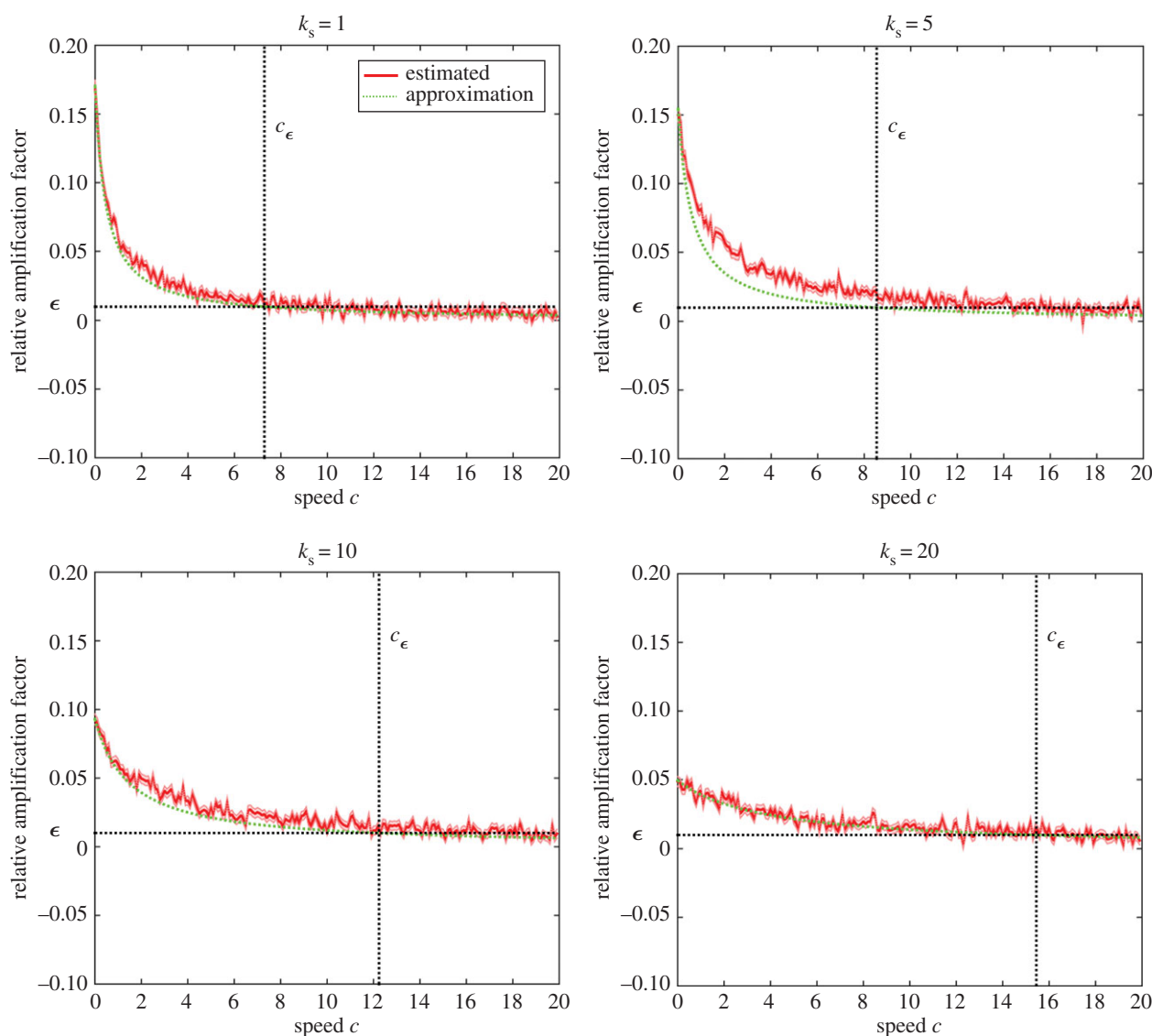
We now explore the effects of changing the levels of noise in the enzymatic activity on the relative amplification factor for the steady-state substrate mean. This noise can be measured using the *coefficient of variation* (CV)<sup>13</sup> of the stationary distribution for the enzyme abundance. As this distribution is Poisson with mean  $k_s/k_d$ , the CV is  $(\sqrt{k_d}/\sqrt{k_s})$  which shows that for a fixed  $k_d$ , we can decrease the relative noise level by simply increasing  $k_s$ . With this in mind, we repeat the above computations (figure 5) for three additional values of  $k_s$ , 5, 10 and 20, and the results are provided in figure 6. For each value of  $k_s$ , the corresponding estimates for the maximum relative amplification factor  $\rho_{\text{max}}$ , the normalized Dirichlet form  $\theta$  and the threshold speed  $c_\epsilon$  (for  $\epsilon = 0.01$ ) are given in table 1.

Recall the discussion at the end of §2.4 on the effects of noise in the enzymatic dynamics. From table 1, it is clear that as expected, decreasing noise (or increasing  $k_s$ ) results in the decline of both  $\rho_{\text{max}}$  and  $\theta$ . These parameters influence the threshold speed  $c_\epsilon$  (see (2.35)) in opposite ways, but their overall effect is to *increase*  $c_\epsilon$ , indicating that as the noise levels go down, the relative enzyme speed needs to be higher and higher for the assumption of deterministic enzymatic activity to be acceptable. In other words, even though noise in enzyme activity causes the stochastic amplification effect, it also helps in eliminating it.

## 4. Discussion

We examined the mathematical properties of a system consisting of a substrate that is degraded through an enzyme with stochastically fluctuating activity levels. Our analysis focused on the effect of enzymatic fluctuations on the mean substrate abundance and its deviations from the deterministic model predictions. It should be pointed out that even if the substrate inflow rate is assumed to be an independent stationary stochastic process with mean  $k_{\text{in}}$ , our results will not be affected.

Whereas a stochastically varying production rate would leave the mean substrate level unaffected and equal to that



**Figure 6.** Stochastic focusing network. Comparison of the relative amplification factors for various values of  $k_s$ . As  $k_s$  gets larger, the level of noise in the enzymatic dynamics decreases. Consequently, the maximum value  $\rho_{\max}$  (achieved at speed  $c = 0$ ) declines, but the values of the normalized Dirichlet form  $\theta$  also decline. This results in lesser convexity of the map  $c \mapsto \rho(c)$  which can be seen from the increasing values of the threshold speeds  $c_\epsilon$  for  $\epsilon = 0.01$ . (Online version in colour.)

**Table 1.** Estimates for  $\rho_{\max}$ ,  $\theta$  and  $c_\epsilon$  for various values of  $k_s$ .

$k_s$	$\rho_{\max}$	$\theta$	$c_\epsilon$
1	0.1703	2.1988	7.2903
5	0.1543	1.6907	8.5349
10	0.0930	0.6782	12.2383
20	0.0494	0.2550	15.4510

of the deterministic model, fluctuations in the removal rate of the substrate result in a system that behaves very differently in the stochastic and deterministic regimes due to the product term in the degradation rate of  $S$ . Our formulae help quantify this discrepancy and study its behaviour as the speed of enzymatic fluctuations varies from zero to infinity. They also provide an interesting connection between the amplification effect and the mixing properties of the Markov process describing the enzymatic activity fluctuations, which allow us to determine the speed above which this amplification becomes negligible for a given system parametrization.

Note that the study of such systems through the use of approximate stochastic models such as the linear noise approximation [42] is particularly challenging, since these methods typically fail to capture the very strong negative correlations between enzyme activity and substrate that can arise at slow enzyme fluctuations (see also discussion in [23]). On the contrary, the results presented here are valid under much milder simplifying assumptions, and can thus accurately reveal the magnitude of the discrepancy between stochastic and deterministic descriptions of the system.

Given the prevalence of enzymatic interactions in cell biology, stochastic fluctuations in enzyme activity and/or abundance are expected to play a large role in shaping the mean intracellular abundances of substrates [15], which could potentially also deviate significantly from the deterministically predicted amounts. Since many enzymes are allosterically regulated [43] by their products, substrates or other small signalling molecules, it would be very interesting to also study the effects of this regulation on the statistics of substrates and products, and examine potential noise reduction [23] or signal amplification strategies. As the sensitivity of single-molecule enzymology experimental techniques

increases, it may soon be possible to study the phenomena described theoretically in this work within living cells.

**Data accessibility.** Proofs of all derived mathematical results are provided in the electronic supplementary material.

**Author's contributions.** A.G. and A.M.-A. conceived the study, carried out the analysis, and wrote the paper with input and supervision from M.K.

**Competing interests.** The authors have no competing interests.

**Funding.** The work was partially supported by Swiss National Science Foundation grant 200021-157129.

## Endnotes

<sup>1</sup>A matrix  $Q$  is called *irreducible* if there does not exist a permutation matrix  $P$  such that the matrix  $PQP^{-1}$  is block upper-triangular.

<sup>2</sup>For a more rigorous definition of a stationary stochastic process, see the electronic supplementary material.

<sup>3</sup>The concentration of any species is its copy number divided by the system volume.

<sup>4</sup>An eigenvalue is said to be simple if its algebraic multiplicity is 1.

<sup>5</sup>A square matrix  $M$  is diagonalizable if it can be written as  $M = P\Lambda P^{-1}$  for some diagonal matrix  $\Lambda$  and some invertible matrix  $P$ . The diagonal entries of  $\Lambda$  are the eigenvalues of matrix  $M$ .

## References

- Lu HP, Xun L, Xie XS. 1998 Single-molecule enzymatic dynamics. *Science* **282**, 1877–1882. (doi:10.1126/science.282.5395.1877)
- Sauer M, Hofkens J, Enderlein J. 2011 *Single-molecule enzymatics*, pp. 241–271. New York, NY: Wiley-VCH.
- Lu HP. 2014 Sizing up single-molecule enzymatic conformational dynamics. *Chem. Soc. Rev.* **43**, 1118–1143. (doi:10.1039/C3CS60191A)
- Lu HP. 2005 Probing single-molecule protein conformational dynamics. *Acc. Chem. Res.* **38**, 557–565. (doi:10.1021/ar0401451)
- English BP, Min W, Van Oijen AM, Lee KT, Luo G, Sun H, Cherayil BJ, Kou SC, Xie XS. 2006 Ever-fluctuating single enzyme molecules: Michaelis-Menten equation revisited. *Nat. Chem. Biol.* **2**, 87–94. (doi:10.1038/nchembio759)
- Xie XS. 2002 Single-molecule approach to dispersed kinetics and dynamic disorder: probing conformational fluctuation and enzymatic dynamics. *J. Chem. Phys.* **117**, 11 024–11 032. (doi:10.1063/1.1521159)
- Zwanzig R. 1990 Rate processes with dynamical disorder. *Acc. Chem. Res.* **23**, 148–152. (doi:10.1021/ar00173a005)
- Dan N. 2007 Understanding dynamic disorder fluctuations in single-molecule enzymatic reactions. *Curr. Opin. Colloid Interface Sci.* **12**, 314–321. (doi:10.1016/j.cocis.2007.08.005)
- Zwanzig R. 1992 Dynamical disorder: passage through a fluctuating bottleneck. *J. Chem. Phys.* **97**, 3587–3589. (doi:10.1063/1.462993)
- Schenter GK, Lu HP, Xie XS. 1999 Statistical analyses and theoretical models of single-molecule enzymatic dynamics. *J. Phys. Chem. A* **103**, 10 477–10 488. (doi:10.1021/jp992324j)
- Kou S, Cherayil BJ, Min W, English BP, Xie XS. 2005 Single-molecule Michaelis-Menten equations. *J. Phys. Chem. B* **109**, 19 068–19 081. (doi:10.1021/jp051490q)
- Reichman DR. 2006 On stochastic models of dynamic disorder. *J. Phys. Chem. B* **110**, 19 061–19 065. (doi:10.1021/jp061992j)
- Min W, English BP, Luo G, Cherayil BJ, Kou S, Xie XS. 2005 Fluctuating enzymes: lessons from single-molecule studies. *Acc. Chem. Res.* **38**, 923–931. (doi:10.1021/ar040133f)
- Tan YW, Yang H. 2011 Seeing the forest for the trees: fluorescence studies of single enzymes in the context of ensemble experiments. *Phys. Chem. Chem. Phys.* **13**, 1709–1721. (doi:10.1039/C0CP02412K)
- Iversen L *et al.* 2014 Ras activation by SOS: allosteric regulation by altered fluctuation dynamics. *Science* **345**, 50–54. (doi:10.1126/science.1250373)
- Grima R, Walter NG, Schnell S. 2014 Single-molecule enzymology à la Michaelis–Menten. *FEBS J.* **281**, 518–530. (doi:10.1111/febs.12663)
- Stefanini M, McKane A, Newman T. 2005 Single enzyme pathways and substrate fluctuations. *Nonlinearity* **18**, 1575–1595. (doi:10.1088/0951-7715/18/4/008)
- Saloff-Coste L. 1997 Lectures on finite Markov chains. In *Lectures on probability theory and statistics: école d'été de probabilités de Saint-Flour XXVI-1996*, pp. 301–413. Berlin, Germany: Springer.
- Elowitz MB, Levine AJ, Siggia ED, Swain PS. 2002 Stochastic gene expression in a single cell. *Science* **297**, 1183–1186. (doi:10.1126/science.1070919)
- Oyarzún DA, Lugagne JB, Stan GBV. 2015 Noise propagation in synthetic gene circuits for metabolic control. *ACS Synth. Biol.* **4**, 116–125. (doi:10.1021/sb400126a)
- Shea LD, Omann GM, Linderman JJ. 1997 Calculation of diffusion-limited kinetics for the reactions in collision coupling and receptor cross-linking. *Biophys. J.* **73**, 2949–2959. (doi:10.1016/S0006-3495(97)78323-1)
- Paulsson J, Berg OG, Ehrenberg M. 2000 Stochastic focusing: fluctuation-enhanced sensitivity of intracellular regulation. *Proc. Natl Acad. Sci. USA* **97**, 7148–7153. (doi:10.1073/pnas.110057697)
- Miliadis-Argeitis A, Engblom S, Bauer P, Khammash M. 2015 Stochastic focusing coupled with negative feedback enables robust regulation in biochemical reaction networks. *J. R. Soc. Interface* **12**, 20150831. (doi:10.1098/rsif.2015.0831)
- Anderson PW. 1954 A mathematical model for the narrowing of spectral lines by exchange or motion. *J. Phys. Soc. Japan* **9**, 316–339. (doi:10.1143/JPSJ.9.316)
- Kubo R. 1954 Note on the stochastic theory of resonance absorption. *J. Phys. Soc. Japan* **9**, 935–944. (doi:10.1143/JPSJ.9.935)
- Goychuk I. 2004 Quantum dynamics with non-Markovian fluctuating parameters. *Phys. Rev. E* **70**, 016109. (doi:10.1103/PhysRevE.70.016109)
- Goychuk I, Hänggi P. 2005 Quantum dynamics in strong fluctuating fields. *Adv. Phys.* **54**, 525–584. (doi:10.1080/00018730500429701)
- Goychuk I. 2005 Rate processes with non-Markovian dynamical disorder. *J. Chem. Phys.* **122**, 164506. (doi:10.1063/1.1884087)
- Norris JR. 1998 *Markov chains*. Cambridge, UK: Cambridge University Press.
- Ethier SN, Kurtz TG. 1986 *Markov processes: characterization and convergence*. Wiley Series in Probability and Mathematical Statistics. New York, NY: John Wiley & Sons Inc.
- Kurtz TG. 1972 The relationship between stochastic and deterministic models for chemical reactions. *J. Chem. Phys.* **57**, 2976–2978. (doi:10.1063/1.1678692)

32. Kallenberg O. 2002 *Foundations of modern probability*, 2nd edn. New York, NY: Springer.
33. Gillespie DT. 1977 Exact stochastic simulation of coupled chemical reactions. *J. Phys. Chem.* **81**, 2340–2361. (doi:10.1021/j100540a008)
34. Asmussen S, Glynn PW. 2007 *Stochastic simulation: algorithms and analysis. Stochastic modelling and applied probability, no. 57*. New York, NY: Springer.
35. Gupta A, Khammash M. 2014 Sensitivity analysis for stochastic chemical reaction networks with multiple time-scales. *Electron. J. Probab.* **19**, 1–53. (doi:10.1214/EJP.v19-3246)
36. Hasenauer J, Wolf V, Kazerooni A, Theis F. 2014 Method of conditional moments (MCM) for the chemical master equation. *J. Math. Biol.* **69**, 687–735. (doi:10.1007/s00285-013-0711-5)
37. Kato T. 1995 *Perturbation theory for linear operators*. Classics in Mathematics. Berlin, Germany: Springer.
38. Hoffman K, Kunze R. 1971 *Linear algebra*, 2nd edn. Englewood Cliffs, NJ: Prentice-Hall Inc.
39. Levin DA, Peres Y, Wilmer EL. 2009 *Markov chains and mixing times*. Providence, RI: American Mathematical Society. With a chapter on coupling from the past by James G. Propp and David B. Wilson. See <http://opac.inria.fr/record=b1128575>.
40. Costarelli D, Spigler R. 2015 How sharp is the Jensen inequality? *J. Inequalities Appl.* **2015**, 69. (doi:10.1186/s13660-015-0591-x)
41. Pincus SM. 1992 Approximating Markov chains. *Proc. Natl Acad. Sci. USA* **89**, 4432–4436. (doi:10.1073/pnas.89.10.4432)
42. Elf J, Ehrenberg M. 2003 Fast evaluation of fluctuations in biochemical networks with the linear noise approximation. *Genome Res.* **13**, 2475–2484. (doi:10.1101/gr.1196503)
43. Liu J, Nussinov R. 2016 Allosteric: an overview of its history, concepts, methods, and applications. *PLoS Comput. Biol.* **12**, e1004966. (doi:10.1371/journal.pcbi.1004966)