State- versus Reaction-Based Information Processing in Biochemical Networks

Anne-Lena Moor,[1,2,*] Age Tjalma,[3] Manuel Reinhardt,[3] Pieter Rein ten Wolde,[3,†] and Christoph Zechner[4,1,2,‡]

[1]Max Planck Institute of Molecular Cell Biology and Genetics, 01307 Dresden, Germany
[2]Center for Systems Biology Dresden, 01307 Dresden, Germany
[3]AMOLF, Science Park 104, 1098 XG Amsterdam, Netherlands
[4]Scuola Internazionale Superiore di Studi Avanzati, 34136 Trieste, Italy

(Received 22 May 2025; accepted 25 March 2026; published 17 June 2026)

Trajectory mutual information is frequently used to quantify information transfer in biochemical systems. Tractable solutions of the trajectory mutual information can be obtained via the widely used linear-noise approximation (LNA) using Gaussian channel theory. This approach is expected to be accurate for sufficiently large systems. However, recent observations show that there are cases, where the mutual information obtained this way differs qualitatively from results derived using an exact Markov jump process formalism, and that the differences remain even in the large copy number regime. In this Letter, we show that these differences can be explained by introducing the notion of reaction- versus state-based descriptions of trajectories. In chemical systems, the information is encoded in the sequence of reaction events, and the reaction-based trajectories of Markov jump processes capture this information. We show that within the Gaussian formalism, trajectories can be defined either based on individual reaction channels, or on a state-based level, where different reaction channels are summarized into a single noise term. While both definitions agree in terms of copy number fluctuations, state-based trajectories contain in general less information than reaction-based trajectories. The commonly used Gaussian mutual information via the linear-noise approximation is consistent with a state-based trajectory notion, which causes a systematic loss of information independent of system size. We show that an alternative, reaction-based variant of the Gaussian mutual information prevents this loss of information. We illustrate the consequences of different trajectory descriptions for two common cellular reaction motifs and discuss their connection with Berg-Purcell and maximum-likelihood sensing.

DOI: 10.1103/nfk6-8x5s

Introduction—Accurately responding to time-varying signals is essential for many natural and engineered systems, ranging from neural and biochemical networks to electronic circuits. The central measure for quantifying information transmission via time-varying signals is the information transmission rate, also called trajectory—or path mutual information rate [1–5]. Mathematically, it is defined as the rate at which the trajectory mutual information between the input and output signal trajectories increases with the duration of these trajectories in the long-time limit. In contrast to other information theoretical measures, such as the instantaneous or time-lagged mutual information [6], the trajectory mutual information not only captures the accuracy of the instantaneous input-output mapping, but also the effect of temporal correlations within the input and output signals.

Cellular signaling networks show that the encoding of information in trajectories can be highly nontrivial. These systems transmit information via chemical reactions between discrete molecules. In the limit that these systems are well-mixed, they are appropriately described as discrete Markov jump processes. Yet, it is commonly believed that Gaussian continuum approximations become accurate when the reactions are linear and the copy numbers are large [7–9]. They can indeed accurately predict the instantaneous mutual information, even when the copy numbers are as small as 10 [7]. Consequently, a Gaussian approach like the linear-noise approximation [10] is frequently used to study information transmission in biochemical systems, where the trajectory mutual information can be conveniently obtained from the temporal correlation functions of molecule counts [2]. However, it has recently been shown that this Gaussian approach not only significantly underestimates the information rate, but can, in some situations,

*Contact author: annemoor96@gmail.com
†Contact author: tenwolde@amolf.nl
‡Contact author: czechner@sissa.it

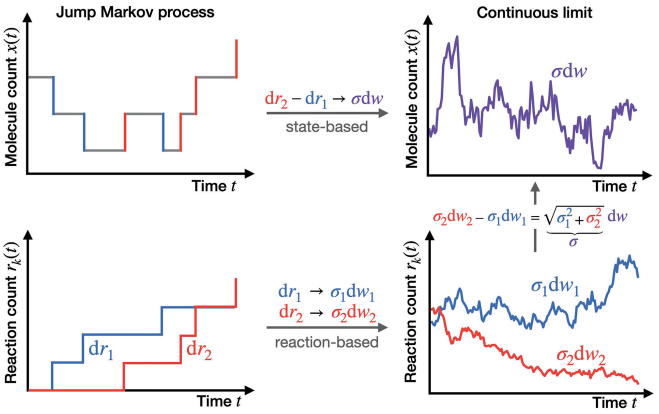

FIG. 1. State- versus reaction-based trajectories. As an example, we consider a birth-and-death process with copy number $x(t)$. The discrete trajectory on the left consists of two types of chemical reactions, a degradation reaction with stoichiometric coefficient $s_{x,1} = -1$ (blue) and a birth reaction with stoichiometric coefficient $s_{x,2} = 1$ (red). The conventional state-based continuous limit summarizes both reactions into effective concentration changes (top), which leads in general to a loss of information in state-based trajectories. In contrast, the suggested reaction-based continuous limit distinguishes birth- and degradation reactions (bottom), thus preserving the information that is encoded in the discrete reaction trajectories.

also lead to qualitatively different behavior, even when the copy numbers are large and the instantaneous mutual information is correctly predicted [11,12]. This observation raises the question of how information is encoded in trajectories and how this information can be quantified.

In this work, we introduce the notion of reaction- versus state-based descriptions of trajectories in the context of information transfer. In chemical systems the information on the input is encoded in the sequence of reaction events of the output, and the discrete trajectories of Markov jump processes capture this encoding. In contrast, the commonly used Gaussian approximation of trajectory mutual information [2] entails a state-based trajectory notion, where individual reaction channels are summarized into effective copy number or concentration fluctuations. This step results in a loss of information as we demonstrate using the data-processing inequality [13]. This explains why the commonly used form of the Gaussian approximation significantly underestimates trajectory mutual information even in the large copy number regime. This observation naturally leads to an alternative, reaction-based variant of the Gaussian approximation, which explicitly keeps track of the individual reaction channels (see Fig. 1). We show that this approach accurately predicts the trajectory mutual information for linear chemical systems at large copy numbers. Finally, we demonstrate that the distinction between reaction- versus state-based trajectories has ramifications for information transmission in cellular systems, by analyzing systems that can harness the information that resides in the reaction events.

Stochastic reaction dynamics—We consider a well-mixed reaction system of size Ω, consisting of M chemical species $Z_1, \ldots, Z_M$. These species interact via K distinct reaction channels. The state of the system is denoted by the vector $z(t)$, which contains the copy numbers of the respective species at time t. To each reaction channel k, we associate a rate function $h_k(z(t))$ and define $h(z) = [h_1(z), \ldots, h_K(z)]^T$ and $H(z) = \mathrm{diag}[h(z)]$. For sufficiently large systems, we can make use of the linear-noise approximation $z(t) = \hat{z}(t) + \delta z(t)$, where $\hat{z}(t)$ is the macroscopic mean copy number of $z(t)$ and $\delta z(t)$ is a random fluctuation that scales with the square root of the system size $\Omega^{1/2}$ [10]. For stationary systems, the dynamics of $\delta z(t)$ satisfies

$$\mathrm{d}[\delta z(t)] = SJ\delta z(t)\mathrm{d}t + SH^{1/2}\mathrm{d}w(t), \tag{1}$$

where S is the stoichiometry matrix with dimension $M \times K$, J is the Jacobian of $h(z)$ evaluated at the stationary macroscopic mean z^*, and $\mathrm{d}w$ denotes a K-dimensional noise vector of independent, standard Wiener processes, with each element $\mathrm{d}w_k(t)$ reflecting an individual reaction channel (see [14] Section A).

Reaction- versus state-based trajectories—The main difference between reaction- and state-based trajectories lies in whether they are defined using a K-dimensional reaction noise vector, or an M-dimensional vector describing the effective copy number noise. In Eq. (1), the matrix $SH^{1/2}$ transforms the K-dimensional reaction noise vector $\mathrm{d}w$ into the M-dimensional species-level noise $\mathrm{d}\xi = SH^{1/2}\mathrm{d}w$, summarizing the net contribution of reactions to the effective change in copy numbers. Since the linear combination of two Gaussian increments is again a Gaussian increment, $\mathrm{d}\xi$ remains a vector of Gaussian increments with a modified covariance structure. For instance, the sum of two independent Gaussian increments $a\mathrm{d}w_1(t) + b\mathrm{d}w_2(t)$ results in a single Gaussian increment $\sqrt{a^2 + b^2}\mathrm{d}w(t)$. This projection leads to the commonly used state-based formulation of the LNA, for which trajectories are defined as $\tilde{z}_0^t = \{\mathrm{d}\xi(s) | 0 \le s < t\}$.

Importantly, the projection from reaction-based onto state-based increments is irreversible. Summarizing reaction channels, as done for state-based trajectories, can only lead to a loss of information as we demonstrate more explicitly in Case Study I. In fact, while the effective noise on a species contains the effect of *all* reactions that modify a species, the information on the input trajectory is, in general, contained only in a subset of reaction channels. The remaining reactions act as a source of uninformative noise, which hampers information transfer.

This observation leads to the formulation of trajectories at the level of reaction channels. In this formalism, each trajectory is described by the individual reaction events, i.e., $z_0^t = \{\mathrm{d}w_k(s), k \in \{1, \ldots, K\} | 0 \le s < t\}$. The reaction-based trajectory distinguishes individual reaction events

and is, hence, analogous to jump processes, where trajectories are unambiguously defined by the sequence of reaction times and types. In the following, we will show that the previously observed discrepancies between the Gaussian and discrete trajectory mutual information can be explained by the difference in defining trajectories in the linear noise regime, i.e., either via the state- or the reaction-based formalism.

Trajectory mutual information for Gaussian processes—The calculation of mutual information between Gaussian trajectories has been studied extensively [19–21]. Let $x(t)$ and $y(t)$ be the copy numbers of two chemical species that are part of the system state $z(t)$. The trajectory mutual information between trajectories of $x(t)$ and $y(t)$ is generally defined as

$$I_t^{xy} = \left\langle \log \frac{\mathrm{dP}^{xy}}{\mathrm{d}(\mathrm{P}^x \times \mathrm{P}^y)} \right\rangle, \tag{2}$$

where the term inside the logarithm is the Radon-Nikodym density of the joint path measure P^{xy} with respect to the product of the marginal path measures $\mathrm{P}^x \times \mathrm{P}^y$ [22]. Using the expressions for the Radon-Nikodym density for Wiener processes, we can evaluate Eq. (2) for both state- and reaction-based trajectories [19–22]. The mutual information between state-based trajectories $\tilde{x}_0^t \subset \tilde{z}_0^t$ and $\tilde{y}_0^t \subset \tilde{z}_0^t$ reads

$$\tilde{I}_t^{xy} = \frac{1}{2}\int_0^t \frac{1}{\sigma_x^2}\Big(\langle \mathrm{Var}[\phi_x(z(s))|\tilde{x}_0^s]\rangle - \langle \mathrm{Var}[\phi_x(z(s))|\tilde{x}_0^s, \tilde{y}_0^s]\rangle\Big) + \frac{1}{\sigma_y^2}\Big(\langle \mathrm{Var}[\phi_y(z(s))|\tilde{y}_0^s]\rangle - \langle \mathrm{Var}[\phi_y(z(s))|\tilde{x}_0^s, \tilde{y}_0^s]\rangle\Big)\mathrm{d}s, \tag{3}$$

where $\phi_x(z(t)) = \sum_{k\in R_{\tilde{x}}} s_{x,k}\lambda_k(z(t))$ and $\phi_y(z(t)) = \sum_{k\in R_{\tilde{y}}} s_{y,k}\lambda_k(z(t))$ are the drifts of $x(t)$ and $y(t)$ with $\lambda_k(z(t)) = \sum_i \partial_{z_i} h_k(z^*)\delta z_i(t)$, $\sigma_x = \sqrt{\sum_{k\in R_{\tilde{x}}} s_{x,k}^2 h_k(z^*)}$ and $\sigma_y = \sqrt{\sum_{k\in R_{\tilde{y}}} s_{y,k}^2 h_k(z^*)}$ denote the noise magnitudes of $x(t)$ and $y(t)$ and $s_{x,k}$ and $s_{y,k}$ are the stoichiometric coefficients of reaction k modifying $x(t)$ and $y(t)$, respectively. The noise magnitudes σ_x and σ_y are determined by the sets of reactions $R_{\tilde{x}}$ and $R_{\tilde{y}}$ that affect $x(t)$ or $y(t)$, respectively.

Evaluating Eq. (2) for reaction-based trajectories $x_0^t \subset z_0^t$ and $y_0^t \subset z_0^t$ instead, yields a different equation for the trajectory mutual information,

$$I_t^{xy} = \frac{1}{2}\int_0^t \sum_{k\in R_x} \frac{\langle \mathrm{Var}[\lambda_k(z(s))|x_0^s]\rangle - \langle \mathrm{Var}[\lambda_k(z(s))|x_0^s, y_0^s]\rangle}{\sigma_k^2} + \sum_{k\in R_y} \frac{\langle \mathrm{Var}[\lambda_k(z(s))|y_0^s]\rangle - \langle \mathrm{Var}[\lambda_k(z(s))|x_0^s, y_0^s]\rangle}{\sigma_k^2}\mathrm{d}s, \tag{4}$$

where $\sigma_k = \sqrt{h_k(z^*)}$ denotes the noise magnitude of reaction k. The sums in Eq. (4) run over the subsets $R_x \subseteq R_{\tilde{x}}$ and $R_y \subseteq R_{\tilde{y}}$, which contain the reactions that affect $x(t)$ or $y(t)$, respectively, but are caused by species other than $x(t)$ or $y(t)$, as discussed further in End Matter. Note that the rate of information transmission in both cases is given by the integrands of Eqs. (3) and (4).

Comparing Eqs. (3) and (4) highlights important differences in the information transfer between state- and reaction-based trajectories. For state-based trajectories, the mutual information between species X and Y involves *all* reactions that change the copy numbers of these species. This is reflected in the denominators of Eq. (3), which contain effective noise magnitudes σ_x and σ_y with contributions from all reactions in $R_{\tilde{x}}$ and $R_{\tilde{y}}$. In contrast, each term in the sums of Eq. (4) is accompanied by a reaction-specific noise magnitude σ_k. This is because in the reaction-based description, only subsets of reactions $R_x \subseteq R_{\tilde{x}}$ and $R_y \subseteq R_{\tilde{y}}$ are considered, namely those that explicitly contribute to information transfer. For example, in the reaction network $Y \to X$, $X \to \varnothing$, both reactions belong to $R_{\tilde{x}}$ and contribute to σ_x, but only the first reaction is also contained in R_x, since only that reaction contributes directly to information transfer from Y to X. Reactions, that are contained in $R_{\tilde{x}}$ but not in R_x can contribute to information transfer, but only implicitly via their effect on the expected conditional variances.

Calculation of the trajectory mutual information—The calculation of the trajectory mutual information for state- and reaction-based trajectories requires knowledge of the expected conditional variances that appear in Eqs. (3) and (4). The conditional variances inside the expectations are the second moments of a so-called filtering distribution, which captures the statistics of a subset of the species at time t, given a complete trajectory of the remaining species between time zero and t. For the calculation of $\mathrm{Var}[\lambda_k(z(t))|x_0^t]$ in Eq. (4), for instance, we need the probability distribution $\pi^x(\bar{z}, t) = p[\bar{z}(t)|x_0^t]$, where $\bar{z}(t)$ contains the copy numbers of all species except $x(t)$. Since the considered dynamics are linear and Gaussian, the required conditional distributions and their second moments are tractable. In the case of state-based trajectories, the expected conditional variances are readily available through the well-known Kalman-Bucy filtering equations [23,24]. In the case of reaction-based trajectories, the calculation of the conditional variances is less straightforward because the standard Kalman-Bucy filter does not apply. We show in [14] Sections C–F how tractable solutions of the expected conditional variances can be obtained using filtering theory [24,25].

Case study (I)—comparison of different network motifs—To illustrate the consequences of using different trajectory descriptions, we calculated mutual information rates for two common reaction motifs [2,26] using reaction- and state-based trajectories (Table I). Even though the studied

TABLE I. Stationary mutual information rate between the trajectories of input species A and output species B of two common reaction motifs obtained via the reaction-based and state-based scheme. All motifs include the reactions $\varnothing \underset{c_2}{\overset{c_1}{\rightleftharpoons}} A$. The motifs have been studied earlier in [2,26], where motif a is referred to as motif III and motif b as motif II. The state-based information rate has been calculated using the formalism by Tostevin *et al.* [2].

Motif	Reaction	State-based rate [2]	Reaction-based rate
(a)	$A \overset{c_3}{\to} A + B,\ B \overset{c_4}{\to} \varnothing$	$-(c_2/2) + \frac{1}{2}\sqrt{c_2(c_2 + c_3)}$	$-(c_2/2) + \frac{1}{2}\sqrt{c_2(c_2 + 2c_3)}$
(b)	$A \overset{c_3}{\to} B \overset{c_4}{\to} \varnothing$	∞	$(c_3/2)$

motifs are quite simple, the results are markedly different. For motif a, we find that the effect of reaction 3 (i.e., with rate constant c_3) is stronger for the reaction-based formalism, where the reaction contributes twice as much to the information rate as in the state-based one. Tracking the occurrence of this factor in the calculation of the trajectory mutual information rate shows that the inclusion of reaction 4 (i.e., with rate constant c_4) in the state-based formalism is responsible for the attenuation of reaction 3. While only the fluctuations caused by reaction 3 carry information about the signal (species A), in the state-based description, distinguishing reaction 3 from 4 is impossible as $\mathrm{d}\xi_b(t) = \sqrt{c_3 a^* + c_4 b^*}\mathrm{d}w_b(t)$, resulting in a loss of information about the reaction events that encode the signal. Instead, in the reaction-based scheme the fluctuations caused by reaction 3 are tracked explicitly, resulting in a higher mutual information rate. The loss of information associated with state-based trajectories can be explained also by the data-processing inequality [13]. Since reaction 4 acts purely as a (noisy) postprocessing step, with no additional information about the signal $a(t)$, the mutual information rate can only decrease when reactions 3 and 4 are combined (see [14] Section H.1 for further discussion).

The discrepancies between reaction- and state-based trajectories can be much more pronounced, as can be seen from motif b. For state-based trajectories, the noise terms are associated with the net change in copy number of the respective species. In motif b, the reaction-specific noise terms $\sqrt{c_1}\mathrm{d}w_1(t)$, $\sqrt{c_2 a^*}\mathrm{d}w_2(t)$, $\sqrt{c_3 a^*}\mathrm{d}w_3(t)$ and $\sqrt{c_4 b^*}\mathrm{d}w_4(t)$ are projected onto the noise terms of species A and B as $\mathrm{d}\xi_a(t) = \sqrt{c_1 + c_2 a^* + c_3 a^*}\mathrm{d}w_a(t)$ and $\mathrm{d}\xi_b(t) = \sqrt{c_3 a^* + c_4 b^*}\mathrm{d}w_b(t)$, respectively. Since reaction 3 affects both $\mathrm{d}\xi_a(t)$ and $\mathrm{d}\xi_b(t)$, these increments are anticorrelated at every moment in time. The trajectory mutual information for instantly correlated copy number dynamics diverges in the state-based formalism. Using the method by Tostevin *et al.*, we obtain an infinite information rate for motif b [2] (see also Ref. [27]). In contrast, the reaction-based formalism keeps track of the individual reactions and the information rate remains finite.

Case study (II)—ligand-receptor binding—Since the reaction-based information rate is, in general, larger than the state-based rate (unless the latter diverges), the question arises whether biological systems are able to access the information stored in individual reactions. To study whether this is possible, we revisit the problem of chemical sensing. In the mechanism of time integration, first analysed by Berg and Purcell (BP), the ligand concentration is estimated from the average receptor occupancy over some integration time T [28]. This receptor occupancy depends on both the ligand binding and unbinding rate. Endres and Wingreen realized, however, that only the binding events provide information on the concentration, since only the binding rate depends on the concentration [29]. This observation led to the scheme of maximum-likelihood (ML) sensing [29], in which the concentration is inferred from the mean duration of the unbound state of the receptor, rather than from its average occupancy. The sensing error of this ML scheme is indeed lower than that of the mechanism of time integration, by precisely a factor of 2 [28–34]. The observation that the information on the ligand concentration is encoded only in the binding reaction, suggests that the ML sensing error is related to the information rate computed for reaction-based signal trajectories. In contrast, the BP sensing error is expected to be related to the information rate of state-based trajectories.

To test this idea, we consider a system consisting of a time-varying ligand concentration and a receptor that senses this ligand by switching between a ligand-bound state and an unbound state. We denote by $l(t), r_b(t), r_u(t)$ the copy numbers of the ligand and the ligand-bound and unbound receptors, respectively. The total number of receptor molecules is given by $r_T = r_b(t) + r_u(t)$ and is constant. In this model, the information rate between the ligand and the ligand-bound receptor state is calculated as

$$i_\alpha^{lr} = -\frac{c_2}{2} + \frac{c_2}{2}\sqrt{\frac{\kappa_\alpha r_T c_- c_+}{c_2 c_- + c_1 c_+} + 1}, \tag{5}$$

with $\alpha = \mathrm{SB}$ and $\kappa_\alpha = \kappa_{SB} = 1$ for state-based trajectories and $\alpha = \mathrm{RB}$ and $\kappa_\alpha = \kappa_{RB} = 2$ for reaction-based trajectories. Interestingly, we can rewrite the expressions for both information rates in terms of the relative sensing error $\eta_\beta^2 = (\delta l_\beta / l^*)^2$, with δl_β referring to the sensing error for maximum-likelihood- and Berg-Purcell sensing, i.e., $\beta \in \{\mathrm{ML}, \mathrm{BP}\}$, respectively and l^* as the stationary average ligand copy number. If the downstream system reads out the receptor over a time T, then the respective errors per receptor

are $\eta_{ML}^2 = 1/\bar{n}_b$ and $\eta_{BP}^2 = 2/\bar{n}_b$, where $\bar{n}_b = Tp/\bar{\tau}_b$ is the average number of binding events during the time T, with $p = r_b^*/r_T$ the binding probability, r_b^* the stationary average copy number of bound receptors, and $\bar{\tau}_b = 1/c_-$ the average binding time [28,29,34]. Since the optimal integration time T is on the order of the correlation time c_2^{-1} of the input signal [35], we take $T = c_2^{-1}$. The information rates can then be rewritten as

$$i_\alpha^{lr} = \frac{c_2}{2}\left(\sqrt{\frac{2r_T}{\eta_\beta^2 l^*} + 1} - 1\right), \tag{6}$$

where $\beta = \mathrm{ML}$ for $\alpha = \mathrm{RB}$ and $\beta = \mathrm{BP}$ for $\alpha = \mathrm{SB}$. Hence, the respective expressions for the information rate differ only by the sensing error. Since the reaction-based information rate is the maximal information rate, this observation supports the idea that ML sensing can optimally extract the information that the binding reactions provide on the input concentration [29]. Noting that for this input the variance $\mathrm{Var}[l(t)]$ is given by the mean l^*, we also see that these expressions lead to the intuitive interpretation that the information rate is determined by the speed at which independent messages are transmitted through the system, set by the correlation time c_2^{-1} of the input, times the number of messages, i.e., concentration levels, that are transmitted reliably per input correlation time, given by $\sqrt{\mathrm{Var}[l(t)]}/(\delta l_\alpha/\sqrt{r_T})$, with $\delta l_\alpha/\sqrt{r_T}$ the absolute sensing error (for additional details see [14] Section I).

Still, the question remains whether cells are able to transfer and access as much information as predicted by Eq. (6) with $\alpha = \mathrm{RB}$. Here, we turn to the scheme proposed by Lang *et al.* [32]. In this scheme, the cell estimates the concentration from the average receptor occupancy over some integration time T as in the classical Berg-Purcell scheme. However, while in the classical scheme receptor unbinding is a one-step process, in the scheme of [32] the receptor, upon ligand binding, goes through a sequence of steps before the ligand is released, such that the waiting time for ligand unbinding becomes narrower, thus reducing the contribution from the unbinding noise to the error in the estimate of the receptor occupancy [Fig. 2(a)]. This leads to the hypothesis that by reading out the receptor *state*, cells are nonetheless able to access the *reaction-based* information rate by eliminating the stochasticity in the unbinding of the ligand.

Calculating the state-based mutual information rate i_{MS}^{lr} between the ligand and the set of all active states of such a multistate receptor (see [14] Section I.4), we find indeed that i_{MS}^{lr} approaches i_{RB}^{lr} in the reaction-based scheme as $N \to \infty$ [Fig. 2(b)]. Increasing the number of receptor states, the unbinding of the ligand becomes less noisy, reducing the information-loss in state-based trajectories. This shows that nonequilibrium reaction cycles can make the information stored in individual chemical reactions

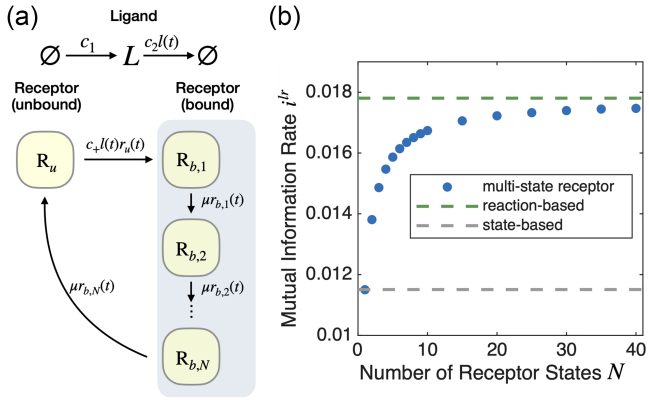

FIG. 2. Multistate receptor system. (a) Sketch of the multistate receptor system. The symbols $\mathrm{R}_{b,i}$, $i \in \{1, \ldots, N\}$ denote the ligand-bound states of the receptor. (b) Stationary mutual information rate i^{lr} between the state-based trajectories $\tilde{l}_0^t$ and $\tilde{r}_0^t$ of the multistate receptor as a function of the number of states N (blue dots). Information rates for the two-state receptor for state-based trajectories (gray dashed line) and reaction-based trajectories (green dashed line) are shown for comparison. Simulations were performed with parameters $\{c_1, c_2, c_+, c_-, \mu, r_T\} = \{1, 0.01, 1, 1, Nc_-, 10\}$.

accessible to the cell. In [14] Section J, we present a minimal model of a linear multi-state signalling process to support the generality of this statement.

Conclusion—We introduced the notion of state-based and reaction-based trajectories and illustrated their impact on information transfer. In the state-based formalism, multiple reaction channels are summarized into effective copy number fluctuations, which, in general, leads to a loss of information about the input signal. The widely used LNA-based approximation of trajectory mutual information relies on state-based trajectories and can therefore substantially underestimate information transmission in chemical systems—even in the high-copy number regime. We showed that a reformulation of the Gaussian mutual information in terms of reaction-based trajectories prevents this loss of information. This suggests that the principal difference between the state-based Gaussian approximation and the Markov jump process is not the discreteness of molecules, but the encoding of information in reaction events rather than concentrations. We demonstrated that the higher information encoded in reaction-based trajectories can be read out through active reaction cycles, which reduce noise stemming from noninformative events. Similar reaction cycles can be found in signal transduction pathways [32], protein degradation systems (e.g., via multistep ubiquitination [36]) and gene promoter networks [37]. It will be interesting to understand where these systems sit on the spectrum from state- to reaction-based information transfer.

Our findings are conceptually related to recent work in stochastic thermodynamics, where entropy production may be underestimated when a discrete system is taken to the

continuum limit under coarse-graining [38,39]. Similarly to the present work, these discrepancies are explained by the fact that individual microscopic currents are no longer distinguished in the coarse-grained continuous dynamics, leading to a loss of information. Gaining a deeper understanding of how these findings relate will be an interesting topic for future work.

Acknowledgments—The authors thank Tommaso Bianucci, Avishek Das, and Andreas Hilfinger for useful insights and critical feedback on this work. This work was supported by core funding of the Max Planck Institute of Molecular Cell Biology and Genetics, is part of the Dutch Research Council (NWO) and was partly performed at the research institute AMOLF. This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (Grant Agreement No. 885065) and was financially supported by NWO through the "Building a Synthetic Cell (BaSyC)" Gravitation Grant (024.003.019).

Data availability—Source codes underlying the numerical analyses of Fig. 2 are openly available [40].

[1] C. E. Shannon, A mathematical theory of communication, Bell Syst. Tech. J. **27**, 379 (1948).
[2] F. Tostevin and P. R. ten Wolde, Mutual information between input and output trajectories of biochemical networks, Phys. Rev. Lett. **102**, 218101 (2009).
[3] R. M. Fano, Transmission of information: A statistical theory of communications, Am. J. Phys. **29**, 793 (1961).
[4] L. Duso and C. Zechner, Path mutual information for a class of biochemical reaction networks, in *2019 IEEE 58th Conference on Decision and Control (CDC)* (2019), pp. 6610–6615.
[5] M. Sinzger, M. Gehri, and H. Koeppl, Poisson channel with binary Markov input and average sojourn time constraint, in *2020 IEEE International Symposium on Information Theory (ISIT)* (2020), pp. 2873–2878.
[6] F. Tostevin and P. R. ten Wolde, Mutual information in time-varying biochemical systems, Phys. Rev. E **81**, 061917 (2010).
[7] W. H. de Ronde, F. Tostevin, and P. R. ten Wolde, Effect of feedback on the fidelity of information transmission of time-varying signals, Phys. Rev. E **82**, 031914 (2010).
[8] E. Ziv, I. Nemenman, and C. H. Wiggins, Optimal signal processing in small stochastic biochemical networks, PLoS One **2**, 1 (2007).
[9] S. Tănase-Nicola and P. R. ten Wolde, Regulatory control and the costs and benefits of biochemical noise, PLoS Comput. Biol. **4**, 1 (2008).
[10] N. G. Van Kampen, *Stochastic Processes in Physics and Chemistry* (Elsevier, New York, 1992), Vol. 1.
[11] A.-L. Moor and C. Zechner, Dynamic information transfer in stochastic biochemical networks, Phys. Rev. Res. **5**, 013032 (2023).
[12] M. Reinhardt, G. Tkačik, and P. R. ten Wolde, Path weight sampling: Exact Monte Carlo computation of the mutual information between stochastic trajectories, Phys. Rev. X **13**, 041017 (2023).
[13] T. M. Cover, *Elements of Information Theory* (John Wiley & Sons, New York, 1999).
[14] See Supplemental Material at http://link.aps.org/supplemental/10.1103/nfk6-8x5s for mathematical details and derivations, which includes Refs. [15–18].
[15] D. T. Gillespie, The chemical Langevin equation, J. Chem. Phys. **113**, 297 (2000).
[16] E. W. J. Wallace, A simplified derivation of the linear noise approximation, arXiv:1004.4280.
[17] Andrew H. Jazwinski, ed., Nonlinear filtering theory, in *Stochastic Processes and Filtering Theory*, Vol. 64, Mathematics in Science and Engineering (Elsevier, New York, 1970), pp. 162–193.
[18] P. B. Warren, S. Tânase-Nicola, and P. R. Ten Wolde, Exact results for noise power spectra in linear biochemical reaction networks, J. Chem. Phys. **125**, 144904 (2006).
[19] T. T. Kadota, M. Zakai, and J. Ziv, Mutual information of the white Gaussian channel with and without feedback, IEEE Trans. Inf. Theory **17**, 368 (1971).
[20] M. Hitsuda, Mutual information in Gaussian channels, J. Multivariate Anal. **4**, 66 (1974).
[21] T. E. Duncan, On the calculation of mutual information, SIAM J. Appl. Math. **19**, 215 (1970).
[22] R. S. Liptser and A. N. Shiriaev, *Statistics of Random Processes: General Theory* (Springer, New York, 1977), Vol. 394.
[23] R. E. Kálmán and R. S. Bucy, New results in linear filtering and prediction theory, J. Basic Eng. **83**, 95 (1961).
[24] A. Bain and D. Crisan, *Fundamentals of Stochastic Filtering* (Springer, New York, 2009), Vol. 3.
[25] A. Kutschireiter, S. C. Surace, and J.-P. Pfister, The Hitchhiker's guide to nonlinear filtering, J. Math. Psychol. **94**, 102307 (2020).
[26] S. Tănase-Nicola, P. B. Warren, and P. R. ten Wolde, Signal detection, modularity, and the correlation between extrinsic and intrinsic noise in biochemical networks, Phys. Rev. Lett. **97**, 068102 (2006).
[27] R. Chétrite, M. L. Rosinberg, T. Sagawa, and G. Tarjus, Information thermodynamics for interacting stochastic systems without bipartite structure, J. Stat. Mech. (2019) 114002.
[28] H. Berg and E. M. Purcell, Physics of chemoreception, Biophys. J. **20**, 193 (1977).
[29] R. G. Endres and N. S. Wingreen, Maximum likelihood and the single receptor, Phys. Rev. Lett. **103**, 158101 (2009).
[30] K. Kaizu, W. de Ronde, J. Paijmans, K. Takahashi, F. Tostevin, and P. R. ten Wolde, The Berg-Purcell limit revisited, Biophys. J. **106**, 976 (2014).
[31] W. Bialek and S. Setayeshgar, Physical limits to biochemical signaling, Proc. Natl. Acad. Sci. U.S.A. **102**, 10040 (2005).
[32] A. H. Lang, C. K. Fisher, T. Mora, and P. Mehta, Thermodynamics of statistical inference by cells, Phys. Rev. Lett. **113**, 148103 (2014).
[33] T. Mora and I. Nemenman, Physical limit to concentration sensing in a changing environment, Phys. Rev. Lett. **123**, 198101 (2019).

[34] P. R. ten Wolde, N. B. Becker, T. E. Ouldridge, and A. Mugler, Fundamental limits to cellular sensing, J. Stat. Phys. **162**, 1395 (2016).
[35] G. Malaguti and P. R. ten Wolde, Theory for the optimal detection of time-varying signals in cellular sensing systems, eLife **10**, e62574 (2021).
[36] N. W. Pierce, G. Kleiger, S.-o. Shan, and R. J. Deshaies, Detection of sequential polyubiquitylation on a millisecond timescale, Nature (London) **462**, 615 (2009).
[37] D. M. Suter, N. Molina, D. Gatfield, K. Schneider, U. Schibler, and F. Naef, Mammalian genes are transcribed with widely different bursting kinetics, Science **332**, 472 (2011).
[38] D. M. Busiello and A. Maritan, Entropy production in master equations and Fokker-Planck equations: Facing the coarse-graining and recovering the information loss, J. Stat. Mech. (2019) 104013.
[39] D. M. Busiello, J. Hidalgo, and A. Maritan, Entropy production for coarse-grained dynamics, New J. Phys. **21**, 073004 (2019).
[40] https://github.com/zechnerlab/PathMI_state_vs_reaction/.

End Matter

Appendix: Reaction sets—In this Appendix, we aim to explain the reaction sets used throughout this letter in detail. We restrict ourselves to the sets $R_z, R_{\bar{z}}, R_{\bar{x}}$, and R_x because $R_{\bar{y}}$ and R_y follow the same rules as $R_{\bar{x}}$ and R_x. Let $z(t)$ be a vector containing the state of all species at time t. The reaction set R_z contains all reactions that modify any species in $z(t)$. For our filtering approach, we divide the system into an observable part and an unobservable part. Let $x(t) \subseteq z(t)$ be the state of the species we observe. Then, we are interested in the conditional expectation value of all species other than X given the history x_0^t of $x(t)$. In this context, $\bar{z}(t) \subseteq z(t)$ denotes the state of all species other than $x(t)$, i.e., the part of $z(t)$ that we do not observe. The subset $R_{\bar{z}} \subseteq R_z$ is the set of all reactions that modify $\bar{z}(t)$. Analogously, the set $R_{\bar{x}} \subseteq R_z$ is the set of all reactions that modify $x(t)$. For instance, the reactions $X \to Z_1 + X$ and $Z_1 \to Z_2$ are contained in $R_{\bar{z}}$, whereas the reaction $X \to \varnothing$ is contained in $R_{\bar{x}}$. As $x(t)$ and $\bar{z}(t)$ can change simultaneously, the intersection of these sets is generally not empty, i.e., $R_{\bar{z}} \cap R_{\bar{x}} \neq \varnothing$. For instance, $X \to Z_2$ is contained in the intersection of these sets. With $R_x \subseteq R_{\bar{x}}$, we denote the set of reactions that modify $x(t)$ and whose rate depends on a species in $\bar{z}(t)$, such as the reactions $Z_3 \to X + Z_3$ or $Z_4 \to X$. Again, this shows that $x(t)$ and $\bar{z}(t)$ can change simultaneously and it follows $R_{\bar{z}} \cap R_x \neq \varnothing$. Note that we exclude systems involving reactions where the species X and Y are modified simultaneously and the corresponding rate function depends on both $x(t)$ and $y(t)$, such as $Y + X \to \varnothing$. This is because the assignment of such reaction to either X or Y is not unique.