FRETboard

semisupervised classification of FRET traces

de Lannoy, Carlos Victor; Filius, Mike; Kim, Sung Hyun; Joo, Chirlmin; de Ridder, Dick

Important note
To cite this publication, please use the final published version (if applicable).
Please check the document version above.

FRETboard: semi-supervised classification of FRET traces

C.V. de Lannoy, M. Filius, S.H. Kim, C. Joo and D. de Ridder

ABSTRACT   Förster resonance energy transfer (FRET) is a useful phenomenon in biomolecular investigations, as it can be leveraged for nano-scale measurements. The optical signals produced by such experiments can be analyzed by fitting a statistical model. Several software tools exist to fit such models in an unsupervised manner, but lack the flexibility to adapt to different experimental setups and require local installations. Here we propose to fit models to optical signals more intuitively by adopting a semi-supervised approach, in which the user interactively guides the model to fit a given dataset, and introduce FRETboard, a web tool that allows users to provide such guidance. We show that our approach is able to closely reproduce ground truth FRET statistics in a wide range of simulated single-molecule scenarios, and correctly estimate parameters for up to eleven states. On *in vitro* data we retrieve parameters identical to those obtained by laborious manual classification in a fraction of the required time. Moreover, we designed FRETboard to be easily extendable to other models, allowing it to adapt to future developments in FRET measurement and analysis.

SIGNIFICANCE   Förster resonance energy transfer (FRET) has seen extensive application as a nano-scale ruler in single-molecule research. The development of increasingly elaborate labeling schemes involving FRET has allowed the interrogation of single-molecule structure and kinetics in an increasing number of cases. However the diversity of produced data makes automated analysis more difficult. With our semi-supervised FRET analysis method and accompanying web tool, FRETboard, we aim to provide the flexibility and user-friendliness required for broad usage under any labeling scheme. With our tool, a user can analyze their data using a statistical model of choice in minutes, solely by providing a few correctly classified examples.

INTRODUCTION

Over the past decades, single-molecule Förster resonance energy transfer (smFRET) experiments have provided fundamental insights in biomolecular structure and many molecular mechanisms (e.g. 1–8). Although all experiments essentially rely on the same principle – that of distance-dependent energy transfer efficiency between fluorescent donor and acceptor dyes – the use of different labeling schemes allows for versatile application. For instance, dyes fixed to two points on a single molecule may provide spatial information valuable in solving its structure (2, 4), or register structural dynamics as the molecule exerts its biological function (1). Fixed to separate molecules, FRET may provide information on the occurrence and nature of molecular interactions (3, 5, 6). If a single dye pair provides insufficient information, a multitude of dye pairs may even be read out simultaneously by making use of stochastically blinking dyes (9). As each labeling scheme produces data of a different nature, it follows that widely applicable smFRET data analysis software should be flexible enough to adapt to these varying natures.

smFRET users currently may choose from a wide array of software packages, which mostly vary in scope and underlying trace analysis algorithms (10–18). The core utility included in all packages is the estimation of FRET efficiency distributions and transition rates given a set of traces, for which most rely on some flavor of hidden Markov model (HMM). While packages differ in how HMMs are fitted, the overall consensus is that the influence of the user in the fitting process should be minimized to safeguard objectivity. However an automated fitting procedure may find one of several good fits of which some may not make sense given the experimental context, a context which the user could provide.

Here we show that an HMM may be given that context for any particular FRET data set using a semi-supervised fitting approach, i.e. by allowing the user to manually curate classification of a limited number of traces to steer the model (Figure 1). Such direct intervention at the classification level makes model fitting a flexible, intuitive and computationally light-weight process. We further increase accuracy by introducing a more elaborate HMM structure that, to our knowledge, has not previously been applied to smFRET data. Using several additional features derived from the original signal further boosts accuracy and increases the flexibility to adapt to data of different labeling schemes. Our method is available for use through our web tool, FRETboard. FRETboard is a smFRET trace analysis solution which also supports data filtering and graphing utilities. As the smFRET field is rapidly developing and diversifying, we designed FRETboard to grow with the needs of the user community; by allowing anyone to easily extend FRETboard with existing or future classification algorithms, our tool may continue to serve as a unifying web front-end for high-level users with both niche and general classification needs.

METHODS

The FRET trace analysis method presented here is different from previous methods in three respects: the semi-supervised training approach, the structure of the models that are fitted and the features on which they were trained. Here we explain and justify our choices in each of these respects.
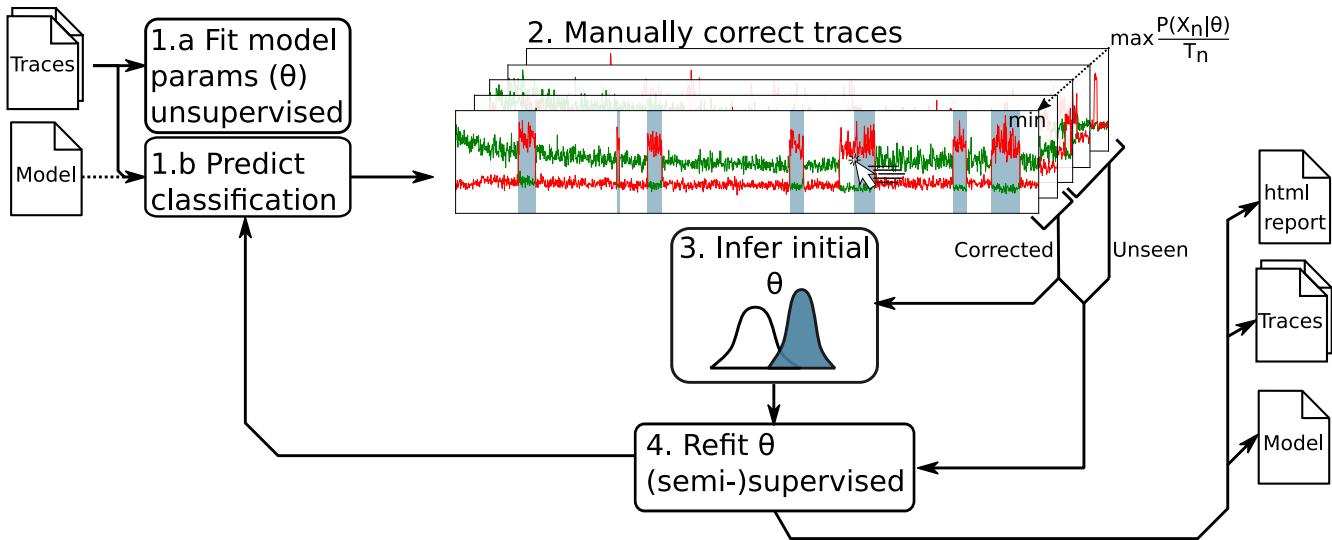
Figure 1: Our semi-supervised classification workflow for FRET trace classification, as implemented in FRETboard, is divided in 4 steps. **[1a]** FRET traces are uploaded to the web server, after which parameters (θ) of an initial model are fitted unsupervised and **[1b]** an initial trace classification is predicted. A suitable model generated during a previous FRETboard run can also be supplied to skip initial fitting. **[2]** The user is then shown the predicted classification of the trace at index n, X_n, with duration T_n, for which model fit was poorest based on duration-normalized trace probability given current model parameters ($P(X_n|\theta)/T_n$), and is asked for manual correction. **[3]** The curated trace is then used to reinitialize the model, which **[4]** is then trained in a (semi-)supervised fashion. Steps 2 to 4 may repeated until model fit is deemed satisfactory by the user.

Semi-supervised model fitting

We introduce semi-supervised fitting of models for the classification of FRET traces. The aim is to utilize the user's insight in the data structure to steer repeated light model fitting procedures, so that the resulting model will match the user's intuition. The fitting procedure is summarized in Figure 1 and a high-level description is given below. A more detailed description is given in Supplementary section C.

The procedure is initialized by fitting an unsupervised HMM on all loaded traces in a traditional manner, using randomly generated initial parameters which are then fitted using an implementation of expectation maximization (EM). The only user-provided guidance at this point is the number of states that should be recognized. Traces are classified and the trace for which the state path probability normalized over sequence length ($P(X_n|\theta)/T_n$ was lowest is presented to the user for manual correction of the classification.

The probability of assigned states for a given trace may be poor due to the presence of noise. If this is the case a user may choose to assign noisy measurements to a state they deem appropriate. However if it is more appropriate to remove the noise, as is the case in bleaching and blinking events, a the user may filter these measurements out by assigning them to a separate state reserved for such events. Such a state may be discarded prior to FRET distribution and transition rate analysis. Alternatively, model fit may suffer if the trace contains more or less states than those included in the current HMM. For example, we show this to be the case for simulated traces containing three states if the model contains only two states (Supplementary figure S3). In that case the user may simply adjust the number of states and adjust classification appropriately.

After applying manual corrections, the first semi-supervised training round on all loaded traces is started. State distributions and transition rates can now be deduced from the corrected trace and be used as initial parameters, after which the HMM is refitted on supervised and unsupervised traces simultaneously using semi-supervised EM. After refitting, traces are reclassified and the trace now marked by the lowest state path probability is presented to the user. The procedure is repeated until the user finds that presented traces are correctly classified.

Features

We trained our models on a combination of four features. The proximity ratio E_{PR} is included as an approximation of FRET efficiency and is defined as:

$$E_{PR} = \frac{F_{A_{em}}^{D_{ex}}}{F_{sum}}$$

Here $F_{D_{em}}^{D_{ex}}$ and $F_{A_{em}}^{D_{ex}}$ are the original donor and acceptor emission and F_{sum} denotes the summed donor and acceptor intensities $F_{D_{em}}^{D_{ex}} + F_{A_{em}}^{D_{ex}} = F_{sum}$. We also include the summed intensity as a separate feature, as it is expected to aid in the detection of bleaching.

Furthermore we used two time-aggregated features that capture the variability of features over a sliding-window of five measurements; the Pearson correlation coefficient between $F_{A_{em}}^{D_{ex}}$ and $F_{D_{em}}^{D_{ex}}$ (C) and standard deviation of F_{sum} ($\sigma_{F_{sum}}$). These features may aid models in capturing feature distributions characteristic for state transitions (Figure 2B). We specifically refrain from using $F_{D_{em}}^{D_{ex}}$ and $F_{A_{em}}^{D_{ex}}$ as features, as systematic variations frequently occur between experiments or even within the same experiment, which decreases the generalizability and re-usability of a trained model.
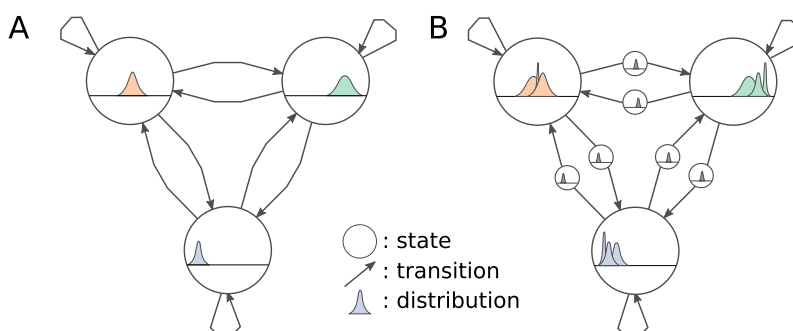
Model structures

Figure 2: Hidden Markov model graph structures used in this work: (**A**) the plain "vanilla" structure, and (**B**) the "GMM-HMM" structure, in which each state contains a Gaussian mixture distribution and additional states for feature distributions at state transitions. Circles denote states with a characteristic feature distribution, arrows denote transitions. Three-state structures are shown here, however similar structures with an arbitrary number of states can be constructed.

We evaluate the performance of two HMM structures in a semi-supervised learning setting (Figure 2). Each structure implements transitions and emissions differently and can be extended to an arbitrary number of states. The "vanilla" structure produces a straight-forward fully-connected HMM sporting no further modifications. The "GMM-HMM" structure models emissions using a Gaussian mixture model (GMM), which add the flexibility to classify noisier distributions as a single state, using multiple Gaussians. The number of Gaussians per GMM is determined per state using a Bayesian information criterion (BIC)-selection procedure. Furthermore this structure adds additional "edge states" between states, which are trained on measurements around a detected transition. Transitions between states may only occur through these edge states. If state transitions are marked by a signature distribution in a certain feature, this distribution is captured by the edge state, which allows for more accurate detection of state transitions. For example, under some labeling schemes transitions between FRET events may be marked by negative correlation between donor and acceptor signal, thus training a GMM-HMM on such data while using pearson correlation coefficient as a feature would make use of this fact.

Implementation

To facilitate application of our method we developed FRETboard, a browser-based graphical user interface (GUI) for semi-supervised training of segmentation and classification algorithms (Supplementary figure S1). In addition to intuitive example supervision, FRETboard offers users the flexibility to choose between model structures and opt which features to include. As we foresee that more suitable supervise-able classifiers may be proposed for the growing number of labeling schemes in the future, we also offer users the option to write custom algorithms and train them through the same FRETboard front-end. However due to the security risk of code injection that is inherent to running such custom code, users are advised to only allow this option on private machines that are not exposed to the public network.

Traces may be loaded in plain text, binary 64-bits or photon-HDF5 (19)-format, and may be corrected for background emission using our DBSCAN-based filter (Supplementary section D).

After the training procedure, the user may generate a report detailing feature distributions per state and transition rates. Transition rates are derived by deriving a transition matrix (A) from classified data, converting from discrete to continuous rates and multiplying by the frame rate f_s, thus arriving at corrected transition rates F (Equation 1) (11).

$$F = I + f_s \cdot \log A \tag{1}$$

Here I is the identity matrix and log denotes the natural matrix logarithm operation. 95% confidence intervals (CIs) of transition rates are estimated by repeatedly extracting transition rates from bootstrapped data. The CI is then reported using the bootstrap standard deviation on each parameter. Note that bootstrapping CIs is applicable to almost any (semi-)supervised model, thus any user-defined algorithm can make use of the same method. FRETboard is available as a web tool (`https://www.bioinformatics.nl/FRETboard`), thus freeing users from the burden of installation and maintenance, but it can also be used and hosted on a private server. FRETboard was written in python 3.7 (`https://www.python.org`). The GUI was implemented using the Bokeh interactive visualization library (v1.4.0)(20) (Supplementary figure S1). Included HMM model structures were implemented using pomegranate (v0.13.4) (21) and scikit-learn (v0.21.2) (22).

RESULTS

Below we validate our analysis method on four *in silico* and four *in vitro* data sets. To demonstrate the flexibility of our method, the different sets were simulated or recorded assuming a variety of realistic labeling schemes. All data sets used here are freely available (`https://git.wageningenur.nl/lanno001/fretboard_data`). All FRETboard runs were performed on a laptop running Ubuntu 18.04, on four CPU cores (Core i7 1.80GHz, Intel Corp.) with 4GB of memory. In total we supervised ten traces for each data set (3% and 10% of the total number of reads for *in silico* and *in vitro* data sets respectively), making use of the four described features (E_{PR}, F_{sum}, C, and $\sigma_{F_{sum}}$).

We assessed how well semi-supervised HMMs were able to reproduce ground truth parameters for simulated state sequences or, in the case of *in vitro* data, parameters acquired by manual labeling. To test whether predicted E_{PR} distributions attain a mean comparable to either reference value, we apply the two one-sided t-tests (TOST) procedure (23). That is, for a given state s the predicted mean of E_{PR} ($\hat{\mu}_s$) and the reference mean (μ_s) are calculated and two one-sided t-tests are employed to test $H_{01} : \hat{\mu}_s - \mu_s \leq -\frac{\delta}{2}$ and $H_{02} : \hat{\mu}_s - \mu_s \geq \frac{\delta}{2}$ versus $H_1 : -\frac{\delta}{2} < \hat{\mu}_s - \mu_s < \frac{\delta}{2}$. A rejection of both null hypotheses implies that the difference between emission means is significantly smaller than δ E_{PR} percentage points. Here we test for a maximum deviation of five or ten percentage points ($\delta = 0.05$ and $\delta = 0.1$ respectively), which we consider sufficiently accurate for many current applications. We report the TOST p-value for a given $\delta = \delta^*$ as $p_{\delta=\delta^*}$. Reported 95% CIs around estimated transition rates were calculated using FRETboard's built-in bootstrapping method, using a bootstrap size of 100.

Performance on *in silico* data

To demonstrate the flexibility of our approach, we simulated realistic FRET traces based on three different labeling schemes and classified them using a semi-supervised vanilla HMM. Briefly, the first two data sets contain two and three FRET states respectively, which are separable based on E_{PR} only (Supplementary figure S2A, B). The third data set contains three states, of which the third is identical to the second in its proximity ratio but has a different transition rate, making it a 'degenerate state' (Supplementary figure S2C). For a full description of the simulation methodology see Supplementary section A.

In all cases, estimated mean E_{PR} significantly differed less than 5 percentage points from the ground truth mean(p =0.05 « 0.001, Figure 3 A-C). Most ground truth transition rates fell well within bootstrapped 95%-CIs around the predicted rates (Figure 3 E-G). If a degenerate third state was present it was identified as such, however occasional misclassification between the two high-FRET states led to transition rates slightly differing from ground truth values (Figure 3G). In general we find that two rounds of semi-supervised training suffices to obtain parameter estimates close to ground truth values, while further rounds account for minor adjustments (Supplementary figure S4). Apart from the manual correction, no further parameter tuning or other user input was required, demonstrating that semi-supervised training provides the expected flexibility while maintaining accuracy.

To stress-test our method on a more difficult case, we generated a data set in which eleven FRET states of differing E_{PR}-levels were present (Supplementary figure S2D). The mean E_{PR} values were distributed such that their corresponding donor-acceptor distances were evenly distributed, thus causing lower and higher ends of the E_{PR} spectrum to be more densely crowded with states. As both visual examination during the training procedure and the produced results indicated poor performance by the vanilla HMM (Supplementary figure S5), we repeated our analysis using the GMM-HMM structure. Six of eleven E_{PR} distribution means significantly differed less than five percentage points from the ground truth means ($p_{\delta=0.05} << 0.001$, Figure 3D), with the remaining five differing less than ten percentage points ($p_{\delta=0.1} << 0.001$) and transition rate estimates were closer to ground truth values than vanilla HMM estimates (Figure 3H). This demonstrates another strength of our method; if the

user discovers upon visual inspection that simpler models cannot capture a user's classification, the training procedure is light enough that a more elaborate model can be selected on the fly, after which the analysis can continue without extra effort.

Finally we assessed how well the semi-supervised approach of FRETboard mitigates the effects of decreased signal-to-noise ratios (SNRs). We find that the GMM-HMM structure correctly estimated transition rates down to an SNR of 4.0, while E_{PR} levels could still be deduced down to an SNR of 2.75 (Supplementary figure S6). Interestingly the accuracy of manual labeling of traces decreased with SNR as well, thus increasingly erroneous supervised examples further contribute to the innate difficulty of classifying low-SNR data.
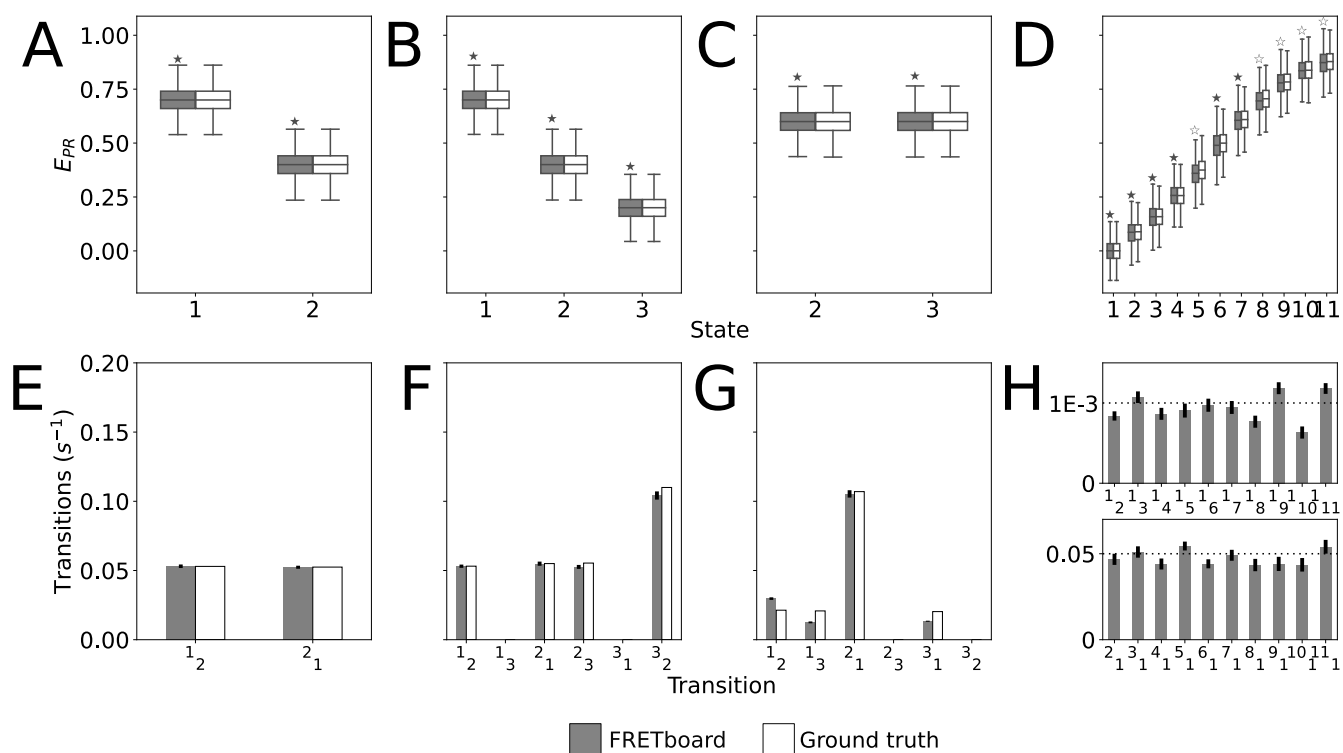

Figure 3: (E_{PR}) distributions per state in boxplots (**A-D**) and transition rates (**E-H**) as estimated by a semi-supervised hidden Markov model on four simulated data sets based on different labeling schemes; (**A,E**) producing two types of spatially separable FRET events, i.e. with different E_{PR} distributions, (**B,F**) three spatially separable states, (**C,G**) three states of which one containing no donor or acceptor signal and the others exhibiting only a kinetic difference, i.e. separable by transition rate and (**D, H**) eleven spatially separable states. Symbols above E_{PR} boxplots denote how much estimated means significantly differ from the ground truth at most (★: 0.05, ☆: 0.1). Solid black lines in **E-H** indicate 95% bootstrapped CIs. Dotted lines in **H** denote ground truth values.

Performance on *in vitro* data

We further validated our method on experimental data generated under immobilization schemes often used in single-molecule FRET, each marked by different classification challenges. Similar to our simulations, our *in vitro* data contains up to two types of FRET events and one ground state, which may be discernible by proximity ratio or transition rate. Lacking knowledge of the state sequence, we manually classified our data sets and used this classification to estimate proximity ratios and transition rates. A more extensive description of experimental methods can be found in Supplementary section B. All experimental data was analyzed using the GMM-HMM model structure, as the vanilla structure did not show a satisfactory increase in classification quality as training progressed (Supplementary figure S7).

First we designed an experiment in which a donor (Cy3)-labeled single-stranded (ss) DNA, containing a target site A, is immobilized through biotin-streptavidin conjugation on a quartz slide (Figure 4A). During measurement, we add (Cy5)-labeled eight-nucleotide 'imager' strands, which upon binding to site A produce FRET events marked by anti-correlated donor and acceptor signals (Figure 4E). Similar labeling schemes have previously seen application in point accumulation for imaging in nano-scale topography (PAINT) methods and the study of on- and off-rates (k_{on} and k_{off}) in biological systems (24, 25). In

this labeling scheme, bleaching of the donor dye due to continuous excitation occurs frequently, which may negatively impact kinetics analysis. Instead of requiring the user to remove bleaching events prior to analysis, we capture them in a separate state while training our GMM-HMM. This bleached state may then be discarded. Following this approach, we find that manually obtained transition rates for ground to high-FRET state and vice versa indeed fall within their respective estimated CIs ($0.119s^{-1}$ versus $CI : (0.095 - 0.122)$ and $0.546s^{-1}$ versus $CI : (0.468 - 0.578)$ respectively, Figure 4M). Estimated E_{PR} values (0.151 and 0.816 for ground and high-FRET states respectively) significantly differ by less than five percentage points from ground truth values ($p_{\delta<0.05} << 0.01$, Figure 4I).

We also performed the reverse experiment, in which the acceptor is immobilized and the donor is attached to the imager strand (Figure 4B). Although this labeling scheme does not suffer from dye bleaching as much, the lack of anti-correlation in the signal is expected to increase the difficulty of classification. Nonetheless, here too the predicted E_{PR} distribution mean of the high-FRET state (0.844) differed from the manually obtained value by less than 5 percentage points ($p_{\delta=0.05} << 0.01$, Figure 4J). As no dye is observable in the ground state under this labeling scheme, its E_{PR} value is meaningless and not analyzed here. The manually obtained transition rate from high-FRET to ground state fell within its predicted 95% CI ($0.367s^{-1}$ versus $CI : (0.333 - 0.429)$), while the rate for ground to high-FRET state was slightly underestimated ($0.012s^{-1}$ versus $CI : (0.007 - 0.011)$, Figure 4N).

Next, we evaluated performance in two scenarios where two FRET states are present. For these experiments we followed the same experimental procedure, but simultaneously flushed in two types of donor-bound free-floating imager strands, at a 1:1 ratio (Figure 4C,D).

In the first experiment, the second imager strand was complementary to a second target site B at 15nt from the acceptor – 10nt further than target site A – where imager strand binding should produce an intermediate E_{PR} (Figure 4G). Upon analysis, the GMM-HMM model found E_{PR} means of 0.85 and 0.72 for states 2 and 3 respectively, matching the manually obtained state means ($p_{\delta<0.05} < 0.01$, Figure 4K). Most transition rate estimates fell within the predicted 95% CIs, except for that from mid-FRET to ground state ($0.439s^{-1}$ versus $CI : (0.483 - 0.780)$, Figure 4O). Upon inspection of traces, we find that some short mid-FRET events had been erroneously detected in noisy ground state stretches – a common occurrence in smFRET analysis and therefore not explicitly accounted for by e.g. removing traces from analysis manually.

In the second experiment, site A was targeted with a second imager strand of 7nt – 1nt shorter than its counterpart – (Figure 4D) which should increase the off-rate and produce a degenerate state (Figure 4H). Here too our GMM-HMM produced parameter estimates close to manually obtained values on traces containing degenerate states, which is surprising given our results on *in silico* data. Predicted transition rates from state 2 to ground state were higher at $2.22s^{-1}(CI : 1.87 - 2.57)$, than that of state $3 - 0.43s^{-1}(CI : 0.382 - 0.479)$ –, which resembled transition rates seen in other *in vitro* experiments (Figure 4O). Presumably, state 2 corresponds to the annealing of the shorter 7nt imager strand. Both are in close agreement with manually obtained rates ($2.24s^{-1}$ and $0.411s^{-1}$ respectively).

Finally, we compared FRETboard estimates for our experiments against those of three other tools, each of which employs a different solution to the classification problem: ebFRET (13), which fits an HMM using a Bayesian approach, MASH-FRET (12, 17), a flexible software suite that includes several other tools and infers transition rates through exponential fitting, and DeepFRET (18), which filters traces from noise using a neural network before classifying them with an HMM. As we focus on transition rate and emission distribution analysis only, we fed the same background-subtracted traces to each tool and forced them to use the correct number of states if it allowed us to do so. We find that for each experiment, FRETboard returns transition rate and E_{PR} distribution estimates that are equally close or closer to the manually derived values as estimates of other tools (Supplementary figure S9). Analysis procedures followed for other tools are detailed in Supplementary section E.

DISCUSSION

We show that semi-supervised classification models, in particular hidden Markov models (HMMs), are capable of capturing properties of FRET events in a wide array of realistic experimental scenarios, using a combination of input features derived from the original donor and acceptor dye emission intensities. We also provide an HMM structure that is better suited for semi-supervised learning than the straight-forward fully connected model, and provide a particular advantage in noisy real-world and complex data sets containing more than two states. To accommodate for the intensive user interaction required for this method we developed FRETboard, an intuitive browser-based tool that allows data filtering, model training, classification and report generation.

We placed our method in the landscape of existing analysis tools by comparing its analysis results to three other often-used tools. We find that FRETboard model parameter estimates were as accurate or more accurate than those of other tools, using manual classification as a reference. Furthermore, FRETboard sports several other features not previously seen in smFRET analysis, including implementation as an analysis server that can be used remotely in the browser and the possibility to run custom algorithms. However it should be noted that FRETboard is focused on trace analysis and does not currently support any
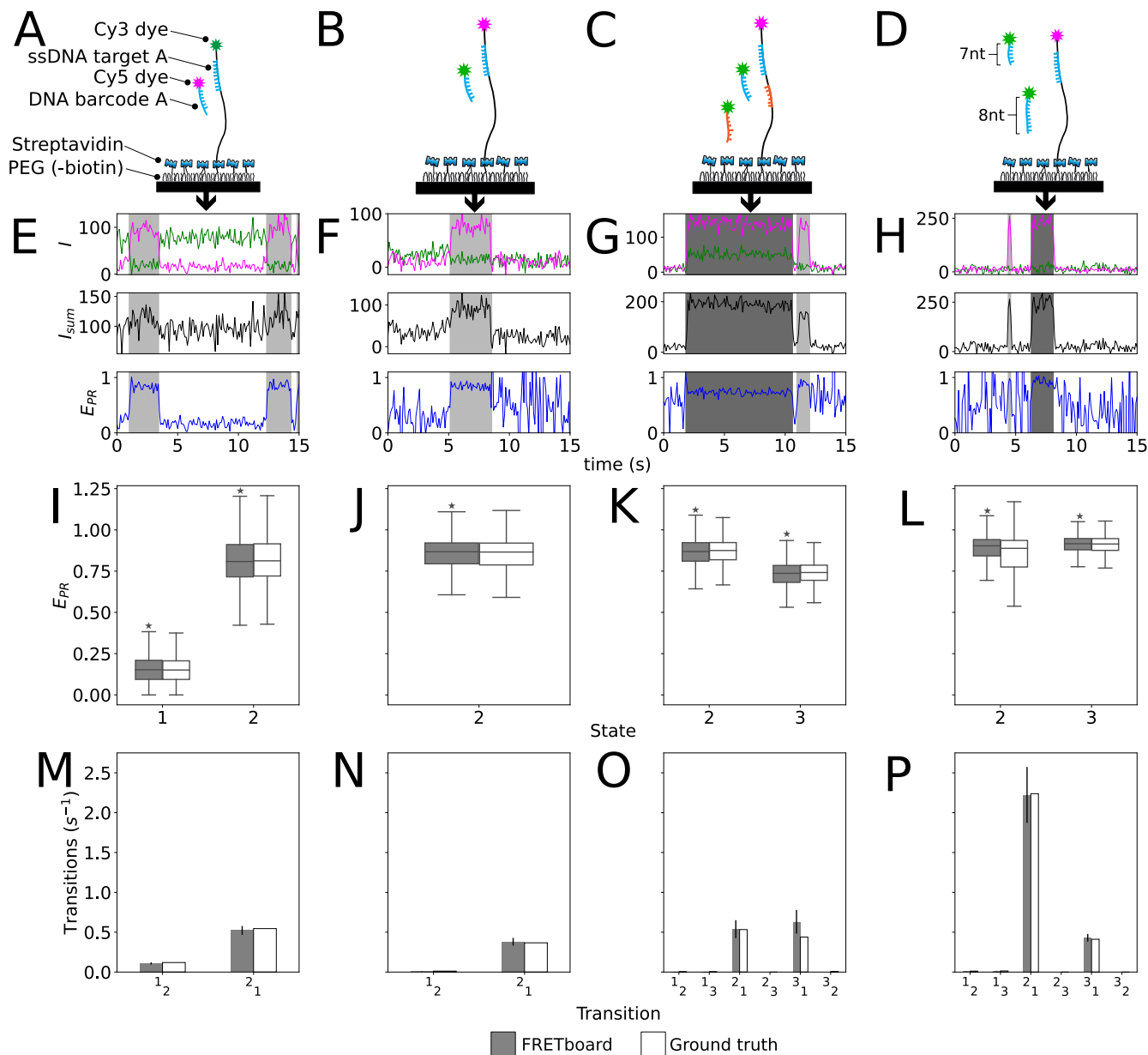
Figure 4: Used labeling schemes (**A-D**), examples of events produced as correctly found by semi-supervised HMM classification (**E-H**), E_{PR} distribution boxplots per state (**I-L**) and estimated transition rates (**M-P**) for four different labeling schemes on which our semi-supervised HMM fitting method was evaluated. From left to right, these labeling schemes were producing single-type FRET events using an immobilized donor (**A**) or an immobilized acceptor (**B**), two types of events producing high- and mid-FRET events (**C**) and two types of kinetically different events (**D**). In **J-L**, the non-FRET state [1] is omitted as E_{PR} values are meaningless for labeling schemes in which the donor is not immobilized. In **I-J**, ⋆ marks E_{PR} distributions for which estimated means significantly differ less than 0.05 from values acquired through manual analysis. In **M-P** solid lines denote bootstrapped 95% CIs.

functionality for trace extraction from microscope images. Several other tools such as MASH-FRET (12, 17), iSMS(14), and SPARTAN (15) do include this functionality, and therefore provide a more complete software solution.

Another important caveat particular to our method is related to the quality of the user's supervision. Contrary to other approaches we give responsibility for proper classification to the user, embracing the pros and cons of user input; on the one hand, it allows for efficient training and yields results that match the user's intuition, on the other hand it matches mistakes that the user may make. Users may derive the knowledge necessary for supervision (e.g. number of states) from their experimental setup as was done in the validation shown here, or start analysis agnostically and estimate such information from the traces itself. To test our method in the latter scenario we have entered FRETboard into the KinSoft challenge (`https://sites.google.com/view/kinsoftchallenge/home`), the first blind assessment of smFRET kinetic analysis tool performance. A publication on the challenge results is pending.

Lastly, we encourage users to design their own classifiers and test them through the FRETboard interface; many more supervise-able HMM flavors and entirely different classifiers exist and may be a better fit than the models currently included for certain experimental data. In consultation with the authors of such custom classifiers, these may also be included in future releases of FRETboard. This would allow it to become a unifying front end for FRET trace analysis, with back end support for the expanding variety of smFRET experimental methods.

AUTHOR CONTRIBUTIONS

C.L., M.F., C.J. and D.R. conceived and designed the project. C.L. developed algorithms and software, and performed analyses. M.F. designed and performed experiments. S.H.K. designed, wrote and performed the simulation procedure. D.R. supervised algorithm development. C.L., M.F. and S.H.K. wrote the manuscript. All authors discussed results and improved the manuscript.

ACKNOWLEDGEMENTS

REFERENCES

1. Kapanidis, A. N., E. Margeat, S. O. Ho, E. Kortkhonjia, S. Weiss, and R. H. Ebright, 2006. Initial transcription by RNA polymerase proceeds through a DNA-scrunching mechanism. *Science* 314:1144–1147.

2. Ferreon, A. C. M., Y. Gambin, E. A. Lemke, and A. A. Deniz, 2009. Interplay of α-synuclein binding and conformational switching probed by single-molecule fluorescence. *Proc. Natl. Acad. Sci. U. S. A.* 106:5645–5650.

3. Börsch, M., and T. M. Duncan, 2013. Spotlighting motors and controls of single FoF1-ATP synthase. *Biochem. Soc. Trans.* 41:1219–1226.

4. Lerner, E., A. Ingargiola, and S. Weiss, 2018. Characterizing highly dynamic conformational states: The transcription bubble in RNAP-promoter open complex as an example. *J. Chem. Phys.* 148:123315.

5. Newton, M. D., B. J. Taylor, R. P. Driessen, L. Roos, N. Cvetesic, S. Allyjaun, B. Lenhard, M. E. Cuomo, and D. S. Rueda, 2019. DNA stretching induces Cas9 off-target activity. *Nat. Struct. Mol. Biol.* 26:185–192.

6. Globyte, V., S. H. Lee, T. Bae, J.-S. Kim, and C. Joo, 2019. CRISPR/Cas9 searches for a protospacer adjacent motif by lateral diffusion. *EMBO J.* 38:e99466.

7. Lerner, E., B. Ambrose, A. Barth, V. Birkedal, S. C. Blanchard, R. Borner, T. Cordes, T. D. Craggs, T. Ha, G. Haran, et al., 2020. The FRET-based structural dynamics challenge–community contributions to consistent and open science practices. *arXiv preprint arXiv:2006.03091* .

8. Hellenkamp, B., S. Schmid, O. Doroshenko, O. Opanasyuk, R. Kühnemuth, S. R. Adariani, B. Ambrose, M. Aznauryan, A. Barth, V. Birkedal, et al., 2018. Precision and accuracy of single-molecule FRET measurements—a multi-laboratory benchmark study. *Nature Methods* 15:669–676.

9. Uphoff, S., S. J. Holden, L. Le Reste, J. Periz, S. Van De Linde, M. Heilemann, and A. N. Kapanidis, 2010. Monitoring multiple distances within a single molecule using switchable FRET. *Nature Methods* 7:831.

10. McKinney, S. A., C. Joo, and T. Ha, 2006. Analysis of single-molecule FRET trajectories using hidden Markov modeling. *Biophys. J.* 91:1941–1951.

11. Greenfeld, M., D. S. Pavlichin, H. Mabuchi, and D. Herschlag, 2012. Single Molecule Analysis Research Tool (SMART): an integrated approach for analyzing single molecule data. *PLOS ONE* 7:30024.

12. König, S., M. Hadzic, E. Fiorini, R. Börner, D. Kowerko, W. Blanckenhorn, and R. Sigel, 2013. BOBA FRET: Bootstrap-Based Analysis of Single-Molecule FRET Data. *PLOS ONE* 8:1–17.

13. van de Meent, J. W., J. E. Bronson, C. H. Wiggins, and R. L. Gonzalez, 2014. Empirical Bayes methods enable advanced population-level analyses of single-molecule FRET experiments. *Biophys. J.* 106:1327–1337.

14. Preus, S., S. L. Noer, L. L. Hildebrandt, D. Gudnason, and V. Birkedal, 2015. iSMS: single-molecule FRET microscopy software. *Nature Methods* 12:593–594.

15. Juette, M., D. Terry, M. Wasserman, R. Altman, Z. Zhou, H. Zhao, and S. Blanchard, 2016. Single-molecule imaging of non-equilibrium molecular ensembles on the millisecond timescale. *Nature Methods* 13:341–344.

16. Schmid, S., M. Götz, and T. Hugel, 2016. Single-molecule analysis beyond dwell times: demonstration and assessment in and out of equilibrium. *Biophys. J.* 111:1375–1384.

17. Hadzic, M. C. A. S., R. Börner, S. L. B. König, D. Kowerko, and R. K. O. Sigel, 2018. Reliable State Identification and State Transition Detection in Fluorescence Intensity-Based Single-Molecule Förster Resonance Energy-Transfer Data. *The Journal of Physical Chemistry B* 122:6134–6147.

18. Thomsen, J., M. B. Sletfjerding, S. Stella, B. Paul, S. B. Jensen, M. G. Malle, G. Montoya, T. C. Petersen, and N. S. Hatzakis, 2020. DeepFRET: Rapid and automated single molecule FRET data classification using deep learning. *bioRxiv* .

19. Ingargiola, A., T. Laurence, R. Boutelle, S. Weiss, and X. Michalet, 2016. Photon-HDF5: An open file format for timestamp-based single-molecule fluorescence experiments. *Biophys. J.* 110:26–33.

20. Bokeh Development Team, 2019. Bokeh: Python library for interactive visualization. `https://bokeh.org/`.

21. Schreiber, J., 2018. Pomegranate: fast and flexible probabilistic modeling in python. *J. Mach. Learn. Res.* 18:1–6.

22. Pedregosa, F., G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and Duchesnay, 2011. Scikit-learn: Machine Learning in Python. *J. Mach. Learn. Res.* 12:2825–2830.

23. Schuirmann, D. J., 1987. A comparison of the two one-sided tests procedure and the power approach for assessing the equivalence of average bioavailability. *J. Pharmacokinet. Biopharm.* 15:657–680.

24. Lee, J., S. Park, W. Kang, and S. Hohng, 2017. Accelerated super-resolution imaging with FRET-PAINT. *Mol. Brain* 10:63.

25. Filius, M., T. J. Cui, A. N. Ananth, M. W. Docter, J. W. Hegge, J. van der Oost, and C. Joo, 2020. High-Speed Super-Resolution Imaging Using Protein-Assisted DNA-PAINT. *Nano Lett.* 20:2264–2270.