The Journal of Chemical Physics

Transients generate memory and break hyperbolicity in stochastic enzymatic networks

Cite as: J. Chem. Phys. **154**, 035101 (2021); https://doi.org/10.1063/5.0031368
Submitted: 30 September 2020 . Accepted: 22 December 2020 . Published Online: 15 January 2021

Ashutosh Kumar, R. Adhikari, and Arti Dua

The Journal of Chemical Physics ARTICLE scitation.org/journal/jcp

Transients generate memory and break hyperbolicity in stochastic enzymatic networks

Cite as: J. Chem. Phys. 154, 035101 (2021); doi: 10.1063/5.0031368
Submitted: 30 September 2020 • Accepted: 22 December 2020 •
Published Online: 15 January 2021

Ashutosh Kumar,[1] R. Adhikari,[2] and Arti Dua[1,a)]

AFFILIATIONS

[1] Department of Chemistry, Indian Institute of Technology, Madras, Chennai 600036, India
[2] DAMTP, Centre for Mathematical Sciences, University of Cambridge, Wilberforce Road, Cambridge CB3 0WA, United Kingdom

Note: This paper is part of the JCP Special Collection in Honor of Women in Chemical Physics and Physical Chemistry.
[a)] **Author to whom correspondence should be addressed:** arti@iitm.ac.in

ABSTRACT

The hyperbolic dependence of catalytic rate on substrate concentration is a classical result in enzyme kinetics, quantified by the celebrated Michaelis–Menten equation. The ubiquity of this relation in diverse chemical and biological contexts has recently been rationalized by a graph-theoretic analysis of deterministic reaction networks. Experiments, however, have revealed that "molecular noise"—intrinsic stochasticity at the molecular scale—leads to significant deviations from classical results and to unexpected effects like "molecular memory," i.e., the breakdown of statistical independence between turnover events. Here, we show, through a new method of analysis, that memory and non-hyperbolicity have a common source in an initial, and observably long, transient peculiar to stochastic reaction networks of multiple enzymes. Networks of single enzymes do not admit such transients. The transient yields, asymptotically, to a steady-state in which memory vanishes and hyperbolicity is recovered. We propose new statistical measures, defined in terms of turnover times, to distinguish between the transient and steady-states and apply these to experimental data from a landmark experiment that first observed molecular memory in a single enzyme with multiple binding sites. Our study shows that catalysis at the molecular level with more than one enzyme always contains a non-classical regime and provides insight on how the classical limit is attained.

I. INTRODUCTION

The Michaelis–Menten equation (MME), describing the hyperbolic dependence of the rate of catalysis on the substrate concentration, is a classical result in enzyme kinetics.[1] It was derived by Michaelis and Menten in 1913 for a network of three elementary reactions, $E + S \rightleftharpoons ES \rightarrow E + P$, describing the reversible binding of enzyme E with substrate S to form complex ES and its irreversible dissociation into product P and regenerated enzyme E.[2–4] The hyperbolic dependence of catalytic rate on substrate concentration is found to hold in enzymatic networks of far greater complexity. It implies a linear relation between the inverse catalytic rate and the inverse substrate concentration and, in this form, is widely used to estimate rate parameters and infer mechanisms from kinetic data.[5] The surprising ubiquity of this equation in chemical and biological processes has recently been rationalized by a graph-theoretical analysis of complex, deterministic, reaction networks.[6]

At the molecular level, however, enzymatic reactions do not proceed deterministically.[7–12] Fluctuations of both the quantum mechanical and thermal origin, termed "molecular noise," influence each step of a chemical reaction such that neither the lifetime of a chemical state nor the state to which it transits can be known with certainty.[13–17] Furthermore, the discrete change in the reactant numbers is comparable to the number of reacting molecules, and a description in terms of continuously varying concentrations is inadmissible.[13–19] In the limit of large numbers of reactants, when both fluctuations and the change in reactants compared to their total number are small, a deterministic description in terms of continuously varying concentrations is recovered.[19] In addition, the Michaelis–Menten equation is obtained when there is a separation of timescales between the (rapid) equilibration of the enzyme and complex and (slow) product formation.[2] This rapid equilibrium approximation is a special case of the steady-state approximation (SSA), in which the rates of complex formation and dissociation are assumed to be equal, as noted by Briggs and Haldane.[20]

The first theoretical study of catalytic fluctuations was undertaken by Bartholomay half a century after the discovery of the Michaelis–Menten (MM) equation.[19] His principal contribution was to show that discrete-state continuous-time Markov processes provide a mathematical framework that incorporates the discrete change in molecular numbers, the effect of molecular noise in each reaction step, and reaction mechanisms of arbitrary complexity. The classical rate equations for concentrations were thus replaced by chemical "master equations" for the probabilities of the (non-negative) number of reactants. Bartholomay obtained the mean and variance of these for the Michaelis–Menten mechanism $E + S \rightleftharpoons ES \rightarrow E + P$. The apparent irreproducibility of experiments that measured the rate of change of concentrations was recognized to be a fluctuation effect, and a method was suggested to estimate the rate constraints from the variances of the concentrations.

The long hiatus of interest that followed this pioneering work was brought to a close by a landmark experiment that directly observed catalytic fluctuations at the single-molecule level.[10,11] As concentrations are not defined for a single molecule, the experiment measured, instead, the times at which the enzyme yielded products, one product at a time. These time series data were analyzed in terms of the interval between consecutive turnovers, defined to be the "waiting time." For repeated experiments under identical conditions, the waiting times showed a distribution and this was attributed to the effect of molecular noise. The analysis of waiting time distributions revealed several remarkable facts. First, the distribution changed character with an increase in substrate concentration, from a single exponential to one that was not. Second, the inverse of the mean waiting time obeyed the Michaelis–Menten equation (MME) at low substrate concentrations. Third, the randomness parameter, the ratio of the variance to the squared mean, was a monotonically increasing function of the substrate concentration bounded below by one. Fourth, the waiting times between consecutive turnovers were found to be statistically dependent with substantial positive correlations, an effect termed "molecular memory." Subsequent experiments in single-nanoparticle catalysis confirmed these empirical facts and established their generality.[21–23] While it was understood that these seemingly disparate observations have their origin in molecular noise, the precise manner in which they emerge from underlying molecular fluctuations and how they are influenced by different reaction mechanisms was not elucidated.

The central theoretical question that needs to be answered in rationalizing such single-molecule temporal data is the following: Can we derive the statistics of *temporal* fluctuations from the chemical master equation (CME), incorporating discreteness, molecular noise, and reaction mechanisms, in the manner that the statistics of *number* fluctuations was derived by Bartholomay? Here, we present a formalism that permits us to answer this question affirmatively. Using this formalism, we are able to make a direct connection between reaction mechanisms and the statistics of waiting times and, thus, explain their puzzling features from a unified point of view.

In Sec. II, we consider a generic stochastic enzymatic reaction network, incorporating conformational fluctuations and parallel pathways to product formation, and present the corresponding chemical master equation (CME). We marginalize the reactant probabilities to obtain the probability of there being n turnovers at any given time and present several experimentally relevant summary statistics. We introduce the probability distributions of the turnover and waiting times and present their relevant summary statistics. We then derive an expression that connects the reactant probabilities of the CME to the distribution of waiting times. This provides the sought-after link between the description in terms of waiting times (point process),[24] in which experimental data are naturally recorded, and the description in terms of reactant numbers (counting process),[25] through which mechanisms are most conveniently expressed.

In Sec. III, we apply this formalism to study a reaction network corresponding to a single enzyme. In such a network, reactant numbers are either zero or one, and a non-zero value of one reactant number implies zero values of all others. We explore the consequences of this "fermionic" character and find that, irrespective of the complexity of the network, turnovers are always statistically independent and identically distributed or, in other words, constitute a renewal process.[26] A single-enzyme network, then, cannot show molecular memory.

In Sec. IV, we consider a network consisting of replicas of single-enzyme networks, corresponding to oligomeric enzymes with independent and identical binding sites. The absence of "fermionic" character in these networks permits turnovers to be statistically dependent and allows them to show molecular memory. The statistical dependence decreases with the number of turnovers and vanishes asymptotically. We characterize this transient with fading memory through the conditional distribution of consecutive turnovers, which we relate to measures of the single-enzyme network. This analysis explains the counter-intuitive appearance of memory in a process whose elementary steps, recalling that the CME describes a Markov process, are memoryless.

In Sec. V, we discuss new statistical measures that, contrary to existing measures,[27–29] do not assume the statistical independence of turnovers. Our measures, then, can be applied uniformly over the entire duration of the catalytic process, both in the transient state with memory and the steady-state in which the memory vanishes. We provide an expression for the enzymatic velocity in terms of turnover times, which reduces to the classical expression in the thermodynamic limit and elucidates how this limit is reached.

In Sec. VI, we compare our theory with the classic experiment on β-galactosidase,[10] a tetrameric enzyme with independent sites, and find excellent agreement with four replicas of a single-enzyme network with conformers and parallel pathways. Saliently, we do not need to assume any *ad hoc* distribution of reaction rates:[30,31] the "dynamic disorder" implied by such a distribution is an emergent feature of our theory.

We summarize our results in Sec. VII with a brief discussion on how our theory compares with stochastic kinetic models that make an *a priori* assumption of dynamic disorder in the reaction pathway. We conclude, in Sec. VIII, with a discussion on how our theory can be extended to non-replica networks corresponding to enzymes with interacting binding sites.

II. STOCHASTIC ENZYMATIC NETWORKS

The stochastic description of chemical reactions begins with a set of k non-negative integers $\boldsymbol{n} = (n_1, \ldots, n_k)$ describing the number

of molecules of the k-th species. Elementary reactions,

$$\sigma\text{-th reaction step: } \boldsymbol{n} \underset{t_\sigma^{\leftarrow}}{\overset{t_\sigma^{\rightarrow}}{\rightleftharpoons}} \boldsymbol{n} + \boldsymbol{r}_\sigma, \tag{1}$$

labeled by the index σ, take the state $\boldsymbol{n}$ to the state $\boldsymbol{n} + \boldsymbol{r}_\sigma$, where $\boldsymbol{r}_\sigma$ is a vector representing the integer changes of each species, as determined by the reaction stoichiometry. The probability per unit time that this reaction takes place is $t_\sigma^{\rightarrow}(\boldsymbol{n})$. The corresponding backward reaction takes the state $\boldsymbol{n} + \boldsymbol{r}_\sigma$ to the state $\boldsymbol{n}$ at the rate $t_\sigma^{\leftarrow}(\boldsymbol{n}+\boldsymbol{r}_\sigma)$. The rates are combinatoric functions that follow from the law of mass action. The probability $P(\boldsymbol{n}, t)$ of being in the state $\boldsymbol{n}$ at time t is governed by the CME,[19,25,32]

$$\begin{aligned}\partial_t P(\boldsymbol{n}) = &\sum_\sigma t_\sigma^{\rightarrow}(\boldsymbol{n}-\boldsymbol{r}_\sigma)P(\boldsymbol{n}-\boldsymbol{r}_\sigma) - t_\sigma^{\leftarrow}(\boldsymbol{n})P(\boldsymbol{n}) \\ &+ \sum_\sigma t_\sigma^{\leftarrow}(\boldsymbol{n}+\boldsymbol{r}_\sigma)P(\boldsymbol{n}+\boldsymbol{r}_\sigma) - t_\sigma^{\rightarrow}(\boldsymbol{n})P(\boldsymbol{n}),\end{aligned} \tag{2}$$

which is a system of coupled ordinary differential equations, equal in number to the number of distinct states of the network.

Here, we consider enzymatic networks that contain the Michaelis–Menten mechanism $E_i + S \rightleftharpoons ES_i \to E_i + P$ as a basic motif while allowing for conformers $i = \alpha, \beta, \ldots$ and parallel pathways to product formation.[10,11] The state is described by the vector $\boldsymbol{n} = (n_{E_\alpha}, n_{ES_\alpha}, n_{E_\beta}, n_{ES_\beta}, \ldots, n)$ of non-negative integers comprising of, in obvious notation, the number of enzyme and complex, of each conformational type, and of products. Examples of such networks for the simplest case of two conformers are shown in Fig. 1 for both parallel and off-pathway kinetics.[33] The corresponding rates are listed in Table I. The bimolecular complexation steps are replaced by pseudo-unimolecular steps with effective rate constants denoted by primes. All rates are then linear in the state vector $\boldsymbol{n}$. It is important to note that the rates do not depend on the number of products.

TABLE I. Elementary reaction steps and their rates for the single-site network labeled (a) in Fig. 1. The networks labeled (b) through (c) are obtained by setting corresponding rate parameters to zero. The forward reaction takes the state $\boldsymbol{n} = (n_{E_\alpha}, n_{ES_\alpha}, n_{E_\beta}, n_{ES_\beta}, n)$ to the state $\boldsymbol{n} + \boldsymbol{r}_\sigma$. The pseudo-first-order rate constants for the forward reaction are $k'_\alpha = k_\alpha[S]$ and $k'_\beta = k_\beta[S]$. All rates are independent of the number of products n.

Step	$\boldsymbol{r}_\sigma$	$t_\sigma^{\rightarrow}(\boldsymbol{n})$	$t_\sigma^{\leftarrow}(\boldsymbol{n})$
$E_\alpha \rightleftharpoons E_\beta$	$(-1, 0, 1, 0, 0)$	$\gamma_{\alpha\beta} n_{E_\alpha}$	$\gamma_{\alpha\beta} n_{E_\beta}$
$ES_\alpha \rightleftharpoons ES_\beta$	$(0, -1, 0, 1, 0)$	$\delta_{\alpha\beta} n_{ES_\alpha}$	$\delta_{\alpha\beta} n_{ES_\beta}$
$E_\alpha \rightleftharpoons ES_\alpha$	$(-1, 1, 0, 0, 0)$	$k'_\alpha n_{E_\alpha}$	$k_{-\alpha} n_{ES_\alpha}$
$E_\beta \rightleftharpoons ES_\beta$	$(0, 0, -1, 1, 0)$	$k'_\beta n_{E_\beta}$	$k_{-\beta} n_{ES_\beta}$
$ES_\alpha \to P + E_\alpha$	$(1, -1, 0, 0, 1)$	$k_{p\alpha} n_{ES_\alpha}$	0
$ES_\beta \to P + E_\beta$	$(0, 0, 1, -1, 1)$	$k_{p\beta} n_{ES_\beta}$	0

It is convenient to partition the state vector into $\boldsymbol{n} = (\boldsymbol{n}^\star, n)$, where $\boldsymbol{n}^\star = (n_{E_\alpha}, n_{ES_\alpha}, n_{E_\beta}, n_{ES_\beta}, \ldots)$ are "hidden" state components unobserved in experiment and n is the "observed" product state visible through fluorescence bursts. The hidden state vector has $2l$ components in a network with l conformers. For a network with v enzymes, or one oligomeric enzyme with v active sites, mass conservation implies that the sum of the number of enzymes in the uncomplexed and complexed states must sum to v: $n_{E_\alpha} + n_{E_\beta} + \cdots + n_{ES_\alpha} + n_{ES_\beta} + \cdots = v$. For a single enzyme (or active site), this implies that $n_{E_\alpha} + n_{E_\beta} + \cdots + n_{ES_\alpha} + n_{ES_\beta} + \cdots = 1$. Therefore, the components of the hidden state vector in a single-enzyme network have a "fermionic" character, where the components only take the values zero or one and only one component is non-zero at any point in time. Mass conservation also implies that the number of hidden states is finite and equal to the number of compositions of v into l parts. For a single enzyme with l conformers, this gives $2l$ states. We shall return to this important property below.

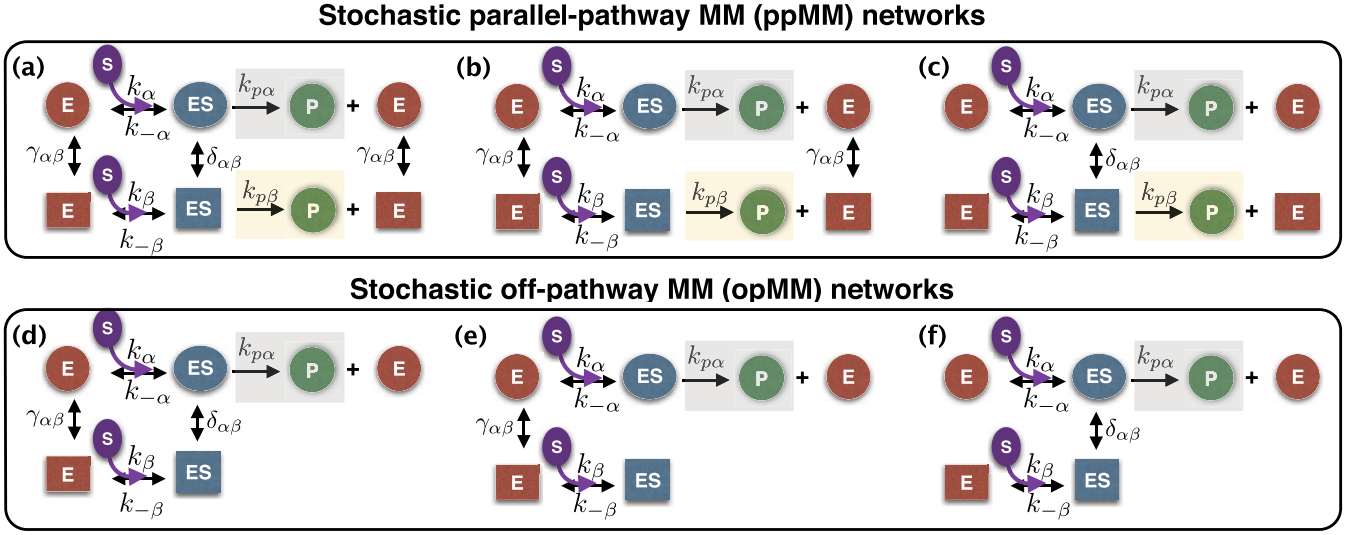

FIG. 1. Enzymatic networks for a single catalytic site including conformational fluctuations and parallel pathways to product formation. The general network in (a) reduces to the special cases in (b) through (f) when rate constants for the corresponding steps are set to zero.

Unlike the hidden components, the number of products n can take values from zero to infinity. The probability of their being n products at time t is obtained by marginalizing the reactant probability over the hidden states,

$$P(n,t) = \sum_{\boldsymbol{n}^\star} P(\boldsymbol{n},t). \tag{3}$$

This is the fundamental probability distribution in the counting process description of turnovers. The expectation with respect to this probability distribution of the mean and variance of n defines the enzymatic velocity and Fano factor,[34]

$$V(t) = \frac{d}{dt}\langle n\rangle, \quad \rho(t) = \frac{\langle n^2\rangle - \langle n\rangle^2}{\langle n\rangle} \quad (t \geq 0). \tag{4}$$

Such quantities have been calculated for a variety of networks beginning with the work of Bartholomay. However, as mentioned in the Introduction, they are not directly relevant to single-enzyme experiments, which record the times T_p at which turnovers occur, rather than the number of turnovers at time t. Here, $p = 1, 2, \ldots$ is the turnover number index. This motivates the study of the point process of turnovers, for which we now introduce the fundamental probability distributions.[24]

We define the turnover time for the p-th product, T_p, to be the smallest value of t such that $n \geq p$, or more precisely, $T_p = \inf\{t > 0: n(t) \geq p\}$. The cumulative distribution of the p-th turnover time T_p is denoted by $P(T_p \leq t)$. This defines the survival probability $P(T_p > t) = 1 - P(T_p \leq t)$ and probability density $w_{T_p}(t)dt = P(t < T_p \leq t+dt)$. The expectation of T_p with respect to the probability density defines the mean turnover time and the randomness parameter for the p-th turnover,

$$\mu_p = \langle T_p\rangle, \quad r_p = p\frac{\langle T_p^2\rangle - \langle T_p\rangle^2}{\langle T_p\rangle^2} \quad (p = 1, 2, \ldots). \tag{5}$$

While the current definition of the randomness parameter is p independent, $r = \langle\tau^2\rangle - \langle\tau\rangle^2/\langle\tau\rangle^2$,[35,36] the factor of p in the above definition of the randomness parameter is introduced for reasons that will become apparent in Sec. V. Higher moments can be studied but, to the best of our knowledge, have not been measured in experiment.

To quantify statistical dependences, it is convenient to define the waiting time between turnovers,

$$\tau_p = T_p - T_{p-1}, \tag{6}$$

and study their joint density distributions, $w(\tau_1, \tau_2, \ldots)$. The marginal distributions $w(\tau_p)$ describe the statistics of individual turnovers, while the joint distributions $w(\tau_p, \tau_q)$ describe the statistics of pairs of turnovers. It is convenient to write this joint distribution as

$$w(\tau_p,\tau_q) = w(\tau_p)g(\tau_p,\tau_q)w(\tau_q) \tag{7}$$

so that statistical dependences are contained in $g(\tau_p, \tau_q)$. The pairs of turnovers are statistically independent if and only if $g(\tau_p, \tau_q) = 1$ for all p and q. The correlation function

$$C_{pq} = \langle\tau_p\tau_q\rangle = c(\tau_p,\tau_q)\langle\tau_p\rangle\langle\tau_q\rangle \tag{8}$$

serves as a second-order statistic for identifying statistical dependences. The expectations are with respect to the joint distribution $w(\tau_p, \tau_q)$. Statistical dependences and molecular memory imply $c(\tau_p, \tau_q) \neq 1$.

The question naturally arises as to how the probability distributions for the counting process, $P(n, t)$, and the point process, $P(T_p \leq t)$, together with their summary statistics, are related to each other and to the underlying CME, which is the generative process that underlies both distributions. We provide the answer below.

From the definitions of the random variables n and T_p, it is clear that at any time t,

$$T_p \leq t \Longleftrightarrow n(t) \geq p, \tag{9}$$

or in other words, these two events are equal in probability. Since the event $n(t) < p$ is their complement, we have

$$P(T_p \leq t) = P(n \geq p, t) = 1 - P(n < p, t). \tag{10}$$

Since the product states are mutually exclusive, we have $P(n < p, t) = \sum_{n=0}^{p-1} P(n, t)$. Combining this with the marginal expression for $P(n, t)$, we obtain

$$P(T_p \leq t) = 1 - \sum_{n=0}^{p-1}\sum_{\boldsymbol{n}^\star} P(\boldsymbol{n},t). \tag{11}$$

This relation between the turnover time distribution and the solution of the CME is the central result of this section. It is applicable to networks of arbitrary complexity and provides the sought-after connection between the statistics of turnovers and reaction mechanisms. The probability density follows upon differentiation,

$$w_{T_p}(t) = -\sum_{n=0}^{p-1}\sum_{\boldsymbol{n}^\star} \partial_t P(\boldsymbol{n},t), \tag{12}$$

and is often more convenient for comparison with experimental data, when the latter is presented in the form of a probability density. A special case of this relation was first obtained in Ref. 13.

We have not been able to find a relation of this generality that relates the waiting time distributions $w(\tau_p)$ and $w(\tau_p, \tau_q)$ to the solution of the CME. However, in particular instances, where the network is "fermionic" or has a "replica" character, relations to the underlying CME can be found, as we show in Secs. III and IV, respectively.

III. RENEWAL STATISTICS IN SINGLE-ENZYME NETWORKS

As we noted above, hidden states in a single-enzyme network have a "fermionic" character: the components of the hidden state vector can only take the values zero or one, and only one component can be non-zero at any time. This implies that immediately after the conclusion of a turnover, say, the p-th, the network is in a state corresponding to a single uncomplexed enzyme. To elaborate, consider the MM network with state vector $\boldsymbol{n} = (n_E, n_{ES}, n)$ and hidden state vector $\boldsymbol{n}^\star = (n_E, n_{ES})$. The two allowed hidden states are (1, 0) and (0, 1) corresponding to the uncomplexed and complexed enzyme. Labeling these by E and ES, the allowed states of the network are

(E, n) and (ES, n). At the conclusion of the p-th turnover at $t = T_p$, the network is in the state (E, p), and so taking limits from above,

$$\lim_{t\to T_p^+} P(E,p,t) = 1. \tag{13}$$

For the two-conformer ppMM network with state vector $\boldsymbol{n} = (n_{E_\alpha}, n_{ES_\alpha}, n_{E_\beta}, n_{ES_\beta}, n)$ and hidden state vector $\boldsymbol{n}^\star = (n_{E_\alpha}, n_{ES_\alpha}, n_{E_\beta}, n_{ES_\beta})$, the four allowed hidden states are (1, 0, 0, 0), (0, 1, 0, 0), (0, 0, 1, 0), and (0, 0, 0, 1). Labeling these by E_α, ES_α, E_β, and ES_β, the allowed states of the network are (E_α, n), (ES_α, n), (E_β, n), and (ES_β, n). At the conclusion of the p-th turnover, the network is either in the state (E_α, p) or in the state (E_β, p), and so

$$\lim_{t\to T_p^+} P(E_\alpha,p,t) + P(E_\beta,p,t) = 1. \tag{14}$$

More generally, for any single-enzyme network, the states can be labeled by the conformation of the enzyme and the number of products, and at the conclusion of a turnover, the network is surely in one of the uncomplexed states with the total probability being partitioned between those states. As we show below, this recurrent return to a fixed subset of the hidden states, together with the structure of the CME for such networks, implies that conditioning on a turnover makes the future independent of the past. This results in turnovers that are statistically independent, with waiting time distributions that are identically distributed. Single-enzyme turnovers, therefore, form a renewal process and cannot show memory.[26]

In the absence of memory, attention can be focussed entirely on the statistics of the waiting time, which, since it is identically distributed for all p, we simply denote by τ. We summarize our two main results before providing explicit results for the MM and ppMM networks. First, we show that for a model with l conformers, the waiting time distribution is a sum of $2l$ exponentials whose time constants are eigenvalues of a $2l \times 2l$ matrix related to the CME. The behavior of the waiting time distribution is related to the spacing of these eigenvalues. For well-separated eigenvalues, the exponential with the lowest time constant is dominant, but for closely spaced eigenvalues, all the $2l$ exponentials contribute. This multiexponentiality leads to variances that are large compared to the squared mean and, hence, to a randomness parameter that exceeds unity. Our analysis thus transparently relates "dynamic disorder" to reaction mechanisms with fixed rate constants, in contrast to fluctuating rate parameters.[30,31] Second, we show that the randomness parameter for τ is a sensitive measure of network topology. It is known that a Markov chain comprising of a linear network of arbitrary complexity always yields $r \le 1$, bounded below by the inverse of the number of rate determining steps σ: $r \ge \frac{1}{\sigma}$.[37] The minimum is attained for a linear sequence of states with equal rate constants, first studied by Erlang.[38] For a single step reaction or a linear network with a single rate determining step, thus, $r = 1$. We find that a branched topology is necessary, but not sufficient, to obtain $r > 1$. Our explicit calculation for the ppMM network shows that r can vary continuously from $r < 1$ to $r > 1$ as the substrate concentration is increased. Thus, networks can be rationally designed to yield a desired value of the randomness parameter.

Consider, now, the CME for the MM network written in terms of the labels E, ES, and n,

$$\begin{aligned}
\partial_t P(E,n) &= -k_a P(E,n) + k_{-1} P(ES,n) + k_2 P(ES,n-1),\\
\partial_t P(ES,n) &= +k_a P(E,n) - k_{-1} P(ES,n) - k_2 P(ES,n),\\
&\qquad n = 0,1,2,\ldots.
\end{aligned} \tag{15}$$

It is understood that states with $n < 0$ have zero probability. This is an infinite system of autonomous linear differential equations whose solution can be obtained using the technique of generating functions. A great simplification results when we recognize the following two features. First, conditioning the system on the p-th turnover at $t = T_p$ collapses the probability on the state (E, p) so that $P(E, p) = 1$ and all other probabilities are zero. Since probabilities only flow into states with increasing number of products, this implies that probabilities of all states with $n < p$ remain zero subsequently. The future is made conditionally independent of the past. Second, the distribution of the $(p + 1)$th turnover T_{p+1}, conditioned on the p-th turnover at T_p, is governed by the same set of equations and initial conditions as the first turnover T_1, conditioned on the initial state at $t = 0$. This implies that T_1 and $T_{p+1} - T_p$ are equal in distribution for all p. Therefore, the waiting times τ_p are independent and distributed identically to T_1.

From Eq. (11), the cumulative distribution of T_1 is

$$P(T_1 < t) = 1 - [P(E,0,t) + P(ES,0,t)], \tag{16}$$

and from the CME, these two probabilities obey

$$\begin{bmatrix} \partial_t P(E,0,t) \\ \partial_t P(ES,,0,t) \end{bmatrix} = \begin{bmatrix} -k_1' & k_{-1} \\ k_1' & -(k_{-1}+k_2) \end{bmatrix} \begin{bmatrix} P(E,0,t) \\ P(ES,0,t) \end{bmatrix}$$

with initial condition $P(E, 0) = 1$ and $P(ES, 0) = 0$ at $t = 0$. This is a system of ordinary differential equations for the vector $\boldsymbol{P}(t) = [P_E(t), P_{ES}(t)]$, where $P(E, 0, t)$ is abbreviated as $P_E(t)$, with the system matrix

$$\boldsymbol{L} = \begin{bmatrix} -k_1' t & k_{-1} t \\ k_1' t & -(k_{-1}+k_2)t \end{bmatrix}.$$

The solution is obtained in terms of the matrix exponential as $\boldsymbol{P}(t) = \exp(\boldsymbol{L}t) \cdot \boldsymbol{P}(0)$ with the explicit result,

$$\begin{aligned}
P_E(t) &= \frac{1}{2A}\left[(A+B-C)e^{-(B+A)t} + (A-B+C)e^{-(B-A)t}\right],\\
P_{ES}(t) &= \frac{k_1'}{2A}\left[e^{-(B-A)t} - e^{-(B+A)t}\right],
\end{aligned}$$

where $k_1' = k_1[S]$, $2A = \sqrt{(k_1' + k_{-1} + k_2)^2 - 4k_1'k_2}$ and $2B = \left[k_1' + k_{-1} + k_2\right]$, and $C = k_{-1} + k_2$. From this, it is clear that the cumulative distribution is a sum of two exponentials whose time constants are determined by the eigenvalues of the 2×2 matrix $\boldsymbol{L}$. The waiting time distribution follows on differentiating the cumulative distribution,

$$w(\tau) = \frac{k_2 k_1'}{2A}\left[e^{(A-B)\tau_1} - e^{-(A+B)\tau}\right], \tag{17}$$

and the corresponding mean and randomness parameter are

$$\langle\tau\rangle = \frac{1}{k_2} + \frac{k_{-1}+k_2}{k_1}\frac{1}{[S]}, \tag{18}$$

$$r = \frac{(k_1[S] + k_{-1})^2 + 2k_2k_{-1} + k_2^2}{(k_1[S] + k_{-1} + k_2)^2}. \tag{19}$$

The variation of these is shown in Fig. 2 as a function of substrate concentration. The mean waiting time has a hyperbolic dependent on the substrate concentration of the Michaelis–Menten form, and the randomness parameter is always less than the unity, in agreement with previous analysis.[35,36]

How, then, are these results altered when the network topology is altered to allow for conformational fluctuations? The master equation for the general network shown in Fig. 1(a) is

$$\begin{aligned}
\partial_t P(E_\alpha, n) &= k_{-\alpha}P(ES_\alpha, n) + \gamma_{\alpha\beta}P(E_\beta, n) - (k'_\alpha + \gamma_{\alpha\beta})P(E_\alpha, n) + k_{p\alpha}P(ES_\alpha, n-1),\\
\partial_t P(E_\beta, n) &= k_{-\beta}P(ES_\beta, n) + \gamma_{\alpha\beta}P(E_\alpha, n) - (k'_\beta + \gamma_{\alpha\beta})P(E_\beta, n) + k_{p\beta}P(ES_\beta, n-1),\\
\partial_t P(ES_\alpha, n) &= k'_\alpha P(E_\alpha, n) + \delta_{\alpha\beta}P(ES_\beta, n) - (k_{-\alpha} + \delta_{\alpha\beta} + k_{p\alpha})P(ES_\alpha, n),\\
\partial_t P(ES_\beta, n) &= k'_\beta P(E_\beta, n) + \delta_{\alpha\beta}P(ES_\alpha, n) - (k_{-\beta} + \delta_{\alpha\beta} + k_{p\beta})P(ES_\beta, n), \qquad n = 0, 1, 2, \ldots.
\end{aligned} \tag{20}$$

Conditioning on a turnover, as before, reduces the CME to a system for four coupled differential equations for the components of the vector,

$$\boldsymbol{P}(t) = [P_{E_\alpha}(t), P_{ES_\alpha}(t), P_{E_\beta}(t), P_{ES_\beta}(t)],$$

in the abbreviated notation introduced above. The solution is given in terms of the exponential of the 4 × 4 matrix system matrix so that the cumulative distribution is a sum of four exponential terms. The expression for the waiting time distribution is obtained by differentiation as before. The expressions, being unwieldy, are provided in

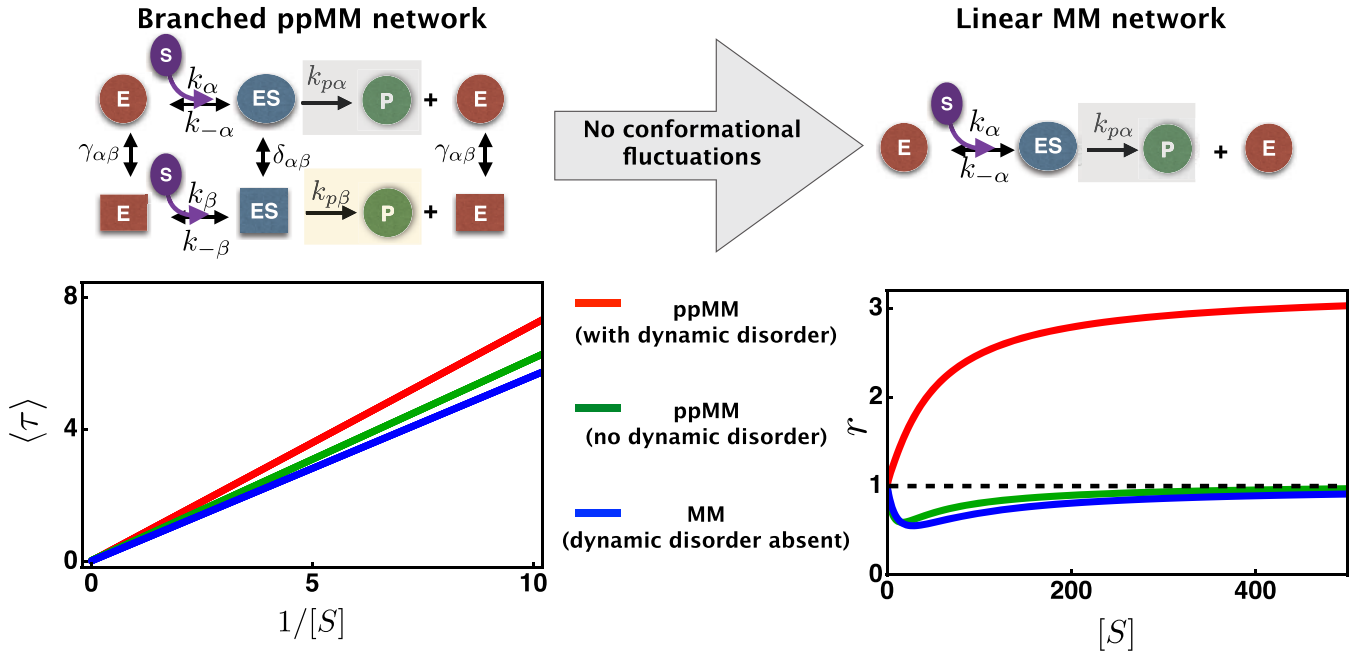

FIG. 2. Distinction between the summary statistics of branched and single-pathway stochastic networks as a function of substrate concentration. The top panel shows how a parallel-pathway MM (ppMM) network reduces to a single-pathway MM network when conformational fluctuations are disallowed. The bottom left panel shows the single-enzyme Lineweaver–Burk plot for the variation of inverse single-enzyme velocity, $\langle\tau\rangle$, with $1/[S]$. The plot is linear, and thus, the variation of single-enzyme velocity with $[S]$ is hyperbolic, irrespective of the network complexity. The bottom right panel shows how the variation of the randomness parameter r with $[S]$ provides a quantitative measure to discern single-enzyme network topologies. For the ppMM network, the rate constant conditions that favor parallel-pathway for product formation can yield $r \geq 1$ (dynamic disorder). The rate constant conditions that disallow the latter always yield $r \leq 1$. For a linear MM network, irrespective of the rate parameter conditions, there is no dynamic disorder as r does not exceed one.

the supplementary material. The results are shown in Fig. 2 for the general network for two sets of rate constants. The mean continues to have a hyperbolic dependence on the substrate concentration, but now the randomness parameter can yield values that are lesser or greater than unity, depending on the choice of rate constants. We find that rate constants that tend to suppress parallel pathways, i.e., to make the system matrix block diagonal, correspond to randomness parameters less than unity. In this limit, there is little to distinguish between linear and branched topologies. On the other hand, the rate constants that promote parallel pathways, i.e., to make the system matrix dense, correspond to randomness parameters greater than unity. This is the regime of dynamic disorder, and our results show that such effects can be obtained without imputing any *ad hoc* fluctuations on the rate constants themselves, but by simply allowing for a change in network topology.

The conditional independence of turnovers, due to the "fermionic" nature of the states, implies that single-enzyme networks can never show molecular memory. We now turn to networks in which the "fermionic" nature is lost, in the simplest possible way, by considering replicas of single-enzyme networks.

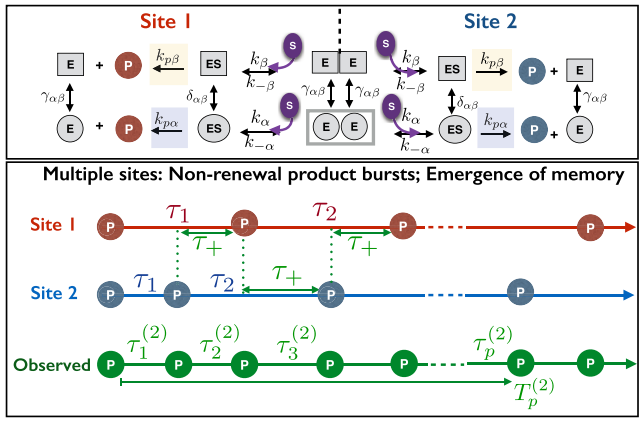

FIG. 3. The sequence of turnovers for a pair of active sites that process substrates in parallel. The waiting times are denoted by τ and the forward recurrence times (see text) are denoted by τ_+. The observed turnovers cannot distinguish which of the two active sites yielded the product, and hence, the turnover process is a pooling of the two independent turnover processes for each site. The generalization to more than two sites is obvious.

IV. REPLICA NETWORKS

Consider now a pair of sites on a single enzyme, as shown in Fig. 3 in red and blue, each governed by identical ppMM mechanisms and catalyzing substrates independently. It is neither possible nor relevant to distinguish such products by their site of production, and the *observed* process of turnover, shown in green, is a "pooling" of the independent turnover processes at each site. The total number of products in the pooled process at time t is the sum of the number of products at each site. Since the latter are independent random variables, the statistics of their sum can be simply obtained from the individual statistics. Therefore, the counting process of pooled turnovers is simple. However, the point process of pooled turnovers has a less simple relation to the individual point processes, as we explain below.

Returning to Fig. 3, assume that both sites start from identical initial conditions of being in uncomplexed states and denote by τ_p the p-th waiting time at any one of the identical sites and $\tau_p^{(2)}$ the p-th waiting time of the pooled process, where the superscript indicates that a pair of processes are pooled. Then, the waiting time $\tau_1^{(2)}$ for the first product of the pooled process is the shorter of the first waiting times τ_1 at each of the sites. For the second and subsequent turnovers, it is necessary to introduce the notion of the *forward recurrence time* τ_+, that is, the waiting time to the *next* product starting at an arbitrary time t. The distribution of τ_+, $P(s < \tau_+ < s + ds|t) \equiv w_+(s|t)ds$, is conditional on the time t, and this conditional dependence is crucial in what follows. In terms of the forward recurrence time, the waiting time $\tau_2^{(2)}$ of the second product is the shorter of the waiting time τ_2 at the site that produced the first product and the forward recurrence time τ_+ of the site that did not. For the example in Fig. 3, these are the first and second sites, respectively. Generalizing, the waiting time τ_p of the p-th product is the shorter of the waiting at one site and the forward recurrence time at the other site, where the recurrence time is measured from the last turnover at $t = T_{p-1}^{(2)}$. Since a waiting time $\tau_p^{(2)}$ exceeding s implies that *both* the waiting time τ_p and the recurrence time τ_+ measured from the last turnover exceed s, that is,

$$\tau_p^{(2)} > s \Longleftrightarrow (\tau_p > s) \text{ AND } (\tau_+ > s), \tag{21}$$

we immediately obtain for the survival probability of $\tau_p^{(2)}$ the relation

$$P(\tau_p^{(2)} > s|T_{p-1}^{(2)}) = P(\tau > s)P(\tau_+ > s|T_{p-1}^{(2)}). \tag{22}$$

This basic result shows that for a pooled process, the p-th waiting time is, in general, conditionally dependent on the time at which the $(p-1)$th turnover takes place. Therefore, the very act of pooling provides a mechanism by which a future waiting time can become conditionally dependent on past waiting times or, in other words, for the emergence of molecular memory.

It is desirable, if possible, to relate this conditional dependence to properties of the renewal process at each site. We now show that this is, indeed, possible. Defining the waiting time distribution of the pooled process as $P(s < \tau_p^{(2)} < s + ds|T_{p-1}^{(2)}) \equiv w_\pi(s|T_{p-1}^{(2)})ds$ and differentiating both sides of the above, we obtain

$$w_\pi(s|T_{p-1}^{(2)}) = w(s)[1 - P(\tau_+ < s|T_{p-1}^{(2)})] + \left[1 - \int_0^s w(s')ds'\right] w_+(s|T_{p-1}^{(2)}). \tag{23}$$

In the above, the distribution of waiting times $w(s)$ is known from the analysis of Sec. III and it only remains to determine the distribution $w_+(s|t)$ of the recurrence time. For a renewal process, the distribution of the forward recurrence time is related to the waiting time distribution and the enzymatic velocity by[26]

$$w_+(s|t) = w(t+s) + \int_0^t V(t-s')w(s+s')ds', \tag{24}$$

and this determines the waiting time distribution of the pooled process completely in terms of quantities that can be calculated from the component renewal processes.

How does the conditional dependence of the waiting time vary with the number of turnovers? Since the entire conditional dependence derives from the distribution of the recurrence time, it is sufficient to examine its conditional dependence. For large times $t \gg T^{\star}$, it is known from both deterministic and stochastic analyses that the enzymatic velocity becomes a constant, that is, $\lim_{t\to\infty} V(t) = V_{ss}$. As a consequence, $\lim_{t\to\infty} w_+(s|t) = V_{ss}[1 - w(s)]$ is independent of t. In this limit, each site is in "equilibrium," and the memory of the initial state of the process, which began with all sites free, is erased. Thus, the conditioning of the recurrence time of one site by the turnover time of another site, together with the deterministic initial condition, provides a mechanism for the statistical dependence between waiting times in the pooled process and of the emergence of molecular memory.

Extending this argument to ν binding sites and denoting pooled quantities with the superscript ν, it is clear that the waiting time $\tau_p^{(\nu)}$ for the p-th product is the shortest of a single waiting time and $\nu - 1$ recurrence times conditioned on $t = T_{p-1}^{(\nu)}$. Since $\tau_p^{(\nu)}$ being longer than s implies both the waiting time and the $\nu - 1$ recurrence times are longer than s, we have

$$P(\tau_p^{(\nu)} > s|T_{p-1}^{(\nu)}) = P(\tau > s)[P(\tau_+ > s|T_{p-1}^{(\nu)})]^{\nu-1}, \tag{25}$$

where the factorizations on the right-hand side follow from independence and identity of the ν binding sites. Defining the waiting time distribution of the pooled process as $P(s < \tau_p^{(\nu)} < s + ds|T_{p-1}^{(\nu)}) \equiv w_\pi(s|T_{p-1}^{(\nu)})ds$ and differentiating both sides of the above yield an explicit expression for $w_\pi(s|T_{p-1}^{(\nu)})$ in terms of the key single-site measures, the enzymatic velocity $V(t)$, the waiting time distribution $w(\tau)$, and the recurrence time distribution $w_+(s|t)$,

$$\begin{aligned} w_\pi(s|T_{p-1}^{(\nu)}) &= w(s)[1 - P(\tau_+ < s|T_{p-1}^{(\nu)})]^{\nu-1} \\ &+ (\nu - 1)\left[1 - \int_0^s w(s')ds'\right] \\ &\times [1 - P(\tau_+ < s|T_{p-1}^{(\nu)})]^{\nu-2} w_+(s|T_{p-1}^{(\nu)}). \end{aligned} \tag{26}$$

In this expression, $\tau_p^{(\nu)}$ is conditionally dependent on $T_{p-1}^{(\nu)}$ and, therefore, on the previous waiting times $\tau_1^{(\nu)} \ldots \tau_{p-1}^{(\nu)}$ as long as $w_+(s|t)$ is dependent on t. Thus, the emergence of memory can now be traced explicitly to the transient in the enzymatic velocity, starting from the deterministic initial condition. Evaluating $w_\pi(s|T_{p-1}^{(\nu)})$ numerically for the MM mechanism for $T_{p-1}^{(\nu)}$ in the transient, crossover, and steady-state regimes quantitatively confirms this qualitative picture (Fig. 1 of the supplementary material).

The waiting times become independent and identically distributed when the enzymatic velocity at each site reaches the steady-state value. Then, inserting the asymptotic form of the recurrence time distribution, we obtain $w_\pi^{(\nu)}(s) = -\frac{d}{ds}\left[P(\tau > s)\left\{V_{ss}\int_s^\infty P(\tau > s')ds\right\}^{\nu-1}\right]$, giving the waiting time distribution in terms of survival probability and the steady-state enzymatic velocity at each site. This is the enzymatic analog of a well-known result in renewal theory.[39]

V. STATISTICAL MEASURES

In classical deterministic enzyme kinetics, the enzymatic velocity of N independent and identical enzymes, $V(t) = d/dt\langle n\rangle$, is a statistical measure of mean rate of product formation. The approach to steady-state is then marked by the asymptotic limit, $V_{ss} = \lim_{t\to\infty} V(t)$, in which the enzymatic velocity reaches its equilibrium value and becomes time-independent. For $N \gg 1$, this asymptotic limit is realized at the onset of the reaction.[1] This implies that the initial mean rate of product formation, i.e., the counting process alone, is sufficient to yield the steady-state enzymatic velocity $V_{ss} = d_t\langle n\rangle|_{t\to 0}$, and the transient regime remains unobserved.[15]

In stochastic enzyme kinetics, in contrast, the statistical measures of counting and point processes for means and fluctuations, introduced in Sec. II, seem to provide an alternative description of product turnover kinetics in the number and time domain, respectively. It is pertinent to ask, then, how these seemingly unrelated statistical measures can be formally linked to demarcate the transient and steady-state regimes in enzyme kinetics at the molecular level, and how these results can be reconciled with the classical results of deterministic enzyme kinetics.

It is clear from Secs. III and IV that the turnover kinetics of single-enzyme networks is a renewal stochastic process with statistically independent waiting times and thus no memory. For replica networks, there exist an initial transient regime with memory and a terminal steady-state without it. The switch from a non-renewal to renewal statistics in replica networks, with increasing turnover number, thus marks a crossover from the transient to steady-state regime. In the steady-state, since waiting times are statistically independent and the governing statistics is renewal, below we use the results of the renewal theorems to formally link the statistical measures of counting and point processes.[26,40,41]

In the steady-state, statistical measures at each site are related to a pooled output, comprising of independent and identically distributed (*iid*) random variables. For the counting process description, the *iid* random variables are the number of products formed at each site, resulting in a pooled output $n = n_1 + n_2 + \cdots + n_\nu$ of n total number of products formed at ν sites. From this, it follows that $V_{ss}^{(\nu)} = \nu V_{ss}$, where $V_{ss} = \lim_{t\to\infty} d/dt\langle n_i\rangle$ with $i = 1, 2, \ldots, \nu$. For the point process description, the *iid* random variables are the waiting times $\tau_p^{(\nu)}$ between consecutive turnovers for ν sites, the sum of which $T_p^{(\nu)} = \tau_1^{(\nu)} + \cdots + \tau_p^{(\nu)}$ yields a pooled output for the p-th turnover time $T_p^{(\nu)}$. Since ν sites are independent and identically distributed, it follows that $\langle T_p^{(\nu)}\rangle = p\langle\tau^{(\nu)}\rangle$ and $\langle\tau^{(\nu)}\rangle = \frac{1}{\nu}\langle\tau\rangle$.

Furthermore, the renewal theorem guarantees that the single-enzyme velocity asymptotes to the inverse mean waiting time, $\lim_{t\to\infty} V(t) = V_{ss} \equiv \langle\tau\rangle^{-1}$.[26] This relates the statistical measures of means for counting and point processes, $V_{ss}^{(\nu)} \equiv \nu\langle\tau\rangle^{-1}$, for replica networks. From this, it follows that the inverse mean waiting time $\langle\tau\rangle^{-1}$ in Eq. (18) can be identified as the single-enzyme velocity.

In the absence of temporal correlations between turnovers, the variance of the sum is the sum of variances, $\sigma^2_{T_p^{(\nu)}} = p\sigma^2_{\tau^{(\nu)}}$. The renewal theorem, then, dictates that the squared coefficient of variation, the randomness parameter, $r^{(\nu)} = \langle(\tau^{(\nu)} - \langle\tau^{(\nu)}\rangle)^2\rangle/\langle\tau^{(\nu)}\rangle^2$ asymptotes to the Fano factor $\rho_{ss}^{(\nu)} = \lim_{t\to\infty}\langle(n^{(\nu)} - \langle n^{(\nu)}\rangle)^2\rangle/\langle n^{(\nu)}\rangle$ for replica networks.[26] A special case of this for $\nu = 1$ was first introduced by Block and Schnitzer[27] in the context of molecular motors, where it has widespread application.[42,43]

The above results show that for replica networks in the steady-state, the description of turnovers in terms of counts, n, and waiting times, τ, is asymptotically equivalent as renewal theorems guarantee that $V_{ss}^{(\nu)} \equiv \nu\langle\tau\rangle^{-1}$ and $\rho_{ss}^{(\nu)} \equiv r^{(\nu)}$.[26,40,41] However, these results are not valid for replica networks in the transient regime where the governing statistics is non-renewal, $\langle T_p^{(\nu)}\rangle \neq p\langle\tau^{(\nu)}\rangle$, and depends on the turnover number p through $T_p^{(\nu)}$. Moreover, the presence of correlations between waiting times clearly suggests that statistical measures in the transient regime should be redefined in terms of $T_p^{(\nu)}$, rather than $\tau_p^{(\nu)}$, as the former naturally contains correlations between waiting times. This motivates the following definitions for the turnover number dependent enzymatic velocity:[14]

$$V_p^{(\nu)} = \frac{p}{\langle T_p^{(\nu)}\rangle} \tag{27}$$

and the randomness parameter associated with $T_p^{(\nu)}$,

$$r_p^{(\nu)} = p\frac{\langle(T_p^{(\nu)} - \langle T_p^{(\nu)}\rangle)^2\rangle}{\langle T_p^{(\nu)}\rangle^2} = p\frac{\sum_i^p\langle(\delta\tau_i^{(\nu)})^2\rangle + \sum_{i\neq j}^p\langle\delta\tau_i^{(\nu)}\delta\tau_j^{(\nu)}\rangle}{\langle\sum_i^p\tau_i^{(\nu)}\rangle^2}, \tag{28}$$

where $\delta\tau_i = \tau_i - \langle\tau_i\rangle$.

The turnover number dependent enzymatic velocity $V_p^{(\nu)}$ and randomness parameter $r_p^{(\nu)}$ provide new statistical measures of means and fluctuations for replica networks that can be used both in transient and steady-state regimes, simply by increasing p. In Eq. (27), the crossover from non-renewal $\langle T_p^{(\nu)}\rangle \neq p\langle\tau^{(\nu)}\rangle$ to renewal $\langle T_p^{(\nu)}\rangle = p\langle\tau^{(\nu)}\rangle = \frac{p}{\nu}\langle\tau\rangle$ statistics with increasing p guarantees that the steady-state enzymatic velocity is asymptotically recovered, $\lim_{p\to\infty} V_p^{(\nu)} = \nu\langle\tau\rangle^{-1} \equiv V_{ss}^{(\nu)}$. In Eq. (28), similarly, the increase in p brings about a switch from non-renewal statistics with statistically dependent waiting times $\sum_{i\neq j}^p\langle\delta\tau_i^{(\nu)}\delta\tau_j^{(\nu)}\rangle \neq 0$ to renewal statistics with statistically independent waiting times, $\sum_{i\neq j}^p\langle\delta\tau_i^{(\nu)}\delta\tau_j^{(\nu)}\rangle = 0$. In the asymptotic limit of large p, thus, $r_p^{(\nu)}$ reduces to the steady-state definition, $r^{(\nu)} = \langle(\delta\tau^{(\nu)})^2\rangle/\langle\tau^{(\nu)}\rangle^2$, which is equivalent to the steady-state Fano factor $\rho_{ss}^{(\nu)}$, as expected from the renewal theorem.[23]

Equations (27) and (28) are the key results of this work as they provide the statistical measures of point process for single-enzyme and replica networks in transient and steady-state regimes. Their link to the counting process, as shown above, relies on the change of statistics from non-renewal at lower p to renewal at higher p. This naturally introduces a critical turnover number p^*, which demarcates the transient $p \ll p^*$ from the steady-state $p \gg p^*$ regime. In the steady-state, $V_p^{(\nu)}$ asymptotes to

$$V_{p\gg p^*}^{(\nu)} = \nu\langle\tau\rangle^{-1} \equiv V_{ss}^{(\nu)}. \tag{29}$$

Similarly, $r_p^{(\nu)}$ in the steady-state asymptotes to

$$r_{p\gg p^*}^{(\nu)} = r^{(\nu)} \equiv \rho_{ss}^{(\nu)}. \tag{30}$$

In the transient regime, $p \ll p^*$, both these equivalences are necessarily violated as the governing statistics is non-renewal.

Equations (29) and (30), while subsuming the results of renewal theorems in the steady-state, provide an empirical test of non-stationarity in experimental data and a diagnostic for the emergence of memory in the transient regime. In Sec. VI, we show how these equalities can be used to determine p^*.

VI. COMPARISON WITH DATA

We now apply the theory developed in Secs. II–V to analyze the data from the landmark experiment in which the molecular memory was first observed.[10] In this experiment, the catalysis of non-fluorescent substrates to fluorescent products by the tetrameric enzyme β-galactosidase over a range of substrate concentrations was monitored using fluorescence spectroscopy. It is to be noted that β-galactosidase has four catalytic sites.[44] It is known to obey the MME in ensemble averaged measurements.[45,46] This suggests that the four well separated catalytic sites are independent of each other.[47]

In this experiment, waiting times were obtained from the primary data of product turnovers as discrete fluorescence bursts, and the distribution $w(\tau)$ and its first two moments were computed for each substrate concentration. While the variation of mean waiting time with $1/[S]$ was linear at low substrate concentrations, the monotonic increase in the randomness parameter with $[S]$, bounded below by one, was a signature of dynamic disorder. The joint distribution, $w(\tau_p, \tau_{p+q})$ of the waiting times, q turnovers apart, revealed that turnover events were not statistically independent, but that a short (or long) first waiting time was more likely to be followed by another short (or long) second waiting time. This was a signature of positive molecular memory. The correlation of waiting times, $C_q = \langle\delta\tau_p\delta\tau_{p+q}\rangle$, remained appreciable and, when expressed in terms of a scaled time $t = q\langle\tau\rangle$, could be collapsed to a single stretched-exponential $C(t) = \exp[-(t/t_0)^\beta]$ with $\beta = 0.45$ and $t_0 = 0.018$ s.

The experimental results reveal statistically dependent waiting times and dynamic disorder in enzyme turnover kinetics of the tetrameric enzyme β-galactosidase with four independent catalytic sites.[10] Following the analysis of Secs. III and IV, this motivates us to select the ppMM network with four replicas as the minimal model to understand the kinetics. For comparison, we also consider a hypothetical single-enzyme ppMM network with identical parameters, but only a single binding site. We use the results of Sec. III and the supplementary material to compute the marginal distribution $w(T_1^{(\nu)})$, where the superscript to $T_p^{(\nu)}$ denotes the results for

a single site ($\nu = 1$) and quadruple sites ($\nu = 4$). From $w(T_1^{(\nu)})$ thus obtained, we analytically compute the mean first turnover time, $\langle T_1^{(\nu)} \rangle$, and the randomness parameter, $r_1^{(\nu)}$, as functions of the eight rate constants. A simultaneous least-squares fit to the experimental data provides us with the maximum-likelihood parameters of each model. These are listed in Table I of the supplementary material, and the corresponding fits are shown in the top four panels of Fig. 4.

The excellent agreement between the model and data for both single- and quadruple-site models leaves little to distinguish between them. Since both models have the same number of parameters and hence equal model complexity, they appear to be equally plausible models for data derived from the marginal distribution. This degeneracy in model space is lifted by using data from the joint distribution, as we now show.

Since we have not found a way to obtain the joint distributions of the Markov chain analytically, we compute them numerically from a time series of turnovers, sampled using the Doob–Gillespie algorithm[48] with chain parameters set to the above least-squares estimates. The distribution of $\tau_p^{(\nu)}$ and the correlation function $C_q^{(\nu)}$ are shown in the bottom four panels of Fig. 4.

There is, now, a clear distinction between single-site and quadruple-site models. The first distinction appears in the distribution of waiting times τ_p. These are identically and independently distributed for the single-site model [panels (c) and (d)], but neither identically nor independently distributed for the quadruple-site model [panels (g) and (h)]. This confirms the results of Secs. III and IV that a model with fermionic hidden states can only yield a renewal process and that multiple binding sites are necessary for molecular memory. Focussing on panel (h), the normalized correlation function has an excellent fit to a stretched exponential function $C_q^{(4)} = \exp[-(q/q_0)^\beta]$ with parameters $q_0 = 0.25$ and $\beta = 0.47, 0.42, 0.39$ for the substrate concentrations $[S] = 20\ \mu$M, $100\ \mu$M, $380\ \mu$M reported in the experiment. Continuing in Fig. 5, we plot the normalized correlation function $C^{(4)}(t)$ in scaled time $t = \langle \tau_1^{(4)} \rangle q$ following experiment. There is a quantitative match between experiment and theory with both following a stretched exponential function $C^{(4)}(t) = \exp[-(t/t_0)^\beta]$ with $t_0 = 0.018$ s and $\beta = 0.45$ for $[S] = 100\ \mu$M. In addition, the pseudocolor plot of the joint distribution of the first and second waiting times, $w(\tau_1^{(4)}, \tau_2^{(4)})$, shows that a short (or long) first waiting time is more likely to be followed by a short (or long) second waiting time (inset), in agreement with the molecular memory observed in experiment.

For comparison, we repeat the above calculations for the MM model and summarize our findings in Table II. The heat map of the joint distribution of successive waiting times, $w(\tau_1, \tau_2)$, shows negative correlations for the MM model, but positive correlations for the ppMM model. In the stationary state, the randomness parameter is negative for the MM model, but positive for the ppMM model. Only the ppMM model, containing both multiple binding sites and conformational fluctuations, is in agreement with experiment, as summarized in the last two rows of the table. While multiple sites alone yield memory, conformational fluctuations are necessary for the

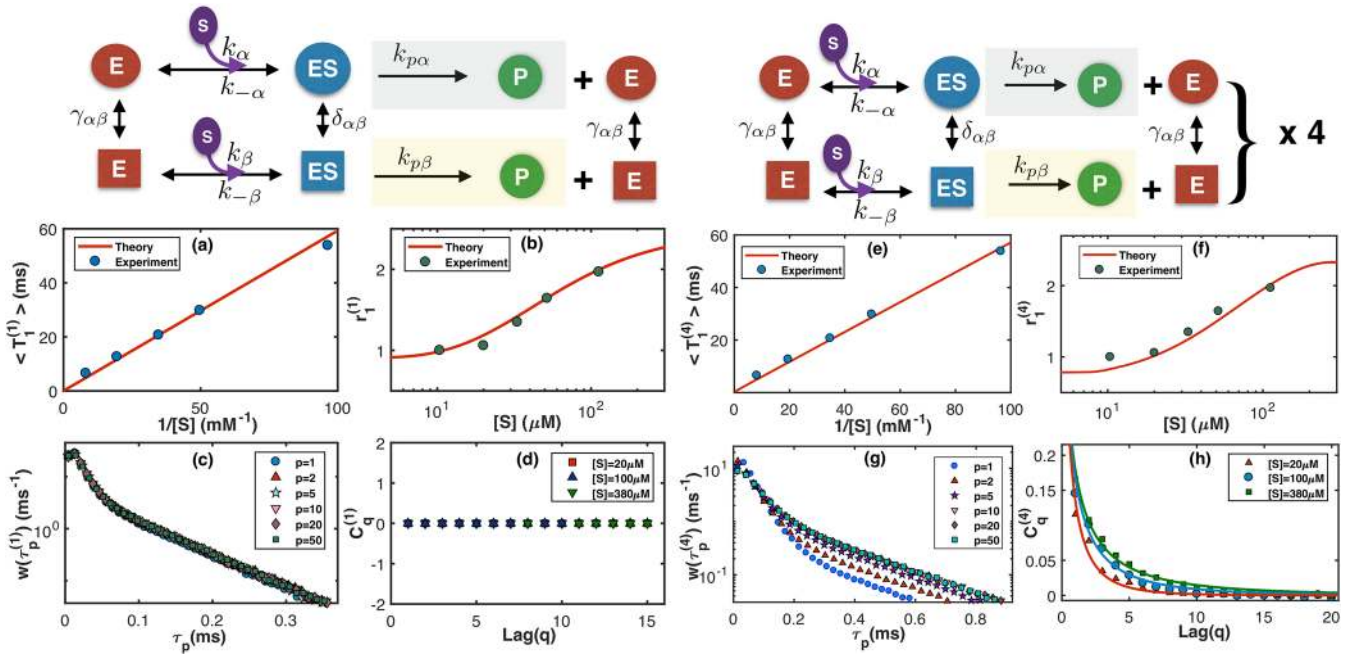

FIG. 4. Quantitative comparison of turnover statistics for single- and quadruple-site ppMM models. The rate constants are estimated by *simultaneously* fitting analytical expressions for $\langle T_1^{(\nu)} \rangle$ and $r_1^{(\nu)}$ to experimental data. Exact numerical sampling with these parameters is used to generate distributions $w(\tau_p)$ and compute correlations C_q. The renewal character of the single-site model and its lack of memory are confirmed in panels (c) and (d). The points are simulation data using best-fit rate parameters and the solid lines are stretched-exponential fits $C_q^{(4)} = \exp[-(q/q_0)^\beta]$ with $q_0 = 0.25$ and $\beta = 0.47, 0.42, 0.39$ for $[S] = 20\ \mu$M, $100\ \mu$M, $380\ \mu$M, respectively.

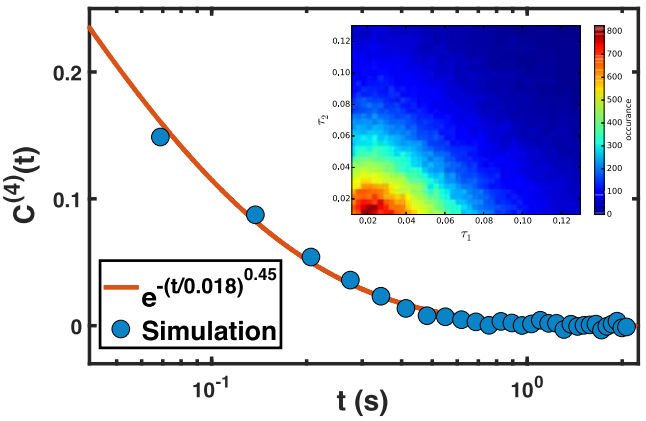

FIG. 5. Normalized correlation $C^{(4)}(t)$ for the quadruple-site PPMM mechanism at $[S] = 100\ \mu$M. The points are simulation data and the solid line is a stretched exponential $e^{-(t/t_0)^\beta}$ with $t_0 = 0.018$ s and $\beta = 0.45$. The inset is the joint probability $w(\tau_1^{(4)}, \tau_2^{(4)})$ of consecutive waiting times, which shows that a short (long) first waiting time is more likely to be followed by a short (long) second waiting time, in agreement with experiment.

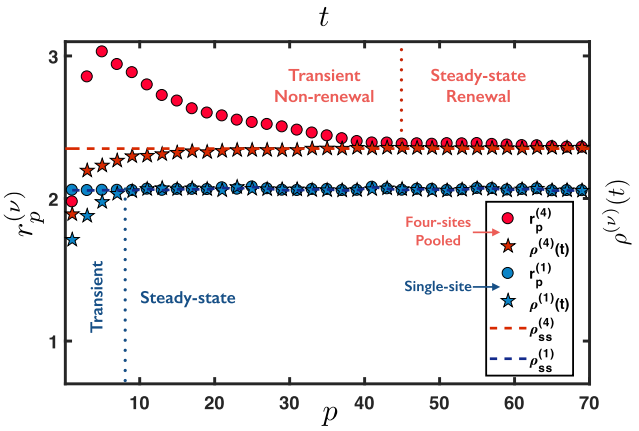

FIG. 6. Randomness parameters and Fano factors for the ppMM obtained from numerically computed trajectories of the single- and quadruple-site models. For the quadruple-site model, these measures of counting and point processes are asymptotically equivalent in the (renewal) steady-state but differ in the (non-renewal) transient state, in agreement with Eq. (30). The memory persists as long as the equality is violated. For the memoryless single-site model, the duration of the transient is determined by the counting process description, i.e., the time required for the Fano factor to reach its steady-state value. The parameter values are listed in Table I of the supplementary material.

correct sign of the correlation function and the correct magnitude of the randomness parameter.

We now use Eqs. (29) and (30) of Sec. V to determine p^* and thus demarcate the transient $p \ll p^*$ and steady-state $p \gg p^*$ regimes for the ppMM model. The variation of $V_p^{(\nu)}$ with substrate concentration is shown in Fig. 2 of the supplementary material. The enzymatic velocity in the transient regime, $V_{p\ll p^*}^{(\nu)}$, deviates from the steady-state value $\nu\langle\tau\rangle^{-1}$, where $\langle\tau\rangle^{-1}$ is the Michaelis–Menten-like (MML) equation [Eq. (8) of the supplementary material]. In the steady-state, $V_{p\gg p^*}^{(\nu)}$ asymptotically approaches both the single-enzyme MML equation and the classical steady-state enzymatic velocity $V_{ss}^{(\nu)}$, in agreement with Eq. (29). A similar analysis for the statistical measures of fluctuations is presented in Fig. 6, which shows the variation of $r_p^{(\nu)}$ with p and $\rho^{(\nu)}(t)$ with t for $\nu = 1, 4$. The comparison shows that the equivalence between the randomness parameter and the steady-state Fano factor is violated in the transient regime, $r_{p\ll p^*}^{(\nu)} \neq \rho_{ss}^{(\nu)}$, but is asymptotically recovered in the steady-state regime $p \gg p^*$, in agreement with Eq. (30).

TABLE II. Qualitative comparison of turnover statistics for MM and ppMM models for single and multiple binding sites. The ppMM model, with conformational fluctuations and four binding sites, best agrees with the experimentally obtained turnovers of β-galactosidase.

Mechanism	Sites	Memory	Correlations	$r_1^{(\nu)}$ vs $[S]$
MM	1	Absent	$C_q^{(1)} = 0$	$r_1^{(1)} \leq 1$
ppMM	1	Absent	$C_q^{(1)} = 0$	$r_1^{(1)} \geq 1$
MM	4	Anti-correlated	$C_q^{(4)} \leq 0$	$r_1^{(4)} \leq 1$
ppMM	4	Correlated	$C_q^{(4)} \geq 0$	$r_1^{(4)} \geq 1$
β-galactosidase	4	Correlated	$C_q \geq 0$	$r \geq 1$

The fading of memory, the convergence of the waiting time distributions, and the equality of the statistical measures all occur at roughly $p^* \approx 50$ turnovers in this model.

VII. SUMMARY AND DISCUSSION

In summary, our statistical analysis for the discrete turnover kinetics of an enzyme with independent active sites shows that in a finite time domain, there always exists a transient regime in which the number average is not equivalent to the time average. In the steady-state, these equivalences are asymptotically recovered. The statistical measures introduced in our work are valid for the Markovian description of stochastic kinetics, described in terms of the CME, with exponentially distributed lifetime of each kinetic state. These measures retain the fundamental first-order character of each elementary step in the mechanism. The dynamic disorder, captured at timescales longer than the lifetime of each kinetic state, is thus an emergent feature of intrinsic number and temporal fluctuations of each kinetic state of the mechanism.

The new statistical measure of temporal fluctuations, the turnover number-dependent randomness parameter r_p, Eq. (28), provides a mechanistic origin of dynamic disorder in the single-molecule turnover data of β-galactosidase (Figs. 4 and 5) in the transient and steady-state regimes. The enzyme β-galactosidase is known to obey the MME in ensemble averaged measurements,[45,46] signifying the absence of temporal correlations in the steady-state. However, the same enzyme at the single-molecule level shows non-hyperbolicity, waiting time correlations, and molecular memory effect.[10] Our work reconciles these seemingly disparate observations for an enzyme with independent sites, using beta-galactosidase as an example, to show that the dynamic disorder ($r_p > 1$) in the

transient regime, with memory, is characteristically distinct from the steady-state regime without memory (Fig. 6).

In particular, the rate parameter values listed in Table I of the supplementary material allow us to make the following inference about the two-state ppMM mechanism for β-galactosidase. The heterogeneity in product formation rates is a necessary condition for the emergence of dynamic disorder in line with the distribution of k_2 captured in Fig. 2(a) of Ref. 10. In the transient regime, the dynamic disorder arises from the correlated action of multiple sites, where the catalytic conversion of substrates to products at each site is governed by the slower rate of interconversion between enzyme–substrate conformers compared to the (faster of the two) product formation rate. For the same rate parameters in the steady-state regime, where catalysis at one site becomes independent of another site, the magnitude of dynamic disorder is mainly determined by the slower rate of off-pathway conformational fluctuations in the bound state of an enzyme compared to the (faster of the two) product formation rate. This, then, opens an unexplored way to elicit kinetic mechanisms from molecular noise by combining the distinct characteristics of intrinsic fluctuations in the transient and steady-states.

For comparison, stochastic kinetic models that assume the dynamic disorder in reaction steps make an *a priori* assumption of non-Markovian dynamics in catalytic turnover cycles.[49–52] These models relax the constraint of exponentially distributed waiting times in each kinetic state and replace them with an arbitrary non-exponential distribution that fits data. At this level of coarse-graining, aggregates of possible bound/intermediate kinetic states are assumed to follow non-exponential kinetics.[51,52] As a result, the waiting time distribution is not an emergent feature of intrinsic fluctuations of individual kinetic states of the reaction mechanism. Inversely, thus, a "coarse-grained" reaction mechanism inferred from such a distribution, and statistical measures such as the Poisson indicator that quantify dynamic disorder from the moments of these distributions, provides physical insights only within the confines of a reduced subspace of variables. Moreover, the implicit assumption of non-Markovian dynamics in these models can yield waiting times correlations in single-site catalysis. This is in contrast to the prediction of the Markovian model for single-site catalysis presented here, where the dynamic disorder (Fig. 6) in the steady-state can arise in the absence of waiting time correlations [Fig. 4(d)].

In general, the MME in the steady-state can be obtained directly from the CME by defining macroscopic variables like the average number of particles in a large N limit[19] or by defining the integrated probability flux in a long time limit.[16,17] In the latter, Cao and co-workers have used the flux-balance method to derive generalized expressions for the MME for stochastic enzymatic networks of arbitrary complexity.[16,17] Their work identifies detailed balance as a sufficient condition for the hyperbolic substrate dependence of the MME, similar to the fast-equilibrium condition in deterministic enzymatic networks.[6]

VIII. CONCLUSION AND FUTURE WORK

The stochastic time-domain approach presented here, when used in combination with the stochastic number-domain approach, namely, CME, provides the most detailed information of enzymatic reactions and can be used to extract mechanistic information that is lost in classical deterministic theories of enzyme kinetics.

Our work shows that the mechanism for the emergence of molecular memory will always be operative through the transient phase in an enzyme with multiple binding sites. It can thus be viewed as a null model (in the sense of a null hypothesis) and should be tested against before embarking on a search for more elaborate models of molecular memory that may require, for instance, interactions between binding sites.

Our statistical analysis of multiple-site catalysis, based on replica networks, assumes that active sites are identical and independent. While single-site catalysis at independent sites is expected to yield a hyperbolic relation between the mean catalytic rate and substrate concentration, deviation from hyperbolicity in replica networks emerges from the concerted action of independent and, hence, non-interacting, multiple binding sites. This form of "cooperativity" is dynamic in nature,[53] which arises from temporal correlations between enzymatic turnovers in the transient regime, and vanishes in the steady-state regime. This contrasts the traditional description of cooperativity in allosteric enzymes in which binding of a substrate at one site affects substrate binding affinity at another site.[1] The allosteric effects manifest as a non-hyperbolic substrate binding response and correlated action at equilibrium. The replica approach presented here can be extended to include interactions between binding sites by replicating, for instance, single enzyme networks of either the concerted or sequential models of allosteric enzymes as a basic motif.[54,55] This, along with rate parameter conditions in substrate binding steps of a replica network, can pave the way for understanding the combined effect of stochasticity and interaction in generating molecular cooperativity in the steady-state.

Our study extends beyond the simple homogeneous mechanisms presented here to more complex mechanisms including inhibitors,[12,56,57] hidden intermediate states[58,59] in single-molecule kinetics and single-molecule photon statistics,[60] and heterogenous catalysis of, for example, nano-particle clusters containing numerous binding sites.[61] In the latter case, the method presented here can be used to estimate the catalytic rate from turnover time data through a new kinetic measure, the heterogeneity index, which can quantify fluctuations from non-identical binding sites.[21–23] Additionally, in the presence of allosteric effects in a single nanoparticle cluster,[21–23] our statistical analysis can be extended to include heterogeneity and allosteric effects, independently and in combination. While these effects can generate temporal correlations in the steady-state, the magnitude of temporal correlations in the transient regime will be over and above the temporal correlations predicted by the present "null-model" for non-interacting multiple sites.

In the context of molecular motors, the renewal theorems have been known to link the mean and variance of the physical distance moved by the motor with the corresponding mean and variance of the number of products formed under the stationary condition.[28,29] Both are linearly proportional to time with proportionality constants being the enzymatic velocity for the measure of mean and the diffusion coefficient for the measure of variance. The relations between the statistical measures of counting and point processes, for the mean and variance, presented here, and their dependence on the turnover number can be used to generalize the corresponding expressions for molecular motors to non-stationary conditions,

where transport coefficients are expected to be turnover number dependent.

Fluctuation statistics of mesoscopic quantum transport in nanoscale devices analyzed in terms of fixed time (counting process) or fluctuating time (point process) description using a master equation framework.[62,63] While the renewal theorems provide formal relations between counting and point process statistics, the violation of renewal statistics and the analysis of non-renewal fluctuation statistics, in which temporal correlations between discrete quantum events are accounted for, are a relatively new premise.[64,65] The generality of our results, for both renewal and non-renewal fluctuation statistics, suggests the search for a transient phase in mesoscopic quantum transport[66–69] and the estimation of the duration of a transient phase in the photo emission statistics of an ensemble of quantum dots.[70,71]

The sensitivity of the waiting time distributions to the reaction mechanism, both in the transient and steady-states, invites the application of the Bayesian probability to calculate the posterior probability $P(\mathcal{M}|\{\tau_p\})$ of a model $\mathcal{M}$ given waiting time data $\{\tau_p\}$ and, thereon, to machine learning reaction mechanisms from turnover data.[72–74]

To conclude, our work shows that complex forms of molecular memory can arise from the combined action of simple memoryless steps, provides a theoretical framework within which such action can be studied systematically, and suggests experiments to test the validity of this generic mechanism.

SUPPLEMENTARY MATERIAL

See the supplementary material for detailed solution of the chemical master equations for single-site and multiple-sites catalysis.

DEDICATION

This work is dedicated to Maud Menten who, together with Leonor Michaelis, made outstanding contributions to enzyme kinetics and discovered the hyperbolic dependence of catalytic rate on substrate concentration, the Michaelis–Menten equation.

ACKNOWLEDGMENTS

We thank King's College for supporting A.D.'s visit to Cambridge, where a part of this work was conceptualized. The work was first presented as an invited lecture in the ninth conference of the Asia-Pacific Association of Theoretical and Computational Chemists (APATCC) at the University of Sydney in 2019.

DATA AVAILABILITY

The data that support the findings are available within the article and its supplementary material.[10]

REFERENCES

[1] A. Cornish-Bowden and A. Cornish-Bowden, *Fundamentals of Enzyme Kinetics* (Wiley-Blackwell Weinheim, Germany, 2012), Vol. 510.
[2] L. Michaelis and M. L. Menten, Biochem. Z. **49**, 333 (1913).
[3] A. Cornish-Bowden, Perspect. Sci. **4**, 3 (2015).
[4] U. Deichmann, S. Schuster, J.-P. Mazat, and A. Cornish-Bowden, FEBS J. **281**, 435 (2014).
[5] H. Lineweaver and D. Burk, J. Am. Chem. Soc. **56**, 658 (1934).
[6] F. Wong, A. Dutta, D. Chowdhury, and J. Gunawardena, Proc. Natl. Acad. Sci. U. S. A. **115**, 9738 (2018).
[7] H. P. Lu, L. Xun, and X. S. Xie, Science **282**, 1877 (1998).
[8] J. Cao, Chem. Phys. Lett. **327**, 38 (2000).
[9] X. Xie and H. Lu, *Single Molecule Spectroscopy* (Springer, 2001), pp. 227–240.
[10] B. P. English, W. Min, A. M. Van Oijen, K. T. Lee, G. Luo, H. Sun, B. J. Cherayil, S. C. Kou, and X. S. Xie, Nat. Chem. Biol. **2**, 87 (2006).
[11] W. Min, B. P. English, G. Luo, B. J. Cherayil, S. C. Kou, and X. S. Xie, Acc. Chem. Res. **38**, 923 (2005).
[12] H. H. Gorris, D. M. Rissin, and D. R. Walt, Proc. Natl. Acad. Sci. U. S. A. **104**, 17680 (2007).
[13] S. Saha, S. Ghose, R. Adhikari, and A. Dua, Phys. Rev. Lett. **107**, 218301 (2011).
[14] A. Kumar, R. Adhikari, and A. Dua, Phys. Rev. Lett. **119**, 099802 (2017).
[15] A. Dua, Resonance **24**, 297 (2019).
[16] J. Cao, J. Phys. Chem. B **115**, 5493 (2011).
[17] D. E. Piephoff, J. Wu, and J. Cao, J. Phys. Chem. Lett. **8**, 3619 (2017).
[18] D. A. McQuarrie, J. Appl. Probab. **4**, 413 (1967).
[19] A. F. Bartholomay, Biochemistry **1**, 223 (1962).
[20] G. E. Briggs and J. B. S. Haldane, Biochem. J. **19**, 338 (1925).
[21] W. Xu, J. S. Kong, Y.-T. E. Yeh, and P. Chen, Nat. Mater. **7**, 992 (2008).
[22] W. Xu, H. Shen, G. Liu, and P. Chen, Nano Res. **2**, 911 (2009).
[23] W. Xu, J. S. Kong, and P. Chen, Phys. Chem. Chem. Phys. **11**, 2767 (2009).
[24] D. J. Daley and D. Vere-Jones, *An Introduction to the Theory of Point Processes: Volume II: General Theory and Structure* (Springer Science, 2007).
[25] C. W. Gardiner, *Handbook of Stochastic Methods for Physics, Chemistry and the Natural Sciences* (Springer-Verlag, Berlin, 2004), Vol. 13.
[26] D. R. Cox, *Renewal Theory* (Wiley, London; Methuen; New York, 1962).
[27] M. J. Schnitzer and S. Block, *Cold Spring Harbor Symposia on Quantitative Biology* (Cold Spring Harbor Laboratory Press, 1995), Vol. 60, pp. 793–802.
[28] K. Svoboda, P. P. Mitra, and S. M. Block, Proc. Natl. Acad. Sci. U. S. A. **91**, 11782 (1994).
[29] K. C. Neuman, O. A. Saleh, T. Lionnet, G. Lia, J.-F. Allemand, D. Bensimon, and V. Croquette, J. Phys.: Condens. Matter **17**, S3811 (2005).
[30] R. Zwanzig, Acc. Chem. Res. **23**, 148 (1990).
[31] R. Zwanzig, J. Chem. Phys. **97**, 3587 (1992).
[32] N. G. Van Kampen, *Stochastic Processes in Physics and Chemistry* (Elsevier, 1992), Vol. 1.
[33] A. Kumar, H. Maity, and A. Dua, J. Phys. Chem. B **119**, 8490 (2015).
[34] U. Fano, Phys. Rev. **72**, 26 (1947).
[35] J. R. Moffitt, Y. R. Chemla, and C. Bustamante, Proc. Natl. Acad. Sci. U. S. A. **107**, 15739 (2010).
[36] J. R. Moffitt and C. Bustamante, FEBS J. **281**, 498 (2014).
[37] A. David and S. Larry, "Communications in statistics," Stochastic Models **3**, 467 (1987).
[38] A. K. Erlang, Post Off. Electr. Eng. J. **10**, 189 (1917).
[39] D. R. Cox and W. L. Smith, Biometrika **41**, 91 (1954).
[40] W. L. Smith, Proc. R. Soc. Edinburgh A: Math. **64**, 9 (1953).
[41] W. L. Smith, Biometrika **46**, 1 (1959).
[42] A. B. Kolomeisky and M. E. Fisher, Annu. Rev. Phys. Chem. **58**, 675 (2007).
[43] D. Chowdhury, Phys. Rep. **529**, 1 (2013).
[44] S. L. Marchesi, E. Steers, Jr., and S. Shifrin, Biochim. Biophys. Acta **181**, 20 (1969).
[45] G. H. Seong, J. Heo, and R. M. Crooks, Anal. Chem. **75**, 3161 (2003).
[46] A. G. Hadd, D. E. Raymond, J. W. Halliwell, S. C. Jacobson, and J. M. Ramsey, Anal. Chem. **69**, 3407 (1997).
[47] B. W. Matthews, C. R. Biol. **328**, 549 (2005).
[48] D. T. Gillespie, Annu. Rev. Phys. Chem. **58**, 35 (2007).
[49] S. Yang and J. Cao, J. Chem. Phys. **117**, 10996 (2002).

[50]S. C. Kou, B. J. Cherayil, W. Min, B. P. English, and X. S. Xie, J. Phys. Chem. B **109**, 19068 (2005).
[51]T. R. Avila, D. E. Piephoff, and J. Cao, J. Phys. Chem. B **121**, 7750 (2017).
[52]J. Cao and R. J. Silbey, J. Phys. Chem. B **112**, 12867 (2008).
[53]A. Kumar, S. Chatterjee, M. Nandi, and A. Dua, J. Chem. Phys. **145**, 085103 (2016).
[54]J. Monod, J. Wyman, and J.-P. Changeux, J. Mol. Biol. **12**, 88 (1965).
[55]D. E. Koshland, G. Némethy, and D. Filmer, Biochemistry **5**, 365 (1966).
[56]P. Mogalisetti, H. H. Gorris, M. J. Rojek, and D. R. Walt, Chem. Sci. **5**, 4467 (2014).
[57]S. Saha, A. Sinha, and A. Dua, J. Chem. Phys. **137**, 045102 (2012).
[58]H. Shen, X. Zhou, N. Zou, and P. Chen, J. Phys. Chem. C **118**, 26902 (2014).
[59]W. J. Ramsay, N. A. W. Bell, Y. Qing, and H. Bayley, J. Am. Chem. Soc. **140**, 17538 (2018).
[60]J. Cao, J. Phys. Chem. B **110**, 19040 (2006).
[61]M. Panigrahy, A. Kumar, S. Chowdhury, and A. Dua, J. Chem. Phys. **150**, 204119 (2019).
[62]S. L. Rudge and D. S. Kosov, J. Chem. Phys. **151**, 034107 (2019).
[63]K. Kaasbjerg and W. Belzig, Phys. Rev. B **91**, 235413 (2015).
[64]K. Ptaszyński, Phys. Rev. E **97**, 012127 (2018).
[65]S. L. Rudge and D. S. Kosov, Phys. Rev. B **99**, 115426 (2019).
[66]A. Kurzmann, P. Stegmann, J. Kerski, R. Schott, A. Ludwig, A. D. Wieck, J. Knig, A. Lorke, and M. Geller, Phys. Rev. Lett. **122**, 247403 (2019).
[67]P. Stegmann, B. Sothmann, A. Hucht, and J. Knig, Phys. Rev. B **92**, 155413 (2015).
[68]E. Kleinherbers, P. Stegmann, and J. König, New J. Phys. **20**, 073023 (2018).
[69]P. Stegmann and J. Knig, Phys. Rev. B **94**, 125433 (2016).
[70]I. Chung, J. B. Witkoskie, J. Cao, and M. G. Bawendi, Phys. Rev. E **73**, 011106 (2006).
[71]I. Chung, J. B. Witkoskie, J. P. Zimmer, J. Cao, and M. G. Bawendi, Phys. Rev. B **75**, 045311 (2007).
[72]A. Zellner, *Bayesian Analysis in Econometrics and Statistics* (Edward Elgar Publishing, 1997).
[73]H. Jeffreys, *The Theory of Probability* (OUP Oxford, 1998).
[74]E. T. Jaynes, *Probability Theory: The Logic of Science* (Cambridge University Press, 2003).