Classification of Single Molecular Motor Events

Anthony E. Butterfield and Mikhail Skliar

***Abstract*—This paper describes the development of two methods for automatic classification of molecular events. The methods are applied to experimental results obtained to study the interaction of a single microtubule with a single Ncd molecular motor. The data were previously obtained using optical trap assay, and have a very low signal-to-noise ratio of approximately 0.1. The first method can be viewed as a syntactic classification. In an alternative approach, radial basis neural networks, trained with simulated data generated by a system of stochastic, Langevin-like differential equations, are used for classification. Following classification, the molecular events are ensemble averaged separately for different types of events. The analysis of the results suggests the existence of Ncd-microtubule interaction events, which deviate from the traditional view on the kinetics of the Ncd-microtubule interactions. This paper employs a novel paradigm, which emphasizes the importance of the model-based, time-resolved filtering of the experimental data on molecular interactions, especially for the case of biological molecules, followed by the analysis of different groups of events, automatically segregated according to features revealed during the single realization analysis of the experimental results.**

***Index Terms*— Single Molecule, Event Characterization, Model-Based Filtering, Pattern Classification**

I. INTRODUCTION

Recent advances in nanotechnology, especially novel instrumentations, such as scanning probe microscopes, optical tweezers and single molecule fluorescent tagging, created unprecedented opportunities in many branches of sciences and engineering. At the same time, probing of processes on the nanometer scales is characterized by significant challenges. First, there is a problem of interference of experimental techniques with the underlying physical and molecular processes. This parallels fundamental limitations imposed by the uncertainty principle on a subatomic level. At the same time, interference from the local environment often cannot be eliminated. The most prominent examples are biological molecules that can function only in aqueous solution. Aqueous solution introduces unavoidable thermal fluctuations. Fundamentally, nanoscience and engineering must rely on experiments that produce measurements with low signal-to-noise ratio, and poor repeatability. A second challenge is presented by the fact that nanoscale processes often limit intuition-based analysis and interpretation of results. On the nanoscale level, there are often no parallels in the macroscopic world. Consequently, the results of nanoscale measurements can often be interpreted based only on models. Such models must necessarily reflect the stochastic nature of nanoevents and probabilistic features of nanoscale experiments.

This paper deals with the interpretation of the experimental data obtained using optical tweezers, and directly reflective of interaction between two biological macromolecules – non-processive molecular motor Ncd, a member of the kinesin family of molecular motors, and its corresponding micro-filament, a microtubule [4]. The experimental data examined in this paper was first reported in [2], and shows the peak-to-peak thermal fluctuations on the order of 80nm, while the motor induced displacement, the primary characteristic of the motor-microtubule interaction of interest to us is on the order of 8nm. Therefore, in this and most other cases of single molecular interactions some sort of data filtering is necessary before interpretation of the experimental results is possible. The traditional approach is to use ensemble averaging of similar molecular events, which requires pre-classification of experimental data into the groups of similar events. Such classification is often subjective, and is usually based on simple statistical characteristics of the experimental data, such as change in the variance. At the same time, it becomes increasingly clear that, at least in the case of the molecular motor—microfilament interactions, a large number of different molecular events can be expected, which cannot be adequately classified using simple statistical segmentation of experimental data. Under these circumstances, we argued [1] that model-based, single realization filtering of the experimental data is needed to generate the hypothesis about different classes of molecular events that are expected to occur. Without knowing what types of events are expected to occur, an adequate classification of the data into different types of molecular events is impossible. In particular, by using model-based single-realization filtering [1] we are able to resolve the action of a single molecular motor, and found, by inspection, that a significant portion of the molecular motor-microtubule interactions is inconsistent with the traditional views on motor kinetics used to guide the analysis of Ncd-microtubule interactions.

The specific emphasis of the current paper is on the development of the molecular event classification methods and their application to the measurements reflective of Ncd-microtubule interaction. We use the results of the model-based Kalman filtering of molecular events to identify the type of events that need to be classified for the experimental data. Two classification methods are proposed to discriminate the single Ncd-microtubule interactions into different types of molecular events. The first method is based on syntactic classification ideas, and is used to discriminate between different molecular events using the state estimation of the Ncd-microtubule system generated by the Kalman filter. Since some of our conclusions deviate from the established view on the Ncd-microtubule interactions, we felt that it is important to confirm our classification results using a completely dissimilar method. The selected method is the neural network (NN) classification with radial NN, trained under supervision. The neural network classification was applied directly to the raw data. The events in the different types of molecular events were then separately ensemble averaged to reduce the effect of the thermal noises, and to obtain the characteristic behavior for different events. The results of the analysis give strong indication that multiple, dissimilar types of the Ncd-microtubule interactions are present.

Authors are with the Chemical and Fuels Engineering Department, University of Utah, Salt Lake City, UT 84112 USA. For color reprints e-mail mikhail.skliar@utah.edu.

0-7803-7896-2/03/$17.00 ©2003 IEEE

Proceedings of the American Control Conference
Denver, Colorado June 4-6, 2003

II. Experimental Methods

The optical trap system used during the experimental study on a single Ncd-microtubule interaction is shown in Figure 1a [2]. A span of biotin-labeled microtubule (approximately 4.3 μm long) was secured between two streptavidin coated silica beads of 1 μm radius. Using optical tweezers, the beads were separated until the microtubule became taut. The bead/microtubule structure was then lowered onto a third bead (2 μm radius), which was immobilized and sparsely coated with Ncd molecular motors. The experimental data were collected at room temperature and with various ATP (adenosine triphosphate) concentrations in solution, ranging from 2 to 50 μM. Motor attachments and subsequent strokes were then detected by a back-focal-plane interferometry using a 4-quadrant position detector running at 5 kHz. The position of a single bead (arbitrarily labeled as the right bead, R) was measured in the x- and y-axis.

Our method of event classification relies on the filtered position of the right bead and the identified motor head position obtained using Kalman filtering. The experimental setup was modeled as a series of rods, springs and masses, leading to the Langevin-like system of equations. Figure 1b shows the graphical representation of the model used during Kalman filtering of the bead position x_R and identification of the motor head displacement, x_{AT}. A simplification of this model was also used, based on a single-mass-on-a-spring model with the overall spring constant K_{total}. The simplified model was used in the extended Kalman filter to estimate K_{total}. Once identified, change in K_{total} was used as an indication of Ncd-microtubule attachment, evidenced by higher values of K_{total}. A third model and the corresponding Kalman filter was used to estimate the motor head displacements in the y-axis, though the results of this filter were not used in classifications.

The ability of the Kalman filters to successfully resolve single motor actions was first demonstrated by their application to simulated data. When applied to experimental data, the model-based filters were able to resolve single motor events, a task which could not be achieved with traditional filtering methods of ensemble averaging.

In analyzing the experimental data it became necessary to determine how a motor event may manifest itself. Attachment of the motor can occur anytime throughout the 80 nm thermal vibration range of the microtubule. The location of the attachment point changes the equilibrium position of the microtubule, leading to a noticeable shift in the mean position of the right bead in the x-axis, as depicted in Figure 2a. We refer to this change as an *equilibrium displacement* and distinguish it from the displacement due to the motor stroke. Once attached, the equilibrium displacements create an apparent change in x_{AT}. When multiplied by the spring constant of the Ncd-microtubule link, x_{AT} is used to model the attachment-point-dependent force introduced by this link. In the hypothetical example of Figure 2a, Ncd attaches when the microtubule is thermally shifted to the right of its mean unattached position. The Ncd attaches to the microtubule at the position, x_{eq}, (Fig. 2a and b), causing the shift in the mean motor head position x_{AT} and, consequently, the shift in the equilibrium position of the microtubule, and finally the change in the position of the right measurement bead, which is directly measured.

According to the previously proposed kinetic model [2,4], in the presence of ATP the Ncd motor action is expected to produce an additional positive displacement (motor stroke) at the end of the Ncd-microtubule attachment event. We refer to the shift in the motor head displacement, just before release, as the *terminal displacement*, x_{tm}, and assume that this displacement is the result of the motor stroke. Figure 2b shows the identified time dependent values of x_{AT} and K_{total} reflecting the following sequence of events: (1) Ncd-microtubule attachment, causing the equilibrium motor head displacement, and an increase in K_{total}; (2) motor stroke resulting in the terminal displacement, and (3) Ncd detachment from the microtubule, and the corresponding return to the original mean motor head position and the decrease in the value of K_{total}.

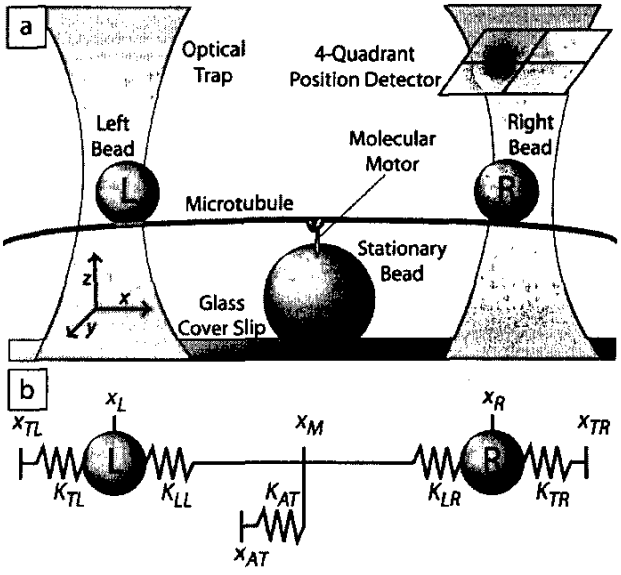

Fig. 1. A) Experimental Setup; B) Phenomocological model of the Ncd-microtubule interactions, used in the Kalman filter.

On comparison of the various types of single experimental motor events, it appeared that a wide range of motor stroke characteristics were present. A minority these events significantly deviated from the expected motor action depicted in Figure 2, such as motor stroke in the direction opposite to the motors ensemble action (the negative direction in this work).

The inclusion of motor events, which deviate from the traditional mechanism into ensemble averaging, may change the conclusion about the properties of a single motor's action. Thus, to account for alternate types of motor actions during ensemble averaging, and to individually quantify the characteristics of the traditional and unexpected motor interactions, each motor event must be individually and objectively characterized.

III. Syntactic Method of Event Characterization

Syntactic method classifies entire events by segmenting the filtered data, obtained using Kalman filtering, into a number of smaller

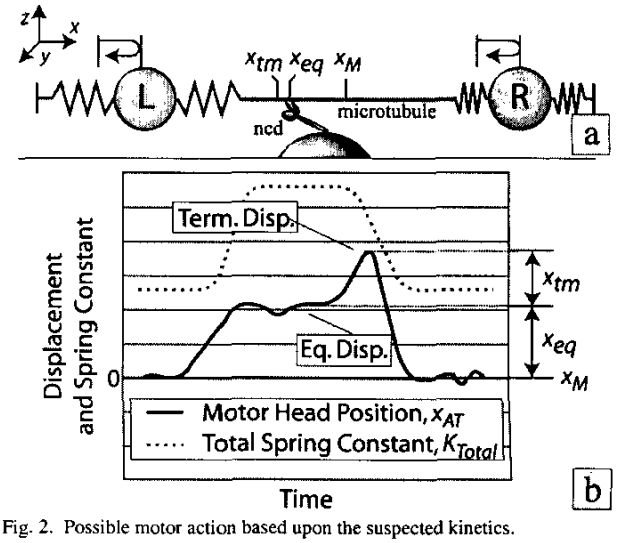

Fig. 2. Possible motor action based upon the suspected kinetics.

Proceedings of the American Control Conference
Denver, Colorado June 4-6, 2003

sections or features that can be more easily categorized [5]. The sequence of these smaller sections may then be used to classify the event into various event types.

Each motor event must first be characterized by determining where within the data set the motor attaches and detaches. These transitions are most easily seen in the total spring constant prediction. Thus the Shewhard control chart was used on the Kalman filter-identified values of K_{total}. The identified K_{total} was smoothed by taking the mean of approximately 1250 data points, which correspond, in real time, to a sliding window of 250 ms. A threshold of approximately 3.75e-5 N/m was used (different data sets required a slightly altered thresholds to account for variations in the link and trap spring constants). When smoothed spring constant crossed the threshold going from lower to higher values, the Ncd-microtubule attachment was identified (the event start time). The instant when K_{total} dropped below the threshold was deemed the event stop time. The described initial segmentation resulted in discrimination of the Ncd-microtubule attachments from the rest of the raw data, collected at 5kHz over the combined experimental time of multiple hours.

Although K_{total} resulted in an adequate event detection at a relatively low computational cost, it was determined, by the type of event ensemble averaging discussed elsewhere in this paper, that detection of attachment and detachment by variance windowing of the bead position gives a greater time resolution (50 milliseconds compared to 200 milliseconds resolution obtained when discrimination is based on the values of K_{total}). Data windowing was used to refine the event start and stop time, with a typical adjustment of approximately 100 milliseconds compared to the initial classification. It should be noted that both methods for detecting motor-microtubule attachments are in a very good agreement.

Once the start and end times are determined, it became necessary to split each event into quantifiable displacements. Visual inspection of detected events revealed that between event start and stop times, the bead may rapidly transition between many equilibrium positions. Thus the next step in classification is to subdivide the event into sections of consecutive displacements. This process starts by determining the rate of change in the predicted motor head position, dx_{AT}/dt, obtained using the polynomial Savitsky-Golay [3] filter. The algorithm determines the derivatives of the Kalman filter-identified x_{AT} by using a progressively fitted polynomial to a window of data. Throughout the event, the temporal positions of the maximum and minimum rates of change in the motor head position were marked off as an indication of the beginning and end of each new equilibrium displacement within the Ncd-microtubule attachment. The threshold on $|dx_{AT}/dt|$ was used to desensitize the algorithm to small variations in dx_{AT}/dt due to measurement noise and thermal fluctuations. In the described manner, each event was divided into a collection of motor head displacements, each beginning and ending with a significant change in the motor head position.

Of particular importance to determination of stroke distance is the motor head displacements, both equilibrium and terminal. Four different cases are depicted in Figure 3. For transition of the type depicted in quadrants 1 and 4 (transition starts and ends with the minimum or maximum $|x_{AT}|$) the segment displacement is calculated as an average x_{AT}. If the displacement begins with a maximum dx_{AT}/dt and ends with a minimum dx_{AT}/dt, as in quadrant 2, then the maximum of x_{AT} is taken as the value of the motor head position assigned to the segment. Similarly, for the case of quadrant 3, the minimum value of x_{AT} is used as the segmental motor head position. Such assignment takes into account the low-pass properties of the Kalman filter used to identify x_{AT}, which implies that the true and unknown maximum (minimum) value of motor head displacement is probably much higher (lower) for the case depicted in the second (third) quadrants of Figure 3. Therefore, the maximum (minimum) values of the identified motor head position can be viewed as an

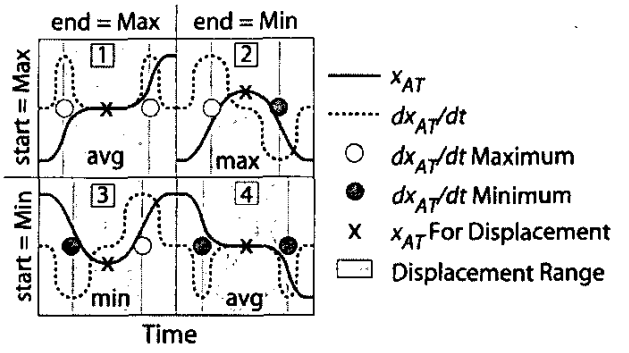

Fig. 3. Determination of the characteristic motor head displacement for different segments of the data.

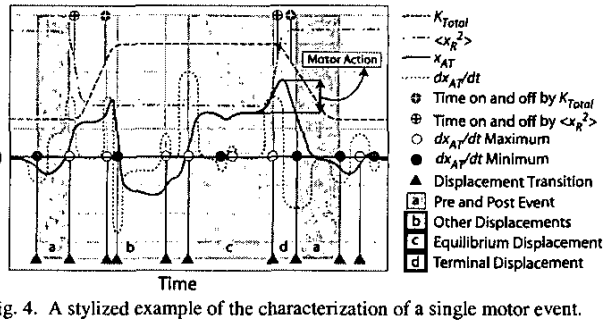

Fig. 4. A stylized example of the characterization of a single motor event.

approximation of an average of the true x_{AT} within the segment. This conclusion was tested and confirmed by applying the developed procedure to the synthetic data, generated by the system of stochastic differential equations, which model the experimental setup depicted in Figure 1. Once the data set had been broken down into events and each event has been broken down into displacements and each displacement had been given a characteristic value of x_{AT}, some sense must be made of each event with regards to the known motor kinetics. Since the motor is expected to perform its stroke and then release from the microtubule, only the terminal equilibrium displacement and the displacement before it, simply named the equilibrium displacement, are used to quantify motor action. Thus, the motor action (pull) is determined to be the difference between the terminal displacement, and the last equilibrium displacement before the motor detaches from the microtubule. The entire process of event characterization is graphically summarized in Figure 4.

IV. Results of Syntactic Characterization

For this analysis, four data sets were used with approximately 1.5 hours of data in total, sampled at the rate of 5kHz. Two were collected at an ATP concentration of 2 μM and two at 5 μM. All data sets were obtained using unique beads, microtubules, and motors, and thus in total at least 4 unique motor-microtubule constructs are involved in this analysis. Because of significant offset and drift in the collected data, 90-second segments of the data are detrended by subtracting a second order polynomial fit. The average length of a motor event is approximately two seconds, thus this detrending should have no effect on the detected motor event characteristics. The results of the event characterization are summarized in Table 1.

Table 1 shows values of the equilibrium displacements, and the terminal motor action. In examining the ratio of positive to negative motor actions, it becomes apparent that the ratio in data set 1 is the inverse of the other sets, indicating the microtubule in this experiment has the opposite experimental orientation from the others. This observation was confirmed by ensemble averaging (discussed

below). Because of this distinction, the sign of raw data in data set 1 is reversed when all events are considered during the syntactic analysis and in the overall ensemble averaging process. In the remainder of the discussion, we will designate the positive direction to be the direction of the conventional motor stroke.

The distribution of equilibrium positions (not shown) is approximately normal and zero-mean centered near the origin, as expected. A slight negative bias in the equilibrium position can be observed from the results of Table 1, which was attributed to negative drift apparent in the raw data, which was not completely removed by the previously described de-trending procedure.

From the distribution of all the terminal motor actions (not shown) it is apparent that motor action towards the minus end of the microtubule (positive bead motion in our case) is predominant. Positive terminal displacements were detected about twice as often as negative (Table 1). The relatively consistent 62% to 32% split of minus-to-plus-end-directed motor action, throughout data sets with different ATP concentrations and microtubule orientations, suggests that we are indeed detecting and characterizing single motor events with a significant degree of accuracy, consistency and objectivity. From the distribution of all terminal motor actions it is also clear that positive terminal events differs significantly from the negative events in both frequency, and length of terminal displacement. Results of Table 1 show that the average of positive terminal motor action is approximately 8.1 nm, while negative terminal action is on average -6.4 nm. The results suggest that the observed negative terminal motor action is not a mirror image of a positive terminal motor action.

Events which are found to end in no discernable motor action (about 6% of the events) are either events in which only one distinct displacement is detected (possibly due to events which occur too quickly for the filter to react), or events in which the difference between the equilibrium and terminal displacement is within the 2 nm threshold used with the filtered results. Due to the reversible nature of the Ncd-microtubule attachment, it is also possible that a motor may attach and detach without completing its kinetic cycle with a mechanical pull.

TABLE I
SYNTACTIC CLASSIFICATION RESULTS

	All	Set 1	Set 2	Set 3	Set 4
ATM (μM)	2, 5	2	5	2	5
Predominant Pull Direction	+	-	+	+	+
Total # of Events	675	48	147	253	227
Positive Equilibrium Displacements	322 (48%)	20 (42%)	64 (44%)	112 (44%)	118 (52%)
Negative Equilibrium Displacements	353 (52%)	28 (58%)	83 (56%)	141 (56%)	109 (48%)
Average Equilibrium Displacements [nm]	-0.8	-1.7	-2.1	-2.5	1.3
Average Positive Equilibrium Displacements [nm]	11.6	13.8	13.4	11.5	10.4
Average Negative Equilibrium Displacements [nm]	-12.13	-12.7	-14.0	-13.5	-8.6
Terminal Motor Actions	634 (94%)	45 (94%)	137 (93%)	246 (97%)	206 (91%)
Positive Terminal Motor Actions	418 (62%)	16 (33%)	89 (61%)	162 (64%)	138 (61%)
Negative Terminal Motor Actions	216 (32%)	29 (60%)	48 (33%)	84 (33%)	68 (30%)
Average Terminal Motor Actions [nm]	3.1	2.4	3.1	3.4	3.0
Average Positive Terminal Motor Actions [nm]	8.1	6.3	8.4	8.6	7.4
Average Negative Terminal Motor Actions [nm]	-6.4	-7.3	-6.5	-6.7	-6.0
Average of Pre-Event [nm]	-1.4	-3.4	-2.6	-2.2	0.7
Average of Post-Event [nm]	-2.2	-0.8	-4.2	-2.1	-1.2

V. RADIAL NEURAL NETWORK METHOD OF EVENT CHARACTERIZATION

To confirm the results obtained with the syntactic method, particularly the existence of previously undetected negative terminal motor stroke, a separate method of event characterization, independent of the Kalman filtering results, was implemented. The alternative method utilized the radial basis neural network classification to segregate each detected event into different categories based on unfiltered, raw experimental data.

The neural network was designed and trained using synthetic data with known molecular motor events, generated from stochastic modeling of the bead-microtubule interactions using Langevin-like equations with additive Brownian perturbations with stochastic characteristics selected to mimic the experimental data. All synthetic events adhere to the following mechanism: motor-microtubule attachment $\rightarrow$ attainment of an equilibrium position $\rightarrow$ motor stroke $\rightarrow$ motor release. Each event was separated by 3 seconds and lasted for 1.8 seconds in total, which does not bias the training set since only the end of each event is used in the neural network training. Equilibrium displacements were chosen to be a random distance for each event between 2 and 20 nm for positive displacements, between –20 and –2nm for negative displacements, and between –2 and 2 nm for zero equilibrium displacements. Equilibrium displacements lasted for an amount of time equal to the event time minus the time of the terminal displacement. A randomly selected terminal displacement was set in the range of 2 to 16 nm with random duration between 100 to 300 ms. Furthermore, a linear displacement from the equilibrium motor displacement to the terminal displacement was found to be more effective in training the NN than a step in displacement. Negative terminal motor displacements were chosen to mimic those observed during synthetic classification. The stroke distanced was randomly chosen to be between –2 and –16 nm. These events were set to occur during the last 100 to 400 ms of the event with the peak negative displacement occurring at the mid point of the terminal displacement.

A set of training data was created containing 93 events of each possible combination of positive, zero, and negative equilibrium and terminal displacements (9 categories in total). These data sets were then used to create and train two separate neural networks, one to categorize equilibrium displacements and the other to categorize terminal displacements. The creation of each neural network was accomplished by repeatedly adding layers of neurons to the system and then adjusting the weights until a zero classification error was achieved. The end result was two radial neural networks with approximately 250 neurons each. A segment of each event starting 1 second before and ending 0.25 seconds after motor-microtubule detachment was used in the NN training. When applied to verification data set of synthetic events not involved in the training of the neural network, categorization was found to be predominantly successful, with approximately 5% of synthetic events incorrectly categorized.

VI. NEURAL NETWORK CLASSIFICATION RESULTS

When applied to experimental data, for 76% of all events the

neural network categorized equilibrium displacements into the same categories as the syntactic classification method. The vast majority of discrepancies occurred with equilibrium displacements of within 10 nm of the origin. Overall, 49% of events were determined to have *positive equilibrium* displacements by the neural network classification. A higher disagreement was observed in the classification results of the terminal displacements, with 65% of events classified into the same category. This lower degree of agreement is expected, since terminal displacements are classified using only short 200 ms segments, compared to the up to 2 second segments used to classify equilibrium displacements. Furthermore, the more quantitative syntactic method allowed for a third category of events with no terminal motor action, and 6% of all events were determined to have no discernible terminal motor action (Table 1). With NN classification, these events are always assigned either positive or negative terminal displacement, contributing to the higher overall disagreement. Neural networks classified 78% of experimental motor events as having positive terminal motor action, confirming the prevalence of positive terminal motor action. When the negative of the same experimental data is sent into the neural network, the percent of positive terminal motor action is 30%, confirming the network's ability to discriminate between positive and negative terminal motor action.

Overall, both classification methods confirmed the existence of negative terminal motor actions. Both methods produced consistent results, though the results of the syntactic classification are more quantitative (summarized in Table 1), and are considered to be more reliable since neural network results were based on unfiltered data with a low signal-to-noise ratio.

VII. Ensemble Averaging

A. Procedure

The primary purpose of introducing single-realization filtering and analysis is to attempt resolving individual motor events and to observe events that may be obscured by ensemble averaging. Once motor events are characterized and segregated into different categories based upon the syntactic and neural network classification methods, ensemble averaging can once again be used to reduce the effect of the thermal fluctuations, and confirm and analyze the unexpected event types by averaging different groups separately. Within each group, the following ensemble averaging procedure was used. Since each event may have a different duration, depending upon the time of arrival and reaction of ATP molecule with Ncd, the events were time normalized with only data adjacent to the event start and stop time retained for the averaging process. After aligning the retained data by the event start and stop times, ensemble averaging was carried out separately for each category of events.

B. Syntactic Results

When all motor events are averaged together indiscriminately, the results may adequately characterize the predominant motor action, but less frequent events will be obscured by averaging. Our single event analysis suggests that, indeed, multiple types of events are possible during Ncd-microtubule interaction. Therefore, to reveal the details of different types of events by averaging out the effect of thermal fluctuations, only ensemble averaging within the same group of events is undertaken. The role of model-based, single realization filtering and subsequent analysis in interpreting experimental data is twofold. The single-realization Kalman filtering and identification results indicate what different types of molecular motor events are present, as well as the salient features of the particular event type. This information forms the basis for segmentation of the raw data or filtered results. The segregated Kalman filtering data can then be ensemble averaged to obtain quantitative characteristics of molecular

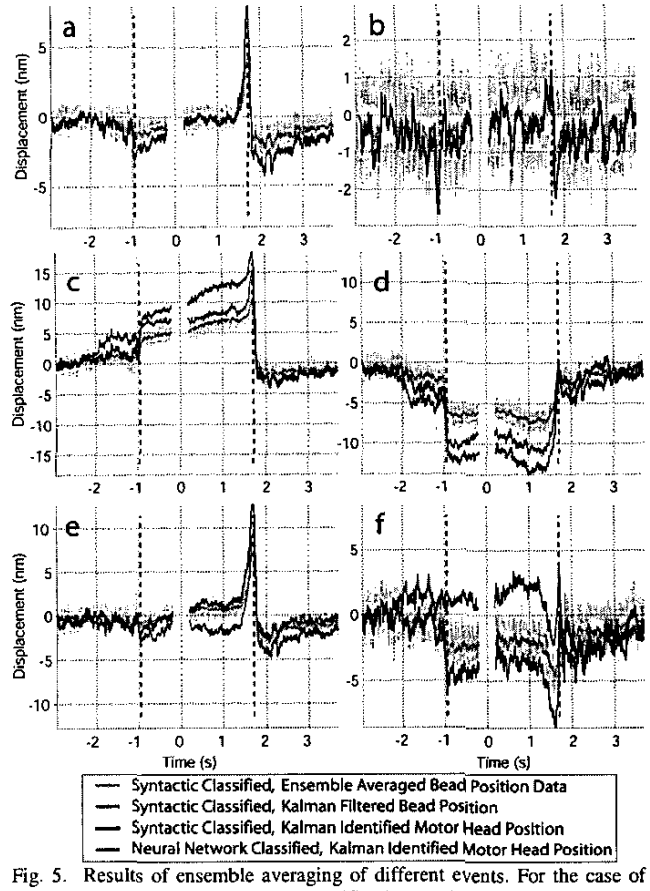

Fig. 5. Results of ensemble averaging of different events. For the case of events segregated using syntactic classification method: The noisy lightest gray line in the background is the ensemble average of the raw data; The next shade of gray (Green) is the Kalman estimated bead displacement which tracks the ensemble of the raw data; Black (Blue) is the estimated motor head displacement based on the ensemble average. For the case of the neural network classification: The darkest Gray (Orange) is the estimated motor head displacement. Dashed vertical lines indicate event start and stop times. a) Results for all 619 events; b) All 619 events in the y-axis; c) All 292 events with a positive equilibrium displacement (333 events for the neural network classification average); d) All 328 events with a negative equilibrium displacement (286 events for the neural network classification average); e) All 393 events with a positive terminal motor action (414 events for the neural network classification average); f) All 196 events with a negative terminal motor action (205 events for the neural network classification average).

interactions as a group. However, double averaging – first in time domain using Kalman filtering, followed by ensemble averaging – produce overly conservative results, which are reported in Table 1.

An alternative procedure is to use ensemble averaging of the raw measurement data pre-classified based on syntactic classification of the Kalman filtering results, or direct classification of the raw data using neural network classification. The results of the direct ensemble averaging can be used as an input to an aggressively tuned (high-bandwidth) Kalman filter to identify motor head displacements x_{AT}. A higher bandwidth Kalman filter can be used with the ensemble averaged experimental measurements because of the reduced effect of thermal fluctuations and therefore higher signal-to-noise ratio of about 2, compared to 0.1 for the filter with the raw measurements as an input.

The results depicted in Figure 5 were obtained following such an alternative approach, and are deemed to be more reflective of the underlying Ncd-microtubule interactions compared to the results of Table 1, obtained with double averaging. Figure 5 shows a collection

of ensemble averages of various types for the events segregated using syntactic classification. The dashed horizontal lines indicate the beginning and end of events as detected by the Kalman identified total spring constant K_{total} and corrected by variance windowing. The result of the ensemble averaged bead position was then filtered with the aggressive Kalman filter to identify the motor head displacements. The Kalman filtered right bead position is also shown in the graphs.

Figure 5a shows the average of the right bead position for all detected events in the x-axis. The event terminates in an 8 nm positive motor head displacement, which is approximately the length of a tubulin unit. This displacement notably peaks at essentially the same time as the end of the event, suggesting an immediate Ncd-microtubule detachment after motor action, which agrees with the previously proposed kinetic model. Note an approximately zero average equilibrium displacement just prior to the motor action, which is expected since both positive and negative equilibrium displacements are equally likely. No interpretation is given to the slight negative displacement immediately following the Ncd-microtubule detachment.

Figure 5b shows the average of all detected events in the y-axis, the axis perpendicular to the microtubule. The y-axis component of the motor actions is close to the noise level when all events are averaged together. A slight, approximately 1.75 nm displacement occurs both at the initiation and termination of the average motor events. Such y-axis displacements may be attributed to imprecise alignment in the x-z plane of the motor and the microtubule. More specifically, assuming that the microtubule lies in the x-z plane, the motor attachment to the immobilized bead can be located on either side of this plane. Therefore, depending on relative motor-microtubule orientation in the x-z plane, the bead displacement due to motor attachment may have a y-component, causing small displacements in the y-direction.

Figure 5c and 5d, shows the results of separate ensemble averaging for positive and negative equilibrium displacements, respectively. Both groups may include events with positive and negative terminal motor actions. In the case of positive equilibrium displacements, we see an averaged equilibrium displacement for the motor head of approximately 11 nm followed by a terminal motor action of approximately 7 nm, followed by a quick return to the origin. In the case of the negative equilibrium displacements, Figure 5d, we find an equilibrium displacement of approximately -12 nm followed by a return to the origin. The direction in which the measurement bead moves to return to an unattached equilibrium position is the same as the direction of the predominant (positive) motor action, which explains why terminal motor action is obscured in this case.

Figures 5e and f show ensemble averaging results for events with terminal motor actions in the positive and negative directions, respectively. A 12 nm positive stroke is obtained for the average positive terminal displacements, significantly larger than a tubulin length, and the previously reported results. Comparison of Figures 5a and 5e shows that 4 nm of terminal motor action is lost when all events, positive and negative, are included in the ensemble averaging. Furthermore, with the single event analysis summarized in Table 1, the average positive motor pull is 8 nm, substantially smaller than the observed 12 nm in Figure 5e. This difference is most likely due to the described double averaging inherent in the procedure used to generate the results of Table 1.

The events with negative terminal motor actions, Figure 5f, are clearly not mirror images of the events terminating in the positive motor stroke, but rather reveal significantly different kinetics. Firstly, the average size of the negative motor head displacement is much smaller (approximately 5 nm). The peak in the average displacement notably occurs approximately 100 milliseconds from the motor-microtubule release, whereas positive terminal motor stroke occurs essentially simultaneously with the detachment. This suggests that the end of the motor action in the negative direction does not coincide with the Ncd-microtubule detachment. Finally this negative stroke takes approximately 600 milliseconds to complete, about twice the time it takes for the positive motor stroke, indicating either a slower stroke or a stroke with a broader distribution of the Ncd-microtubule detachment after the motor action. The totality of observations indicates that negative terminal displacements are caused by motor action different from the previously reported typical motor stroke.

C. *Neural Network Results*

To confirm the ensemble averaging results of events segregated with the syntactic classification method, the same ensemble averaging procedure as described above was repeated using the neural network classifications. The resulting averages were found to follow those found by the syntactic method. When an aggressively tuned Kalman filter was applied to the groups of events classified using neural networks, the estimated motor strokes were consistent with the results obtained with syntactic classification, but with 1 to 3 nm shorter average motor strokes for different cases considered in Figure 5.

VIII. Discussion

The classification procedure described in this paper is a step in the sequence of the model-based approach to the interpretation of molecular interactions on a single molecule basis, which starts with (a) Model development of the molecular interaction, followed by (b) Model-based, time-resolved filtering of experimental data; (c) Identification of different types of molecular events suggested by the time-resolved analysis; (d) Automatic classification of the data and filtering results into different types of events; (e) Separate ensemble averaging of different events, and the analysis of the results. This process culminates in (f) The refinement of the original model to account for improved understanding of the molecular interactions, and direct experimental verification of the results produced by the described analysis. When steps (a)-(e) are applied to the experimental data on the interaction between motor protein Ncd with the microtubule, the results deviate from the traditional view on the kinetics of the Ncd-microtubule interaction, suggesting that Ncd-generated mechanical strokes in both directions takes place. Direct experimental confirmation is desirable, which may require the development of the experimental optical trap assay equipped with the real-time feedback control system capable of automatic recognition of molecular events, and the corresponding adjustment of the optical trap operation.

References

[1] A.E. Butterfield, R.J. Stewart, C. F. Schmidt and M. Skliar, "Model Based Filtering in the Analysis of Single Ncd Motor Action", *Biophysical Journal*, submitted., 2003.

[2] M. J. deCastro, R. M. Fondecave, L. A. Clarke, C. F. Schmidt, and R. J. Stewart, "Working Strokes by Single Molecules of the Kinesin-Related Microtubule Motor Ncd," *Nature Cell Biology*, 2:724-729, 2000.

[3] A. Savitzky and M. J. E. Golay, "Smoothing and Differentiation of Data by Simplified Least Squares Procedures," *Analytical Chemistry*, **36**:1627-1639, 1964

[4] R. J. Stewart, J. P. Thaler, et al. "Direction of Microtubule Movement is an Intrinsic Property of the Motor Domains of Kinesin Heavy Chain and Drosophilia NCD Protein," *Proceedings of the National Academy of Science*, **90**: 5209-5213, 1993.

[5] T. Y. Young, *Handbook of Pattern Recognition and Imaging Processing*. New York: Academic Press, 1986.