PNAS RESEARCH ARTICLE | BIOPHYSICS AND COMPUTATIONAL BIOLOGY

Optimal inference of molecular interaction dynamics in FRET microscopy

Keita Kamino[a,b,c,d,1], Nirag Kadakia[a,b,e], Fotios Avgidis[f], Zhe-Xuan Liu[g], Kazuhiro Aoki[h,i,j], Thomas S. Shimizu[f], and Thierry Emonet[a,b,k,1]

Edited by Kimani C. Toussaint, Brown University, received July 9, 2022; accepted February 10, 2023 by Editorial Board Member John A. Rogers

Intensity-based time-lapse fluorescence resonance energy transfer (FRET) microscopy has been a major tool for investigating cellular processes, converting otherwise unobservable molecular interactions into fluorescence time series. However, inferring the molecular interaction dynamics from the observables remains a challenging inverse problem, particularly when measurement noise and photobleaching are nonnegligible—a common situation in single-cell analysis. The conventional approach is to process the time-series data algebraically, but such methods inevitably accumulate the measurement noise and reduce the signal-to-noise ratio (SNR), limiting the scope of FRET microscopy. Here, we introduce an alternative probabilistic approach, B-FRET, generally applicable to standard 3-cube FRET-imaging data. Based on Bayesian filtering theory, B-FRET implements a statistically optimal way to infer molecular interactions and thus drastically improves the SNR. We validate B-FRET using simulated data and then apply it to real data, including the notoriously noisy in vivo FRET time series from individual bacterial cells to reveal signaling dynamics otherwise hidden in the noise.

FRET | statistical inference | live-cell imaging | cell signaling | information theory

Significance

Signal-to-noise ratio (SNR) limits what we can learn from data. In fluorescence microscopy, SNR is set by the number of photons acquired from a sample and the efficiency with which these photon are used in data analysis. Experimental configurations that determine the former tend to be highly optimized, whereas analysis methods to maximize the latter remain comparatively underexplored. This is the case for intensity-based, time-lapse fluorescence resonance energy transfer (FRET) microscopy, a powerful method for quantifying the dynamics of molecular interactions inside cells. Here, we develop an information-theoretically optimal method to estimate molecular interaction from such FRET data that maximizes SNR. Like bright fluorescent proteins and sensitive photodetectors, the method expands the scope of FRET microscopy by significantly improving SNR.

Molecular interactions inside cells underlie essentially all biological functions (1–3). Leveraging the principle of FRET (4, 5), intensity-based time-lapse FRET microscopy based on the so-called 3-cube imaging method (6–11) has been widely used to investigate the dynamics of molecular interactions inside cells. With growing interest in quantifying cell-to-cell (or compartment-to-compartment) variations and temporal fluctuations in cellular processes (12–19), time-lapse FRET measurements are increasingly being conducted on single cells or subcellular compartments (20–37). In so-called ensemble FRET (as opposed to single-molecule FRET) measurements, which we are interested in here, fluorescence is integrated from a number of fluorescence proteins (FPs) in a compartment; however, cells and subcellular compartments often contain small numbers of FPs and thus exhibit relatively large photon shot noise, which limits the signal-to-noise ratio (SNR). Attempts to reduce the shot noise by increasing fluorescence-excitation power are confronted with higher phototoxicity and photobleaching rates, both of which limit the achievable measurement durations in live cells (Fig. 1). This trade-off between SNR and measurement duration in turn places limits on what we can learn from FRET data, as demonstrated by recent FRET-based investigations of signaling dynamics in single bacterial cells (24–27).

From an information-theoretic perspective, the trade-off between SNR and measurement duration can only be mitigated by improving the efficiency of information usage (Fig. 1). In a typical FRET experiment designed to study the temporal dynamics of molecular interactions, the focus of this paper, image analysis is applied to multicolor time-lapse movies to extract a set of fluorescence-intensity time series from a compartment (e.g., a single cell). Via FRET, these time-series data *encode* information about specific molecular interactions within the compartment. However, because these time series also reflect processes other than FRET, such as photobleaching, careful data analysis is needed to *decode* the relevant information about the specific molecular interactions of interest from the time series. To date, researchers have primarily focused on increasing the efficiency of the encoding process that generates the time series by, e.g., optimizing FRET pairs (28, 38–40) and microscope setups (41, 42). However, equally important—and yet vastly underexplored—is the efficiency of the decoding process: Data processing inevitably incurs information loss (43), and data-analysis methods that do not explicitly consider informational aspects tend to incur unnecessarily large losses, effectively discarding valuable information contained in the original time-series data (44).

Numerous methods to analyze intensity-based time-series FRET data have been proposed (6–8, 45–50). Qualitative methods, such as simple ratiometry (45), neglect spectral crosstalk and photobleaching. Consequently, the results depend on the specific instruments and are

Author contributions: K.K. designed research; K.K., N.K., and K.A. performed research; K.K., F.A., Z.-X.L., T.S.S., and T.E. analyzed data; and K.K., N.K., F.A., and T.E. wrote the paper.

The authors declare no competing interest.

This article is a PNAS Direct Submission. K.C.T. is a guest editor invited by the Editorial Board.

[1]To whom correspondence may be addressed. Email: kkamino@gate.sinica.edu.tw or thierry.emonet@yale.edu.

This article contains supporting information online at https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2211807120/-/DCSupplemental.

Published April 4, 2023.

biased by photobleaching (7, 51), making them unsuitable for quantitative studies. Quantitative methods such as E-FRET (7) and lux-FRET (49) provide algebraic formulae that correct for both spectral crosstalk and photobleaching and infer instrument-independent quantities that can be interpreted from molecular terms. However, as previously pointed out (51), such algebraic calculations inevitably build up noise in data and thus lower the SNR. To see this, consider a case where a fluorescence intensity I_1 is subtracted from another I_2, as often done to correct for spectral cross talk (7–9, 46, 49): The resulting intensity is smaller than I_2 (i.e., $I_2 - I_1 < I_2$), but its noise measured by variance is larger than that of I_2 (i.e., $\mathrm{Var}(I_2 - I_1) = \mathrm{Var}(I_2) + \mathrm{Var}(I_1) > \mathrm{Var}(I_2)$). This implies that such algebraic methods implement suboptimal decoding methods. Previous works (52, 53) have explored and implemented non-algebraic, principled methods for *single-time-point* FRET data, resulting in better performance than algebraic methods. However, no such method exists for the analysis of *time-series* FRET data. The decoding of FRET dynamics from time-series data is complicated by other time-varying factors, such as photobleaching kinetics and measurement noise. Some of the confounding factors can be avoided or mitigated with more specialized lifetime-based FRET measurements (54), but the presence of such confounding factors in intensity-based FRET measurements does not mean the measurement method cannot encode sample information efficiently. Rather, it could be said that intensity-based FRET data encode information in a more intricate manner that calls for more careful decoding.

Here, we develop a general computational framework, B-FRET, that implements a principled and statistically optimal way to extract information about molecular-interaction dynamics from fluorescence-intensity time-series data. Using filtering theory (55, 56), B-FRET directly learns from data the photophysical model (46, 49, 52, 57) that maps the concentrations of FPs to fluorescence intensities. This allows B-FRET to take into account all the confounding factors, such as spectral crosstalk, photobleaching, and measurement noise, and to systematically propagate these into the estimation of molecular interaction dynamics. Consequently, B-FRET not only drastically increases the SNR but also enables the computation of the estimates' statistical uncertainties—an aspect absent in previous algebraic methods (58). We use B-FRET to analyze noisy time-lapse FRET data both from computational models and from single live cells and show that it estimates molecular interaction dynamics at unprecedented precision.

Results

The B-FRET Framework and Learning Algorithm. A FRET sample contains FPs whose states (fluorescent or photobleached and free or complexed) change in time. An intensity-based time-lapse FRET measurement is a (noisy) map from a configuration of FPs in various states to observable fluorescence-intensity time series (Fig. 2*A*). The goal of quantitative FRET-data analysis is to infer, from the observables, the dynamics of the degree of donor–acceptor interaction [i.e., protein–protein interaction for bimolecular FRET or protein conformational changes for unimolecular FRET (54)]. Ideally, the estimate should be interpretable in molecular terms and independent of instrumentation and photobleaching. The degree of interaction can be defined in various ways depending on the purpose of an experiment (7, 46), and we call such a user-defined degree of interaction a FRET index E. *B-FRET is a computational framework to infer the FRET index E in a statistically optimal manner*. For concreteness, we consider the general case of bimolecular FRET, in which the donor and acceptor are on different carrier molecules; hence the stoichiometry of the acceptor to the donor is unknown. However, the B-FRET algorithm can also be used for unimolecular FRET (*Discussion* and *SI Appendix*, SI Texts 1 and 2). To demonstrate its performance, we compare the results of B-FRET with those of E-FRET (7), arguably the most widely used (25, 59, 60) quantitative method that corrects for both spectral crosstalk and photobleaching (*Materials and Methods* and *SI Appendix*, SI Text 1).

In a bimolecular FRET system, the donor (D) and acceptor (A), fused to two different molecules, can form a molecular complex leading to FRET from the donor to the acceptor. The system is characterized by the concentrations of eight chemical species: D^*, D, A^*, A, D^*A^*, D^*A, DA^*, and DA, where fluorescent and nonfluorescent molecules are indicated by the presence and absence of a star ($*$) respectively, and donors and acceptors can be free (e.g., D^*) or complexed (e.g., D^*A^*) (Fig. 2*A*). Inferring the temporal evolution of all eight chemical concentrations from the three

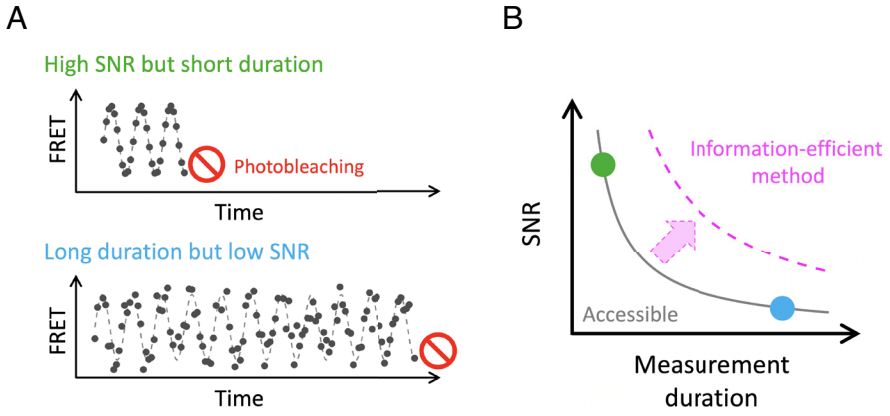

Fig. 1. Trade-off between signal-to-noise ratio (SNR) and total measurement duration in time-lapse FRET measurements. (*A*) Schematics showing FRET-signal time series obtained in two measurement conditions. True FRET values (dashed lines) are not directly accessible but are estimated through measurements (gray dots). High SNR can be achieved by increasing the power of fluorescence excitation, but it also increases the rate of photobleaching, which shortens the measurement duration (*Top*). Long measurement duration can be achieved by reducing the power of fluorescence excitation, but it lowers SNR (*Bottom*). (*B*) The fundamental trade-off between SNR and measurement duration constrains the accessible data quality. This trade-off can only be mitigated (magenta) either by increasing the number of photons one gets from a sample (encoding efficiency) or by improving the efficiency with which the acquired photons are used to estimate FRET (decoding efficiency). The former has been thoroughly explored previously by, e.g., optimizing FRET pairs. Here, we address the latter by developing an algorithm to optimally extract the information of FRET from a set of time-lapse FRET data.

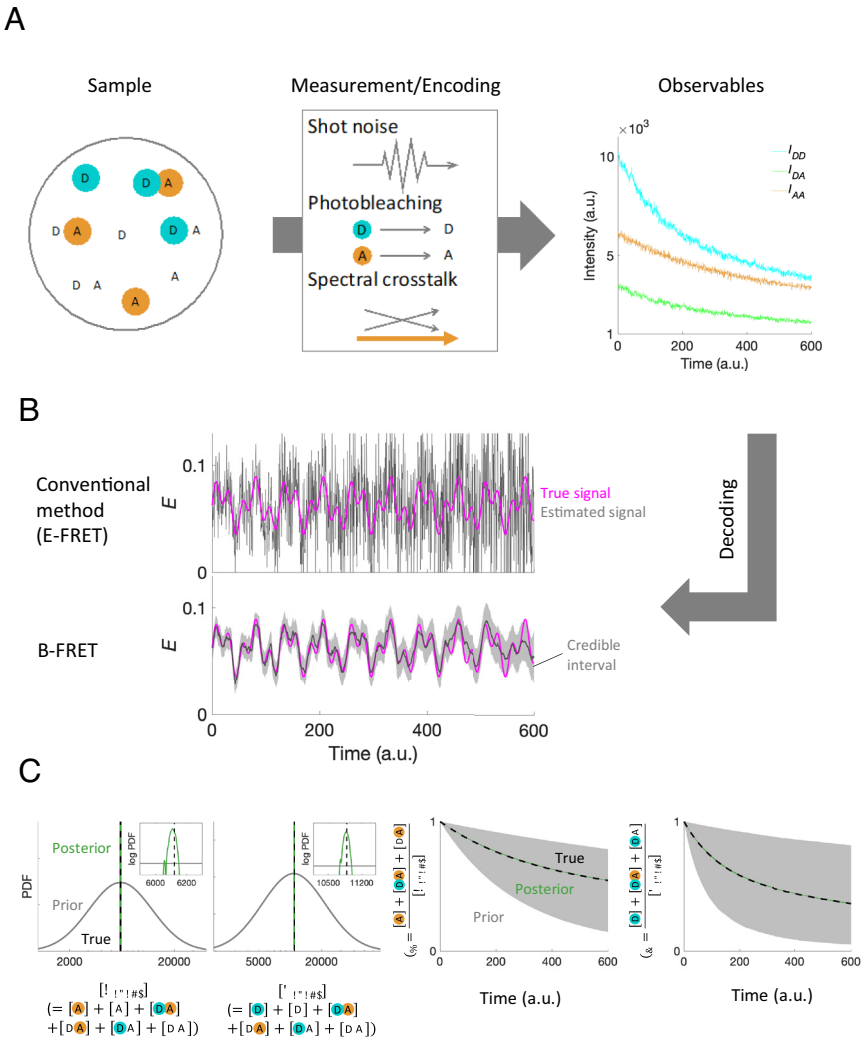

Fig. 2. Optimal decoding of the degree of molecular interaction by B-FRET. (*A*) A schematic of 3-cube time-lapse FRET imaging. A FRET sample (*Left*) consists of 8 chemical species (the schematic implies bimolecular FRET, where the donor and acceptor can be isolated from each other. See *SI Appendix*, SI Text 1 for unimolecular FRET): fluorescent donor (D; cyan) and acceptor (A; orange), photobleached D and A (gray) and 4 different D–A complexes. A time-lapse FRET measurement encodes the molecular information into the three fluorescence time series (*Right*), I_{AA} (acceptor excitations and acceptor emissions), I_{DD} (donor excitations and donor emissions), and I_{DA} (donor excitations and acceptor emissions). The encoding subjects to measurement noise, photobleaching, and spectral crosstalk (*Middle*). A photophysical model relates the sample state to the observables, and the model is learned from the observables via B-FRET. (*B*) A FRET index E, which quantifies the degree of molecular interaction, is decoded from the synthetic data in *A* by E-FRET (*Top*) and B-FRET (*Bottom*). True (magenta) and estimated E (gray line) are shown. For B-FRET, 95% credible intervals are shown by gray shade. (*C*) Prior and posterior distributions of the unknown parameters of the photophysical model. From *Left* to *Right*: the total concentrations of the acceptor ($[A_{total}]$), the donor ($[D_{total}]$), the fraction of intact acceptors (f_A), and donors (f_D). Insets in the left two panels are magnified posterior distributions (green) plotted in log scale for both *X* and *Y* axes. In the right two panels, 2.5 and 97.5 percentiles of f_A and f_D sampled from prior or posteriors are shown, respectively. Posterior distributions are generally well confined, implying the high efficiency with which B-FERT extracts the information from the data.

available time series (see below) is ill-conditioned and intractable. However, we do not necessarily need a general solution to this problem; instead, we only need a solution that would apply to the types of phenomena we are interested in, such as cell signaling, which has a characteristic timescale. All existing quantitative FRET analysis methods exploit the fact that the biological timescale of interest studied using FRET measurements is much longer than that of individual protein binding/unbinding events but shorter than the timescale over which the copy number of constituent FPs changes significantly (*SI Appendix*, SI Text 1). Accordingly, in what follows, we focus on the same timescale and seek an optimal solution within that regime. In previous FRET analysis methods, some of these assumptions were implicit or unclear. One of the advantages of the Bayesian framework described below is that it forces us to make all assumptions mathematically explicit, as described in *Materials and Methods* and *SI Appendix*, SI Text 1, thus improving clarity.

The B-FRET algorithm is independent of how a user defines a FRET index E (*Discussion*). For concreteness, here we consider a standard measure of the degree of interaction (7, 24, 25, 27, 46, 61, 62):

$$E(t) = \frac{E_{max}[DA_{total}](t)}{[D_{total}]}, \qquad [1]$$

where $[D_{total}] \equiv [D^*](t) + [D](t) + [D^*A^*](t) + [D^*A](t) + [DA^*](t)[DA^*](t) + [DA](t)$ is the total concentration of the donor molecule, and $[DA_{total}](t) \equiv [D^*A^*](t) + [D^*A](t) + [DA^*](t) + [DA](t)$ is the total concentration of the donor-acceptor complex; E_{max} is the specific FRET efficiency of the complex, defined as the probability of energy transfer from the donor to the acceptor in the donor–acceptor complex per donor excitation event and is a constant given an experimental condition (38, 54). The FRET index E defined by Eq. **1** is independent of instrument-specific

parameters and the degree of the photobleaching of the fluorescent molecules and is linearly dependent on the fractional occupancy of the donor, making it an ideal measure of the degree of molecular interaction (7, 46).

At its core, B-FRET is an application of filtering theory or Bayesian inference for so-called *state space models* (55, 56). In this framework, one infers the temporal evolution of hidden (i.e., unobservable) dynamical state variables from noisy observations. A state-space model consists of a *dynamic model*, which describes the temporal evolution of hidden state variables, and a *measurement model*, which is a static function mapping the hidden variables at time t to observables at time t. We discuss these in turn.

In B-FRET, the hidden dynamic variable is the product of the specific FRET efficiency and the total concentration of the complex, i.e., $\chi(t) = E_{max}\,[DA_{total}](t)$ (*Materials and Methods*). The dynamic model links $\chi(t)$ at two consecutive times via a probability distribution called the *process noise* described by a parametric probability distribution such as a Gaussian distribution (*Materials and Methods* and *SI Appendix*, SI Text 2). The assumption of a dynamical model is a central feature of B-FRET: It allows us to exploit temporal correlations in the hidden variable over small times (55, 56), which always exist, and yet algebraic methods fail to utilize. On the other hand, process noise introduces additional parameters (e.g., the standard deviation of a Gaussian distribution). These are estimated as part of the B-FRET algorithm, as described below and in more detail in *Materials and Methods* and *SI Appendix*, SI Text 2.

In addition to the dynamic model, B-FRET requires a measurement model, which describes the photophysical processes by which the hidden variable $\chi(t)$ is converted into observables (46, 49, 52, 57). In the standard 3-cube FRET imaging setup, the observables consist of three fluorescence-intensity time series I_{AA}, I_{DD}, and I_{DA} (Fig. 2*A*). These are, respectively, fluorescence measured at the acceptor emission band during excitation of the acceptor band; fluorescence measured at the donor emission band during excitation of the donor band; and fluorescence measured at the acceptor emission band during excitation of the donor band. Other than excitations of and emissions from FPs, the photophysical processes involved include photobleaching of FPs, spectral crosstalk (i.e., bleed-through of donor emission to the acceptor emission band, and cross-excitation of the acceptor at the donor excitation band), energy transfer from the donor to acceptor due to FRET, and measurement noise (Fig. 2*A*). These are incorporated into a single probabilistic measurement model (*Materials and Methods* and *SI Appendix*, SI Text 1). The measurement model is another central feature of B-FRET: This not just allows us to exploit our knowledge about the photophysical processes but also makes all the assumptions involved in the decoding process mathematically explicit. This model also has unknown parameters, one of which is $[D_{total}]$ (Eq. **1**). We estimate them as described below.

Given these two ingredients—the dynamical model and measurement model—our goal is to estimate the FRET index $E(t) = \chi(t)\,/\,[D_{total}]$ (Eq. **1**) from observables. In the framework of Bayesian inference (44, 63), this amounts to computing the posterior distribution of $E(t)$, $p(E(t)|\boldsymbol{\mathcal{D}}, \boldsymbol{\mathcal{M}})$, which quantitatively describes how well the possible values of $E(t)$ are confined given all data $\boldsymbol{\mathcal{D}}$ and model $\boldsymbol{\mathcal{M}}$. Since this distribution contains all the information one can theoretically have, computing the distribution ensures the statistically optimal inference of the FRET index. Because model parameters are also unknown, they must also be inferred from data. Thus, the computation of the posterior distribution of the FRET index is decomposed into the evaluations of two distributions: the posterior distribution of $E(t)$ *given* specific model parameter values and the posterior distribution of the model parameters themselves (*Materials and Methods*). These two distributions can be evaluated using filtering algorithms (*SI Appendix*, SI Text 2). Once these distributions are determined, the posterior distribution $p(E(t)|\boldsymbol{\mathcal{D}}, \boldsymbol{\mathcal{M}})$ is computed using a Monte Carlo approach (*Materials and Methods*).

B-FRET Efficiently Learns from Data. To see how much the B-FRET algorithm improves the SNR of the estimated FRET index, we compared the FRET index E computed by B-FRET with that computed by E-FRET. We first generated a synthetic (bimolecular) FRET dataset by simulating oscillatory dynamics of FRET signals and all the confounding factors present in real data, namely spectral crosstalk, photobleaching, and measurement noise (Fig. 2*A* and *SI Appendix*, SI Text 4). With relatively large measurement noise, the oscillatory FRET dynamics are hard to see in the raw time-series data (Fig. 2*A*). Consequently, the FRET index computed by E-FRET is highly noisy, and the true oscillatory dynamics are obscured (Fig. 2*B*). However, we found that the FRET index computed by B-FRET estimates the true signal substantially more precisely, as evidenced by the comparison of estimation errors (Fig. 3*B*). Furthermore, unlike E-FRET, B-FRET naturally provides statistical uncertainty of the estimated FRET index as a credible interval (CI) at each time point. As expected, the width of 95% CIs (*Materials and Methods* for definition) increases over time because of the decreasing data quality resulting from photobleaching (gray shadow in Fig. 2*B*). Consistent with the precise estimation of E, the posterior distributions of the model parameters are highly confined around their true values used to generate the synthetic data (Fig. 2*C*). These observations demonstrate that the raw fluorescence time-series data, despite high levels of noise, contain rich information about molecular interactions, and B-FRET successfully exploits the available information to better constrain the possible values of the FRET index and model parameters.

B-FRET Is Robust to the Variation in FRET Temporal Patterns. To see how much the precision of FRET-index estimation is affected by the underlying temporal pattern of molecular interactions, we next generated synthetic data in which the FRET signal exhibits random dynamics (*SI Appendix*, SI Text 4). Unlike the case of oscillatory dynamics (Fig. 2), the random signal is aperiodic and contains a broad range of frequencies, including those comparable to or higher than the data-sampling frequency, which precludes algorithms that exploit regular patterns in a signal. Despite this, we found that the FRET index computed by B-FRET is more precise and less noisy than that computed by E-FRET (Fig. 3*A* and *SI Appendix*, Fig. S1*A*).

The above two cases, oscillatory and random, were successfully analyzed with Gaussian process noise, a standard choice for the process noise for its flexibility in capturing a broad class of dynamics (55, 56). However, for highly non-Gaussian dynamics, e.g., ones that remain unchanged most of the time but exhibit abrupt step changes only occasionally, it is known that non-Gaussian process noise can perform better than Gaussian process noise (56). Although B-FRET is computationally cheaper with Gaussian process noise since many calculations can be executed analytically, the algorithm can be adapted to other process noise statistics by replacing the analytical calculations with numerical ones (*SI Appendix*, SI Text 2). To test the performance of B-FRET for non-Gaussian dynamics, we generated a synthetic FRET signal consisting of discrete steps (*SI Appendix*, SI Text 4) and modeled the process noise using a Student's t distribution (*Materials and Methods*). Indeed, the FRET index computed by B-FRET precisely captures the dynamics (Fig. 3*A* and *SI Appendix*, Fig. S1*B*).

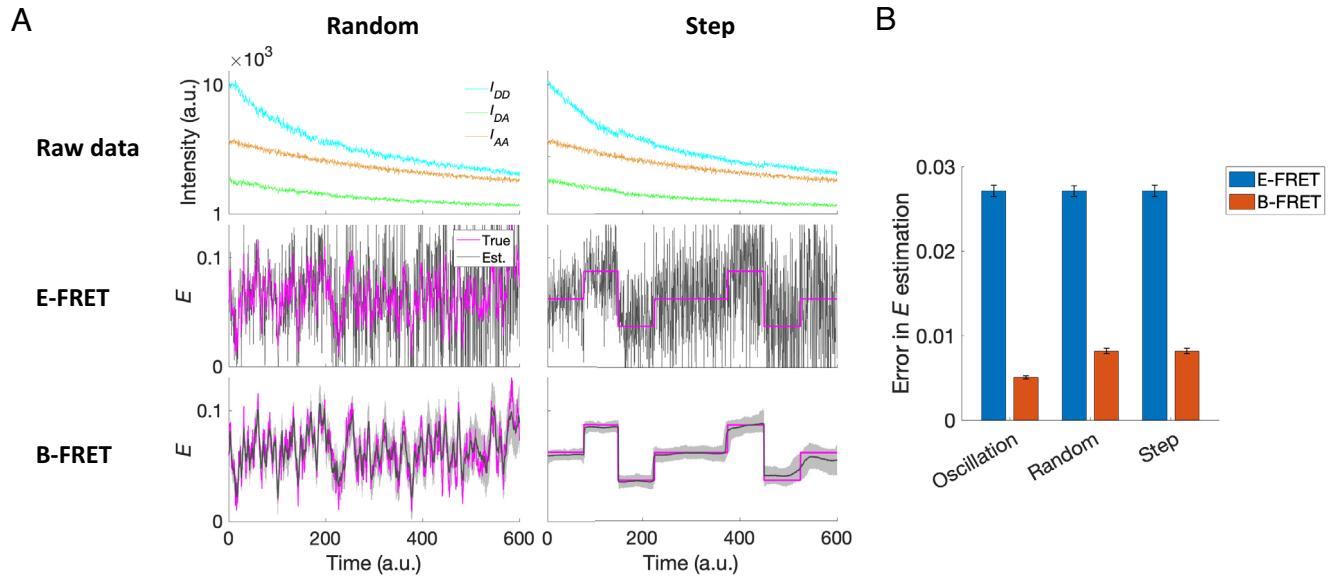

Fig. 3. B-FRET exhibits high signal-estimation performance irrespective of underlying molecular-interaction dynamics. (*A*) Synthetic data with random (*Left*) and step (*Right*) FRET signals. (*Top*) Fluorescent-intensity time series. (*Middle*) The FRET index E computed by E-FRET (gray; E_{est}) and its true values (magenta; E_{true}). (*Bottom*) The FRET index E computed by B-FRET (gray; E_{est}) and its true values (magenta). The shade shows ~95% credible intervals. (*B*) A bar chart quantifying the error in E estimation defined as $\langle |E_{est} - E_{true}| \rangle$, where the angle bracket is temporal average. The error bars are SDs over 5 datasets with identical FRET signal dynamics but different realizations of measurement noise.

We note that B-FRET, combined with the framework of model selection, does not require a user to know in advance which model (e.g., Gaussian or non-Gaussian process noise) to use to analyze a set of data. By computing the Bayes Information Criterion (BIC; *Materials and Methods*), B-FRET enables users to automatically select a model that is best evidenced by the data. Applying this, we confirmed that the step data support the choice of non-Gaussian process noise, while the oscillatory and random data do not (*SI Appendix*, Fig. *S2*).

B-FRET Outperforms Conventional Methods Irrespective of the Measurement Conditions. To see how the relative performance of B-FRET to E-FRET depends on the specific conditions of time-lapse imaging, such as the levels of measurement noise and sampling intervals, we investigated the signal-estimation errors of both methods for various measurement conditions. We first generated sets of synthetic data in which the degree of donor–acceptor interaction follows Gaussian random statistics over time with a correlation time τ_c, and fluorescence signals I_{AA}, I_{DD}, and I_{DA} were measured with many different relative sampling frequencies $\tau_c/\Delta t$ and levels of measurement noise (*SI Appendix*, SI Text 4). For each dataset, we then estimated the FRET index E using both B-FRET and E-FRET methods.

Fig. 4*A* shows representative results for B-FRET. As the SNR of raw fluorescent time-series data increases, B-FRET detects more subtle changes in E with lower statistical uncertainties (95% CIs are shown by gray shades in Fig. 4*A*). Also, the higher the sampling frequency relative to the FRET dynamics, the more precise the B-FRET estimation. This is because faster sampling increases the correlation between successive fluorescence signals, and these correlations are exploited by B-FRET. Meanwhile, the effect of reduced sampling frequency (Fig. 4 *A*, *Lower Left*) can be compensated by increasing the data SNR (Fig. 4 *A*, *Upper Left*).

We quantified the average error in FRET signal estimation as the root-mean-square error normalized by the magnitude of the fluctuation of the true signal, $\sqrt{\langle (E_{est} - E_{true})^2 \rangle} / \mathrm{Std}(E_{true})$, for both E-FRET and B-FRET in various measurement conditions (Fig. 4 *B* and *C*). B-FRET (red) outperforms E-FRET (blue) in all conditions explored; importantly, even if E-FRET signals are smoothed with median filters with an optimal window size in terms of error reduction (gray)—which requires knowledge of the true FRET signal that an experimenter does not usually have access to—E-FRET still significantly underperforms compared with B-FRET. This can be understood by noting that, in the E-FRET method, some information about the true FRET signals contained in the raw fluorescent time series is already lost upon the algebraic computation to obtain E, and no smoothing after that computation can recover the lost information. As a result, even in the limit of low data-sampling frequency (Fig. 4*B*; $\tau_c/\Delta t < 0.1$), where there is almost no correlation to exploit between two consecutive time points, B-FRET outperforms E-FRET. Thus, irrespective of specific measurement conditions, B-FRET utilizes a larger amount of information in the raw observables and achieves a more precise estimation of E without requiring knowledge about molecular-interaction dynamics.

B-FRET Improves Signal Estimation of Real Data. To test the performance of B-FRET on real data, we first applied the method to a previously developed bimolecular FRET system that reports the kinase activity of the *Escherichia coli* chemotaxis signaling pathway (24, 61, 62). Recent FRET analyses of this pathway at the single-cell level have revealed fundamental features of cell signaling that are inaccessible by a population-level assay, such as spontaneous fluctuation in the pathway activity (24), environment-dependent dynamic modulation of the degree of cell-to-cell variability (27), and the high efficiency with which cells use information acquired by the pathway (25). However, the FRET data from single *E. coli* cells are noisy, severely suffering from the trade-off discussed in Fig. 1. This has limited further characterizations of the signaling pathway.

The *E. coli* chemotaxis signaling pathway is a two-component signal transduction system (64), where the receptor-associated kinase CheA phosphorylates the response regulator CheY, which is then dephosphorylated by the phosphatase CheZ. The binding

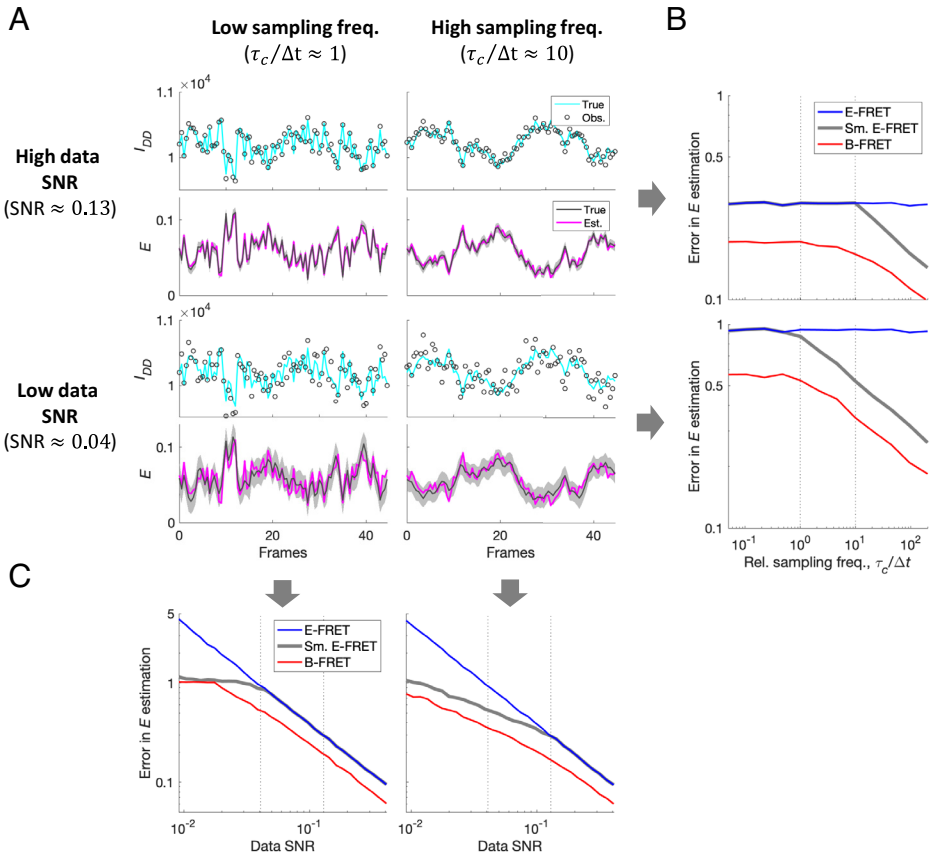

Fig. 4. B-FRET exhibits high signal-estimation performance irrespective of measurement conditions. (*A*) Representative simulated data and estimated values of a FRET index *E* in four different measurement conditions. I_{DD} with (dark gray) and without (cyan) measurement noise and estimated (dark gray) and true (magenta) *E* are shown. In the simulation, the correlation time of the molecular-interaction dynamics τ_c, sampling interval Δt, and the level of measurement noise (SNR) were varied systematically (*SI Appendix*, SI Text 4). (*B*) Errors in *E* estimation defined as $\sqrt{\langle (E_{est} - E_{true})^2 \rangle} / \mathrm{Std}(E_{true})$ for B-FRET (red), E-FRET (blue), and E-FRET combined with optimal median filtering (gray) were plotted against relative sampling frequency, $\tau_c/\Delta t$. Cases of high data SNR (=0.13; *Top*) and low data SNR (=0.04; *Bottom*) are shown. To obtain the gray curves, *E* computed by E-FRET was median filtered with a window size that minimizes the error. Such optimal smoothing requires knowledge about true signals, which an experimenter cannot access in a real experiment, and therefore cannot be implemented in practice. B-FRET outperforms this optimally-filtered E-FRET in all conditions. (*C*) Error in *E* estimation plotted against data SNR. Cases of under-sampling ($\tau_c/\Delta t \approx 1$) and over-sampling ($\tau_c/\Delta t \approx 10$) are shown.

of chemoattractant molecules to the receptors changes the propensity for the receptor, and hence the kinase, to be active. Opposing this propensity is the feedback regulation by methylation and demethylation enzymes. These two mechanisms together produce a steady-state kinase activity that is independent of the background chemoattractant concentration, a ubiquitous phenomenon in cell signaling called response adaptation (20, 65–67). The pathway's activity can be read out by quantifying the FRET between the donor (mYFP; monomeric yellow fluorescent protein) fused to CheZ, and the acceptor (mRFP; monomeric red fluorescent protein) fused to CheY, which binds to CheZ when phosphorylated by CheA. It has been well established that, upon a step increase in a chemoattractant concentration, the kinase activity and the concentration of phosphorylated CheY (and hence the level of FRET) decrease rapidly before response adaptation. In contrast, a step decrease in a chemoattractant concentration causes the opposite response (24, 61).

We measured fluorescence time series, I_{AA}, I_{DA}, and I_{DD} from single *E. coli* cells using a 3-cube FRET measurement setup (Fig. 5*A* and *Materials and Methods*). In this setup, we delivered fast-switching (~0.1 s) step-like changes of α-methyl-aspartate (MeAsp), a nonmetabolizable analog of the chemoattractant aspartate, using a recently developed microfluidic system (25, 27). Large step changes in MeAsp (100% changes or higher) were delivered to cells at the beginning and end of the measurement to define the dynamic range (i.e., minimal and maximal FRET levels) of each cell. Several small step changes in MeAsp (20% changes) that cause subsaturating responses, on average, were also applied in the middle of the measurement (Fig. 5*B*). First, we extracted the FRET index *E* using the E-FRET method (Fig. 5 *B*, *Left*). As expected, the large noise prevented us from discerning single responses to the subsaturating (20%) step stimuli. Quantifying responses from such noisy data requires some form of data averaging, as was done before (25, 27); however, it unavoidably masks the properties of individual responses. Next, we analyzed the same dataset using B-FRET (Fig. 5 *B*, *Right*). B-FRET drastically improved the SNR and disclosed the cell-to-cell and temporal variations in the signaling dynamics more vividly: Some cells respond to small step signals faithfully, whereas others neglect the same signals; some cells fluctuate vigorously, whereas others are more stable. Such variations could be functionally important for a cell population to deal with environmental uncertainties, as recent studies have suggested (13, 14, 19, 68, 69). Furthermore, B-FRET does not just make some subsaturating responses clearly discernible by eye; it also enables us to determine whether the changes in FRET are statistically significant or not (red boxes in Fig. 5*B*). Finally, as with synthetic data (Fig. 2*C*), the posterior distributions of the model parameters are highly confined (Fig. 5*C*), demonstrating that real experimental data also contain sufficient information to confine the photophysical model.

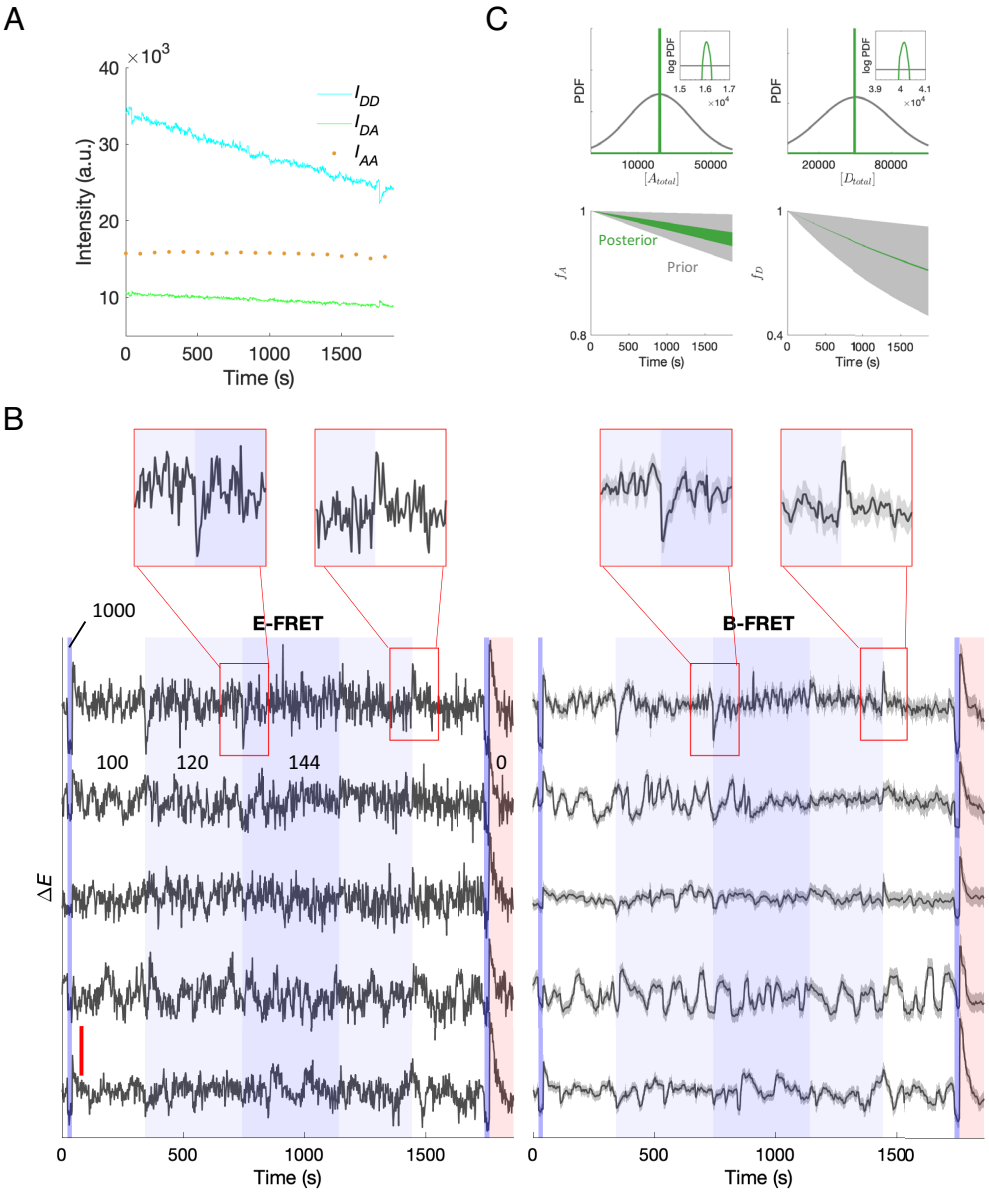

Fig. 5. B-FRET shows high signal-estimation performance with real bimolecular FRET data. (*A*) Three observables, I_{AA}, I_{DD}, and I_{DA} were acquired from a bimolecular FRET system in single *E. coli* cells. In this system, the FRET level changes depending on the activity of a kinase that governs the chemotaxis behavior of *E. coli*. (*B*) Using FRET datasets from individual *E. coli* cells, FRET index *E* was estimated by E-FRET (*Left*) and B-FRET (*Right*). Five representative cells are shown. Blues and red in the background indicate different concentrations (shown in the unit of μM) of a chemoattractant MeAsp delivered to the cells using a microfluidic device. The red vertical line in the *Left* panel corresponds to the change in the FRET index $\Delta E = 0.05$. For B-FRET, 95% credible intervals are shown by the gray shade. The noise reduction by B-FRET reveals temporal and cell-to-cell variation in the FRET dynamics more clearly. The data in *A* corresponds to the cell shown at the bottom. (*C*) Prior and posterior distributions of the model parameters for the cell shown at the bottom, plotted in the same way as Fig. 2*C*.

Next, we tested the performance of B-FRET on real unimolecular FRET data. We used HeLa cells stably expressing cAMP FRET biosensor and dopamine receptor D1, and we analyzed cAMP responses to 0.1 μM dopamine (Fig. 6*A* and *Materials and Methods* and *SI Appendix*, SI Text 6). In the same way as the bimolecular FRET data, B-FRET significantly improved the SNR compared with E-FRET (Fig. 6*B*), and the posterior distributions of the model parameters were greatly confined by the data (Fig. 6*C*). Together, these results demonstrate that B-FRET can greatly improve the quality of extracted FRET signals and therefore help experimenters reveal dynamic features of cellular processes.

Discussion

Inefficient decoding of molecular interactions from FRET data amounts to wasting hard-won information from detected photons. Here, we propose a computational framework, B-FRET, to decode the information with theoretically maximal efficiency. A conventional way to improve the SNR in live FRET imaging has been to aggregate signals from many samples (e.g., cells) and compute their average (61, 62); however, this method fails to capture variations and asynchronous dynamics across samples. B-FRET reduces the need for such averaging, as we demonstrated here by analyzing signaling dynamics in single cells (Figs. 5 and 6), thereby providing a powerful aid to studies of biological variation—both across cells within a population and across time within a single cell—that would be lost in averaging.

B-FRET is of practical use even to experimenters who do not necessarily aim to reduce SNR per se: To achieve a given SNR, B-FRET requires fewer photons, reducing the need for high-power illumination and therefore the unwanted effects of photobleaching and phototoxicity (Fig. 1). Thus, B-FRET *computationally* extends

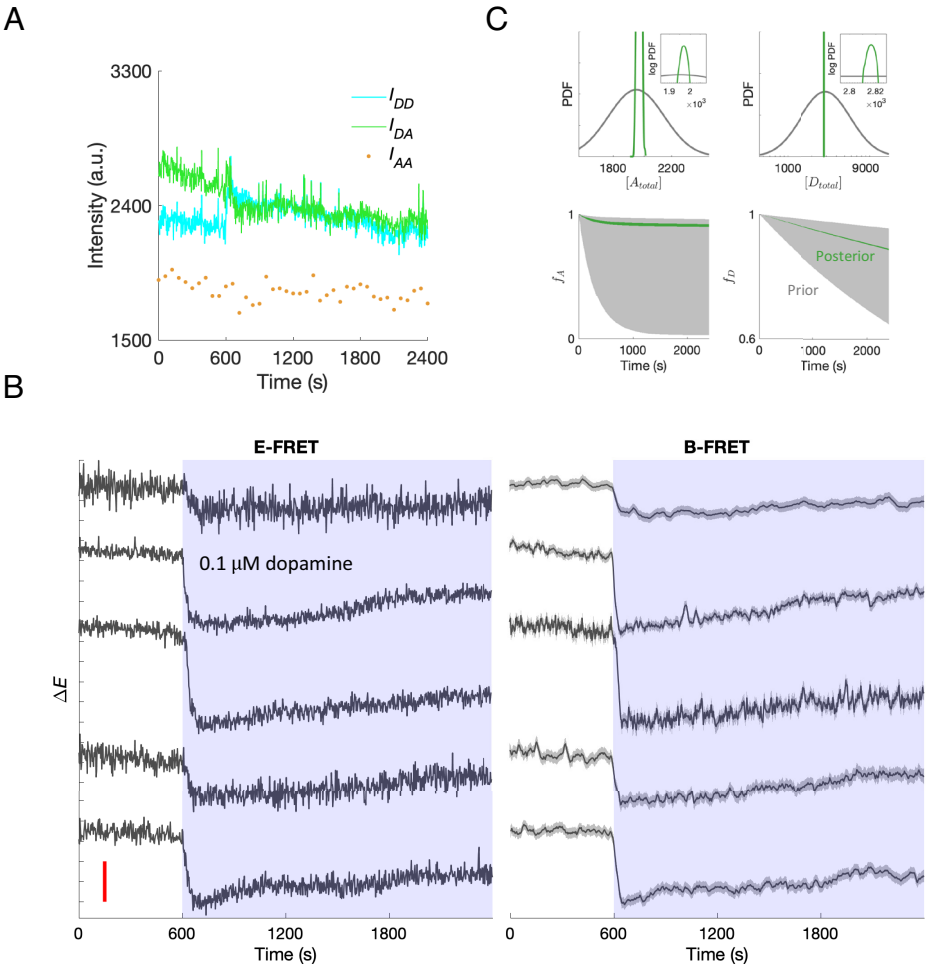

Fig. 6. B-FRET shows high signal-estimation performance with real unimolecular FRET data. (*A*) Three observables, I_{AA}, I_{DD}, and I_{DA} were acquired from a unimolecular FRET system in single HeLa cells that express a cAMP FRET biosensor and dopamine receptor D1 (DRD1). (*B*) Using FRET datasets from individual HeLa cells, FRET index E was estimated by E-FRET (*Left*) and B-FRET (*Right*). Five representative cells are shown. Blue in the background indicates 0.1 μM dopamine delivered to the cells. The red vertical line in the *Left* panel corresponds to the change in the FRET index $\Delta E = 0.05$. For B-FRET, 95% credible intervals are shown by the gray shade. The data in *A* corresponds to the cell shown at the bottom. (*C*) Prior and posterior distributions of the model parameters for the cell shown at the bottom, plotted in the same way as Fig. 2*C*.

the scope of FRET analyses by increasing the SNR and hence requiring fewer photons, in much the same way as brighter FRET pairs or more sensitive photodetectors *experimentally* enhance them.

A prominent feature of B-FRET is that it directly learns the entire photophysical model (46, 49, 52, 57)—the governing equations of a FRET measurement— rather than a single variable in the system. As a consequence, B-FRET provides a single algorithm for computing *any* measure of the degree of molecular interactions. For example, since B-FRET estimates $\chi = E_{max}[DA_{total}]$, $[D_{total}]$, and $[A_{total}]$ simultaneously, it is straightforward to estimate an acceptor-denominated FRET index $E_A(t) = \frac{E_{max}[DA_{total}](t)}{[A_{total}]}$ (46) rather than the more widely used donor-denominated FRET index $E(t) = \frac{E_{max}[DA_{total}](t)}{[D_{total}]}$ (Eq. **1**). In doing so, one only needs to note that a 3-cube FRET measurement makes two different conversions between the concentrations of FPs and fluorescence intensities, corresponding to the usage of the two different bands of excitation wavelengths (49, 70). Specifically, $E_{max}[DA_{total}]$ and $[D_{total}]$ are both converted by the same coefficient C_{DD} (Eq. **2** in *Materials and Methods*) to an observable, and so B-FRET determines the concentrations of them in the same, but generally unknown, unit; on the other hand, $[A_{total}]$ is converted by another coefficient C_{AA} (Eq. **2** in *Materials and Methods*) to an observable, and so B-FRET determines $[A_{total}]$ in a different unit from $E_{max}[DA_{total}]$ and $[D_{total}]$. Thus, computing the acceptor-denominated E_A requires the measurement of the ratio C_{DD}/C_{AA}, which is readily done using standard samples with known concentrations of FPs (49), while computing the donor-denominated E does not require such an additional measurement. The key point is that, because B-FRET explicitly considers the governing equations, it is always straightforward to identify the additional experiments necessary to obtain a certain measure of molecular interaction.

Bayesian methods have been developed for the analyses of data from other FRET experimental paradigms, such as lifetime-based (71) and single-molecule FRET (72, 73). However, the encoding methods in those other paradigms are qualitatively different, and therefore the Bayesian methods used in those studies are not applicable to the intensity-based, ensemble, time-lapse FRET measurements we focused on here. B-FRET provides a statistically optimal data-analysis method for this important class of FRET experiments, by exploiting temporal correlations present in time series and disentangling confounding factors associated with them.

As with all other quantitative FRET methods, the applications of B-FRET are naturally limited by our understanding of the photophysical processes involved in FRET measurement. For example, photoconversion of fluorescent proteins, which has been reported in certain conditions (54, 74, 75), can produce another chemical species apart from the donor and acceptor. If

such secondary processes are present in data but absent in the photophysical model (*Materials and Methods*), then B-FRET, like other methods, yields misleading results. Also, even with B-FRET, it is impossible to distinguish between a false negative and a true negative because there is no FRET signal to exploit in both cases. Thus, although it is possible to incorporate such processes as photoconversion into our photophysical model once they are characterized, B-FRET does not alleviate the necessity for careful selection of a FRET pair and for control experiments to validate the basic assumptions involved in the data analysis. The distinct advantage of B-FRET is that it makes all assumptions explicit—helping experimenters identify necessary controls and tailor their experiments accordingly.

Materials and Methods

Photophysical Model. Here, we consider the case of bimolecular FRET systems discussed in the main text; however, essentially the same argument applies to unimolecular FRET systems (*SI Appendix*, SI Text 1).

First, we define the FRET dataset. Time-lapse measurements of I_{DD}, I_{DA}, and I_{AA} are conducted at discrete time points. We assume that FRET from the donor to acceptor affects I_{DD} and I_{DA}, but not I_{AA} (See below and *SI Appendix*, SI Text 1). Thus, the sampling frequency of I_{DD} and/or I_{DA} limits the temporal resolution of an estimated FRET signal. In practice, I_{DD} and I_{DA} are measured (almost) simultaneously to better exploit the FRET-induced changes in I_{DD} and I_{DA}. Thus, we designate the same time points for the I_{DD} and I_{DA} measurements, and the set of the time points are written as $t^D_{1:N_D} \equiv \left\{ t^D_1, t^D_2, \cdots, t^D_{N_D} \right\}$, where N_D is the total number of measurements. I_{AA} is generally acquired at different time points from I_{DD} and I_{DA}, and thus we designate the time points for I_{AA} as $t^A_{1:N_A} \equiv \left\{ t^A_1, t^A_2, \cdots, t^A_{N_A} \right\}$, where N_A is the total number of measurements, and generally $N_D \neq N_A$. The entire set of the time-lapse fluorescence intensity data is $\boldsymbol{D} = \left\{ I_{AA,1:N_A}, I_{DD,1:N_D}, I_{DA,1:N_D} \right\}$, where

$$I_{AA,1:N_A} = \left\{ I_{AA}\left(t^A_1\right), I_{AA}\left(t^A_2\right), \cdots, I_{AA}\left(t^A_{N_A}\right) \right\},$$

$$I_{DD,1:N_D} = \left\{ I_{DD}\left(t^D_1\right), I_{DD}\left(t^D_2\right), \cdots, I_{DD}\left(t^D_{N_D}\right) \right\},$$

$$I_{DA,1:N_D} = \left\{ I_{DA}\left(t^D_1\right), I_{DA}\left(t^D_1\right), \cdots, I_{DA}\left(t^D_{N_D}\right) \right\}.$$

Next, we construct a photophysical model $\mathcal{M}$ to be learned from data $\boldsymbol{D}$. Under a standard 3-cube FRET-microscopy setup, the (background-subtracted) observables I_{AA}, I_{DD}, and I_{DA} are generally linked to the concentrations of the chemical species as follows (46, 49, 52, 57):

$$\begin{aligned} I_{AA}(t) &= C_{AA}\left([A^*](t) + \left[D^*A^*\right](t) + \left[DA^*\right](t)\right) + \xi_{AA}(t), \\ I_{DD}(t) &= C_{DD}\left(\left[D^*\right](t) + \left[D^*A\right](t) + \left(1 - E_{max}\right)\left[D^*A^*\right](t)\right) + \xi_{DD}(t), \\ I_{DA}(t) &= a\,\overline{I_{AA}(t)} + d\,\overline{I_{DD}(t)} + C_{DD}\,G\,E_{max}\left[D^*A^*\right](t) + \xi_{DA}(t), \end{aligned} \quad [2]$$

where ξ_{AA}, ξ_{DD}, and ξ_{DA} describe the measurement noise of corresponding fluorescent channels, and we assume they follow the zero-mean Gaussian distributions, i.e., $\xi_{AA}(t) \sim \mathrm{N}\left(0, \sigma^2_{AA}(t)\right) \xi_{DD}(t) \sim \mathrm{N}\left(0, \sigma^2_{DD}(t)\right)$ and $\xi_{DA}(t) \sim \mathrm{N}\left(0, \sigma^2_{DA}(t)\right)$ where σ^2_{AA}, σ^2_{DD}, and σ^2_{DA} are time-dependent variances and are determined from the data (Note, in ensemble FRET measurements, the number of photons collected from a sample is large enough to approximate the underlying Poisson distribution of photon counts by a continuous Gaussian distribution; see *SI Appendix*, SI Text 3 for more detail); $\overline{I_{AA}}$ and $\overline{I_{DD}}$ are, respectively, the expectation values of I_{AA} and I_{DD}, and thus $a\,\overline{I_{AA}}$ and $d\,\overline{I_{DD}}$, respectively, represent the cross-excitation of the acceptor by the donor excitation wavelengths and the bleed-through of the donor emission into the acceptor emission filter (6–9); C_{AA}, C_{DD}, a, d, and G are parameters dependent on imaging systems and the photophysical properties of the donor and acceptor, which are defined as

$$C_{AA} \equiv \nu_A \epsilon_{AA} Q_A L_A S_A t_{AA},$$

$$C_{DD} \equiv \nu_D \epsilon_{DD} Q_D L_D S_D t_{DD},$$

$$a \equiv \frac{\nu_D \epsilon_{DA} t_{DA}}{\nu_A \epsilon_{AA} t_{AA}},$$

$$d \equiv \frac{L_A S_A t_{DA}}{L_D S_D t_{DD}},$$

$$G \equiv \frac{Q_A L_A S_A t_{DA}}{Q_D L_D S_D t_{DD}},$$

where ν_D (ν_A) is the intensity of illumination reaching the sample through the donor (acceptor) excitation filter, ϵ_{DD} the absorption coefficient of the donor, ϵ_{DA} (ϵ_{AA}) the absorption coefficient of the acceptor at the donor-excitation (acceptor-excitation) wavelength, Q_D (Q_A) the quantum yield of donor (acceptor), L_D (L_A) the throughput of the donor (acceptor) emission light-path, S_D (S_A) the quantum sensitivity of the camera for donor (acceptor) emission, and t_{DA}, t_{AA}, and t_{DD}, respectively, the exposure time for the FRET, acceptor, and donor channels (7, 49, 52). The parameters a, d, and G can be determined by independent measurements (7, 11, 25). C_{DD} and C_{AA} do not necessarily need to be determined for determining the donor-denominated FRET index E (Eq. **1** and *Discussion*). The model (Eq. **2**) is general, only assuming that the acceptor fluorescence is not detectable through the donor emission filter and that the acceptor excitation light does not excite the donor, which are easily achieved by selecting appropriate filter sets (7, 25).

We introduce the following set of assumptions, which are satisfied in a typical FRET experiment and used also in E-FRET (7) (*SI Appendix*, SI Text 1). i) The total amount of donor and acceptor molecules are conserved during the course of a measurement. ii) The photobleaching locally follows a first-order decay process, i.e., the rate of change of the amount of intact (i.e., fluorescent) donor (acceptor) is proportional to its concentration, although the proportionality constants may change over time. iii) The system is in a quasi-steady state at each time point with the timescale of photobleaching being much larger than other relevant timescales (e.g., that of binding and unbinding of the fluorescently labeled proteins). See *SI Appendix*, SI Text 1 for how these are expressed mathematically.

Without loss of generality, we set $C_{AA} = C_{DD} = 1$, because these parameters only affect the units of the concentrations of the chemical species (*Discussion*). Also, we introduce a label $\chi(t) = E_{max}\left[DA_{total}\right](t)$ because we are not necessarily interested in decomposing E_{max} and $\left[DA_{total}\right](t)$, which only appear as a product in our definition of the FRET index E. Under these assumptions, Eq. **2** is reduced to

$$\begin{aligned} I_{AA}(t) &= f_A(t)\left[A_{total}\right] + \xi_{AA}(t), \\ I_{DD}(t) &= f_D(t)\left[D_{total}\right] - f_A(t) f_D(t)\chi(t) + \xi_{DD}(t) \\ I_{DA}(t) &= a f_A(t)\left[A_{total}\right] + d f_D(t)\left[D_{total}\right] + \left(G - d\right) f_A(t) f_D(t)\chi(t) + \xi_{DA}(t), \end{aligned} \quad [3]$$

where $[A_{\mathrm{total}}]$ and $[D_{total}]$ are the total concentrations of the acceptor and donor, respectively; $f_A(t)$ and $f_D(t)$ are arbitrary functions that express the intact fractions of the acceptor and donor at time t, respectively, and hence take values between 0 and 1 (see *SI Appendix*, SI Text 1 for derivation). To learn the model from data, $f_A(t)$ and $f_D(t)$ need to be expressed by parametric functions, whose parameters, as well as other parameters, are estimated by the inference algorithm described below. Any parametric functions can be used depending on the data in principle. See *SI Appendix*, SI Text 4 for the specific functions used to analyze data in this paper.

The presence of the hidden variable $\chi(t)$ in the equations for I_{DD} and I_{DA} (Eq. **3**) makes the learning of the model less straightforward. To deal with this, we rewrite the equations for I_{DD} and I_{DA} using the framework of the state-space model (55, 56):

$$\mathbf{x}_k = \mathbf{x}_{k-1} + \mathbf{q}_{k-1},$$

$$\mathbf{y}_k = \mathbf{H}_k \mathbf{x}_k + \mathbf{r}_k.$$

The first line, the dynamic model, describes the time evolution of the state $\mathbf{x}_k = \left(1,\ \chi(t_k^D)\right)^T$. The process noise $\mathbf{q}_{k-1}$ governs the transition between two consecutive states. For example, Gaussian process noise can be written as:

$$\mathbf{q}_{k-1} \sim \mathrm{N}\left(\mathbf{0}, \mathbf{Q}\left(\sigma_\chi\right)\right),$$

where the covariance matrix $\mathbf{Q}\left(\sigma_\chi\right)$ is defined as

$$\mathbf{Q}\left(\sigma_\chi\right) = \begin{pmatrix} 0 & 0 \\ 0 & \sigma_\chi^2 \end{pmatrix}.$$

For non-Gaussian dynamics, one can use Student's *t* distribution (56, 63), which can be written as

$$\mathbf{q}_{k-1} = \begin{pmatrix} 0 \\ q \end{pmatrix},$$

and

$$q \sim \mathrm{St}\left(q|\,\sigma_\chi, \nu\right) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\pi\nu\sigma_\chi^2}\,\Gamma\left(\frac{\nu}{2}\right)}\left(1+\frac{1}{\nu}\frac{q^2}{\sigma_\chi^2}\right)^{-\frac{\nu+1}{2}},$$

where σ_χ is the scale parameter and $\nu > 0$ is called the degree of freedom. When $\nu = 1$, the *t* distribution reduces to the Cauchy or Lorentz distribution, while for $\nu \gg 5$, it approaches a Gaussian distribution $N(0, \sigma_\chi^2)$.

The second line, the measurement model, describes the relationship between the observables $y_k = \left(I_{DD}\left(t_k^D\right),\ I_{DA}\left(t_k^D\right)\right)^T$ given the state $\mathbf{x}_k$. The measurement model matrix $\mathbf{H}_k$ at time t_k^D is defined as

$$\mathbf{H}_k = \begin{pmatrix} f_D(t_k^D)\left[D_{total}\right] & -f_A(t_k^D)\, f_D(t_k^D) \\ a\left[A_{total}\right] f_A(t_k^D) + d[D_{total}] f_D(t_k^D) & (G-d) f_A(t_k^D)\, f_D(t_k^D) \end{pmatrix}.$$

The Gaussian measurement noise $\mathbf{r}_k$ at time t_k^D is written as

$$\mathbf{r}_k \sim \mathrm{N}\left(\mathbf{0}, \mathbf{R}_k\right),$$

where the covariance matrix $\mathbf{R}_k$ is defined as

$$\mathbf{R}_k = \begin{pmatrix} \sigma_{DD}^2\left(t_k^D\right) & 0 \\ 0 & \sigma_{DA}^2\left(t_k^D\right) \end{pmatrix}.$$

The variances of measurement noise can be determined from data (*SI Appendix*, SI Text 3).

Learning Algorithm. To compute the posterior distribution of the FRET index E_k, $p\left(E_k|\mathcal{D},\ \mathcal{M}\right)$, where $E_k \equiv E(t_k^D)$, we rewrite it in term of model parameters $\boldsymbol{\theta}$:

$$p\left(E_k|\mathcal{D},\ \mathcal{M}\right) = \int p(\boldsymbol{\theta}|\mathcal{D},\ \mathcal{M}) p\left(E_k|\boldsymbol{\theta}, \mathcal{D},\ \mathcal{M}\right) d\boldsymbol{\theta}.$$

The right-hand side is the expectation of a function of model parameters $\boldsymbol{\theta}$, $p\left(E_k|\boldsymbol{\theta}, \mathcal{D},\ \mathcal{M}\right)$, with respect to the (posterior) distribution of $\boldsymbol{\theta}$, $p(\boldsymbol{\theta}|\mathcal{D},\ \mathcal{M})$. Thus, we evaluate the integral on the right-hand side by a Monte Carlo approach, drawing many samples from the two probability distributions as described below.

First, we evaluate $p(\boldsymbol{\theta}|\mathcal{D},\ \mathcal{M})$. Using Bayes' rule, this can be written as

$$p(\boldsymbol{\theta}|\mathcal{D},\ \mathcal{M}) = \frac{p(\boldsymbol{\theta}|\mathcal{M}) p(\mathcal{D}|\boldsymbol{\theta}, \mathcal{M})}{Z},$$

where $p(\boldsymbol{\theta}|\mathcal{M})$ is the prior distribution of the model parameters, which are usually wide in width to express one's ignorance about the parameter values (See *SI Appendix*, SI Text 2 for how to design prior distributions for each model parameter and *SI Appendix*, SI Text 4 for the actual distributions used to analyze data presented in this paper); $p(\mathcal{D}|\boldsymbol{\theta}, \mathcal{M})$ is the likelihood function, which describes the probability of observing a set of data $\mathcal{D}$ as a function of model parameters $\boldsymbol{\theta}$; and Z is the normalization constant, which one does not have to evaluate for the purpose of drawing samples from $p(\boldsymbol{\theta}|\mathcal{D},\ \mathcal{M})$. The prior distribution $p(\boldsymbol{\theta}|\mathcal{M})$ is given by the user of B-FRET, and the likelihood function $p(\mathcal{D}|\boldsymbol{\theta}, \mathcal{M})$ is evaluated by using the Bayesian filtering algorithm (55, 56) (*SI Appendix*, SI Text 2). Then, using a sampling method (44, 63), one can draw a set of samples $\left\{\boldsymbol{\theta}^i\right\}_{i=1}^R$ from the distribution, where $\boldsymbol{\theta}^i \sim p(\boldsymbol{\theta}|\mathcal{D},\ \mathcal{M})$ and $R\ (\gg 1)$ is the number of samples. Samples were drawn either directly from the distribution using a Markov chain Monte Carlo method [e.g., slice sampling (44, 63)] or from an approximated Gaussian distribution obtained by Laplace's method (44, 63). In drawing many samples, the latter is computationally much cheaper, and thus we adopted it upon confirmation that the bias introduced by the approximation is negligible (*SI Appendix*, Fig. S4).

Second, we evaluate the (posterior) distribution of the FRET index $p\left(E_k|\boldsymbol{\theta}^i, \mathcal{D}, \mathcal{M}\right)$ based on Bayesian smoothing algorithm (55, 56), using the sampled parameter set $\left\{\boldsymbol{\theta}^i\right\}_{i=1}^R$ (*SI Appendix*, SI Text 2). This enables to draw samples $\left\{E_k^i\right\}_{i=1}^R$ from the distribution, where $E_k^i \sim p\left(E_k|\boldsymbol{\theta}^i, \mathcal{D},\ \mathcal{M}\right)$. Using the samples, we can evaluate the integral as

$$\int p(\boldsymbol{\theta}|\mathcal{D},\ \mathcal{M}) p\left(E_k|\boldsymbol{\theta}, \mathcal{D},\ \mathcal{M}\right) d\boldsymbol{\theta} \approx \frac{1}{R}\sum_{i=1}^{R} \delta\left(E_k - E_k^i\right),$$

where $\delta(x)$ is the Dirac delta function. With sufficiently large R samples from $p\left(E_k|\mathcal{D},\ \mathcal{M}\right)$, one can quantify any properties of the distribution $p\left(E_k|\mathcal{D},\ \mathcal{M}\right)$. We used the median of the sample as representative values of the estimates and the interval between 2.27 and 97.73 percentiles, each of which corresponds to $\mu \pm 2\sigma$, respectively, for a Gaussian distribution $\mathrm{N}(\mu, \sigma)$, as a measure of the statistical uncertainty of the estimation and called it a "95% CI".

E-FRET Method and the Effect of Error in Optical Parameter Estimation. The E-FRET method (7) provides a formula for a FRET index E_{corr} that gives an estimate of Eq. **1**. This reads

$$E_{corr}(t) = \frac{I_{DA}(t) - a\overline{I_{AA}(t)} - dI_{DD}(t)}{I_{DA}(t) - a\overline{I_{AA}(t)} + (G-d)I_{DD}(t)} \frac{\overline{I_{AA}(t=0)}}{\overline{I_{AA}(t)}} \qquad [4]$$

where the optical parameters a, d, and G are defined in *Photophysical Model*. For the variables with bars, e.g., $\overline{I_{AA}(t)}$, their expected (or smoothed) values can be used as opposed to raw intensity values. It can be shown that under the assumptions described in *Photophysical Model* and in the limit of zero measurement noise, this quantity converges to the FRET index defined by Eq. **1** (*SI Appendix*, SI Text 1). The optical parameters a, d, and G are measured from independent measurements but only with finite precision. The errors in the estimations of these parameters introduce some biases in the computed FRET index, whose effect grows as more fluorescent proteins are photobleached, which can be corrected under some assumptions (25) (See *SI Appendix*, SI Text 3 for more detail).

Model Selection. In case a user of B-FRET is not sure about what model to use (e.g., Gaussian or non-Gaussian process noise), the framework of model selection enables to select, among a set of candidate models, a model that is best evidenced by a set of data. For this purpose, B-FRET computes the BIC defined as

$$\mathrm{BIC} = N_{\boldsymbol{\theta}} \log N_{\mathcal{D}} - 2\log p\left(\mathcal{D}|\boldsymbol{\theta}_{MAP}, \mathcal{M}\right),$$

where $N_{\boldsymbol{\theta}}$ and $N_{\mathcal{D}}$ are the numbers of the model parameters and data points, respectively, and $\boldsymbol{\theta}_{MAP}$ is the parameter values that maximize the likelihood function $p(\mathcal{D}|\boldsymbol{\theta}, \mathcal{M})$. A model with the lowest BIC value is selected as the best model among a set of candidates (*SI Appendix*, Fig. S2).

Data, Materials, and Software Availability. Data will be made available upon publication. Codes written in MATLAB and Python; numerical data in MATLAB format (MAT files) and MATLAB scripts data have been deposited in Github and Dryad (https://github.com/emonetlab; https://doi.org/10.5061/dryad.w3r2280w2) (76, 77).

ACKNOWLEDGMENTS. We thank H. H. Mattingly and Y. Kojima for commenting on the text. We acknowledge R. Gomez-Sjoberg, Microfluidics Lab, for providing information and software to control the solenoid valves in the microfluidic setup. This works was supported by NIH awards R01GM106189 (K.K., F.A., T.S.S., and T.E.) and R01GM138533 (K.K. and T.E.); JST, PRESTO Grant Number JPMJPR21E4, Japan (K.K.); a postdoctoral fellowship through the Swartz Foundation for Theoretical Neuroscience and postdoctoral fellowships NIH F32MH118700 and NIH K99DC019397 (N.K.); JSPS KAKENHI Grants, no. 19H05798 (K.A.); Netherlands Organization for Scientific Research (Nederlandse Organisatie voor Wetenschappelijk Onderzoek) (T.S.S.).

Author affiliations: [a]Department of Molecular, Cellular, and Developmental Biology, Yale University, New Haven, CT 06511; [b]Quantitative Biology Institute, Yale University, New Haven, CT 06511; [c]Institute of Molecular Biology, Academia Sinica, Taipei 115, Taiwan; [d]PRESTO, Japan Science and Technology Agency, Kawaguchi-shi, Saitama 332-0012, Japan; [e]Swartz Foundation for Theoretical Neuroscience, Yale University, New Haven, CT 06511; [f]AMOLF Institute, Amsterdam 1098 XG, Netherlands; [g]Institute of Physics, National Yang Ming Chiao Tung University, Hsinchu 30010, Taiwan; [h]Exploratory Research Center on Life and Living Systems, National Institutes of Natural Sciences, Okazaki, Aichi 444-8787, Japan; [i]National Institute for Basic Biology, National Institutes of Natural Sciences, Okazaki, Aichi 444-8585, Japan; [j]Department of Basic Biology, School of Life Science, SOKENDAI (The Graduate University for Advanced Studies), Okazaki, Aichi 240-0193, Japan; and [k]Department of Physics, Yale University, New Haven, CT 06511

1. J. E. Purvis, G. Lahav, Encoding and decoding cellular information through signaling dynamics. *Cell* **152**, 945–956 (2013).
2. U. Alon, *An Introduction to Systems Biology: Design Principles of Biological Circuits* (Chapman and Hall/CRC, 2006).
3. W. Lim, B. Mayer, T. Pawson, *Cell Signaling* (Taylor & Francis, 2014).
4. R. Y. Tsien, Building and breeding molecules to spy on cells and tumors. *FEBS Lett.* **579**, 927–932 (2005).
5. C. Albrecht, R. Joseph, *Lakowicz: Principles of Fluorescence Spectroscopy* (Springer, 2008).
6. G. W. Gordon, G. Berry, X. H. Liang, B. Levine, B. Herman, Quantitative fluorescence resonance energy transfer measurements using fluorescence microscopy. *Biophys. J.* **74**, 2702–2713 (1998).
7. T. Zal, N. R. J. Gascoigne, Photobleaching-corrected FRET efficiency imaging of live cells. *Biophys. J.* **86**, 3923–3939 (2004).
8. D. C. Youvan *et al.*, Calibration of fluorescence resonance energy transfer in microscopy using genetically engineered GFP derivatives on nickel chelating beads. *Biotechnology* **3**, 1–18 (1997).
9. A. Miyawaki, R. Y. Tsien, "Monitoring protein conformations and interactions by fluorescence resonance energy transfer between mutants of green fluorescent protein" in *Methods in Enzymology* (Elsevier, 2000), **vol. 327**, pp. 472–500.
10. M. G. Erickson, B. A. Alseikhan, B. Z. Peterson, D. T. Yue, Preassociation of calmodulin with voltage-gated Ca2+ channels revealed by FRET in single living cells. *Neuron* **31**, 973–985 (2001).
11. H. Chen, H. L. Puhl, S. V. Koushik, S. S. Vogel, S. R. Ikeda, Measurement of FRET efficiency and ratio of donor to acceptor concentration in living cells. *Biophys. J.* **91**, L39–L41 (2006).
12. A. J. Waite, N. W. Frankel, T. Emonet, Behavioral variability and phenotypic diversity in bacterial chemotaxis. *Annu. Rev. Biophys.* **47**, 595–616 (2018).
13. J. P. Moore, K. Kamino, T. Emonet, Non-genetic diversity in chemosensing and chemotactic behavior. *Int. J. Mol. Sci.* **22**, 6960 (2021).
14. M. Ackermann, A functional perspective on phenotypic heterogeneity in microorganisms. *Nat. Rev. Microbiol.* **13**, 497–508 (2015).
15. A. Guillemin, M. P. H. Stumpf, Noise and the molecular processes underlying cell fate decision-making. *Phys. Biol.* **18**, 011002 (2021).
16. C. Loos, J. Hasenauer, Mathematical modeling of variability in intracellular signaling. *Curr. Opin. Syst. Biol.* **16**, 17–24 (2019).
17. N. Eling, M. D. Morgan, J. C. Marioni, Challenges in measuring and understanding biological noise. *Nat. Rev. Genet.* **20**, 536–548 (2019).
18. S. Mitchell, A. Hoffmann, Identifying noise sources governing cell-to-cell variability. *Curr. Opin. Syst. Biol.* **8**, 39–45 (2018).
19. M. Niepel, S. L. Spencer, P. K. Sorger, Non-genetic cell-to-cell variability and the consequences for pharmacology. *Curr. Opin. Chem. Biol.* **13**, 556–561 (2009).
20. K. Kamino *et al.*, Fold-change detection and scale invariance of cell–cell signaling in social amoeba. *Proc. Natl. Acad. Sci. U.S.A.* **114**, E4149–E4157 (2017).
21. T. Gregor, K. Fujimoto, N. Masaki, S. Sawai, The onset of collective behavior in social amoebae. *Science* **328**, 1021–1025 (2010).
22. P. Conlon, R. Gelin-Licht, A. Ganesan, J. Zhang, A. Levchenko, Single-cell dynamics and variability of MAPK activity in a yeast differentiation pathway. *Proc. Natl. Acad. Sci. U.S.A.* **113**, E5896–E5905 (2016).
23. S. L. Spencer, S. Gaudet, J. G. Albeck, J. M. Burke, P. K. Sorger, Non-genetic origins of cell-to-cell variability in TRAIL-induced apoptosis. *Nature* **459**, 428–432 (2009).
24. J. M. Keegstra *et al.*, Phenotypic diversity and temporal variability in a bacterial signaling network revealed by single-cell FRET. *eLife* **6**, e27455 (2017).
25. H. H. Mattingly, K. Kamino, B. B. Machta, T. Emonet, Escherichia coli chemotaxis is information limited. *Nat. Phys.* **17**, 1426–1431 (2021).
26. R. Colin, C. Rosazza, A. Vaknin, V. Sourjik, Multiple sources of slow activity fluctuations in a bacterial chemosensory network. *eLife* **6**, e26796 (2017).
27. K. Kamino, J. M. Keegstra, J. Long, T. Emonet, T. S. Shimizu, Adaptive tuning of cell sensory diversity without changes in gene expression. *Sci. Adv.* **6**, eabc1087 (2020).
28. C. D. Harvey *et al.*, A genetically encoded fluorescent sensor of ERK activity. *Proc. Natl. Acad. Sci. U.S.A.* **105**, 19264–19269 (2008).
29. C. Janetopoulos, T. Jin, P. Devreotes, Receptor-mediated activation of heterotrimeric G-proteins in living cells. *Science* **291**, 2408–2411 (2001).
30. K. Aoki *et al.*, Stochastic ERK activation induced by noise and cell-to-cell propagation regulates cell density-dependent proliferation. *Mol. Cell* **52**, 529–540 (2013).
31. R. Tany, Y. Goto, Y. Kondo, K. Aoki, Quantitative live-cell imaging of GPCR downstream signaling dynamics. *Biochem. J.* **479**, 883–900 (2022).
32. J. Roux *et al.*, Fractional killing arises from cell-to-cell variability in overcoming a caspase activity threshold. *Mol. Syst. Biol.* **11**, 803 (2015).
33. B. Ponsioen *et al.*, Quantifying single-cell ERK dynamics in colorectal cancer organoids reveals EGFR as an amplifier of oncogenic MAPK pathway signalling. *Nat. Cell Biol.* **23**, 377–390 (2021).
34. S. De, C. Campbell, A. R. Venkitaraman, A. Esposito, Pulsatile MAPK signaling modulates p53 activity to control cell fate decisions at the G2 checkpoint for DNA damage. *Cell Rep.* **30**, 2083–2093.e5 (2020).
35. H. Imamura *et al.*, Single-cell dynamics of pannexin-1-facilitated programmed ATP loss during apoptosis. *eLife* **9**, e61960 (2020).
36. V. Lebrec, M. Poteau, J.-P. Morretton, O. Gavet, Chk1 dynamics in G2 phase upon replication stress predict daughter cell outcome. *Dev. Cell* **57**, 638–653.e5 (2022).
37. A. Papagiannakis, B. Niebel, E. C. Wit, M. Heinemann, Autonomous metabolic oscillations robustly gate the early and late cell cycle. *Mol. Cell* **65**, 285–295 (2017).
38. B. Bajar, E. Wang, S. Zhang, M. Lin, J. Chu, A guide to fluorescent protein FRET pairs. *Sensors* **16**, 1488 (2016).
39. A. J. Lam *et al.*, Improving FRET dynamic range with bright green and red fluorescent proteins. *Nat. Methods* **9**, 1005–1012 (2012).
40. N. Komatsu *et al.*, Development of an optimized backbone of FRET biosensors for kinases and GTPases. *Mol. Biol. Cell* **22**, 4647–4656 (2011).
41. J. W. Lichtman, J.-A. Conchello, Fluorescence microscopy. *Nat. Methods* **2**, 910–919 (2005).
42. M. J. Sanderson, I. Smith, I. Parker, M. D. Bootman, Fluorescence microscopy. *Cold Spring Harb. Protoc.* **2014**, pdb.top071795 (2014).
43. T. M. Cover, J. A. Thomas, *Elements of Information Theory* (Wiley, 2006), p. 774.
44. D. J. C. MacKay, *Information Theory, Inference, and Learning Algorithms* (Cambridge University Press, 2003), p. 640.
45. A. Miyawaki *et al.*, Fluorescent indicators for Ca2+based on green fluorescent proteins and calmodulin. *Nature* **388**, 882–887 (1997).
46. A. Zeug, A. Woehler, E. Neher, E. G. Ponimaskin, Quantitative intensity-based FRET approaches–A comparative snapshot. *Biophys. J.* **103**, 1821–1827 (2012).
47. C. Berney, G. Danuser, FRET or No FRET: A quantitative comparison. *Biophys. J.* **84**, 3992–4010 (2003).
48. C. Thaler, S. V. Koushik, P. S. Blank, S. S. Vogel, Quantitative multiphoton spectral imaging and its use for measuring resonance energy transfer. *Biophys. J.* **89**, 2736–2749 (2005).
49. J. Wlodarczyk *et al.*, Analysis of FRET signals in the presence of free donors and acceptors. *Biophys. J.* **94**, 986–1000 (2008).
50. M. Yamao *et al.*, Two new FRET imaging measures: Linearly proportional to and highly contrasting the fraction of active molecules. *PLOS One* **11**, e0164254 (2016).
51. A. Woehler, J. Wlodarczyk, E. Neher, Signal/noise analysis of FRET-based sensors. *Biophys. J.* **99**, 2344–2354 (2010).
52. C. A. Lichten, P. S. Swain, A Bayesian method for inferring quantitative information from FRET data. *BMC Biophys.* **4**, 10 (2011).
53. H. Chen, H. L. Puhl, S. R. Ikeda, Estimating protein-protein interaction affinity in living cells using quantitative Förster resonance energy transfer measurements. *J. Biomed. Opt.* **12**, 054011 (2007).
54. A. Miyawaki, Development of probes for cellular functions using fluorescent proteins and fluorescence resonance energy transfer. *Annu. Rev. Biochem.* **80**, 357–373 (2011).
55. S. Sarkka, *Bayesian Filtering and Smoothing* (Cambridge University Press, 2013), 10.1017/CBO9781139344203.
56. G. Kitagawa, *Introduction to Time Series Modeling* (Chapman and Hall/CRC, 2010).
57. R. A. Neher, E. Neher, Applying spectral fingerprinting to the analysis of FRET images. *Microsc. Res. Tech.* **64**, 185–195 (2004).
58. S. S. Vogel, C. Thaler, S. V. Koushik, Fanciful FRET. *Sci. STKE* **2006**, re2 (2006).
59. H. Babel *et al.*, Ratiometric population sensing by a pump-probe signaling system in Bacillus subtilis. *Nat. Commun.* **11**, 1176 (2020).
60. S. M. Emrich *et al.*, Omnitemporal choreographies of all five STIM/Orai and IP3Rs underlie the complexity of mammalian Ca2+ signaling. *Cell Rep.* **34**, 108760 (2021).
61. V. Sourjik, H. C. Berg, Receptor sensitivity in bacterial chemotaxis. *Proc. Natl. Acad. Sci. U.S.A.* **99**, 123–127 (2002).
62. V. Sourjik, A. Vaknin, T. S. Shimizu, H. C. Berg, "In vivo measurement by FRET of pathway activity in bacterial chemotaxis" in *Methods in Enzymology* (Elsevier, 2007), **vol. 423**, pp. 365–391.
63. C. M. Bishop, *Pattern Recognition and Machine Learning* (Springer, 2006).
64. J. S. Parkinson, G. L. Hazelbauer, J. J. Falke, Signaling and sensory adaptation in Escherichia coli chemoreceptors: 2015 update. *Trends Microbiol.* **23**, 257–266 (2015).
65. U. Alon, M. G. Surette, N. Barkai, S. Leibler, Robustness in bacterial chemotaxis. *Nature* **397**, 4 (1999).
66. K. Kamino, K. Fujimoto, S. Sawai, Collective oscillations in developing cells: Insights from simple systems: Collective oscillations in developing cells. *Dev. Growth Differ.* **53**, 503–517 (2011).
67. K. Kamino, Y. Kondo, Rescaling of spatio-temporal sensing in eukaryotic chemotaxis. *PLOS One* **18**, e0164674 (2016).
68. O. Karin, U. Alon, Temporal fluctuations in chemotaxis gain implement a simulated-tempering strategy for efficient navigation in complex environments. *iScience* **24**, 102796 (2021).
69. N. W. Frankel *et al.*, Adaptability of non-genetic diversity in bacterial chemotaxis. *eLife* **3**, e03526 (2014).
70. A. Hoppe, K. Christensen, J. A. Swanson, Fluorescence resonance energy transfer-based stoichiometry in living cells. *Biophys. J.* **83**, 3652–3664 (2002).
71. B. Kaye, P. J. Foster, T. Y. Yoo, D. J. Needleman, Developing and testing a Bayesian analysis of fluorescence lifetime measurements. *PLOS One* **12**, e0169337 (2017).
72. J. E. Bronson, J. Fei, J. M. Hofman, R. L. Gonzalez, C. H. Wiggins, Learning rates and states from biophysical time series: A bayesian approach to model selection and single-molecule FRET data. *Biophys. J.* **97**, 3196–3205 (2009).

73. S. Preus, S. L. Noer, L. L. Hildebrandt, D. Gudnason, V. Birkedal, iSMS: Single-molecule FRET microscopy software. *Nat. Methods* **12**, 593–594 (2015).
74. N. Malkani, J. A. Schmid, Some secrets of fluorescent proteins: Distinct bleaching in various mounting fluids and photoactivation of cyan fluorescent proteins at YFP-excitation. *PLoS One* **6**, e18586 (2011).
75. M. T. Kirber, K. Chen, J. F. Keaney, YFP photoconversion revisited: Confirmation of the CFP-like species. *Nat. Methods* **4**, 767–768 (2007).
76. K. Kamino, N. Kadakia, B-FRET. Github. https://github.com/emonetlab. Deposited 5 July 2022.
77. K. Kamino *et al.*, Data for: Optimal inference of molecular interaction dynamics in FRET microscopy. Dryad. https://datadryad.org/stash/dataset/; doi:10.5061/dryad.w3r2280w2. Deposited 23 March 2023.