Challenges in Inferring the Directionality of Active Molecular Processes from Single-Molecule Fluorescence Resonance Energy Transfer Trajectories

Aljaz Godec* and Dmitrii E. Makarov*

ABSTRACT: We discuss some of the practical challenges that one faces in using stochastic thermodynamics to infer directionality of molecular machines from experimental single-molecule trajectories. Because of the limited spatiotemporal resolution of single-molecule experiments and because both forward and backward transitions between the same pairs of states cannot always be detected, differentiating between the forward and backward directions of, e.g., an ATP-consuming molecular machine that operates periodically, turns out to be a nontrivial task. Using a simple extension of a Markov-state model that is commonly employed to analyze single-molecule transition-path measurements, we illustrate how irreversibility can be hidden from such measurements but in some cases can be uncovered when non-Markov effects in low-dimensional single-molecule trajectories are considered.

INTRODUCTION

Nonequilibrium phenomena have direction or “arrow of time”; that is, a movie of a nonequilibrium process played backward will look statistically different from the original movie. The directionality or time irreversibility of life is evident at the macroscopic scale. If, however, we zoom in on the microscopic motion of molecules that constitute a living organism, such directionality becomes far from evident. For example, the Maxwell–Boltzmann distribution of molecular velocities, a fingerprint of equilibrium and thus of the time-reversible state of matter, not only enables us to discuss our body’s temperature but also ensures the validity of the Arrhenius law that biochemists unabashedly use to describe the rates of elementary steps in nonequilibrium biochemical networks.

The question of whether specific components of biomolecular machinery undergo equilibrium or nonequilibrium dynamics has been the subject of some controversy. For example, while a molecular motor walking along its track clearly steps more often in one direction than the other, it has been argued that individual steps are time-reversible and follow the same (average) path in the forward and backward direction, yet we know that such microscopic reversibility must be violated at some length scale to give rise to directional persistence. For example, according to the “scalloping theorem”, a bacterium could not swim if its movements were time-reversible (at least in incompressible fluids). Another obvious example is cell division.

The problem of detecting directionality at a molecular level is twofold. First, adequate experimental tools are needed. Because molecular machines usually operate under nonequilibrium steady-state conditions (that is, concentrations of various molecules in the cell do not change over time), single-molecule experiments are usually required to study their dynamics. Despite single-molecule experiments becoming nearly routine tools of molecular biology, only a small number of studies have attempted to address the question of directionality of biochemical phenomena and/or employed nonequilibrium models to describe single-molecule data except when directionality is self-evident (e.g., the rotation or walking of a molecular motor along its track). Second, proper data analysis is needed. There have been many recent efforts to identify and quantify nonequilibrium effects in living systems, yet this has proven to be difficult even in mesoscopic systems where the dynamics of the system can often be observed directly and quite accurately, e.g., with a fast camera. Inferring and quantifying irreversible behavior at smaller, molecular scales pose additional difficulties, as the time and length scales of the probes employed in single-molecule experiments often overlap with the time and length scales of the phenomena probed. In this Perspective, we argue that those limitations are especially

Received: October 26, 2022
Accepted: December 14, 2022
important in single-molecule studies that probe nonequilibrium steady-state phenomena, describe challenges faced when applying ideas of stochastic thermodynamics to single-molecule data, and outline promising theoretical and experimental approaches to overcoming these challenges.

LIMITATIONS OF SINGLE-MOLECULE MEASUREMENTS IN APPLICATION TO ACTIVE MOLECULAR PROCESSES

Single-Molecule Measurements Report on Projected Dynamics. Single-molecule experiments have the ability to probe conformational changes in individual molecules in real time. Such measurements, however, often suffer from limited spatial and temporal resolution and inevitably entail a coarse-grained view of the process. In fluorescence resonance energy transfer (FRET) studies, for example, the ability to resolve a microscopic state of a molecule is limited by the number of individual photons that can be detected while the molecule resides in this state; moreover, experimentalists usually have only a limited number of distinct photon colors (usually two) available to differentiate among multiple states. Likewise, in single-molecule force spectroscopy studies, the dynamics of the molecular coordinate of interest, such as the distance between two residues in a protein, must be inferred from the displacement of a sluggish force probe.

These limitations are especially important when single-molecule techniques are applied to molecular machines that operate away from equilibrium. Sometimes the ensuing directionality (i.e., time irreversibility) in their motion is immediately evident from data. For example, a molecular motor can be observed walking in a particular direction, but there are also many cases in which this directionality is more subtle. For example, the dynamics of HSP90 is driven by ATP hydrolysis, yet the motion is periodic and observed along a one-dimensional reaction coordinate; thus, its time irreversibility is not straightforward to discern. To illustrate the difficulties that arise from limited spatiotemporal resolution, consider a minimal hypothetical three-state kinetic scheme of the type that has, for example, been used to describe the operation of the disaggregation machine ClpB (Figure 1).

The observed process is a conformational rearrangement between two mesoscopic states A and B, but the microscopic underlying process is described by a three-state kinetic scheme driven by, e.g., coupling to ATP hydrolysis. According to Kolmogorov’s cycle criterion, the dynamics of the system is time-reversible only if, for the 1 → 2 → 3 → 1 cycle, the product of the forward rate coefficients is the same as the product of the backward rate coefficients, i.e.,

\[ \epsilon^X = \frac{k_{1 \rightarrow 2}k_{2 \rightarrow 3}k_{3 \rightarrow 1}}{k_{2 \rightarrow 1}k_{3 \rightarrow 2}k_{1 \rightarrow 3}} \]  

is equal to 1. Here \( k_{i \rightarrow j} \) denotes the (pseudo)first-order rate coefficient for the transition from \( i \) to \( j \) if this transition is coupled to, e.g., ATP hydrolysis, then \( k_{i \rightarrow j} \) would be proportional to the ATP concentration such that \( \epsilon^X = 1 \) only when the concentrations of molecules participating in the process assume their equilibrium values. The quantity \( X \) (which, in the language of stochastic thermodynamics, is sometimes called the affinity of the 1 → 2 → 3 → 1 cycle) quantifies the thermodynamic force driving the process. When \( X > 0 \), the steady state of the system entails overall motion in the clockwise direction. This directional motion can be further characterized by a non-zero flux \( J \), which is equal to the total number of overall cycles completed per unit time.

The degree of irreversibility of the dynamics is usually quantified by the mean entropy production rate, which is a measure of both the heat dissipated to the environment and of how statistically different the forward process is from its time reverse. For Markovian dynamics, the entropy production rate is given by

\[ \langle S \rangle = \lim_{\tau \to \infty} \frac{1}{\tau} \left( P[x(t)] \right) \]  

where \( P[x(t)] \) is the probability of a forward path \( x(t), t \in [0, \tau] \), \( P[x(\tau - t)] \) is the probability of its time reverse, Boltzmann’s constant is set to 1, and the angular brackets indicate averaging over the ensemble of trajectories. For continuous-time Markov jump dynamics, the entropy production is generally given by

\[ \langle S \rangle = \frac{1}{2} \sum_{i \neq j} (k_{i \rightarrow j}P_{i} - k_{j \rightarrow i}P_{j}) \ln \frac{k_{i \rightarrow j}P_{i}}{k_{j \rightarrow i}P_{j}} \]  

where \( P_{i} \) is the steady-state probability of finding the system in state \( i \). In the case of a single cycle as in Figure 1, this becomes

\[ \langle S \rangle = JX \]  

We note that, more generally, Kolmogorov’s cycle criterion demands that the product of the forward rates be equal to the product of the backward rates and thus \( X = 0 \) for any cycle present in the system for its dynamics to be time-reversible.

We now suppose that the details about transitions among the three microscopic states are hidden from the observer, who has access to only two coarser states A and B. In Figure 1, it is assumed that the observer cannot differentiate between states 2 and 3, which are therefore lumped together as state B. What can we say about the underlying process (and particularly about its irreversible, driven nature) by observing transitions between A and B? It is intuitively obvious that the direction of the process is much easier to discern from the sequence of the “microscopic” three states that it visits than from the corresponding coarse two-
state sequence. For example, the sequence 123123123123... is obviously irreversible (i.e., distinct from its temporal reverse 321321321321...) while the corresponding sequence ABABABABABAB... is reversible. For an arbitrary random sequence of states A and B, any transition from A to B is followed by a transition from B to A, so the corresponding unidirectional fluxes (i.e., mean numbers of transitions per unit time), \( J_{A \rightarrow B} \) and \( J_{B \rightarrow A} \), are identical, with zero overall flux \( J \). Using these fluxes, one can introduce the rate coefficients \( k_{A \rightarrow B} = J_{A \rightarrow B} / P_A \) and \( k_{B \rightarrow A} = J_{B \rightarrow A} / P_B \) where \( P_{A(B)} \) is the fraction of time spent in A(B), and model the two-state trajectory with jump rates \( k_{A \rightarrow B} \) and \( k_{B \rightarrow A} \); at this level of description, the process is manifestly reversible, and the measured entropy production rate is zero. Any information about the irreversible character of the original process is lost. This argument, however, does not imply that the irreversible character of the true dynamics cannot be recovered from the coarse-grained two-state trajectory, only that the Markov description of this trajectory based on the average fluxes or average dwell times in each state does not suffice. Non-Markov effects (i.e., memory) must be included in the model. Given the experimental limitations further outlined below, however, experimental detection of memory effects is far from straightforward.

The fact that the A \( \rightleftharpoons \) B Markovian kinetic scheme is fundamentally time-reversible has further implications. Usually, the Markov approximation is justified for the observables that are slow, slower than all “hidden” degrees of freedom that equilibrate much faster. Upon application to the kinetic scheme in Figure 1 (top left), states 2 and 3 that constitute compound state B could be lumped together and the scheme could be reduced to the two-state scheme (Figure 1, bottom) if the 2 \( \rightleftharpoons \) 3 interconversion dynamics within B is fast. However, the detailed balance may still be violated by the full kinetic model, while it is not violated by the reduced two-state model; in this case, the irreversibility of the true dynamics is hidden from observation.

We will see another example of this kind in Experimental Limitations in Transition-Path Measurements.

**Only Some Transitions Are Observable.** In the interpretation of single-molecule measurements, it is often beneficial to consider the combined dynamics of the molecule of interest and of the experimental probe. Here we illustrate this idea for a FRET experiment (further discussed below) in which the internal state of the molecule (1, 2, or 3 in Figure 2) is inferred indirectly from the colors and arrival times of photons emitted by two fluorescent labels called the donor (D) and acceptor (A). Photon-by-photon analysis of such an experiment, which is especially important when there is no clear separation between photophysical and biochemical time scales, can be accomplished by considering the combined states of the system specified by the molecule’s state and the state of the donor and the acceptor, each of which can be in either the ground state (D or A) or the excited state (D* or A*). In the resulting kinetic network of combined photophysical and molecular states (Figure 2), some transitions can be observed directly while others cannot. For example, a transition such as a 1D*A \( \rightarrow \) 1DA or 3DA* \( \rightarrow \) 3DA transition is detected because the donor or acceptor emits a photon of a specific color. A FRET transition such as a 1D*A \( \rightarrow \) 1DA* transition, in which excitation energy is transferred from the donor to the acceptor, cannot, however, be observed directly. It can be deduced only from the subsequent emission of a photon by the acceptor. Likewise, a transition in which the donor is excited by absorbing a photon from a light source is unobservable.

![Figure 2. Combined network representing dynamics of a molecule (here with three internal states) along with the photophysics of the two fluorescent dyes (donor D and acceptor A) that report on the molecule’s state. Some of the possible transitions within such a network are shown, including laser excitation, FRET (transfer of excitation from the excited donor D* to the acceptor), and emission by either the donor or the acceptor. Note that the arrow lengths do not necessarily correspond to the respective transition rates.](https://doi.org/10.1021/acs.jpclett.2c03244)

Detection of directionality and estimation of entropy production in such “partially observed” kinetic networks have been the subject of considerable recent interest in stochastic thermodynamics. A particularly promising approach to this problem, focusing on the waiting times between observed transitions (here, photon emission events) rather than on dwell times in network states or first passage times to network states, can be used to infer quantities such as entropy production. Unfortunately, the use of such an approach requires that, if transitions from one network state to another are observed, the reverse transitions also be observed, which, strictly speaking, is not true for FRET experiments. For example, while emission of a photon is observed, the reverse transition (i.e., photoexcitation) is not.

We therefore conclude that inferring directionality of molecular phenomena from single-molecule data poses both fundamental and practical open problems. On the fundamental side, we need to understand how much information about the directionality of the underlying process is retained in projected/partially observed dynamics; on the practical side, we need workable tools for estimating the entropy production and related quantities from such partially observed dynamics. While progress toward these goals has been achieved in the field of stochastic thermodynamics, it is not clear whether such existing approaches can be directly applied to single-molecule FRET data given their limitations described above.
**Transition Paths: Definition and Importance.** In what follows, we focus on the more specific problem of inferring the directionality of molecular processes via measurements of transition paths. Transition paths have recently become a subject of considerable interest in single-molecule biophysics. A transition path is a short segment of a molecular trajectory that accomplishes a successful conformational transition (Figure 3). More precisely, one considers the dynamics along a molecular reaction coordinate \( x \) of interest [such as the distance between two parts of a biopolymer labeled with two fluorescent dyes of different colors (Figure 3)] and associates state A with all configurations with \( x < x_A \), state B with configurations with \( x > x_B \), and a "transition region" with \( x_A < x < x_B \). Here \( x_A \) and \( x_B \) are the transition region boundaries that can (in principle) be chosen arbitrarily. A transition path enters the transition region through one boundary and exits through the other without escaping the transition region between them. The transition-path time thus does not include the time spent outside the transition region or the temporal duration of failed attempts to cross the transition region; it is the temporal duration of the successful transition itself. Importantly, while within the standard picture of chemical kinetics the transitions are deemed to be instantaneous, a kinetic description that accounts for transition-path times is more general than a kinetic model with instantaneous jumps between states.

Transition paths encapsulate the transition mechanisms that biochemists and biophysicists strive to glean from experimental data. For instance, transition-path times contain information about whether the dynamics of the experimental observable is Markovian, whether the transition involves multiple pathways that cannot be observed directly, or whether an on-pathway transient intermediate is visited by transition paths. For nonequilibrium systems, examination of transition paths informs one about the (ir)reversibility of the process in question. For example, one may ask whether the pathway taken by a molecular motor during a forward step is different from that taken during a backward step, regardless of whether the motor is making more steps in one direction than the other and thereby undergoing overall unidirectional motion. At the same time, the relatively short temporal duration of transition paths illustrates the challenge encountered when there is no clear time-scale separation between the dynamics of the probes employed (e.g., photophysics of the FRET donor–acceptor pair) and the intrinsic biochemical dynamics that one desires to measure (see Only Some Transitions Are Observable).

**Experimental Limitations in Transition-Path Measurements.** The detailed molecular trajectory of interest, \( x(t) \), is often inaccessible experimentally (at least with an unlimited spatiotemporal resolution) but, instead, is deduced from a different experimental observable. In two-color FRET experiments, for example, one infers \( x(t) \) from the arrival times of photons of two colors; in general, there is no one-to-one correspondence between the current value of \( x \) and the color of the photon emitted, and extraction of transition-path times from photon sequences usually involves maximum likelihood analysis using a discrete model in which the continuous trajectory is replaced by dynamics with only three states, A, B, and a "transition region" intermediate TR.

---

**Figure 3.** Transition paths are segments of a molecular trajectory that traverse a specified transition region staying continuously in this region (blue). Experimentally, the transition region can be defined by requiring that the molecular reaction coordinate of interest \( x \) belong to a specified interval \((x_A, x_B)\). Such a coordinate, however, is often not directly observable. In two-color FRET experiments, for example, molecular trajectory \( x(t) \) is deduced from the arrival times of photons of two colors; in general, there is no one-to-one correspondence between the current value of \( x \) and the color of the photon emitted, and extraction of transition-path times from photon sequences usually involves maximum likelihood analysis using a discrete model in which the continuous trajectory is replaced by dynamics with only three states, A, B, and a "transition region" intermediate TR.
resulting three-state system that consists of states A, B, and TR is no longer Markovian, and the distributions of transition-path times reflect this fact.

The distributions of transition-path times from A to B and from B to A are generally given by

\[
P_{A \rightarrow B}(t) = \frac{2\alpha}{\alpha + 1}[k_1 G_{11}(t) + k_2 G_{12}(t)] + \frac{2\alpha}{\alpha + 1}[k_2 G_{21}(t) + k_1 G_{22}(t)]
\]

and

\[
P_{B \rightarrow A}(t) = \frac{2}{\alpha + 1}[k_1 G_{21}(t) + k_2 G_{22}(t)] + \frac{2\alpha}{\alpha + 1}[k_2 G_{12}(t) + k_1 G_{11}(t)]
\]

where \(k_1, k_2\) is the rate of escape from the corresponding intermediate to either side (Figure 4), and where

\[
\begin{bmatrix}
G_{11}(t) & G_{12}(t) \\
G_{21}(t) & G_{22}(t)
\end{bmatrix} = \exp\left(\begin{bmatrix} -2k_1 & \gamma \\ \gamma & -2k_2 - \gamma \end{bmatrix}t\right)
\]

is the “transition region Green’s function”. \(^{39}\) We will first consider the case in which there are no transitions within the transition region (\(\gamma = 0\) in Figure 4). For transition paths proceeding via intermediate 1 (2), the (conditional) distribution of the transition-path time is exponential

\[
p_{1(2)}(t) = 2k_{1(2)} G_{11}(t) = 2k_{1(2)} e^{-2k_{1(2)} t}
\]

For the transitions from A or B into the transition region, the clockwise transitions take place with a rate \(k\alpha\) and counterclockwise transitions with a rate \(k\), the parameter \(\alpha\) being a measure of the nonequilibrium driving force. As a result, for the \(A \rightarrow 1 \rightarrow B \rightarrow 2 \rightarrow A\) cycle, the product of the forward rate coefficients differs from that of the backward ones, with an overall driving force

\[
\epsilon^X = \alpha^2
\]

and with equilibrium attained only when \(\alpha = 1\).

The distributions of transition-path times from A to B and B to A are linear combinations of the contributions from each pathway. Using the fact that the probability to accomplish a transition from A to B via the pathway 1 is \(\alpha/(\alpha + 1)\), etc., the distributions of transition-path times in each direction are

\[
P_{A \rightarrow B}(t) = \frac{\alpha p_1(t) + p_2(t)}{\alpha + 1}
\]

and

\[
P_{B \rightarrow A}(t) = \frac{p_1(t) + \alpha p_2(t)}{\alpha + 1}
\]

a result that also follows directly from eqs 5 and 6. Several observations follow from eqs 7 and 9. First, unless one has \(k_1 = k_2\), the distributions of transition-path times are always broader than exponential. Specifically, the coefficients of variation \(C_{A \rightarrow B}\) and \(C_{B \rightarrow A}\) equal to the ratios of the distribution variances to their means, e.g.

\[
C_{A \rightarrow B} = \sqrt{\frac{\langle t^2_{A \rightarrow B} \rangle - \langle t_{A \rightarrow B} \rangle^2}{\langle t_{A \rightarrow B} \rangle^2}}, \quad \langle t_{A \rightarrow B} \rangle = \int_0^\infty dt p_{A \rightarrow B}(t)
\]

are greater than 1, a value expected for a single-exponential distribution. As was shown recently, \(^{38,41}\) such a broad distribution is impossible for any linear Markov kinetic scheme and, in particular, for the single-intermediate pathway model (Figure 3) commonly used to infer experimental transition-path times for the TR. \(^{50,55}\) Thus, the analysis of experimental transition-path times with a broad distribution in terms of a single-pathway mechanism would be internally inconsistent. On the contrary, we note that experimental observation of such a broad distribution would suggest that more than one pathway exists even if the two intermediate states, 1 and 2, have the same FRET signature and thus cannot be distinguished directly. Second, the forward/backward symmetry is broken \([p_{A \rightarrow B}(t) \neq p_{B \rightarrow A}(t)]\) when the system is out of equilibrium, e.g., when \(\alpha \neq 1\) (Figure 4). Again, such a situation cannot be captured if the experimental distributions of transition-path times \(p_{A \rightarrow B}(t)\) and \(p_{B \rightarrow A}(t)\) are interpreted in terms of a single-intermediate model or any linear Markov model such as the one shown on the right side of Figure 3. Third, notice that increased driving (larger \(\alpha\)) implies greater irreversibility (Figure 4). In particular, in the limit of strong driving, \(\alpha \gg 1\), the A-to-B transition preferentially proceeds via intermediate 1 and the B-to-A transition proceeds via intermediate 2, with the corresponding distributions approaching \(p_{A \rightarrow B}(t) \rightarrow p_1(t)\) and \(p_{B \rightarrow A}(t) \rightarrow p_2(t)\). As we will show below, however, a stronger driving does not necessarily imply greater asymmetry in transition-path times.

Consider now the possibility that the system can switch between pathways 1 and 2 while in transit between A and B. This is captured by a non-zero rate \(\gamma\) of switching between the two intermediates (Figure 4). Notice that even if \(\alpha = 1\), non-zero switching in general breaks the time reversibility, as the Kolmogorov criterion is violated for the A to B transition and B to A transition paths: it is evident from eq 5 that the forward/backward symmetry is preserved, \(p_{A \rightarrow B}(t) = p_{B \rightarrow A}(t)\), no matter how strong the driving is (i.e., regardless of switching rate \(\gamma\)), as the contributions of the two pathways to A-to-B and B-to-A transitions are the same and independent of \(\gamma\). However, the non-Markov character of the observed dynamics would still be discernable because the distributions \(p_{A \rightarrow B}(t)\) and \(p_{B \rightarrow A}(t)\), while identical in this case, would be nonexponential (cf. eq 9 with \(\alpha = 1\)), with a coefficient of variation exceeding 1, and thus incompatible with a Markov model with a linear topology. \(^{57}\)
Consider now the case in which \( \alpha \neq 1 \). In the limit \( \gamma \ll k, k_\alpha, k_1, k_2 \), the distributions of the transition-path times approach the results of eq 9. The opposite limit where switching is the fastest time scale of the system, \( \gamma \gg k, k_\alpha, k_1, k_2 \), is also easy to understand. In this case, many switches take place while the system remains in the transition region, and the transition region is effectively a single-intermediate TR with apparent transition rates to each side equal to the average rate \( (k_1 + k_2)/2 \). Thus, the two-pathway scheme can be replaced with a single-pathway one (Figure 3), with an exponential distribution of transition-path times given by

\[
p_{A \rightarrow B}(t) = p_{B \rightarrow A}(t) = (k_1 + k_2)e^{-(k_1+k_2)t}
\]  
(11)

Forward/backward symmetry is therefore recovered for the \( \gamma \to \infty \) step despite the nonequilibrium character of the process.\(^\text{57}\)

In this case, the entropy production is increasing with exchange rate \( \gamma \) (Figure 6), yet the observed distributions of the transition-path time evolve from two distinguishable distributions (eq 9) to two identical distributions (eq 11) (see Figure 5). Moreover, the observed distribution approaches a single-exponential one, consistent with the assumption that the transition region consists of a single intermediate state. The system becomes increasingly indistinguishable from an inherently time-reversible, Markovian three-state system with a linear topology as long as states 1 and 2 are indistinguishable experimentally. Note, however, that even for high values of \( \gamma \), short-time behavior of \( p_{A \rightarrow B}(t) \) and \( p_{B \rightarrow A}(t) \) is always different (Figure 5). Indeed, it follows from eqs 5 and 6 that \( p_{A \rightarrow B}(0) \) and \( p_{B \rightarrow A}(0) \) are independent of \( \gamma \). Therefore, given a sufficiently high (and increasing with \( \gamma \)) time resolution, it is still in principle possible to discern the time-irreversible behavior of the system.

**SUMMARY AND FUTURE DIRECTIONS**

To summarize, here we argue that it may be surprisingly difficult to determine, from a trajectory observed at a single-molecule level, whether it represents an active (e.g., driven by ATP hydrolysis) process or is a manifestation of equilibrium fluctuations of the observed molecule. The reason is ultimately the low-dimensional character of most single-molecule observ-
ables, which often forces an analysis of data in terms of simple models with a linear topology. When such models are assumed to obey Markovian dynamics, they, fundamentally, are equilibrium models that cannot capture active processes. Therefore, a successful fit of data to a linear hidden Markov model does not necessarily imply that the true dynamics is reversible.

A solution, on the experimental side, is a development of multidimensional techniques. On the data analysis side, irreversibility is often encoded in non-Markov effects. Such effects, fundamentally, cannot be neglected, and they can be detected with appropriate data analysis. If discovering and quantifying the directionality of the observed dynamics are the objectives, it may also be desirable to estimate entropy production directly from raw experimental time series (e.g., the photon sequence in Figure 3) rather than after postprocessing the data using, e.g., hidden Markov models. In this regard, histogram entropy estimators and compression algorithm-based estimators appear to be promising, although, to the best of our knowledge, they have not yet been applied to photon sequences. On the theory side, there have been recent developments addressing the question of how measures of irreversibility such as the entropy production can be deduced from partial observations such as some but not all transitions; application of those ideas to experimental trajectories is a promising new direction.

### AUTHOR INFORMATION

**Corresponding Authors**

Aljaž Godec — Mathematical bioPhysics Group, Max Planck Institute for Multidisciplinary Sciences, 37077 Göttingen, Germany; Email: agodec@impinat.mpg.de

Dmitrii E. Makarov — Department of Chemistry and Oden Institute for Computational Engineering and Sciences, The University of Texas at Austin, Austin, Texas 78712, United States; Email: makarov@cm.utexas.edu

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jpcl.2c03244

**Notes**

The authors declare no competing financial interest.

**Biographies**

Aljaž Godec obtained his Ph.D. in theoretical physics in 2012 from the University of Ljubljana (Ljubljana, Slovenia) and afterward moved to a postdoc at the University of Potsdam (Potsdam, Germany) as an Alexander von Humboldt fellow. Since 2017, he has been head of the research group “Mathematical bioPhysics” at the Max-Planck-Institute for Multidisciplinary Sciences in Göttingen, Germany. His research focuses on the statistical and mathematical physics of out-of-equilibrium systems, in particular on stochastic thermodynamics and emergent phenomena in soft matter and biophysics.

Dmitrii E. Makarov received a Ph.D. in theoretical physics from the Institute for Chemical Physics (Moscow, Russia) in 1992. Since 2001, he has been on the faculty at the Department of Chemistry at The University of Texas at Austin. He is also a core faculty at the Oden Institute for Computational Engineering & Sciences. His research interests are in the field of computational and theoretical chemical physics, with current research topics ranging from quantum reaction rate theory to molecular biophysics.

### ACKNOWLEDGMENTS

A.G. was supported by German Research Foundation (DFG) through the Emmy Noether Program (GO 2762/1-2). D.E.M. was supported by the Robert A. Welch Foundation (Grant F-1514), the National Science Foundation (Grant CHE 1955552), and an Erwin Neher fellowship of the Max Planck Institute for Multidisciplinary Sciences. D.E.M. also thanks the Max Planck Institute for Multidisciplinary Sciences and particularly his co-author for the hospitality during his stay there. Discussions with Gilad Haran, Hagen Hofmann, Anatoly Kolomeisky, and Benjamin Schuler are greatly appreciated.

### REFERENCES

(59) Kim, J. Y.; Chung, H. S. Disordered proteins follow diverse transition paths as they fold and bind to a partner. Science 2020, 368 (6496), 1253–1257.