# Quantification of microtubule stutters: dynamic instability behaviors that are strongly associated with catastrophe

Shant M. Mahserejian[a,b,†,*], Jared P. Scripture[c,†], Ava J. Mauro[c,d,*], Elizabeth J. Lawrence[e], Erin M. Jonasson[c,f], Kristopher S. Murray[c], Jun Li[a], Melissa Gardner[g], Mark Alber[a,h], Marija Zanic[e,i,j], and Holly V. Goodson[c,*]

[a]Department of Applied and Computational Mathematics and Statistics and [c]Department of Chemistry and Biochemistry, University of Notre Dame, Notre Dame, IN 46556; [b]Pacific Northwest National Laboratory, Richland, WA 99352; [d]Department of Mathematics and Statistics, University of Massachusetts Amherst, Amherst, MA 01003; [e]Department of Cell and Developmental Biology, Vanderbilt University, Nashville, TN 37240; [f]Department of Natural Sciences, Saint Martin's University, Lacey, WA 98503; [g]Department of Genetics, Cell Biology, and Development, University of Minnesota, Minneapolis, MN 55455; [h]Department of Mathematics, University of California, Riverside, Riverside, CA 92521; [i]Department of Chemical and Biomolecular Engineering, Vanderbilt University, Nashville, TN 37235; [j]Department of Biochemistry, Vanderbilt University, Nashville, TN 37205

**ABSTRACT** **Microtubules (MTs) are cytoskeletal fibers that undergo dynamic instability (DI), a remarkable process involving phases of growth and shortening separated by stochastic transitions called catastrophe and rescue. Dissecting DI mechanism(s) requires first characterizing and quantifying these dynamics, a subjective process that often ignores complexity in MT behavior. We present a Statistical Tool for Automated Dynamic Instability Analysis (STADIA) that identifies and quantifies not only growth and shortening, but also a category of intermediate behaviors that we term "stutters." During stutters, the rate of MT length change tends to be smaller in magnitude than during typical growth or shortening phases. Quantifying stutters and other behaviors with STADIA demonstrates that stutters precede most catastrophes in our in vitro experiments and dimer-scale MT simulations, suggesting that stutters are mechanistically involved in catastrophes. Related to this idea, we show that the anticatastrophe factor CLASP2γ works by promoting the return of stuttering MTs to growth. STADIA enables more comprehensive and data-driven analysis of MT dynamics compared with previous methods. The treatment of stutters as distinct and quantifiable DI behaviors provides new opportunities for analyzing mechanisms of MT dynamics and their regulation by binding proteins.**

**Monitoring Editor**
Claire Walczak
Indiana University

Received: Jun 2, 2020
Revised: Nov 1, 2021
Accepted: Jan 5, 2022

This article was published online ahead of print in MBoC in Press (http://www.molbiolcell.org/cgi/doi/10.1091/mbc.E20-06-0348) on February 2, 2022.

*Address correspondence to: Shant M. Mahserejian (shant.mahserejian@pnnl.gov), Ava J. Mauro (avamauro@bu.edu), Holly V. Goodson (hgoodson@nd.edu).
†Co–first authors.

Abbreviations used: DI, dynamic instability; fps, frames per second; GDP, guanosine diphosphate; GTP, guanosine triphosphate; k, number of clusters; MT, microtubule; MTBP, microtubule binding protein; PF, protofilament; STADIA, Statistical Tool for Automated Dynamic Instability Analysis; TIRF, total internal reflection fluorescence.



## 1. INTRODUCTION

Microtubules (MTs) are protein-based biological polymers that have a central role in fundamental eukaryotic processes including cellular organization, chromosome separation during cell division, and intracellular transport (Goodson and Jonasson, 2018). Crucial to the function of MTs in these processes is a well-known behavior termed dynamic instability (DI), where the polymers switch stochastically between periods of growth and shortening as seen in traditional MT length-history plots (Figure 1, A and B) (Mitchison and Kirschner, 1984; Desai and Mitchison, 1997). Accurate quantification of MT DI behavior is essential for understanding its significance and mechanism and also for investigating the activities of DI-regulating proteins and pharmaceutical agents (e.g., chemotherapy drugs, fungicides).

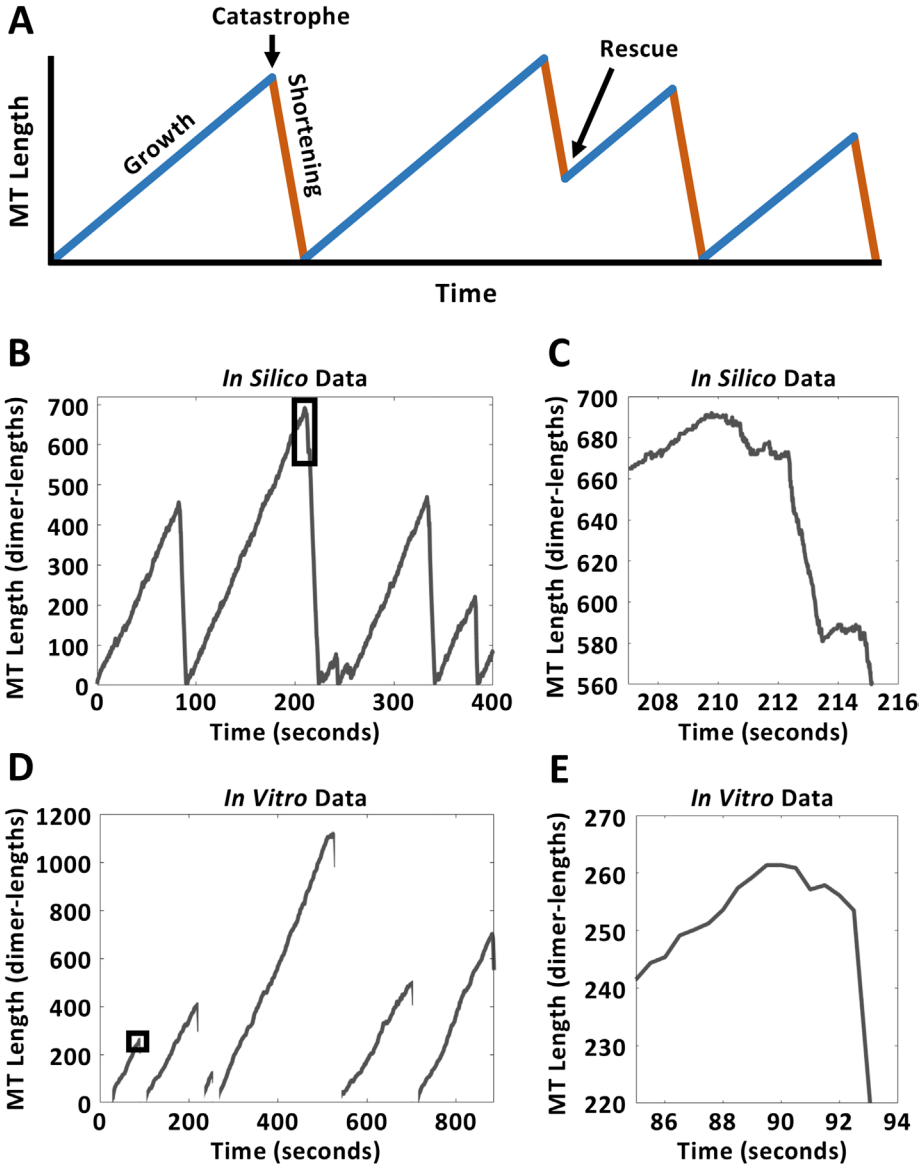


FIGURE 1: Qualitative examples of MT behaviors that do not fit the two-state (growth-shortening) framework. (A) An illustration of the classically recognized two-state representation of dynamic instability (DI), in which behavior is classified as either growth or shortening phases, with instantaneous transitions known as catastrophe and rescue events. (B, D) Zoomed-out length-history plots of simulation data (B, dimer-scale 13-protofilament model, temporal resolution of ~1650 dimer-scale events per second per MT, *Materials and Methods* Section 5.2) and experimental data (D, temporal resolution of 2 frames per second, note that depolymerizations were not tracked in their entirety in these experiments, *Materials and Methods* Section 5.1). Black rectangles in B and D indicate the zoomed-in portions shown in C and E, respectively. (C, E) Closer inspection of transitions shows ambiguous behavior that cannot clearly be categorized as either growth or shortening.

### 1.1. Traditional DI measurements

Traditionally, MTs have been treated as two-state polymers; that is, MTs have been considered to be either growing or shortening, with abrupt, instantaneous transitions called catastrophes and rescues between these two phases (Figure 1, A, B, and D). In this framework, MT behavior is characterized by four quantities called DI parameters (Walker *et al.*, 1988):

- $V_{growth}$—velocity of growth, commonly measured as the mean of the growth rates as averaged over the set of growth phases
- $V_{short}$—velocity of shortening, commonly measured as the mean of the shortening rates as averaged over the set of shortening phases
- $F_{cat}$—frequency of catastrophe, commonly measured as the number of catastrophes (transitions from growth to shortening) per time in growth
- $F_{res}$—frequency of rescue, commonly measured as the number of rescues (transitions from shortening to growth) per time in shortening

The specific procedures for measuring these DI parameters have varied among research groups, but methods typically begin with the user specifying the minimal values (i.e., thresholds) of length change, time duration, and/or velocity required for recognizing phases of growth and shortening; sometimes pauses are also allowed, as discussed more below in Section 1.2. Then the length-history plot is partitioned into growth and shortening segments (Figures 1A and 2, A and B). The endpoints of the segments are assumed to correspond to the events of catastrophe and rescue, and the slopes of the segments provide the growth or shortening velocities. In other words, the velocity of an individual growth or shortening phase is typically determined as the slope of a line drawn between the points of catastrophe and rescue (e.g., Zanic, 2016).

### 1.2. Limitations of common methods for quantifying dynamic instability

While determination of DI parameters as described above is a standard way to quantify MT behavior (see, e.g., Portran *et al.*, 2017; Zwetsloot *et al.*, 2018; Kapoor *et al.*, 2019, for recent examples), there are aspects of MT behavior that are not captured using this approach. First, it has long been recognized that both growth and shortening rates are variable. This variability occurs both with and without MT binding proteins (MTBPs), and it is observed both within and between individual growth phases and similarly for shortening phases (e.g., Gildersleeve *et al.*, 1992; Pedigo and Williams, 2002; Schek *et al.*, 2007; Lawrence *et al.*, 2018). Spectral analysis of such variability in growth and shortening rates has suggested that the two-state (growth and shortening) model approximation agrees well with experimentally observed MT behavior when frequencies in the length-history data are analyzed at timescales longer than ~1 min but underestimates the observed variability at timescales shorter than ~1 min (Odde *et al.*, 1996). These observations raise the concern that DI analysis methods that categorize an entire period between nucleation (or rescue) and catastrophe as a single growth or shortening phase could cause functionally significant details of MT behavior to be missed.

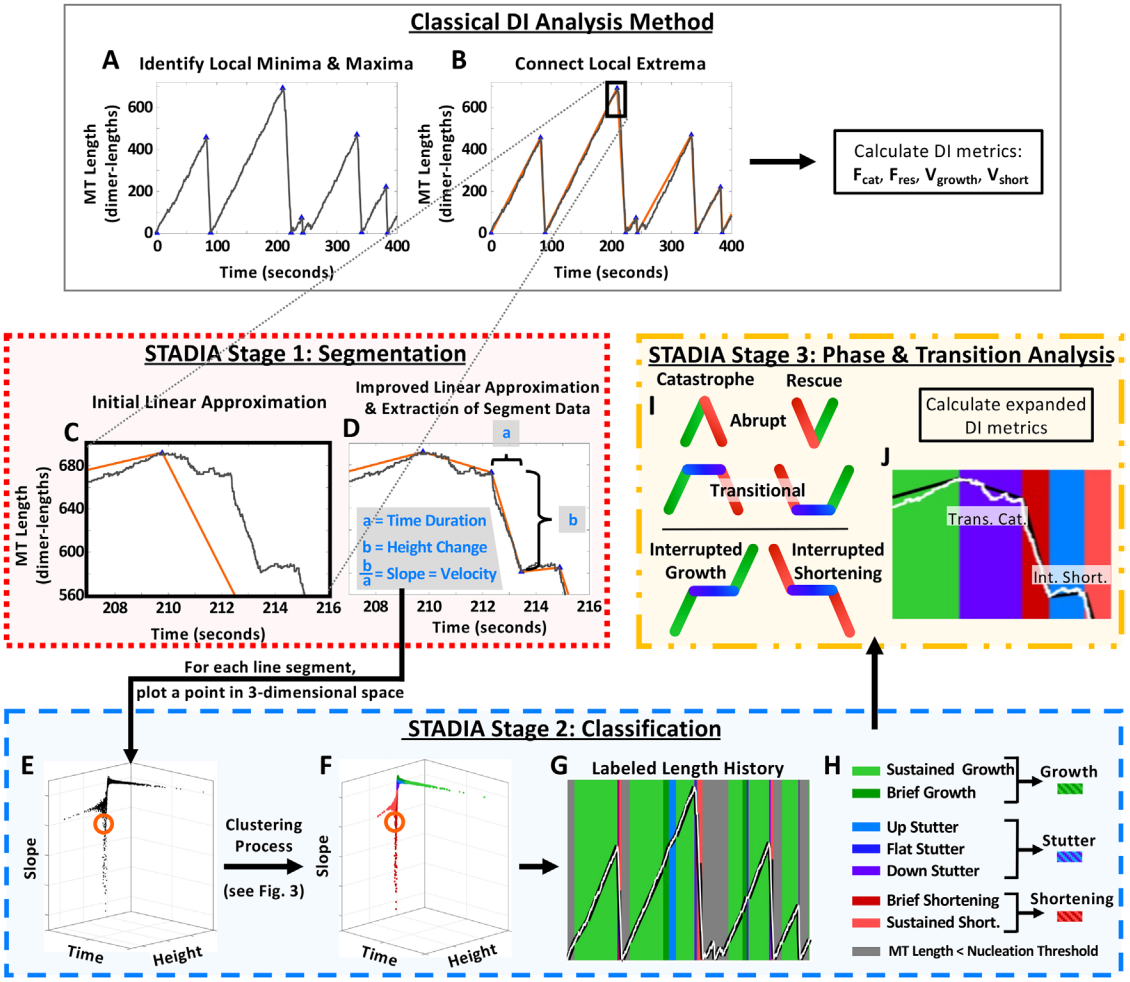


FIGURE 2: Comparison of classical DI analysis method and STADIA. Classical DI analysis method (A, B): Major peaks and valleys (blue triangles) are first identified (A), and these are defined as catastrophes and rescues (or nucleation events), respectively. Each period from a nucleation event or rescue to a catastrophe is defined as a growth phase, and each period from a catastrophe to a complete depolymerization or rescue is defined to be a shortening phase (B). Then, $V_{growth}$ and $V_{short}$ are calculated from the slopes of straight line segments plotted between the transitions (B, orange lines) or alternatively, from regression lines fitted to the data points in each period. $F_{cat}$ and $F_{res}$ are calculated from the number of catastrophes or rescues divided by the total time in growth or shortening, respectively. STADIA (C–J): An initial approximation of inputted length-history data is produced by connecting major peaks and valleys with line segments (C, similar to classical methods). STADIA then iteratively adds segment endpoints to improve the approximation (D, this iterative process is regulated by user-defined parameters Maximum Error Tolerance and Minimum Segment Duration). The time duration, height change, and slope (velocity) of each line segment are measured (D) and visualized as a point in 3-dimensional space (E). The orange circles in (E, F) denote the approximate location of the data point corresponding the example line segment in (D). The line segments are grouped into "clusters" (see Figure 3), as indicated by the colors in the plots (F, G, J; key in left column of H). The clusters are named and grouped into larger behavior classes based on their average features (e.g., average slope) (H). STADIA then identifies multiple types of transitions (I), allowing the calculation of various metrics including (and expanding beyond) the traditional $V_{growth}$, $V_{short}$, $F_{cat}$, and $F_{res}$. In (G, J), the white lines represent the raw length-history data, and the black lines represent the line segment approximation. Results Section 2.2 contains a more thorough overview of STADIA's analysis procedure, and full details are provided in *Materials and Methods* Sections 5.4–5.6.

Second, pauses, attenuation phases, and other intermediate states have been observed in experiments and proposed in models, but the way these behaviors have been identified and defined has varied. Pauses are commonly observed in vivo (e.g., Sammak and Borisy, 1988; Schulze and Kirschner, 1988; Waterman-Storer and Salmon, 1997; Gierke *et al.*, 2010; Kamath *et al.*, 2010; Applegate *et al.*, 2011). Pauses have also been observed in vitro in the presence of MTBPs (e.g., Moriwaki and Goshima, 2016), cell extracts (e.g., Keller *et al.*, 2008), and drugs (e.g., Toso *et al.*, 1993), and occasionally for purified tubulin (e.g., Walker *et al.*, 1988). Recognition of states other than growth and shortening has led various authors to consider theoretical three- or four-state models in which the additional states are pauses or an intermediate state (Odde *et al.*, 1995; Tran *et al.*, 1997; Jánosi *et al.*, 2002; Maly, 2002; Keller *et al.*, 2008; Smal *et al.*, 2010; Blackwell *et al.*, 2017). Thus, it is clear that many researchers are interested in methods for identifying states beyond growth and shortening in data and the inclusion of such states in the development of theory. However,

there is not a general consensus on how these states should be defined.

In particular, as noted in Section 1.1, identification of growth, shortening, and pause phases in length-history data frequently relies on fixed thresholds for velocity, length change, and/or time duration. For example, it has been common to require a length-change threshold of at least 0.5 microns to recognize a growth or shortening phase, but the exact way in which this threshold was applied to data has varied among research groups (e.g., compare Sammak and Borisy, 1988; Dhamodharan and Wadsworth, 1995; Rusan *et al.*, 2001; Kamath *et al.*, 2010; and Fees *et al.*, 2017). Others have used combinations of thresholds on the speed of length change (e.g., in pixels per frame or microns per minute), length change itself, and/or number of data points involved (e.g., compare Panda *et al.*, 1996; Gierke *et al.*, 2010; Kiris *et al.*, 2010; Matov *et al.*, 2010; Yenjerla *et al.*, 2010; Mahrooghy *et al.*, 2015; and Moriwaki and Goshima, 2016). It is important to be aware that thresholds have differed between analyses, because it is well-established (but perhaps not widely recognized) that thresholds can have dramatic effects on measurements of MT dynamics (e.g., Odde *et al.*,1996; Gierke *et al.*, 2010; Matov *et al.*, 2010; Smal *et al.*, 2010; Prahl *et al.*, 2014; Guo *et al.*, 2018).

Finally, recent improvements in imaging technology have enabled acquisition of MT DI data with both high temporal and high spatial resolution, which allows for the possibility of analyzing length-history data at finer scales (e.g., Maurer *et al.*, 2014; Andrecka *et al.*, 2016; Mickolajczyk *et al.*, 2019). These data have verified the intrinsic variability of MT behavior. They have also demonstrated that both growth and shortening phases can include significant time periods (e.g., a few seconds in duration or longer) during which the growth or shortening velocity slows significantly (Figure 1, C and E; see also Maurer *et al.*, 2014; Duellberg *et al.*, 2016a,b; Rickman *et al.*, 2017). These slowdown periods likely overlap with pauses discussed above, though it is important to note that "bona fide" pauses are often considered to be time periods "during which no polymerization or depolymerization occurs" (Gierke *et al.*, 2010) and so are separable from periods of slowed growth or shortening, at least in principle.

Significantly, these slowdown periods can also occur in association with catastrophe (Maurer *et al.*, 2014; Duellberg *et al.*, 2016a,b; see also predictions based on simulations in Margolin *et al.*, 2012), making it difficult to determine with reasonable precision where transitions between phases begin and end. To illustrate this problem, consider the zoomed-out length-history plots that are typically used for DI analysis (Figure 1, B and D). Examination of these plots can make the task of determining when transitions occur look trivial. However, the zoomed-in views made possible by high-resolution data acquisition (Figure 1, C and E) demonstrate the difficulty of identifying the points of transition and/or categorizing DI behaviors.

Thus, many researchers have recognized that MT DI behavior is more complex than a simple two-state system of growth and shortening with abrupt transitions. The four traditional DI parameters ($V_{growth}$, $V_{short}$, $F_{cat}$, and $F_{res}$) would be sufficient to quantify such a two-state system but are not sufficient to quantify all aspects of observed MT DI as discussed above. One previous approach to dealing with the existence of slowdown periods has been to exclude them from quantification of DI parameters, because including these slowdown periods in either growth or shortening phases would reduce the magnitude of measured values of $V_{growth}$ and $V_{short}$ (e.g., Rickman *et al.*, 2017). However, entirely excluding these behaviors from analysis could result in the loss of information critical for understanding the mechanisms of phase transitions or their regulation by MTBPs. Furthermore, although previous publications have quantified some aspects of the slowdown periods (e.g., time durations [Maurer *et al.,* 2014]), none of these to our knowledge have presented a set of velocities and transition frequencies that expands beyond the traditional four DI parameters. Capturing and quantifying behaviors in addition to growth and shortening would be a key step toward further dissecting the recognized variations in growth and shortening rates, improving the precision of DI metrics, and elucidating mechanisms of DI.

To study MT dynamics more comprehensively than is possible with standard DI approaches, we developed the Statistical Tool for Automated Dynamic Instability Analysis (STADIA), an automated tool that uses established statistical methods to characterize and quantify MT behavior without prior assumptions about the number or characteristics of the behaviors detected. As shown in the *Results* below, STADIA can be used with both simulation- and experiment-generated data, and it is compatible with a wide range of data acquisition rates.

### 1.3. Summary of conclusions

Applying STADIA to in silico and in vitro MT length-history data demonstrated the prevalence of a category of intermediate behaviors that we propose calling "stutters." Stutters share similar characteristics with each other and are distinguishable from typical growth and shortening. The primary distinguishing factor is that during stutters the overall rate of change in MT length is markedly smaller in magnitude compared with the velocities of classically recognized growth and shortening phases. Stutters are also distinguishable from pauses in that during true pauses "no polymerization or depolymerization occurs" (Gierke *et al.*, 2010). In contrast, during stutters dimer-scale dynamics continue, and during most stutters measurable length changes do occur although at slower velocities than during typical growth and shortening. Stutters, as recognized and quantified by STADIA, overlap with previously observed behaviors such as precatastrophe slowdowns (e.g., Maurer *et al.*, 2014) and events that have been called "pauses" despite length changes occurring (e.g., Kamath *et al.,* 2010; Matov *et al.,* 2010; Guo *et al.,* 2018). The relationship of our results to previous work is further covered in *Discussion* Section 3.3.

Analysis of length-history data using STADIA leads to two major observations regarding the relationship between stutters and catastrophes:

- Stutters precede most catastrophes in our in vitro control and in silico data sets.
- The MT stabilizing protein CLASP2γ reduces catastrophe in vitro by increasing the fraction of stutters that return to growth rather than entering shortening phases. Specifically, CLASP2γ reduces the frequency of growth-to-stutter-to-shortening (which we term transitional catastrophe) and increases the frequency of growth-to-stutter-to-growth (which we term interrupted growth).

These results indicate that STADIA is able to recognize and quantify behaviors that are missed by classical methods of analyzing MT length-history data. Furthermore, these results suggest that stutters play a mechanistically significant role in the process of catastrophe. We conclude that identification of stutters as distinct from growth, shortening, or pause warrants their future inclusion in DI analyses and serves as a necessary step forward in gaining a better understanding of MTs, their dynamics, and their regulation by MTBPs.

## 2. RESULTS

For ease of navigation and to allow readers to focus on the information that is most relevant to them, we have divided the Results below into six sections. Section 2.1 introduces the in vitro and in silico data sets used in this work. Section 2.2 provides a general overview of our new tool, STADIA. Sections 2.3, 2.4, and 2.5 present the results of using STADIA to analyze our data sets. More specifically, in Section 2.3, we use STADIA to identify and characterize MT behaviors, including a category of intermediate behaviors that we term "stutters." In Section 2.4, we use STADIA to quantify characteristics of the behaviors identified in Section 2.3. This quantification includes studying the relationship between stutters and phase transitions, which shows that stutters are strongly associated with catastrophe. In Section 2.5, we further test the functional significance of stutters in catastrophe and demonstrate the utility of STADIA in studying MTBPs. More specifically, we use STADIA to analyze the dynamics of in vitro MTs growing in the presence of the anticatastrophe factor CLASP2γ, thus examining for the first time its effect on stutters. In Section 2.6, we test the effects of varying the values of STADIA's input parameters and demonstrate the robustness of the conclusions drawn in Sections 2.3–2.5.

### 2.1. Data sets used in this work: in vitro and in silico

In the analysis below (Sections 2.3–2.6), we used STADIA to analyze length-history data sourced from both laboratory experiments (in vitro) and computational simulations (in silico). We provide a brief overview of the data sets here, with additional information in *Materials and Methods* Sections 5.1 and 5.2.

We analyzed two in vitro data sets: a control with purified tubulin + EB1 and a treatment data set with purified tubulin + EB1 + CLASP2γ. The data sets were obtained using total internal reflection fluorescence (TIRF) microscopy with images taken at 2 frames per second (fps). A subset of the experimental data used here was previously analyzed using other methods (in Lawrence *et al*., 2018). The in vitro data sets enabled us to test STADIA on data from physical experiments and to test STADIA's utility in analyzing the effects of a MTBP on DI behavior.

The in silico data set was obtained using our dimer-scale 13-protofilament (PF) kinetic Monte Carlo model of MT dynamics (Margolin *et al*., 2012). The model simulates attachment/detachment of tubulin dimers to/from PFs, formation/breaking of lateral bonds between dimers in neighboring PFs, and hydrolysis converting GTP-bound dimers to GDP-bound dimers. The values of kinetic rate constants governing these biochemical events are input by the user. The observed DI behavior is an emergent property that arises as a consequence of the dimer-scale events. The input parameters for the model were tuned based on experimental measurements from Walker *et al*. (1988). The MT length-history data outputted by the simulation have spatial resolution at the scale of individual tubulin dimers (8 nm in length) and temporal resolution at the scale of the biochemical events described above (>1000 events per second per MT for the parameters used here).

Including the in silico data in our analysis is useful for several reasons. First, our in silico data allow us to test STADIA on a data set that has quantitatively different DI behavior as compared with the in vitro data sets. Note that the in silico data are not intended to replicate any numerical values from the in vitro data sets used here. Rather, using STADIA on quantitatively different data sets provides a test of the generality of the qualitative conclusions that we draw. Further, the detection of stutters in the in silico data demonstrates that stutters can arise as an emergent property of the dimer-scale biochemical events described above. Additionally, because the in silico data are recorded at the scale of the addition and loss of individual tubulin dimers, the in silico data have higher resolution than is currently possible in in vitro experiments. Therefore, comparison of the in silico and in vitro data sets demonstrates that STADIA is able to process data from a wide range of spatial and temporal resolutions. Relatedly, the high temporal resolution makes the in silico data set ideal for testing the robustness of STADIA to changes in data acquisition rates, because the full-resolution in silico data can be compared with data with imposed slower acquisition rates (Section 2.6).

### 2.2. STADIA: a novel tool for characterizing and quantifying MT dynamics

***2.2.1. Goals of STADIA.*** To meet the goal of identifying, categorizing, and quantifying the range of MT behaviors in length-history data more precisely than with previous methods, we created STADIA. Specific aims for the development of STADIA were that it have the following attributes: 1) Automated to create a consistent and reproducible method with minimal user input; 2) Impartial such that it does not presuppose that MT dynamics are restricted to two states (i.e., limited to growth and shortening); 3) Adaptive to handle data from systems with qualitatively and quantitatively different DI behaviors (e.g., different types of tubulin and/or the presence of MTBPs); 4) Compatible with classical DI analysis, enabling comparison to and continuity with previous work; 5) Capable of analyzing data from a range of spatial and temporal resolutions (e.g., from computational simulations or laboratory experiments). The features of STADIA that collectively satisfy these goals are described in the remainder of Section 2.2.

***2.2.2. Brief summary of STADIA.*** STADIA's analysis procedure consists of three major stages: segmentation, classification, and phase and transition analysis. In the segmentation stage, STADIA approximates inputted length-history data with a series of straight-line segments that are connected to each other at their endpoints (Figure 2, C and D) and then measures characteristics of each line segment (Figure 2D). In the classification stage, STADIA uses the characteristics of the line segments from the segmentation stage (Figure 2, E and F) to identify how many distinguishable DI behaviors exist in the set of line segments and then groups line segments into named DI behaviors (Figure 2H). For visualization purposes, color labels corresponding to the named behaviors from the classification stage are applied to each line segment in the length-history approximation (Figure 2G). In the phase and transition analysis stage, STADIA measures aggregate phase metrics (e.g., total time in growth) as well as the frequencies of various transitions beyond typical catastrophe and rescue (Figure 2I). These three stages are summarized here in Sections 2.2.3–2.2.5, with more details provided in *Materials and Methods* Section 5.5. Limitations of STADIA and guidance for users are also discussed in *Materials and Methods* Section 5.6.

***2.2.3. Segmentation stage.*** The first step in the segmentation stage is to generate an initial approximation of inputted length-history data by identifying major peaks and valleys (Figure 2C), similar to more classical DI analysis methods (Figure 2, A and B). However, unlike classical methods, STADIA does not go directly from this step to calculating DI parameters. Instead, to improve the initial linear approximation, STADIA implements an iterative process to add new segment endpoints, which mark where the MT velocity changes as shown in Figure 2D. The user regulates how closely the segments approximate the raw data through the values of two input parameters: 1) the maximum error allowed between the line segments and

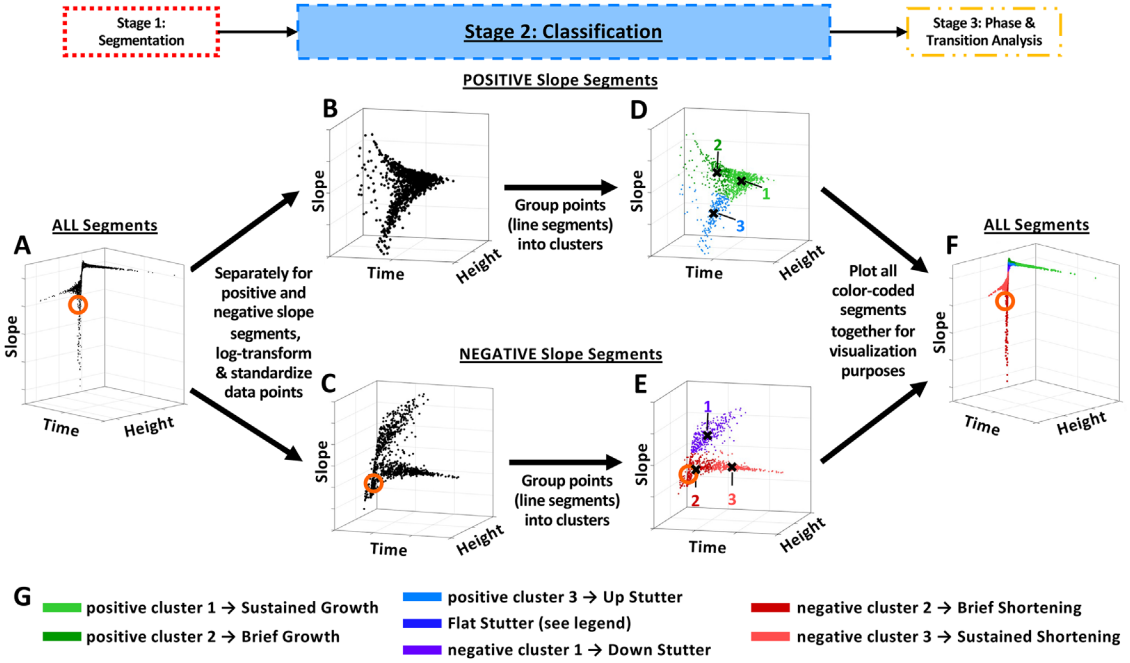


FIGURE 3: Overview of the clustering process in the Classification Stage of STADIA. As shown in Figure 2C–D, STADIA begins by approximating inputted length-history data with a series of line segments. For each line segment, three features are measured: time duration, height change, and slope (Figure 2D). Using these three features, a data point corresponding to each line segment is plotted in 3-dimensional space (A, replotted from Figure 2E). The orange circles in (A, C, E, F) denote the approximate location of the data point corresponding the example line segment in Figure 2D. Then, line segments that share similar values of time duration, length change, and slope (i.e., data points that are near each other in the 3-dimensional space) are grouped together into "clusters". As initial preparation for this grouping process, segments with slopes that are very near zero are identified by user-defined thresholds and assigned to one group (called "flat stutters"; see Section 5.5.2.1 for more information). In the next step of the process, applied separately to the remaining positive and negative slope segments (B,C), STADIA uses established statistical methods (described in Box 1 and *Materials and Methods* Sections 5.4.3 and 5.5.2.2) to determine the number of distinguishable clusters and then to assign each line segment to a cluster (D, E). After the segments are grouped into clusters, the average features (e.g., average slope) of the segments in each cluster are used to label each cluster with a named DI behavior (G).

the raw length-history data at each timepoint; 2) the minimum time duration for any line segment. Together these two parameters enable the user to avoid overfitting or underfitting the data relative to the scale of the dynamics the user wishes to study. As demonstrated in Section 2.6, proper tuning of these parameters enables STADIA to be compatible with data sets that have a wide range of temporal resolutions.

Note that the segmentation process imposes no restrictions on the slope of the segments and makes no assumptions about the number or type of behaviors present. These attributes are in contrast to the more traditional DI analysis methods described above, which use thresholds (e.g., on length change or velocity) to seek out segments corresponding to predefined behaviors (e.g., growth, shortening, pauses).

Effectively, the approximation produced by the segmentation stage of STADIA resembles the raw data more closely (Figure 2D) than does the approximation from classical methods (Figure 2, A–C). In particular, there are two fundamental differences between the segmentations resulting from classical methods and Stage 1 of STADIA. First, an individual segment of growth or shortening as identified by classical methods (Figure 2, A and B) may be identified as multiple segments of various slopes in the STADIA analysis (Figure 2D). Second, STADIA's refined approximation identifies segments of shallower slope that are not separated out from longer growth and shortening segments in classical methods.

***2.2.4. Classification stage.*** This stage identifies the number of behavior types observed in the output from the segmentation stage and bins similar segments into behavior classes. To do this, STADIA measures three key features of each line segment (namely length change, time duration, and velocity; Figure 2D) and then plots a data point corresponding to each line segment in three-dimensional (3-D) space (Figure 2E). Line segments that share similar values of the three features (i.e., data points that are near each other in the 3-D space) are grouped into "clusters." STADIA uses established statistical methods (Box 1) to determine the optimal number of distinguishable clusters and to assign each line segment to a cluster (Figures 2F and 3). Note that this clustering step avoids assuming that any cluster corresponds to a predetermined DI phase/behavior.

After the line segments are assigned to clusters, the average features (e.g., average slope) of the segments in each cluster are used to assign each cluster to a named DI behavior (Figure 3G). Clusters containing segments with similar slopes are "bundled" into DI phase/behavior classes (Figure 2H). In particular, clusters of shallow-slope segments are bundled into "stutters," clusters of steep positive slope segments are bundled into "growth," and clusters of steep negative slope segments are bundled into "shortening."

***2.2.5. Phase and transition analysis stage.*** For each cluster identified in the classification stage, STADIA calculates the following metrics: total number of segments (counts obtained from the piecewise linear approximation) in each cluster, percent time spent in each

**Box 1:**
For interested readers, Box 1 summarizes how established statistical methods are used in the clustering step of STADIA's Classification Stage (more details in *Materials and Methods* Sections 5.4.3 and 5.5.2.2).

We first note that STADIA can be run in two modes: Diagnostic Mode (used to inform the number of distinguishable clusters in the data set), and Automated Mode (used for performing full DI analysis after Diagnostic Mode work is complete). Automated Mode performs all three stages: segmentation, classification, and phase and transition analysis (Figure 2, C–I; workflow diagram in Supplemental Figure S1.1). Diagnostic Mode stops after a modified version of the classification stage.

The clustering process, applied separately to the positive slope segments and the negative slope segments, uses an established algorithm called *k*-means clustering (Macqueen, 1967; Lloyd, 1982). *K*-means groups together data points that share similar characteristics, that is, data points that are near each other in a relevant feature space (in our case, the log-transformed and standardized 3-D space defined by segment time duration, height change, and slope [Figure 3, B–E]).

The *k*-means algorithm requires that the number of clusters, *k*, be provided in advance. The value of *k* is informed by running STADIA in Diagnostic Mode. When run in Diagnostic Mode, STADIA repeats the *k*-means clustering process with the value of *k* set equal to each integer from 1 to 12. For each of these *k*-values, STADIA calculates a measurement called the gap statistic, which quantifies how well the data can be separated into *k* many clusters (Tibshirani *et al.*, 2001). Analysis of the gap statistic data from Diagnostic Mode as well as visual examination of the clusters plotted in the feature space (Supplemental Figures S1.4 and S1.5) informs the choice of the optimal *k*-value, which the user then inputs into Automated Mode.

cluster, percent height change corresponding to each cluster and average velocity of each cluster. These metrics can also be calculated for each of the larger bundled phase/behavior classes (i.e., growth, shortening, stutters.

Next, STADIA examines the chronological occurrences of the phases in the length-history data to identify all examples of transitions to/from growth and shortening, with or without stutters. Specifically, STADIA automatically categorizes the following types of phase transitions (Figure 2, I and J):

- "Abrupt Catastrophe": growth ➔ shortening directly
- "Abrupt Rescue": shortening ➔ growth directly
- "Transitional Catastrophe": growth ➔ stutter ➔ shortening
- "Transitional Rescue": shortening ➔ stutter ➔ growth
- "Interrupted Growth": growth ➔ stutter ➔ growth
- "Interrupted Shortening": shortening ➔ growth ➔ shortening

Similar chronological orderings of phases have previously been considered with pauses in experiments performed in the presence of cell extracts (Keller *et al.*, 2008).

After identifying all occurrences of the above transitions, STADIA calculates the frequency of each type of transition (see *Materials and Methods* Section 5.5.3 for formulas). For continuity with previous methods, the traditional transition frequencies can be calculated: the total catastrophe frequency, $F_{cat}$, is the sum of the frequencies of abrupt and transitional catastrophes, and the total rescue frequency, $F_{res}$, is the sum of the frequencies of abrupt and transitional rescues.

The STADIA process as outlined here enables extraction of traditional DI parameters as well as information about more complex behaviors and transitions. STADIA thus characterizes and quantifies MT dynamics without predefined assumptions about the number of behaviors or their defining attributes.

### 2.3. MT behaviors identified and characterized using STADIA

***2.3.1. STADIA identifies multiple types of behavior within the groups of positive and negative slope segments.*** If MT growth and shortening each corresponded to one behavior (with variation), one would expect that the positive slope line segments from the approximation of the length-history plot would all fall into one cluster (i.e., one group of line segments); similarly, all the negative slope segments would be expected to fall into one cluster.

Contrary to these expectations, STADIA identified three clusters within the positive slope segment data of each data set (i.e., the in silico data set and the in vitro control and CLASP2γ data sets; Supplemental Figure S1.4). Examination of the characteristics of the three clusters shows that they can be described as follows (Figure 4, A–C, and Supplemental Figures S1.4, B, D, F, and H, and S1.8, A and B):

- segments with steep slopes and long time durations (positive slope cluster 1);
- segments with steep slopes and short time durations (positive slope cluster 2);
- segments with shallow slopes and short time durations (positive slope cluster 3).

Similarly, when analyzing the negative slope segments from the in silico data, STADIA identified three clusters (Supplemental Figure S1.5, A–D), which have the following characteristics (Figure 4, D and F, and Supplemental Figures S1.5, B and D, and S1.8, C and D):

- segments with shallow slopes and short time durations (negative slope cluster 1);
- segments with steep slopes and short time durations (negative slope cluster 2);
- segments with steep slopes and long time durations (negative slope cluster 3).

For technical reasons, the in vitro data sets contain the beginnings of shortening phases (Figure 1D), but not full depolymerizations of MTs to near-zero length as were present in the simulation data set (Figure 1B). Consistent with this information, STADIA's analysis of the in vitro negative slope segments (Figure 4E) did not find a cluster of long-time-duration segments (i.e., no cluster analogous to negative slope cluster 3 in the in silico data in Figure 4D). Nonetheless, STADIA did find evidence for two distinguishable clusters of short-duration negative-slope segments in the in vitro data: negative slope cluster 1 with shallow slope segments, and negative slope cluster 2 with steep slope segments (Figure 4, E and F, and Supplemental Figures S1.5, E–H, and S1.8, C and D). For illustration purposes, a "ghost" region was added to Figure 4E, where we expect the missing third negative slope cluster would reside if full depolymerization events had been captured in the experiments.

In summary, our simulations and experiments lead to a similar conclusion: the data argue against the idea that MT DI can be characterized as a two-state process consisting of only growth and shortening with instantaneous transitions. Instead, the results

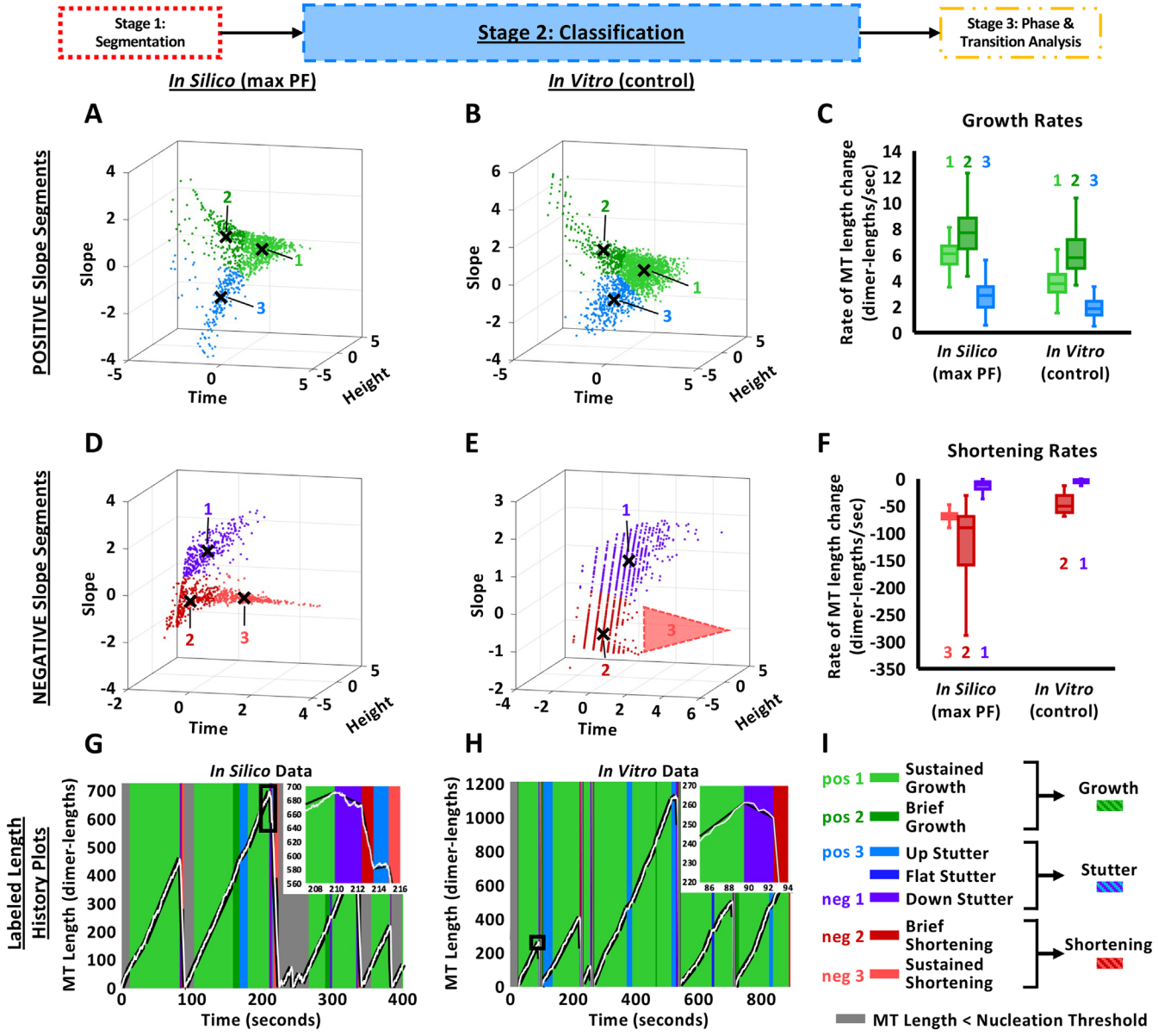


FIGURE 4: Results of STADIA's Classification analysis of in silico and in vitro datasets. (A, B, D, E) Color-coded clustering results for the in silico data (A, D) and in vitro control data (B, E); the clustering results for the in vitro CLASP2γ dataset are in Supplemental Figures S1.4 H, S1.5 H. Each data point in these plots corresponds to one line segment from the length-history approximations (see Figures 2D–F, 3). The scales of each axis reflect log-transformation and standardization of the data (see Figure 3). (C, F) Box plots of growth rates (C) and shortening rates (F) (i.e., segment slopes) for segment clusters as indicated. Outliers were excluded from the box plots (but not from the cluster plots) using the default definition in MATLAB (i.e., any value more than 1.5 times the interquartile range away from the bottom or top of the box is considered an outlier). (G, H) MT length-history plots with each segment labeled according to its assigned cluster. Zoomed-in portions of previously ambiguous length-history data are now clearly labeled (compare to Figure 1B–E). (I) Clusters with similar average slopes (panels C, F) bundled (grouped together) into larger behavior classes (see also Supplemental Figure S1.8). Notes: The raw length-history data have temporal resolution of ~1650 events per second per MT in silico and 2 frames per second in vitro. *Materials and Methods* Section 5.4.3 and Supplemental Figures S1.4, S1.5 provide justification for identifying three clusters each of positive and negative slope segments in most of our datasets (two clusters were used for the in vitro negative slope segments because complete depolymerizations to the seeds were not captured).

provide evidence for considering more complexity, including multiple types of behavior within both the positive and negative slope segments.

In the next two sections (Sections 2.3.2 and 2.3.3), we examine the average characteristics of the length-history segments in each cluster to determine how these clusters might correspond to recognizably different DI behaviors.

***2.3.2. Growth and shortening phases consistent with classical DI analysis are among the multiple types of behavior identified by STADIA.*** Examining the average characteristics of the segments in each cluster (Figure 4, Supplemental Figure S1.8, and Table 1) shows that, for both the in silico and in vitro data, some of the clusters correspond to the well-recognized growth and shortening phases of DI. More specifically, two of the positive segment clusters

| *In Silico* Data (Max PF Length; ~1650 events per second) | | | | | | |
|---|---|---|---|---|---|---|
| **Analysis Method** | **# of Cat.** | **# of Res.** | **$F_{cat}$ ($min^{-1}$)** | **$F_{res}$ ($min^{-1}$)** | **$V_{growth}$ (nm/s)** | **$V_{short}$ (nm/s)** |
| Classical Method | 355 | 123 | 0.659 | 2.483 | 46.1 ± 5.1 | 540.0 ± 47.9 |
| STADIA limited to growth and shortening: *k*=1 for pos & neg | 449 | 214 | 0.912 | 4.391 | 46.4 ± 18.4 | 530.4 ± 556.0 |
| STADIA limited to growth, shortening, and flat stutters: *k*=1 for pos & neg | 429 | 195 | 0.870 | 4.098 | 47.2 ± 17.6 | 547.2 ± 556.8 |
| All behaviors identified by STADIA: *k*=3 pos, *k*=3 neg | 298 | 75 | 0.660 | 1.944 | 48.0 ± 7.2<br>63.2 ± 19.2<br>22.4 ± 8.8 | 552.8 ± 87.2<br>1016.8 ± 717.6<br>107.2 ± 72.8 |

| *In Vitro* Data - Control (2 fps) | | | | | | |
|---|---|---|---|---|---|---|
| **Analysis Method** | **# of Cat.** | **# of Res.** | **$F_{cat}$ ($min^{-1}$)** | **$F_{res}$† ($min^{-1}$)** | **$V_{growth}$ (nm/s)** | **$V_{short}$ (nm/s)** |
| Classical Method | 802 | 40 | 0.717 | N.D. | 29.5 ± 12.7 | 330.1 ± 136.5 |
| STADIA limited to growth and shortening: *k*=1 for pos & neg | 856 | 83 | 0.777 | N.D. | 32.0 ± 24.8 | 216 ± 199.2 |
| STADIA limited to growth, shortening, and flat stutters: *k*=1 for pos & neg | 846 | 76 | 0.760 | N.D. | 32.8 ± 24.8 | 227.2 ± 198.4 |
| All behaviors identified by STADIA: *k*=3 pos, *k*=2 neg | 734 | 18 | 0.756 | N.D. | 30.4 ± 7.2<br>60.0 ± 44.8<br>15.2 ± 5.6 | 373.6 ± 143.2‡<br>39.2 ± 25.6 |

| *In Vitro* Data - CLASP2γ (2 fps) | | | | | | |
|---|---|---|---|---|---|---|
| **Analysis Method** | **# of Cat.** | **# of Res.** | **$F_{cat}$ ($min^{-1}$)** | **$F_{res}$† ($min^{-1}$)** | **$V_{growth}$ (nm/s)** | **$V_{short}$ (nm/s)** |
| Classical Method | 99 | 62 | 0.500 | N.D. | 43.1 ± 34.4 | 155.1 ± 77.6 |
| STADIA limited to growth and shortening: *k*=1 for pos & neg | 142 | 94 | 0.720 | N.D. | 46.4 ± 41.6 | 96.0 ± 84.0 |
| STADIA limited to growth, shortening, and flat stutters: *k*=1 for pos & neg | 131 | 87 | 0.676 | N.D. | 48.0 ± 41.6 | 108 ± 82.4 |
| All behaviors identified by STADIA: *k*=3 pos, *k*=2 neg | 86 | 52 | 0.498 | N.D. | 37.6 ± 11.2<br>100.0 ± 57.6<br>16.0 ± 7.2 | 158.4 ± 68.8‡<br>32.8 ± 17.6 |

**Sustained Growth** (positive slope cluster 1)

**Brief Growth** (positive slope cluster 2)

**Up Stutter** (positive slope cluster 3)

**Sustained Shortening** (negative slope cluster 3)

**Brief Shortening** (negative slope cluster 2)

**Down Stutter** (negative slope cluster 1)

**TABLE 1:** Comparison of DI measurements from classical two-state analysis, STADIA two-state analysis (i.e., STADIA with $k$ = 1), and STADIA analysis with full classification. *Top row of each subtable*: classical two-state analysis method (*Materials and Methods* Section 5.3) performed by identifying only major peaks and valleys (Figure 2 A–B). *Second row of each subtable*: STADIA analysis with classification limited to two states: only growth and shortening. *Third row of each subtable*: STADIA analysis with classification limited to growth, shortening, and flat stutters. *Bottom row of each subtable*: STADIA analysis using full results of the classification stage (Figure 4; *Results* Section 2.3). All STADIA analyses used the fine-grained length-history approximation generated by the segmentation stage of STADIA (Figure 2D) but differed in the settings for the classification stage. These data show that there is general, but not exact, agreement between the analysis methods as applied to each dataset. $V_{growth}$ and $V_{short}$ measurements are listed as mean ± standard deviation over the set of all segments identified in each type of behavior. See Supplemental Figure S1.9 for the number of segments in each cluster from the STADIA full analysis. See *Materials and Methods* Sections 5.1.4 and 5.2.2 for the number of MTs and total observation times in each dataset.
†,‡: Because depolymerizations in the in vitro datasets were not captured in their entirety (see examples in Figure 1D), rescue frequencies are not reported (†), and negative slope segments were separated into only two clusters, yielding only two $V_{short}$ measurements in the full STADIA analysis (‡), instead of three as seen with the in silico data.

(positive slope clusters 1 and 2 from Figure 4, A and B) have slopes (rates of length change) similar to growth rates reported in classical DI analysis (compare STADIA results in Figure 4C and Table 1 to classical analysis results in Table 1). Similarly, negative slope cluster 2 (in silico and in vitro, Figure 4, D and E) and negative slope cluster 3 (in silico, Figure 4D) have slopes similar to shortening rates reported in classical DI analysis (compare Figure 4F and Table 1). Based on this information, STADIA classifies length-history segments as "growth" if they belong to one of the clusters with a steep positive slope (positive slope cluster 1 or 2 in Figure 4, C and I) and as "shortening" if they belong to a cluster with a steep negative slope (negative slope cluster 2 or 3 in Figure 4, F and I).

The two clusters of steep positive slope segments (and of steep negative slope segments for the in silico data) differ primarily by time duration, so we refer to them as "brief" or "sustained" (Figure 4I and Supplemental Figure S1.8, B, D, and E). It is also notable that the brief growth/shortening clusters have greater variation in slope than the sustained growth/shortening clusters (Figure 4, C and F, and Supplemental Figure S1.8, A and C), which suggests that the most rapid velocities are not sustainable over long periods of time.

These observations may be evidence of different behaviors of tapered or split tips relative to the rest of the MT (e.g., as observed by Coombes *et al.*, 2013; Doodhi *et al.*, 2016; and Aher *et al.*, 2018); such structures might be able to extend or retract faster than the bulk MT lattice in the absence of lateral bonds. Future work is needed to investigate whether the differences between brief and sustained growth (or shortening) relate to tip structure.

***2.3.3. STADIA detects and characterizes "stutters": a category of dynamic behaviors distinct from growth, shortening, and pause.*** Examination of Figure 4, A–F, shows that, in addition to clusters of segments with slopes that correspond to rates of length change seen in classical growth or shortening behaviors, STADIA also identifies clusters of segments with much shallower slopes (positive slope cluster 3 and negative slope cluster 1 in Figure 4, A–F; Table 1). Moreover, the segments in these shallow-slope clusters have time durations shorter than typical segments classified as sustained growth and sustained shortening, though typically longer than those classified as brief growth and brief shortening segments (Supplemental Figure S1.8). We term these clusters of shallow-slope segments "stutters" to convey the idea that these sections of length-history data exhibit high-frequency, low-amplitude fluctuations throughout which the overall rate of MT length change is slow from a macro-level perspective. Within the category called "stutters," we name the clusters based on their slopes (Figure 4I): "up stutters" (positive slope cluster 3), "down stutters" (negative slope cluster 1), and "flat stutters" (relatively rare near-zero slope segments identified before analyzing the positive and negative slope segments as described in the Figure 3G legend and Section 5.5.2.1).

In summary, stutters are a category of intermediate behaviors that share similar characteristics with each other and are distinguishable from typical growth and shortening. Distinguishing the various behaviors described above involved the use of segment slope, time duration, and height change, as explained in *Materials and Methods* Section 5.5. For any one of these three features individually, there is overlap between different clusters identified in the data (Figure 4, A–F, and Supplemental Figures S1.3 and S1.8). Of the three segment features, slope is the primary feature distinguishing stutters from typical growth and shortening (Figure 4C and F, and Supplemental Figure S1.8, A and C). In other words, the rate of change in MT length tends to be slower during stutters than during growth and shortening. In regard to time durations, up and down stutters, respectively, have similar or somewhat longer time durations than brief growth and shortening segments, but shorter time durations than sustained growth and shortening segments (Supplemental Figure S1.8, B and D).

***2.3.4. Stutters overlap with previously observed slowdown periods but are distinguishable from pauses.*** Note that most stutters are distinguishable from previously identified "pauses" during which the MT neither grows nor shortens detectably (e.g., Gierke *et al.*, 2010; Yenjerla *et al.*, 2010). In contrast to pauses, MT lengths do indeed change measurably during most periods identified as stutters (for examples, see insets in Figure 4, G and H), with a net rate of change that is small but nonzero (Figure 4, C and F). In addition, it is notable that events categorized as pauses are typically described as being rare (<1% of total experiment time duration) in the absence of MT stabilizing proteins (e.g., Walker *et al.*, 1988; Moriwaki and Goshima, 2016). In contrast, stutters are relatively common, as discussed more below (Section 2.4.3 and Supplemental Figure S1.9). These observations support the conclusion that most stutters are different from events previously classified as pauses, though there is likely some overlap, particularly between the relatively rare flat stutters (Supplemental Figure S1.9) and cases where pauses were allowed to be short in duration (e.g., Walker *et al.*, 1988; Guo *et al.*, 2018). Stutters as described above likely do encompass the periods of slowed growth or shortening previously noted (but not quantified or characterized in detail) in recent DI data of in vitro MTs acquired at high spatiotemporal resolution (e.g., Maurer *et al.*, 2014; Duellberg *et al.*, 2016a,b; Rickman *et al.*, 2017; see also Margolin *et al.*, 2012). In contrast to this previous work, here we have quantified spontaneously occurring stutters and examined their relationship to other DI behaviors.

***2.3.5. Negative control: two-state growth-shortening model.*** As a negative control to verify that the observation of stutters is not an artifact of STADIA's analysis process, we ran STADIA on length-history simulation data from a model designed to have only two states: growth and shortening. As would be expected, STADIA analysis of the length-history data from the two-state model did not identify behaviors comparable to the stutters detected in our main data sets (i.e., the dimer-scale simulation data and the in vitro data). For a description of the two-state simulations and the analysis results, please see Supplemental Material Section S4, "Negative Control: Simulations of a Two-State (Growth-Shortening) Model," and Supplemental Figures S4.1, S4.2, and S4.3.

### 2.4. Quantification of MT dynamics using STADIA

***2.4.1. Comparison of the traditional DI parameters as measured by STADIA versus a classical DI analysis method.*** For each of the three data sets, Table 1 contains a comparison of results obtained using STADIA with three different sets of conditions in the classification stage to results from a classical DI analysis method (top row of each subtable; the classical analysis procedure is described in *Materials and Methods* Section 5.3). In the first set of STADIA conditions (second row), meant to approximate classical DI analysis, we restricted the classification stage of STADIA to recognizing only growth and shortening, that is, all positive slope segments were classified as growth and all negative slope segments were classified as shortening. In the next set of conditions (third row), we allowed STADIA to separate out near-zero slope segments as flat stutters but constrained the classification of the remaining positive and negative slope segments to one cluster each of growth and shortening. The final STADIA analysis (bottom row) utilized all clusters identified in the classification results in Section 2.3.

Within each data set in Table 1, the measured values of the standard DI parameters ($V_{growth}$, $V_{short}$, $F_{cat}$, and $F_{res}$) are similar across the different analysis approaches. The differences in measured values between the classical analysis and STADIA constrained to only growth and shortening occur because of differences in the segmentation. More specifically, the line-segment approximation produced by the segmentation stage of STADIA (Figure 2D) resembles the raw length-history data more closely than does the approximation from the classical method (Figure 2, A–C). Thus, STADIA can produce measurements of the traditional DI parameters but does so by using a finer linear approximation of the length-history data than the classical analysis, resulting in differences in the measured values of the DI parameters (Table 1).

***2.4.2. Quantification of velocities and transition frequencies beyond the traditional DI parameters.*** While STADIA can provide measurements of the four traditional DI parameters that are similar to the measurements obtained from classical approaches, the multiple clusters (i.e., multiple types of behaviors) detected in the

full classification results (Section 2.3) indicate that the traditional DI parameters alone are inadequate to capture the full range of MT dynamics. Expanding beyond the traditional $V_{growth}$ and $V_{short}$, the full STADIA analysis provides quantification of the intrinsic variability in growth and shortening rates by separately measuring the velocities for each cluster (Table 1, bottom row of each subtable). Expanding beyond the traditional $F_{cat}$ and $F_{res}$, STADIA measures the frequencies of additional types of phase transitions (Figure 2I; Supplemental Figure S1.10).

***2.4.3. MTs spend a significant fraction of time in stutters.*** We begin to investigate the significance of stutters by first examining the fraction of time that MTs spend in stutters. As one might expect, both in silico MTs and physical MTs spend the majority of their time in growth phases. However, in both the simulations and experiments, MTs spend a substantial amount of time in behaviors categorized as stutters. Notably, in our in silico data sets, the MTs spent more time in stutters (8%) than in shortening (6%) (Figure 5A; Supplemental Figure S1.9). The in vitro MTs spent a substantial amount of the time in stutters (Supplemental Figure S1.9), but direct comparison to time spent in shortening phases is not conclusive because depolymerizations were not fully captured. These observations indicate that stutters contribute appreciably to MT behavior as assessed in length-history plots.

***2.4.4. Catastrophes are usually preceded by stutters in silico and in vitro.*** To investigate the functional significance of stutters, we used STADIA to examine how transitions between phases occur. More specifically, STADIA considers all possible transitions into and out of growth or shortening, with or without stutters (see Figure 2I for schematic, Figure 5, D–I, for in silico examples, Figure 6 for in vitro examples with corresponding kymographs, Figure 7, D–I, for additional in vitro examples, and Supplemental Figure S1.10 for frequencies).

Notably, in both the simulation data and the experimental control data, the majority of catastrophes involved a stutter between the growth and shortening phases (i.e., they were transitional catastrophes). In particular, 78% of the catastrophes in the simulation data were transitional (Figure 5B). A related observation in the simulation data is that almost half (44%) of stutters that occurred after a growth segment ended in catastrophe as opposed to returning to growth (i.e., they occurred as part of a transitional catastrophe as opposed to interrupted growth; Figure 5C). A similar but more dramatic association between stutters and catastrophe was observed in the in vitro control data: 86% of catastrophes involved a stutter (Figure 7A), and 75% of stutters from growth ended in a catastrophe (Figure 7B).

In contrast to catastrophes, rescues as observed in the in silico data set rarely occurred with stutters. More specifically, only 5% of in silico rescues were transitional (i.e., few rescues involved a stutter) (Figure 5B), and only 8% of stutters that occurred during shortening resulted in a rescue (Figure 5C). Because we do not have sufficient data for rescues in vitro, we cannot make strong conclusions on the correlation between stutters and rescue in physical MTs. However, these results do suggest that catastrophe and rescue are not simply the mechanistic opposites of each other.

### 2.5. Dissecting the effects of a MT binding protein: CLASP2γ reduces the frequency of catastrophe by increasing the prevalence of interrupted growth

To further test STADIA's utility in analyzing DI and to examine both the prevalence and the significance of stutters, we compared the control in vitro data set to the in vitro data set with the MTBP CLASP2γ, which has been previously characterized as an anticatastrophe factor (Aher *et al.*, 2018; Lawrence and Zanic, 2019). CLASP2 proteins are of interest to the biomedical community because they have been implicated in functions as diverse as kinetochore attachment (Girão *et al.*, 2020), nervous system development (Dillon *et al.*, 2017), and the insulin response (Kruse *et al.*, 2017).

Recall that the clustering results, including detection of stutters, are similar for the control and CLASP2γ data sets (Supplemental Figures S1.4 and S1.5). However, dramatic differences in the transition frequencies between the CLASP2γ data and control in vitro data were observed when these data were examined quantitatively by STADIA.

First, the overall frequency of catastrophe in the presence of CLASP2γ was significantly reduced (Figure 7C and Supplemental Figure S1.10). This observation itself is not surprising, given that previous work has shown that CLASP2γ reduces the frequency of catastrophe (e.g., Sousa *et al.*, 2007; Aher *et al.*, 2018; Lawrence *et al.*, 2018; Majumdar *et al.*, 2018). However, STADIA provides additional insight by distinguishing transitional catastrophes (growth-stutter-shortening) from abrupt catastrophes (growth-shortening). In particular, our results demonstrate that the reduction in overall catastrophe frequency was due to a large decrease in transitional catastrophe frequency, while the abrupt catastrophe frequency actually increased somewhat (Figure 7A and Supplemental Figure S1.10).

Second, CLASP2γ slightly reduced the frequency of growth-to-stutter occurrences (i.e., $F_{TransCat}$ + $F_{IntGrowth}$; Figure 7C) but not enough to account for the large decrease in transitional catastrophe frequency.

Third, and most striking, CLASP2γ increased the frequency of interrupted growth (growth-stutter-growth) (Supplemental Figure S1.10). More specifically, among transitions that began as growth-to-stutter, CLASP2γ increased the proportion of transitions that resulted in interrupted growth (growth-stutter-growth) and decreased the proportion of transitions that proceeded to transitional catastrophes (growth-stutter-shortening) (Figure 7B). This change in proportions is the factor that accounts for most of the decrease in transitional catastrophe frequency.

Taken together, these data demonstrate that STADIA analysis provides information about CLASP2γ function not supplied by traditional analysis and indicate that CLASP2γ suppresses catastrophe at least in part by enabling stuttering MTs to reenter growth (i.e., CLASP2γ tends to convert would-be transitional catastrophes into interrupted growths). This idea is supported by recent reports that MTs can withstand greater growth rate variability without undergoing catastrophe in the presence of CLASP2γ (Lawrence *et al.*, 2018; Lawrence and Zanic, 2019) and that CLASP2γ can protect against catastrophe in the presence of lagging PFs (Aher *et al.*, 2018).

### 2.6. Robustness of conclusions over varied values of input parameters and data acquisition rates

*Note that this section assumes that readers are familiar with STADIA's analysis procedure as described in Section 2.2 and Box 1.*

The results in the preceding sections led to the following main conclusions: 1) stutters (previously observed but not quantified in detail) are distinguishable from typical growth and shortening (Section 2.3); 2) stutters are strongly associated with spontaneously occurring catastrophes, both in silico and in vitro (Section 2.4.4); 3) the anticatastrophe factor CLASP2γ reduces catastrophe by increasing the fraction of stuttering microtubules that return to growth rather than entering shortening phases (Section 2.5). An important remaining question is whether these conclusions are robust to

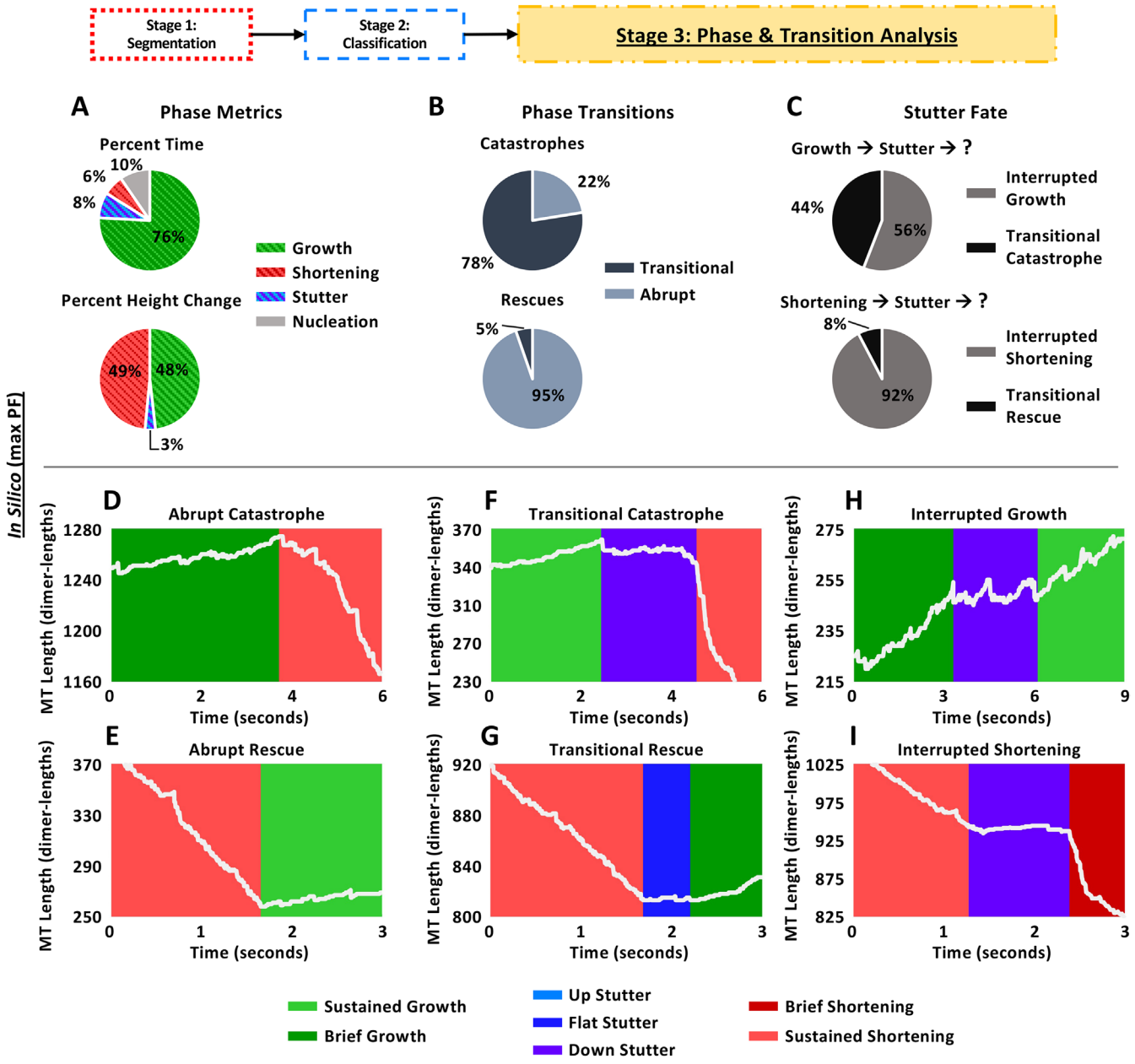


FIGURE 5: Results of STADIA's Phase and Transition Analysis of the dimer-scale in silico data. (A) Percent time spent in each class of phases/behaviors (top) and percent height (MT length) change occurring during each class of phases/behaviors (bottom). These data show that a large majority of time is spent in growth. Notably, the in silico MTs spend more time in stutters than in shortening, emphasizing the importance of studying stutter behaviors. Most height change occurs during growth and shortening phases, as expected. (B) Percentages of catastrophes (top) and rescues (bottom) that are transitional or abrupt (see Figure 2I and Section 2.2.5 for transition definitions). These data show that most catastrophes are transitional, whereas rescues are overwhelmingly abrupt. (C) Examination of stutter fate. These data show that when growth-to-stutter occurs (top), interrupted growth is slightly more likely than transitional catastrophe. However, when shortening-to-stutter occurs (bottom), interrupted shortening is much more likely than transitional rescue. (D–H) Examples of abrupt/transitional catastrophes (D, F), abrupt/transitional rescues (E, G), and interrupted growth/shortening (H, I). As noted earlier, the in silico dataset has temporal resolution of approximately 1650 events per second per MT (see *Materials and Methods* Section 5.2 for more information).

variations in the STADIA input parameters. To address this question, we performed sensitivity analyses, which are summarized here, with full details provided in Supplemental Material Sections S2 and S3.

***2.6.1. STADIA input parameter sensitivity analysis.*** To test the effects of STADIA's input parameters, we analyzed all three data sets (i.e., in silico as well as in vitro with and without CLASP2γ), using a range of values for each of the key user-defined segmentation parameters in STADIA, namely the Minimum Segment Duration and the Maximum Error Tolerance. Directly, the values of these parameters determine how closely the segmentation stage's continuous piecewise linear approximation matches the raw length-history data inputted into STADIA. Indirectly, these parameters have downstream effects on the results of the classification stage and the

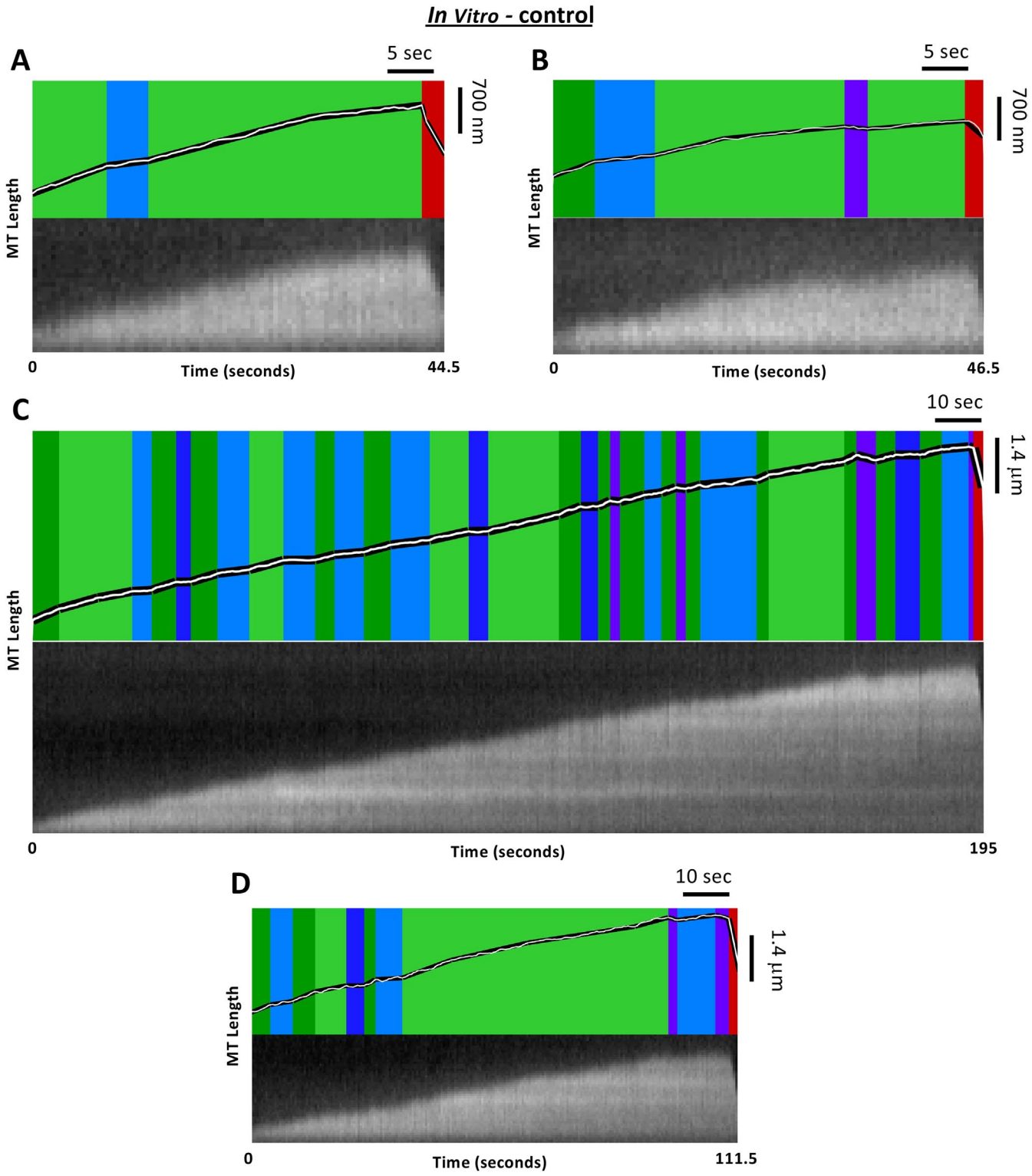


FIGURE 6: Alignment of STADIA length-history plots (top of each panel) and their corresponding kymographs (bottom of each panel) from the in vitro control dataset (with STADIA colors as in Figure 4 and scale bars as indicated). (A, B) Examples of abrupt catastrophes, where a growth phase (green) is followed directly by a shortening phase (red). (C, D) Examples of transitional catastrophes, where one or more types of stutter (blue, purple) occurs between a growth phase (green) and a shortening phase (red). Note also the numerous stutters (blue, purple) that interrupt growth phases (green). These length histories include examples of all three types of stutters that we distinguish based on slope: up stutters (light blue), flat stutters (dark blue), and down stutters (purple). The kymographs and the length-history traces inputted into STADIA were generated from the in vitro imaging data (2 fps) as described in *Materials and Methods* Section 5.1. The movies corresponding to each kymograph are provided as Supplemental Materials. Note that the movies are presented at 3.5 × real time (i.e., 3.5 × the time labeled on the kymographs and length-history plots).

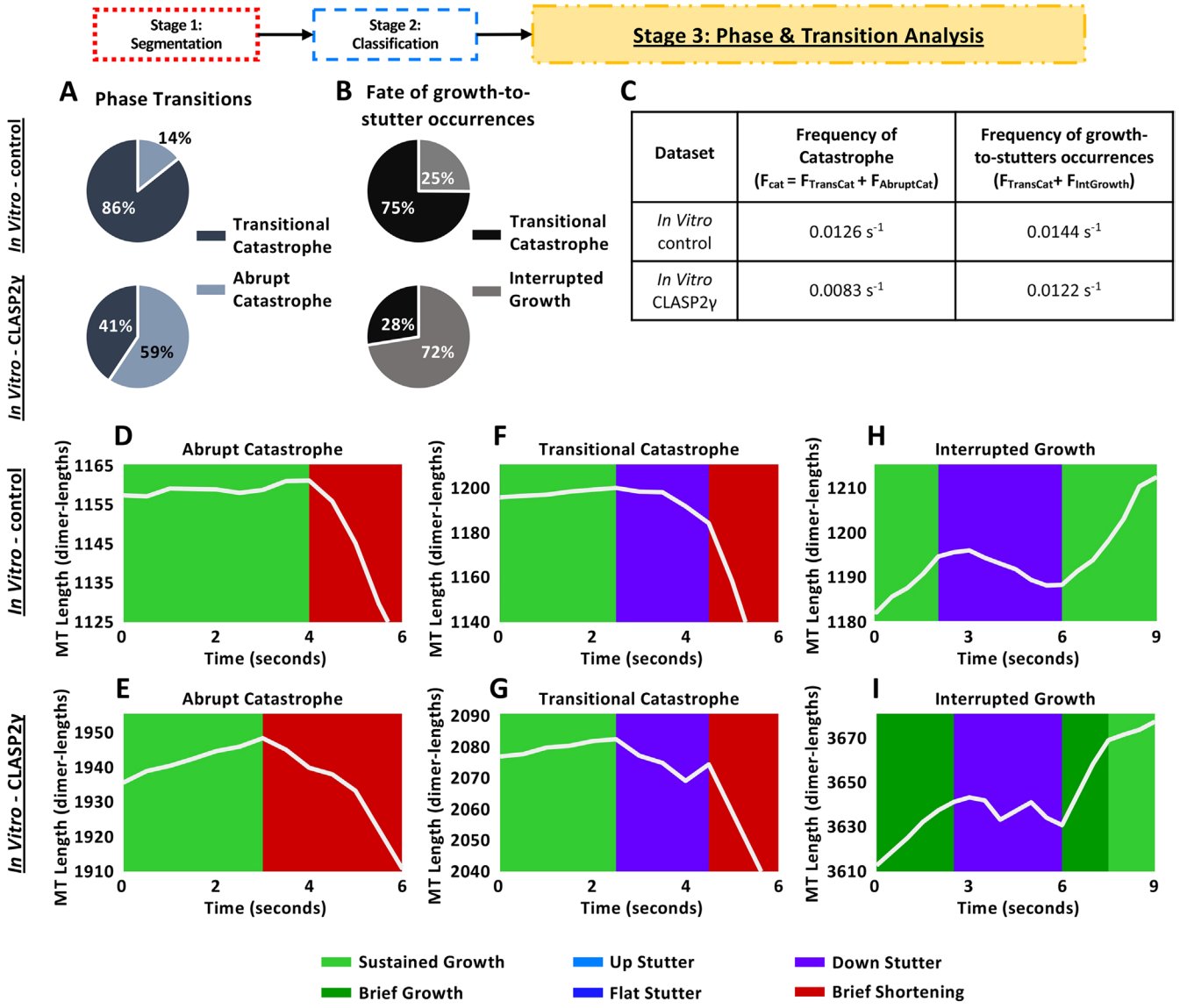


FIGURE 7: Results of STADIA's transition analysis of in vitro data (2 fps): Effect of CLASP2γ on the nature of catastrophes and the fate of stuttering MTs. See Section 2.2.5 for transition definitions. (A) The majority of catastrophes for in vitro MTs without CLASP2γ are transitional (top). However, introduction of CLASP2γ increases the fraction of catastrophes that are abrupt (bottom). (B) Most growth-to-stutter occurrences for the in vitro MTs without CLASP2γ (top) result in catastrophe. Addition of CLASP2γ (bottom) decreases the probability that a growth-to-stutter occurrence will proceed to shortening and increases the probability of returning to growth. (C) CLASP2γ decreases the overall frequency of catastrophe without greatly reducing the frequency of stutter-to-growth occurrences. Taken together (A–C), these data indicate that CLASP2γ reduces catastrophes by promoting growth following stutters. More specifically, transitions that would have been transitional catastrophes without CLASP2γ tend to become interrupted growths with CLASP2γ. (D–I) Examples of transitions as observed for the in vitro MTs both without CLASP2γ (top) and with CLASP2γ (bottom).

phase and transition analysis stage (for overview of STADIA stages, see Section 2.2, Figure 2, and Supplemental Figure S1.1). The results of the sensitivity analysis show how changes to the above input parameters impact STADIA outputs (e.g., number of behaviors detected, classification of segments as illustrated by labeled length-history plots, transition frequencies).

Briefly, results of this analysis (Supplemental Material Section S2) indicate that the three main conclusions enumerated above are indeed robust as long as the user-defined parameters are kept within ranges relevant to the scale of the dynamics being studied. In the first part of the sensitivity analysis, we examined how changing the user-defined parameters affects the number of behaviors detected (as determined by the number of clusters, i.e., *k*-values as described in Box 1 and *Materials and Methods* Sections 5.4.3 and 5.5.2.2). The results show that to detect stutter clusters, one must use spatial and temporal parameters that are sufficiently fine to distinguish the multiple behaviors detected (see the classification results in Sections 2.3.1–2.3.3) but not so fine that the analysis is affected by experimental noise or MT length fluctuations below the scale of the dynamics being studied. For our in silico data set, ideal ranges for these parameters were empirically determined to be from 15 to 25 dimer lengths (i.e., 120–200 nm) for the Maximum Error Tolerance

and 2 s or less for the Minimum Segment Duration (Supplemental Figures S2.1, S2.3, S2.5, and S2.7). For our in vitro control data set, ideal ranges were determined to be from 15 to 20 dimer lengths (120–160 nm) for the Maximum Error Tolerance and 1 s or less for the Minimum Segment Duration (Supplemental Figures S2.2, S2.4, S2.6, and S2.8). Note that while these ranges meet strict standards for detecting all of the behaviors studied, the analysis can tolerate a wider range of parameter values and still detect multiple types of behaviors including stutters within the positive and negative slope segments, thus demonstrating further robustness (see Supplemental Section S2).

In the next part of the sensitivity analysis, we examined the transition frequencies measured from running STADIA with varied values of the Minimum Segment Duration and the Maximum Error Tolerance while using the number of behaviors detected in the full classification results in Sections 2.3.1–2.3.3 (i.e., the analysis used the *k*-values selected in *Materials and Methods* Section 5.4.3). For the in silico data, both the values of the frequencies of abrupt and transitional catastrophe and the ratio between them are relatively insensitive to changes in the Maximum Error Tolerance in the range of 10–40 dimer lengths (i.e., 80–320 nm) (Supplemental Figure S2.9). Moreover, the conclusion that most catastrophes are transitional is robust for Minimum Segment Duration values of 1 s or less. However, for Minimum Segment Duration values of 1.5 s or greater, abrupt catastrophes become more common (Supplemental Figure S2.9), likely due to fewer stutters being detected and therefore fewer catastrophes being recognized as transitional. This observation is illustrated in examples of the labeled length-history plots (Supplemental Figure S2.7). Notably, the overall catastrophe frequency (the sum of abrupt and transitional) is less sensitive to changes in Minimum Segment Duration and Maximum Error Tolerance than are the abrupt and transitional catastrophe frequencies. The situation is similar, though somewhat noisier, for the in vitro data (Supplemental Figures S2.8, S2.10, and S2.11).

Significantly, the conclusion that CLASP2γ reduces catastrophe by promoting the growth of stuttering MTs is robust to changes across wide ranges of both Minimum Segment Duration and Maximum Error Tolerance (Supplemental Figures S2.10, S2.11, and S2.12). More specifically, for almost all parameter combinations tested, the presence of CLASP2γ decreased the frequency of transitional catastrophe and increased the frequency of interrupted growth relative to the in vitro control, even as the values of the frequencies themselves changed with varying the segmentation parameters (Supplemental Figures S2.10, B and C, S2.11, B and C, and S2.12, C–H). These results are particularly relevant to demonstrating STADIA's usefulness in studying the effects of MTBPs.

***2.6.2. Data acquisition rate sensitivity analysis.*** To test the effect of varying the acquisition rate of length-history data inputted into STADIA, we took the original full-resolution in silico data set and resampled the length-history data at varied fixed data acquisition time steps (Supplemental Material Section S3). Examining a wide range of data acquisition rates is feasible because the in silico data set records every dimer-scale biochemical event (bond formation/breaking, hydrolysis; on the scale of >1000 data points per second per MT; see *Materials and Methods* Section 5.2). For comparison, frame rates in physical experiments vary from more than 100 fps (e.g., Mickolajczyk *et al.*, 2019) to fewer than 0.3 fps (e.g., Gierke *et al.*, 2010).

The resulting analysis (Supplemental Figures S3.1–S3.9) shows that the conclusion that "up stutters" exist is robust for data acquisition time steps up to 3 s, and similarly for "down stutters" at data acquisition time steps up to 1 s, assuming reasonable choices for Maximum Error Tolerance and Minimum Segment Duration (see Section 2.6.1). However, even when stutters are detected as distinct clusters, the number of stutter segments detected generally decreases for larger data acquisition time steps (i.e., slower data acquisition rates). This observation is not surprising because some stutters, particularly down stutters for the in silico data, have time durations on the order of 1 s or less (Supplemental Figure S1.8), and with frame rates slower than 1 s, such stutters would be undetectable.

Note that the in vitro data set was obtained using a frame rate of 2 fps, which was determined to be near the slower end of the range of acceptable data acquisition rates for some of the conclusions. As continued technological improvements allow physical experiments to have faster frame rates, in vitro data may tolerate a wider range of STADIA parameters (similar to the in silico data set). Significantly, short data acquisition time steps do not introduce problems (indeed, they are ideal, as seen with the full-resolution in silico data set) because the Maximum Error Tolerance and Minimum Segment Duration parameters prevent the segmentation (i.e., the continuous piecewise linear approximation) of MT length-history data from containing arbitrarily short segments.

## 3. DISCUSSION

Here we have presented STADIA, a data-driven, automated tool for performing DI analysis using length-history data as input. Using STADIA, we have quantified stutters and their associated transitions (Figures 4–7, Table 1, and Supplemental Figures S1.8–S1.10). Stutters are a set of dynamic behaviors that can be distinguished from typical growth or shortening; the primary differentiating factor is that stutters on average have slower rates of MT length change (Section 2.3). Stutters are also distinguishable from pauses in that a pause is typically described as a period of time when the MT neither grows nor shortens. Our analysis shows that stutters (i.e., slowdown periods, previously observed but not quantified in detail) are strongly associated with spontaneously occurring catastrophes, both in silico and in vitro (Section 2.4.4). Our STADIA analysis also indicates that the anticatastrophe factor CLASP2γ reduces catastrophe by increasing the fraction of stutters that return to growth rather than entering shortening phases (Section 2.5).

Importantly, we have shown that these results are robust across a range of STADIA parameter values (Section 2.6.1) and are compatible with data acquired across a range of temporal resolutions (Section 2.6.2). More specifically, as shown by the temporal resolution sensitivity analysis and the fact that our full-resolution simulation data set includes every subunit attachment and detachment event, STADIA is compatible with data acquired at temporal resolutions ranging from those produced by iSCAT (which can be as high as 1000 fps) to those used in TIRF experiments (e.g., 2 fps). This sensitivity analysis further shows that STADIA can also be used with data at lower temporal resolutions, but the ability to quantify stutters is reduced.

### 3.1. Mechanisms of stutters and implications for the process of catastrophe

What causes stutters, especially those that disrupt growth, and why are they associated with catastrophe? A fundamental component to answering this question comes from recognizing that when transitioning from growth to stutter, there is a net decrease in the number of subunits (tubulin dimers) that are incorporated into the MT per unit time. This net decrease could occur because new subunits attach to the tip less frequently than during normal growth, or because bound subunits leave the tip more frequently than during growth, or a combination of these two.

While simple stochastic fluctuations in subunit arrival or departure could potentially contribute to stutters, examination of length-history plots (Figures 4–7) suggests that the stochastic fluctuations, which occur throughout growth, shortening and stutter segments, are too short in duration to account for the sustained decrease in growth rate that occurs when going from growth to stutter. Alternatively, changes in rates of attachment and detachment could also result from changes in tip structure. However, one could argue that the rate of subunit attachment should not vary with tip structure: assuming that longitudinal bonds form first, there are always 13 landing sites for new subunits (Castle and Odde, 2013). Therefore, we suggest that stutters following growth segments likely result from a situation where an unusually large fraction of incoming subunits detach from the tip structure without being fully incorporated into the lattice (e.g., because tip taper or other structural features like the presence of GDP-tubulin make it difficult for lateral bonds to form). In other words, we suggest that stutters occur when the structure of the tip is such that the subunit detachment rate is unusually high compared with the average detachment rate during growth.

This reasoning provides a potential explanation for the correlation between stutters and catastrophe: if the fraction of incoming subunits incorporated into the lattice is smaller than during normal growth periods, the stabilizing cap of GTP-tubulin at the MT end will become smaller, the likelihood of exposing GDP-tubulin subunits will increase, and the possibility of complete cap loss (catastrophe) will rise. At present, these ideas are speculation, but future work may be able to shed light on these hypotheses (see also related discussions in VanBuren *et al.*, 2005; Howard and Hyman, 2009; Gardner *et al.*, 2011; Margolin *et al.*, 2012; Coombes *et al.*, 2013; Zakharov *et al.*, 2015; and McIntosh *et al.*, 2018).

Furthermore, the mechanisms could vary for different types of stutters. As demonstrated in the results, STADIA distinguishes up, down, and flat stutters and distinguishes stutters that occur as part of interrupted growth, interrupted shortening, transitional catastrophe, and transitional rescue. Thus, as a tool for comprehensively identifying multiple types of stutters, STADIA lays the groundwork for future mechanistic studies.

### 3.2. Comparison of the in silico and in vitro results

The behaviors observed in the dimer-scale simulation data and the experimental data are qualitatively similar. In particular, both types of data support the prevalence of stutters throughout length histories and the role of stutters in catastrophes. The differences in the particular numerical values of measured quantities are not surprising, because the simulation parameters were tuned in Margolin *et al.* (2012) based on an experimental data set (Walker *et al.*, 1988) different from the experimental data sets used here (a subset of which was used in Lawrence *et al.*, 2018). The qualitative similarities between the results from the different data sets provide evidence that the observed trends are not specific to one experimental preparation or one type of tubulin (e.g., 10 μM pure porcine tubulin in Walker *et al.*, 1988, vs. 12 μM bovine tubulin with EB1 and with or without CLASP2γ here and in Lawrence *et al.*, 2018). Furthermore, comparison with the negative control (two-state growth-shortening model; Supplemental Material Section S4) demonstrates that the existence of stutters in the dimer-scale simulations and the in vitro data is not manufactured by STADIA.

### 3.3. Relationship to previous work

***3.3.1. Distinguishing stutters and previously identified pauses.*** Pauses have most frequently been observed in vivo (see citations in the *Introduction*) and are likely caused by MTBPs (Moriwaki and Goshima, 2016) and other factors external to the MTs themselves (e.g., reaching the cell edge [Rusan *et al.*, 2001; Komarova *et al.*, 2002]). Furthermore, in vitro pauses in the absence of drugs or MTBPs are rare (Walker *et al.*, 1988). In contrast, the observation that stutters are prevalent in both our in silico and in vitro data sets suggests that stutters are an intrinsic component of DI itself.

Gierke *et al.* (2010) have described "bona fide pauses" as phases "during which no polymerization or depolymerization occurs." Due to physical detection limits, true pauses would be indistinguishable from periods of very slow polymerization or depolymerization that do not meet the detection threshold (Gierke *et al.*, 2010). Particularly in older data sets with large thresholds (e.g., a length-change threshold of 0.5 microns), some stutters may have been considered pauses while others may not have been separated out from larger growth or shortening phases at all. With newer imaging technology, data can be obtained at higher temporal and spatial resolution (e.g., Maurer *et al.*, 2014; Duellberg et al., 2016a,b; Rickman *et al.*, 2017; Guo *et al.*, 2018; Mickolajczyk *et al.*, 2019), which can enable the distinction of stutters and pauses.

For most stutters, a measurable net length change does occur over the course of the stutter segment: up and down stutters occur much more often than flat stutters. Using this information about stutters from our results and definitions of pauses already existing in the literature, we propose the following operational criteria for distinguishing pauses and stutters: pauses are periods during which no detectable length change occurs, whereas stutters are periods during which the MT structure changes but with slower net rates of length change than typical growth and shortening phases. In data sets that contain both stutters and pauses, the current version of STADIA would classify "bona fide pauses" as flat stutters. Future work is needed to determine whether it would be meaningful to apply criteria to distinguish flat stutters, which generally have very short time durations, from pauses.

***3.3.2. Previously observed behaviors that are similar to particular types of stutters.*** Maurer *et al.* (2014) observed short episodes of pause or slow growth before catastrophes in experiments with EB1. These precatastrophe slowdowns are analogous to transitional catastrophes in our terminology. Pauses or slowdowns before catastrophe have also been observed in cases where the catastrophe is induced by outside factors such as mechanical force (Janson and Dogterom, 2004) or reduction in tubulin concentration (Duellberg *et al.*, 2016a,b), similar to predictions based on simulations in Margolin *et al.* (2012). In contrast, the catastrophes in our data sets occur spontaneously as part of DI; in the in vitro data sets, EB1 and CLASP2γ affect the frequency of catastrophe, but the catastrophes still occur stochastically over time, as opposed to being induced by an experimenter at a particular moment.

The episodes of slow growth in Rickman *et al.* (2017) bear some similarity to stutters interrupting growth as identified by STADIA. However, the slow growth episodes of Rickman *et al.* occurred rarely (two to five occurrences; ~0.26% to ~6.1% of the time analyzed, depending on tubulin concentration). These episodes appear to correspond to the most extreme of our stutters, meaning the stutters with the longest time durations or with the most variability in length during the stutter.

Based on analysis of variability in growth rates in experimental data, Odde *et al.* (1996) proposed a model with multiple substates of growth and "near catastrophes," which are similar to stutters interrupting growth. They suggested that the largest of the "near catastrophes" may correspond to previously observed pauses and that

the smaller "near catastrophes" would not be easily detected by eye (the time between data points in their analysis was ~3 s).

Building beyond this previous work, STADIA provides a comprehensive method for detecting and quantifying multiple types of stutters and distinguishing phase transitions that include stutters from those that do not.

***3.3.3. Differences between STADIA and previous segmentation/classification methods.*** Although STADIA identifies DI phases at a finer scale than many existing DI analysis methods, it differs from methods that simply consider individual displacements between frames and label them as growth, shortening, or pause using thresholds on the length change (e.g., Komarova *et al.*, 2002; Guo *et al.*, 2018). In contrast to such methods, STADIA identifies larger-scale segments during which a MT exhibits a consistent behavior.

Similar to STADIA, many existing time-series analysis methods that have been used in other applications (e.g., identifying runs and pauses in the transport of organelles along MTs by motor proteins [Zaliapin *et al.*, 2005]) involve a segmentation step (e.g., Zaliapin *et al.*, 2003) that is often followed by a classification step (e.g., Fu, 2011). To our knowledge, such methods have not been previously applied to MT DI data. In contrast, many DI analysis methods essentially perform classification before segmentation, by setting thresholds for classifying growth, shortening, and possibly pause or slowdown periods, and then applying the thresholds to identify segments in the data (e.g., Dhamodharan and Wadsworth, 1995; Kiris *et al.*, 2010; Fees *et al.*, 2017). Additionally, unlike existing methods that use predefined thresholds on segment features (length change, time duration, and/or slope), STADIA uses a data-driven approach to identify emergent clusters in the segment feature data (e.g., stutters have shallow slopes, but shallow is relative to the slopes of other segments in a given data set).

***3.3.4. Differences between STADIA and previous phase transition analysis.*** In regard to phase transition analysis, several previous articles grouped their pauses with growth when defining catastrophe and rescue; by their definitions, a catastrophe is a transition from growth or pause to shortening, and a rescue is a transition from shortening to growth or pause (Dhamodharan *et al.*, 1995; Dhamodharan and Wadsworth, 1995; Panda *et al.*, 1996; Rusan *et al.*, 2001; Kamath *et al.*, 2010; Kiris *et al.*, 2010; Yenjerla *et al.*, 2010; Moriwaki and Goshima, 2016). By these definitions or analogous definitions with stutter in place of pause, an episode of interrupted shortening would be labeled as a rescue followed by a catastrophe, whereas an interrupted growth would not be distinguished from *un*-interrupted growth.

STADIA improves upon typical transition analysis by considering all possible transitions between growth, shortening, and stutters (similar to the transitions among growth, shortening, and pause that were considered in Keller *et al.*, 2008). Such transition analysis enables more in-depth investigation of the mechanisms of DI and DI-regulating proteins. For example, the observation that CLASP2γ tends to convert would-be transitional catastrophes into interrupted growths would not have been possible without a method that is able to identify transitional catastrophes and interrupted growths.

## 4. CONCLUSIONS

Our work has four major conclusions: 1) STADIA can quantify and examine "stutters," a previously observed category of behaviors during which MTs undergo slow rates of overall length change compared with growth or shortening phases; 2) stutters are strongly associated with catastrophe in dimer-scale in silico and TIRF-imaged in vitro data; 3) the anticatastrophe factor CLASP2γ reduces catastrophe by increasing the fraction of stutters that return to growth rather than enter shortening phases; 4) STADIA provides an improved analytical tool for quantification of MT behavior, as exemplified by the first three points. Our results concerning the detection of stutters differ from those of previous work in that STADIA comprehensively and systematically identifies all types of stutters (up stutter, flat stutter, down stutter) across length-history data and considers all possible transitions among growth, shortening, and stutters. We suggest that quantification of stutters in future DI analysis through STADIA or similar tools will enable improved analysis of MT dynamics that is more complete, precise, and reproducible. The clearer picture that results from this analysis will facilitate investigation of the mechanisms of catastrophe and rescue as well as the activities of the MTBPs that regulate these transitions.

## 5. MATERIALS AND METHODS

The methods are presented in the following order: Sections 5.1 and 5.2, respectively, describe the acquisition of the in vitro and in silico data sets. Section 5.3 summarizes our classical DI analysis method, used for comparison with STADIA. Section 5.4 outlines our use of STADIA to analyze the data sets in this article. Section 5.5 describes STADIA's analysis procedure in more detail than the overview in *Results* Section 2.2. Section 5.6 provides guidance for users of STADIA.

### 5.1. Data acquisition: in vitro microtubule experiments

The in vitro MT data were obtained from two sets of conditions: a control group (tubulin + EB1) and a group with the MTBP CLASP2γ (tubulin + EB1 + CLASP2γ). A subset of these data was previously published in Lawrence *et al.* (2018).

***5.1.1. Protein preparation.*** His-CLASP2γ and His-EB1 were purified as previously described (Zanic *et al.*, 2013; Lawrence *et al.*, 2018). Bovine brain tubulin was purified using the high-molarity method (Castoldi and Popov, 2003). Tubulin was labeled with TAMRA and Alexa Fluor 488 (Invitrogen) according to the standard protocols, as previously described (Hyman *et al.*, 1991).

***5.1.2. TIRF microscopy.*** Imaging was performed using a Nikon Eclipse Ti microscope with a 100×/1.49 n.a. TIRF objective; an Andor Neo sCMOS (complementary-metal–oxide–semiconductor) camera; 488- and 561- solid-state lasers (Nikon Lu-NA); a Finger Lakes Instruments HS-625 high-speed emission filter wheel; and standard filter sets. An objective heater was used to maintain the sample at 35°C. Microscope chambers were constructed as previously described (Gell *et al.*, 2010). In brief, 22 × 22 mm and 18 × 18 mm silanized coverslips were separated by strips of Parafilm to create a narrow channel for the exchange of solution (Gell *et al.*, 2010). Images were acquired using NIS-Elements (Nikon).

***5.1.3. Dynamic MT assay.*** GMPCPP-stabilized MTs were prepared according to standard protocols (Hyman *et al.*, 1992; Gell *et al.*, 2010). Dynamic MT extensions were polymerized from surface-immobilized GMPCPP-stabilized templates as described previously (Gell *et al.*, 2010). The imaging buffer consisted of BRB80 supplemented with 40 mM glucose, 40 µg/ml glucose oxidase, 16 µg/ml catalase, 0.5 mg/ml casein, 100 mM KCl, 10 mM dithiothreitol, and 0.1% methylcellulose. Purified proteins and 1 mM GTP were added to the imaging buffer, and the solution was introduced into the imaging chamber. Dynamic MTs were grown with 12 µM Alexa 488–labeled tubulin and 200 nM EB1 with or without 400 nM CLASP2γ

and imaged at 2 fps using TIRF microscopy as described above (pixel size of 70 nm). Alexa 488–labeled tubulin was used at a ratio of 23% of the total tubulin.

***5.1.4. In vitro MT length-history data.*** Length-history data for in vitro MTs were obtained from 30-min experiments using both a control group and a group with the stabilizing MTBP, CLASP2γ. Kymographs of dynamic microtubules (examples in Figure 6) were generated using the KymographClear macro for ImageJ, and the dynamic MT tip positions as a function of time were determined in KymographClear, using a thresholding-based, edge-detection method that can trace the microtubule tip position in kymographs with subpixel accuracy (Mangeol *et al.*, 2016). Note that long shortening phases were not well-captured by this process for technical reasons including photobleaching. Therefore, the position-time data from a given MT were broken into samples that typically consisted of a growth phase followed by an initial depolymerization and then termination of that observation (e.g., Figure 1D).

The control group data set was acquired from 68 MT seeds, from which 776 individual traces were observed. The group with CLASP2γ was acquired from 29 MT seeds, from which 85 individual traces were observed. The control group and the group with CLASP2γ yielded total time durations of more than 21 h and 3.5 h, respectively. The in vitro MT lengths were measured in nanometers and then divided by 8 nm per dimer length to convert to units of dimer lengths.

### 5.2. Data acquisition: in silico microtubule experiments

This section outlines the details regarding the acquisition of the dimer-scale simulation MT data, analyzed in the *Results* and in Supplemental Sections S1, S2, and S3, including information about both the model and the parameters used.

***5.2.1. The computational model: stochastic model for simulating 13-protofilament MTs.*** The computational MT model used in this paper to generate the in silico length-history data is an updated version of the detailed, stochastic 13-PF MT model published in Margolin *et al*. (2012) and utilized in Margolin *et al.*, 2011; Gupta *et al.*, 2013; Li *et al.*, 2014; Duan *et al.*, 2017; Mauro *et al.*, 2019; Jonasson *et al.*, 2020). The model tracks the state of individual subunits (representing tubulin dimers bound to either GTP or GDP) in the entire 13-PF MT structure. The events that occur in the model are attachment/detachment of subunits to/from a PF, formation/breaking of lateral bonds between subunits in neighboring PFs, and hydrolysis of GTP subunits to GDP subunits. The values of the biochemical kinetic rate constants for each type of event are user inputs and depend on the state (GTP-bound or GDP-bound) of the subunits involved in the event. To carry out the simulation, the event that occurs at each step and the times between events are sampled using the Gillespie algorithm (Gillespie, 1976, 1977), which is a kinetic Monte Carlo algorithm. At each event, the simulation outputs the time of the event and the length of the MT. The DI behavior, including stutters, and the values of DI parameters are emergent properties that arise as a consequence of the subunit-scale events. This feature is in contrast to two-state growth-shortening DI models, where the four traditional DI parameters are inputs (e.g., negative control in Supplemental Section S4; Verde *et al.*, 1992; Dogterom and Leibler, 1993).

A key difference between the previous versions of our 13-PF MT computational model and the current implementation is strict adherence to the assumption that only one of the many possible biochemical events occurs at a time. The previous detailed-level 13-PF MT model approximated hydrolysis events by allowing several subunits to hydrolyze simultaneously after one of the other four reaction events (lateral bonding/breaking or subunit gain/loss) has occurred. Individual hydrolysis events are now considered as a possible event in the same way that the other events are handled. This modification resulted in very little change in macro-level behavior of in silico MTs, but the ability to output dedicated observations of each dimer-level event provides a more accurate representation of MT biochemistry. The overall result of the simulation is in silico MTs that exhibit macro-level DI behaviors in agreement with those observed previously (Margolin *et al.*, 2012).

***5.2.2. Simulation setup and parameters.*** The dimer-scale kinetic parameters used in this study to simulate a 13-PF MT using the model described above were tuned in Margolin *et al*. (2012) based on in vitro DI measurements from Walker *et al*. (1988); a detailed list of parameters can be found in Supplemental Table S1.2. For the purposes of this analysis, a single MT was simulated at a constant [free tubulin] of 10 μM for 10 h of simulation time. For the kinetic parameters and tubulin concentration used here, approximately 1650 subunit-scale reaction events occurred per second on average over the course of the simulation.

To generate the length-history data passed into STADIA, we used either the max PF length (i.e., the length of the longest of the 13 PFs) or the mean PF length (i.e., the mean of the 13-PF lengths) as the length of the MT. Comparisons of results using the mean or max PF length are shown in Supplemental Figures S1.4, S1.5, S1.6, S1.9, and S1.10. The clustering profiles in Supplemental Figure S1.4 show better agreement with the in vitro data used here when using the max PF length instead of the mean PF length. Thus, all the in silico results are presented for the max PF length unless otherwise indicated. Each dimer has a length of 8 nm. The max and mean PF lengths are both reported in units of dimer lengths; this is not the same as the number of dimers in the MT, which would be 13 times the mean PF length.

### 5.3. Classical DI analysis

For purposes of comparison to STADIA, we used our implementation of classical DI analysis, a custom program written in MATLAB and described in the Supplemental Methods of Jonasson *et al.* (2020). Briefly, this method segments growth and shortening phases by first identifying major peaks and valleys in the length-history data using the MATLAB function "findpeaks." Then the ascent to each major peak is classified as a growth segment, and the descent from the peak is classified as a shortening segment. Each major peak is classified as a catastrophe, where the end of growth and the start of shortening are identified as occurring at the same time point. A major valley is classified as a rescue only if the MT length at the time of the major valley is greater than or equal to a user-defined value called the "minimum rescue length," in which case the end of shortening and the start of growth are identified as the same point. For a major valley that occurs below the minimum rescue length, the end of shortening can be identified as an earlier point in time than the start of growth, in which case the time between these points would correspond to a nucleation period near the MT seed (see Supplemental Methods of Jonasson *et al.*, 2020, for additional details).

For the classical DI analysis in this paper, the minimum prominence for major peaks (i.e., minimum height change between a major peak and the nearest major valley) in the classical method was set equal to the same value that we used for the Maximum Error Tolerance in STADIA (Supplemental Table S1.1). The minimum peak

height and the minimum rescue length in the classical method were each set equal to the sum of the values of the Nucleation Height Threshold plus the Maximum Error Tolerance in STADIA (Supplemental Table S1.1).

In the classical method results shown in Table 1, the $V_{growth}$ and $V_{short}$ calculations relied on linear regressions fitted to each growth or shortening segment. $V_{growth}$ was calculated as the arithmetic mean of the slopes of the regression lines for all growth segments whose linear regression had an $R^2$ value of at least 95%. $V_{short}$ was calculated in the same manner using the shortening segments. $F_{cat}$ was calculated as the total number of catastrophes divided by the total time spent in growth phases. Similarly, $F_{res}$ was calculated as the total number of rescues divided by the total time spent in shortening phases. Note that linear regressions and $R^2$ values are used here in our classical analysis but not in STADIA.

### 5.4. Using STADIA for the analyses in this article

***5.4.1. Data input and preprocessing.*** The simulation data (Section 5.2) were inputted into STADIA as one long length-history time series from an individual MT. For each of the two in vitro data sets (Section 5.1), individual length-history traces (which typically consisted of a growth phase followed by the beginning of a shortening phase; e.g., Figure 1D) were inputted from multiple MTs recorded over shorter periods of observation. As described below (Section 5.5.1.1), when multiple length-history traces are inputted into STADIA, STADIA "stitches" the traces into a single time-series representation, but with separators between the traces to avoid artifactually introducing rescues, catastrophes, or any other transitions. Thus, for our inputted in vitro data, STADIA automatically stitched all of the traces for all of the MTs within each experiment into a single time-series representation (one with CLASP2γ and one without).

***5.4.2. Input parameter values (Supplemental Table S1.1).*** In both Diagnostic and Automated Modes, STADIA analysis requires that the user provide values for the following five user-defined parameters: Minimum Segment Duration, Maximum Error Tolerance, Nucleation Height Threshold, Maximum Height Change Magnitude for Flat Stutters, and Maximum Slope Magnitude for Flat Stutters. The role of each of these parameters in STADIA is further described in Sections 5.5.1 and 5.5.2.1 below.

The Minimum Segment Duration and Maximum Error Tolerance parameters regulate the accuracy of the continuous piecewise linear approximations. For all analyses in the main text and Supplemental Material Section S1, they were set to the following values: Minimum Segment Duration = 0.5 s; Maximum Error Tolerance = 20 dimer lengths. These segmentation parameters were varied over a range of values in Supplemental Material Sections S2 and S3 for the purposes of the sensitivity analysis, which is summarized in Section 2.6.

The Nucleation Height Threshold sets the minimum MT length required for further DI analysis. Segments where the MT length is entirely below the Nucleation Height Threshold are classified as "nucleation" at the beginning of the classification stage and then omitted from analysis thereafter. For all analyses in this article, we set the Nucleation Height Threshold to 75 dimer lengths.

STADIA identifies a segment as a flat stutter if the absolute value of its net height change is less than the user-input Maximum Height Change Magnitude for Flat Stutters and/or the absolute value of its slope is less than the Maximum Slope Magnitude for Flat Stutters. In our analyses, we set the Maximum Height Change Magnitude for Flat Stutters to 3 dimer lengths and the Maximum Slope Magnitude for Flat Stutters to 0.5 dimer lengths/second.

***5.4.3. Determination of the number of clusters (i.e., values of k) for k-means clustering.*** In Automated Mode, STADIA requires that the user provide the number of clusters (i.e., values of *k*) for the *k*-means clustering step. As discussed more in Section 5.5.2, the value of *k* is set separately for the positive and negative slope line segments of each length-history data set and is informed by first running STADIA in Diagnostic Mode. Briefly, the Diagnostic Mode of STADIA outputs gap statistic plots and cluster plots (Supplemental Figures S1.4 and S1.5), which provide information that aids in choosing the optimal number of clusters (i.e., *k*-values) to input into Automated Mode. The gap statistic is a quantity that is calculated at each possible value of *k* to provide a measure of how well the data can be described by *k* clusters. Though there are various ways to interpret gap statistic plots, one rule of thumb is to choose the *k*-value corresponding to the first local maximum of the gap statistic plot (Maechler, 2021) (see also Tibshirani *et al.*, 2001; Hastie *et al.*, 2009, for related information). However, because these plots can have ambiguities, visual examination of the gap statistic plots and cluster plots is useful for interpreting the results in the context of the particular application and thus determining an appropriate number of clusters.

In our analyses of the Diagnostic Mode outputs for each data set, the *k*-value corresponding to the first local maximum of the gap statistic plot was usually chosen as the optimal number of clusters. In particular, for the simulation data, we chose $k = 3$ for each of the positive and negative slope segment groups, as indicated by the first local maximum of the gap statistic plots (Supplemental Figures S1.4C and S1.5C). However, the situation was more complicated for the in vitro data. First, for the negative slope segments in each in vitro data set, we chose $k = 2$, consistent with the first local maximum of the gap statistic plots (Supplemental Figure S1.5, E and G); as discussed in *Results* Section 2.3.1, $k = 2$ for negative slope segments was appropriate for these data sets (in contrast to the $k = 3$ for negative slope segments in the in silico data set) because full depolymerizations to the seed were not captured in the in vitro data sets for technical reasons. For the positive slope segments in the in vitro CLASP2γ data set, we chose $k = 3$, again consistent with the first local maximum of the gap statistic plot (Supplemental Figure S1.4G). However, for positive slope segments in the in vitro control data set, we chose the second local maximum ($k = 3$) instead of the first local maximum ($k = 1$) based on qualitative inspection of the cluster profiles (Supplemental Figure S1.4, E and F). More specifically, the cluster profile for the in vitro control positive slope segments (Supplemental Figure S1.4F) displays three appendages, similar to the cluster profiles for the positive slope segments in the in vitro CLASP2γ data set (Supplemental Figure S1.4H) and the in silico data set (Supplemental Figure S1.4D), where the gap statistic plots indicate $k = 3$ (Supplemental Figure S1.4, C and G).

After choosing the *k*-values based on the results of running STADIA in Diagnostic Mode, we inputted these *k*-values into Automated Mode to perform the full STADIA analysis.

### 5.5. STADIA'S analysis procedure

This section provides an in-depth description of the three major stages of STADIA analysis (Segmentation, Classification, and Phase and Transition Analysis; Figure 2; Supplemental Figure S1.1). For readers interested in a shorter overview of STADIA's analysis procedure, please see *Results* Section 2.2.

***5.5.1. Segmentation stage.*** In the segmentation stage, STADIA takes MT length-history data and generates a continuous piecewise linear approximation of the MT length-history plot (Figure 2, A–D). The approximation is a series of straight-line segments (i.e., the

approximation is "piecewise linear"), where the endpoint of each line segment coincides with the start point of the next line segment (i.e., there are no discontinuities in the approximation). The segmentation stage includes a preprocessing step that prepares the user's length-history data for input into STADIA and a postprocessing step that prepares the results of the segmentation stage for classification.

*5.5.1.1. Preprocessing of input length-history data*. As an initial step, STADIA automatically formats the inputted MT length-history data into a single time series of length-history data points. MT length-history data can be inputted into STADIA either as a long-time observation of a single MT (possible with simulations) or as a series of length histories of multiple MTs (common with experimental data). In the latter case, STADIA automatically connects, or "stitches," the data from multiple MTs (with separators in between) into a single time-series representation of MT length-history data (e.g., Figure 1D). Note that special treatment of the stitching separator between observations allows STADIA to avoid misclassification of stitch boundaries as transitions. This preprocessing step allows STADIA to conduct analysis for both simulation data and experimental data in a similar and consistent manner.

*5.5.1.2. Segmentation process.* STADIA takes the single time-series length-history graph produced by the preprocessing step and performs segmentation as an adaptive, iterative process. As described in this section, how closely the segmentation fits the length-history plot is regulated by two user-defined parameters: Maximum Error Tolerance and Minimum Segment Duration.

The segmentation process begins by identifying major peaks and valleys (i.e., local extrema) in MT length-history data using the "findpeaks" function in MATLAB. The "findpeaks" function uses inputs of minimum peak prominence (i.e., minimum vertical distance between a major peak and nearest major valley) and minimum peak height. The values that STADIA uses for the minimum peak prominence and the minimum peak height in "findpeaks" are the same values, respectively, as the user-input values of the Maximum Error Tolerance and the Nucleation Height Threshold.

Consecutive extrema are connected by line segments to form an initial linear approximation of the length-history data (Figure 2C). An initial list of vertices is defined by these peaks and valleys.

New vertices are added to mark the locations where the MT length crosses the user-input Nucleation Height Threshold, generally chosen to be near the lower limit of observation in experimental conditions. When a MT crosses from below to above the threshold (i.e., a growing MT), the vertex is added at the last data point less than or equal to the Nucleation Height Threshold. When a MT crosses from above to below the threshold (i.e., a shortening MT), the vertex is added at the first data point less than or equal to the Nucleation Height Threshold. At the classification stage described below (Section 5.5.2), segments that are entirely below the Nucleation Height Threshold are excluded from further analysis because these segments are generally not experimentally detectable (note that in our in vitro data sets none of the tracked lengths are below the nucleation threshold).

Then, the iterative process seeks to include new vertices to define line segment endpoints. This improves the approximation accuracy by constructing a continuous piecewise linear approximation that satisfies the user-defined parameters of Maximum Error Tolerance and Minimum Segment Duration mentioned above (Figure 2D). Note that the segmentation algorithm implemented in STADIA is similar, but not identical, to the "top-down" category of algorithms reviewed in Keogh *et al*. (2001).

STADIA's segmentation algorithm can be explained in the following steps:

1. Let $\{x^1, x^2, \ldots, x^N\}$ represent the initial list of vertices (i.e., segment endpoints), where $x^1$ and $x^N$ are the first and last points of the length-history data, respectively, and $\{x^2, \ldots, x^{N-1}\}$ are the consecutive peaks, valleys, and nucleation threshold points described above.
2. For any $i = 1, \ldots, N-1$, define the $i$th region as the interval between the consecutive pair of initial vertices, $x^i$ and $x^{i+1}$. Construct a line segment with endpoints as $x_1^i = x^i$ and $x_2^i = x^{i+1}$ such that the vertices corresponding to the $i$th region are $\{x_1^i, \cdots, x_M^i\}$, where initially $M = 2$, but we seek to grow this list in the following steps.
3. For $j = 1 \ldots, M-1$, consider the $j$th line segment in the $i$th region defined by $x_j^i$ and $x_{j+1}^i$. Calculate the error (absolute value of the difference) between this line segment and the corresponding points in the original length-history data.
   - If the maximum error is greater than the user-defined Maximum Error Tolerance, then the error condition is not satisfied, and an additional data point needs to be included in the vertex list. Proceed to step 4.
   - If the maximum error from this segment is less than the user-defined Maximum Error Tolerance, then the error condition is satisfied for the $j$th line segment in the $i$th region. Proceed to step 6.
4. Define the data point where the greatest error occurs in step 3 as $x_{new}^i$
   - If $x_{new}^i$ violates the user-defined Minimum Segment Duration, attempt to choose the closest point in the length-history data that would satisfy both the Maximum Error Tolerance and Minimum Segment Duration.
5. Include the newly identified vertex into the list of vertices for the $i$th region. This will require reindexing to preserve ordering. For example, for the first new vertex added to the $i$th region, the original single segment in the $i$th region is now broken into two segments, and the list of vertices corresponding to the $i$th region is now defined as

   $$\{x_1^i, x_2^j, x_3^j\} = \{x_1^i, x_{new}^i, x_2^j\}$$

   where the vertex list on the right side is indexed according to the preceding iteration and the updated vertex list on the left side replaces the list defined in step 2, such that $x_j^i < x_{j+1}^i$ for all $j = 1, \ldots, 1, M-1$.
6. Repeat steps 3–5 until the error condition is satisfied without adding more vertices into the $i$th region.
7. Repeat steps 2–6 for all $i \leq N-1$.

The final result is a continuous piecewise linear approximation of the inputted length-history data set (excerpts of the full length-history approximation are illustrated in Figure 2D, orange lines, and Figure 2, G and J, black lines). The vertices of the piecewise linear approximations provide line segments with endpoints at moments where significant changes in slope occur in length-history plots. Thus, the activity covered by each segment between endpoints represents a consistent period of MT length-history behavior that can be identified as belonging to a DI phase in the classification stage.

*5.5.1.3. Justification for segmentation method.* To create a more accurate approximation of MT length-history data as compared with more classical methods that identify segment endpoints only at peaks and valleys (Figure 2, A–C), STADIA employs the iterative approach described above in Section 5.5.1.2 to create an improved continuous piecewise linear approximation of the MT length-history data. The resulting approximation satisfies the user-defined Maximum Error Tolerance and Minimum Segment Duration (Figure 2D). We chose this approach because it provides a simple method for identifying points that may not necessarily be peaks or valleys, but where a change from one sustained MT behavior to another occurs. Through the Maximum Error Tolerance choice, the user is able to regulate the accuracy of the linear approximation. Through the Minimum Segment Duration choice, the user is able to perform the analysis of MT length-history data at the desired timescale. An assumption of performing segmentation in this manner is that MT behavior follows a linear trend at the timescale being analyzed. Finally, we note that this segmentation method in STADIA produces a *continuous* piecewise linear approximation, whereas some other segmentation methods produce *discontinuous* approximations (e.g., Zaliapin *et al.*, 2003).

*5.5.1.4. Postprocessing to prepare for classification.* For each line segment from the continuous piecewise linear approximation, STADIA measures the slope, time duration, and height change of the segment (Figure 2D); this set of measurements provides a 3-D feature space where the segments reside (Figure 2E).

*5.5.1.5. Justification for using all three of slope, time duration, and height change in the classification feature space.* Mathematically, knowing the values of any two of the segment variables (time duration, height change, and slope) provides sufficient information to calculate the value of the remaining third variable. However, we use all three variables in the clustering step because some data points that are well separated in the 3-D space would become indistinguishable for all practical purposes if only two of the variables were used (Supplemental Figures S1.2 and S1.3). Additionally, which data points become indistinguishable would depend on which pair of variables was used (time duration and height change, time duration and slope, or height change and slope).

The slope = height/time surface (Supplemental Figures S1.2A and S1.3E) could be parameterized with only two variables in a way that would maintain the separation present in the 3-D space. However, these two new variables would be some combination of the original three variables, and these combinations would not necessarily have clear physical meanings. We therefore chose to use all three of the basic variables (time duration, height change, and slope) to maintain a more direct connection to the biology.

The inclusion of nonlinear combinations of variables (i.e., interaction terms) is not uncommon in statistics (e.g., Rawlings *et al.*, 1998; Karaca-Mandic *et al.*, 2012; Matuschek and Kliegl, 2018). Additional combinations of our three basic variables as well as other variables may be worth exploring in future work that aims to further dissect MT length-history behaviors. For the purposes of the present work, the three basic variables are sufficient for verifying the existence of distinguishable clusters within the positive and negative slope groups.

***5.5.2. Classification stage.*** The purpose of the classification stage in STADIA is to group the segments from the segmentation stage into subsets that share similar attributes. In the classification procedures, each segment from the approximation of the MT length-history data is represented as a point in the 3-D space generated by segment time duration, height change, and slope (Figure 2, D and E). The classification stage is where differences arise between the two modes of STADIA: Diagnostic Mode aids the user in selecting the number of clusters to use but ends after the clustering step, which is described below in Section 5.5.2.2; Automated Mode requires that the number of clusters be provided as input but performs all other stages of the analysis.

*5.5.2.1. Classification first steps—identification of nucleation segments and flat stutters.* First, segments that are entirely below the user-input Nucleation Height Threshold described above (Section 5.5.1.2) are classified as "nucleation." These nucleation segments are excluded from further analysis and therefore are excluded from the 3-D plots of segment features (e.g., Figure 3).

Next, STADIA identifies any segments that satisfy either or both of the following criteria:

- the absolute value of the segment's net height change is less than the user-input Maximum Height Change Magnitude for Flat Stutters;
- the absolute value of the segment's slope is less than the Maximum Slope Magnitude for Flat Stutters.

These near-flat segments clearly lack the qualities characteristic of traditionally recognized growth or shortening and thus already qualify as a subset of points that share attributes different from the remaining points requiring classification. Therefore, STADIA assigns them into a class labeled "flat stutters."

We remark that in comparison to the up stutters and down stutters that are identified by the next step of classification (Section 5.5.2.2), flat stutters are relatively rare, in terms of both number of segments and total time spent in each type of segment (Supplemental Figure S1.9). Thus, flat stutters account for only a small share of all stutter behaviors detected.

Removing flat stutters from the rest of the collection of points creates a clear boundary between points that represent positive and negative slope segments. However, we do not simply label the remaining segments as growth and shortening. Instead, further analysis is warranted for two reasons. First, attempting to execute the rest of the classification procedures on the positive and negative slope segments together fails to produce conclusive results (Supplemental Figure S1.6), suggesting that the positive and negative groups should be analyzed separately. Second, complex geometric structures of distinguishable appendages observed in both the positive and negative slope point groups (Figure 3, B and C) suggest that multiple types of behaviors are present within each subset.

*5.5.2.2. k-means clustering step of the classification stage.* To continue the classification stage, STADIA takes the segments that are now segregated into positive and negative slope line segments and analyzes them using *k*-means clustering (Macqueen, 1967; Lloyd, 1982), where the number of clusters, the *k*-value, is suggested by the gap statistic (Tibshirani *et al.*, 2001).

Justification for using k-means clustering. As an unsupervised clustering algorithm commonly used in machine learning, *k*-means does not require prior knowledge of the characteristics of the clusters to be found in order for the algorithm to identify boundaries that separate them. Rather, *k*-means groups together data points that share similar characteristics (i.e., data points that are near each other in a relevant feature space). The *k*-means algorithm also has

the advantages of its ease of use and interpretability. Ideal data sets for *k*-means have globularly shaped clusters (i.e., each cluster would follow a Gaussian distribution). Although the clusters resulting from our data are not Gaussian per se, *k*-means still provides an objective methodology to find substructures in the overall data structure. The observation that *k*-means enables us to identify and quantify stutters (behaviors that have been noted previously but not quantified in detail) indicates that it provides a useful methodology for categorization and quantification of MT behavior.

**Preprocessing of segment data for input into k-means clustering.** *k*-means clustering uses Euclidean distance (i.e., straight-line distance) between points in the feature space (3-D space for our data) as the primary measurement in its algorithm to classify data. Therefore, all features should exist on the same scale to give each feature equal weight in the *k*-means classification process. To meet this requirement, the segment features (slope, height change, and time duration values) are transformed by first being log-scaled and then standardized with respect to each feature's statistics (i.e., by subtracting the mean and dividing by the SD) (Figure 3, B and C). Scaling and standardizing the data in this way is a common practice for analysis utilizing *k*-means clustering (Hastie *et al*., 2009).

**Determining appropriate number of clusters for each data set.** As noted in Section 2.2.1, one of the goals for STADIA development was that it be impartial in determining the number of behaviors exhibited by MTs, thus avoiding any assumptions about MT dynamics being restricted to two behaviors (i.e., only growth and shortening). The *k*-value (i.e., number of clusters to use in *k*-means) is determined for positive and negative slope segments separately and is informed by running the Diagnostic Mode of STADIA.

Though various approaches exist for determining the *k*-value with which to perform the clustering (reviewed by Pham *et al*., 2005; Steinley, 2006), STADIA utilizes a quantity called the gap statistic, which is calculated at each potential value of *k* (Supplemental Figures S1.4 and S1.5, left column) (Tibshirani *et al*., 2001). The gap statistic aids in answering the question, "what number of clusters results in the best separation between the clusters?" More technically, the gap statistic measures the within-cluster dispersion compared with a null reference distribution.

When examining the values of the gap statistic at different values of *k* to seek the optimal number of clusters that best separates the data, higher values of the gap statistic indicate better separation between clusters. However, a significant increase in the value of the gap statistic is generally considered necessary to justify using an additional cluster. Tibshirani *et al*. (2001) formalized this idea with the following criterion: choose the smallest value of *k* such that

$$\text{Gap}(k) \geq \text{Gap}(k+1) - \text{one standard error of Gap}(k+1).$$

In words, this criterion means choose the smallest value of *k* such that the value of the gap statistic does not increase by more than one standard error when going to the next value of *k*. Other possible criteria include choosing the first local maximum of the gap statistic plot or the smallest *k*-value such that the gap value is within one standard error of the first local maximum (Hastie *et al*., 2009; Maechler, 2021). Depending on the particular data set, the different criteria may yield the same *k*-value as each other or different *k*-values.

The Diagnostic Mode of STADIA outputs the *k*-value chosen by the Tibshirani *et al*. (2001) criterion. However, when choosing *k*-values to input into Automated Mode, it is also recommended for the user to examine the gap statistic plots and cluster profiles (Supplemental Figures S1.4 and S1.5) to check how well the number of clusters suggested by the gap statistic describes the data set qualitatively. For example, in some cases, qualitative inspection of the data may suggest that the second local maximum of the gap statistic plot describes the data better than the first local maximum (e.g., as seen in Supplemental Figure S1.4. E and F).

**Measuring the gap statistic in Diagnostic Mode.** For the purposes of informing the optimal *k*-value for use in *k*-means clustering, the Diagnostic Mode of STADIA repeats the clustering procedure for each potential value of *k* ranging from 1 through 12, using 100 random starts for each value (a single run of *k*-means clustering does not necessarily converge to a global optimum, so multiple starts are required to determine optimal centroid locations). Using the clustering results at each *k*-value, STADIA measures the value of the gap statistic for each value of *k* (Supplemental Figures S1.4 and S1.5).

**k-means clustering in Automated Mode.** As noted above, the purpose of *k*-means clustering is to group together data points that share similar characteristics (i.e., data points that are near each other in the feature space of segment slope, height change, and time duration). Once the optimal number of clusters is determined for both positive and negative slope segments using the Diagnostic Mode of STADIA, the user inputs these *k*-values and runs STADIA in Automated Mode. In Automated Mode, STADIA performs *k*-means clustering, on the positive and negative slope segments separately, using 500 random starts. Centroid locations that attain the lowest sum of squared distances between the centroids and each point in their respective clusters are chosen for further analysis. The chosen centroid locations are indicted by x-symbols in the cluster plots (e.g., Figure 3, D and E).

*5.5.2.3. Phase/behavior bundling step of the classification stage.* After *k*-means clustering is performed on the log-transformed and standardized data, the resulting cluster assignments are applied to the original segment data (i.e., the data before applying log-transformation and standardization; Supplemental Figure S1.7). Statistics such as average slopes, average time duration, and average height change are calculated for each cluster (slopes in Figure 4, C and F; slopes and time durations in Supplemental Figure S1.8) and then utilized for naming the clusters (Figure 3G). Clusters with similar average slopes are bundled together to form larger groups, which we refer to as "phase classes" or "behavior classes" (Figures 2H and 4I). Groups of clusters with large positive slopes are classified as growth, while those with large negative slopes are classified as shortening. The remaining clusters with segment slopes considerably smaller in magnitude (i.e., flatter) are grouped into the category of behaviors called "stutters" (along with the "flat stutters," which were separated out before the clustering process).

At this point, every segment identified during the segmentation stage has been classified as growth, shortening, stutter, or nucleation. Applying these phase class labels to each segment in the length-history plot is illustrated in Figure 2, G and J.

***5.5.3. Phase and transition analysis stage.*** After classifying segments into clusters and then bundling the clusters into larger phase/behavior classes as described above (Section 5.5.2), classical methods of calculating DI metrics are adapted to account for stutters in addition to growth and shortening.

*5.5.3.1. Phase analysis.* For each cluster, STADIA calculates the average velocity of the segments in the cluster (Table 1, bottom row of each subtable). STADIA also calculates the following cluster attributes (Supplemental Figure S1.9):

- total number of segments (counts obtained from the piecewise linear approximation) in each cluster,
- percent time spent in each cluster $= \left( \frac{\text{sum of segment time durations in cluster}}{\text{total time of data set}} \times 100\% \right)$,
- and percent height change corresponding to each cluster $= \left( \frac{\text{sum of segment height change magnitudes in cluster}}{\text{sum of all segment height change magnitudes}} \times 100\% \right)$.

These attributes can be determined for each of the larger phase/behavior classes (i.e., growth, shortening, stutters) by combining the measurements for the clusters in each class (Figure 5A).

*5.5.3.2. Transition analysis.* Transition frequencies are calculated in a manner similar to what has been done classically. However, when considering stutters in addition to growth and shortening, there are additional transitions to quantify (Figure 2I). In particular, it is necessary to determine whether catastrophes and rescues are or are not directly preceded by stutters. Catastrophes and rescues are identified as either abrupt (occurring without detectable stutters) or transitional (occurring via a stutter) (Figures 5–7). Additionally, our analysis quantifies interrupted growth (growth → stutter → growth) (Figures 5H, 6, and 7, H and I) and interrupted shortening (shortening → stutter → shortening) (Figure 5I).

As mentioned above in Section 5.5.2.1, MTs shorter than the user-defined Nucleation Height Threshold are considered to be in "nucleation" phases. Transitions into or out of nucleation phases are not analyzed by the current version of STADIA because such MTs would be difficult to detect in experiments, and their behavior might be influenced by proximity to the seed.

In agreement with what has been done in classic DI analyses, frequencies of catastrophe and rescue are calculated as the ratio of the number of catastrophe or rescue events to the total time spent in growth or shortening, respectively (Table 1). For the additional types of transitions identified by STADIA (Figure 2I), the frequencies are calculated in a similar manner: the frequency of each type of transition out of growth or shortening is calculated as the ratio of the number of transition events of that type to the total time spent in growth or shortening, respectively (Supplemental Figure S1.10). More specifically,

$$F_{\text{AbruptCatastrophe}} = \frac{\text{\# of abrupt catastrophes}}{\text{total time spent in growth}},$$

$$F_{\text{TransitionalCatastrophe}} = \frac{\text{\# of transitional catastrophes}}{\text{total time spent in growth}},$$

$$F_{\text{InterruptedGrowth}} = \frac{\text{\# of growth interruptions}}{\text{total time spent in growth}},$$

$$F_{\text{AbruptRescue}} = \frac{\text{\# of abrupt rescues}}{\text{total time spent in shortening}},$$

$$F_{\text{TransitionalRescue}} = \frac{\text{\# of transitional rescues}}{\text{total time spent in shortening}},$$

$$F_{\text{InterruptedShortening}} = \frac{\text{\# of shortening interruptions}}{\text{total time spent in shortening}}.$$

The total $F_{cat}$ equals $F_{\text{AbruptCatastrophe}} + F_{\text{TransitionalCatastrophe}}$ (Figure 7C; Supplemental Figure S1.10), and the total $F_{res}$ equals $F_{\text{AbruptRescue}} + F_{\text{TransitionalRescue}}$ (Supplemental Figure S1.10). Similarly, the total frequency of growth-to-stutter transitions equals $F_{\text{TransitionalCatastrophe}} + F_{\text{InterruptedGrowth}}$ (Figure 7C), and the total frequency of shortening-to-stutter transitions equals $F_{\text{TransitionalRescue}} + F_{\text{InterruptedShortening}}$.

### 5.6. Guidance for users: expectations for input data and effect of thresholds

STADIA is ideally intended for use on data sets with moderate or high temporal resolution, for example, at least 2 fps. For lower resolution data sets, we suggest that STADIA will provide more systematic analysis than manual methods, but the resolution of the data themselves will be a limiting factor in what conclusions can be supported.

We expect that the most common difficulty will be obtaining a total amount of data that is large enough for effective clustering during the classification stage. The clustering process performs better as the amount of data increases; more specifically, determining the optimal number of clusters and assigning segments to the appropriate cluster is done more accurately when there are more data points in the segment feature space (e.g., Figure 3). The total time duration of MT length-history data required will generally be on the order of hours, not minutes. To determine whether one has a sufficient quantity of data, we recommend two possible tests. First, users should examine the error bars in the gap statistic plots generated by Diagnostic Mode; if the error bars are too large to conclusively choose an optimal *k*-value, then more data may be needed. Second, we suggest that users run STADIA on their entire data set and on half of their data set; if both cases yield similar clustering results, this indicates that the user has a sufficient quantity of data. If one has an insufficient amount of data for effective clustering, STADIA can still be used to perform segmentation, detection of flat stutters with user-defined parameters, and clustering with $k = 1$ (one cluster each for positive and negative slope segments), because these analyses do not depend on the number of data points in the feature space; on the contrary, the number of data points in the feature space depends on the number of segments in the segmentation. However, DI metrics resulting from sparse data sets should be treated with caution.

As one specific example of the amount of data needed, in the in silico results presented here, we used 10 h of simulation time to ensure that enough segments were generated for effective clustering. Testing different total time durations of data yielded consistent results for simulations that ran for 7.5 h or longer when using STADIA in Diagnostic Mode. However, using significantly shorter length-history data sets (e.g., 2.5 h) did not provide acceptable clustering results. On the other hand, if the number of clusters (i.e., *k*-values) is preestablished (e.g., from a similar but larger data set), then STADIA can be used in Automated Mode to calculate DI metrics from significantly fewer data (e.g., at least 2.5 h).

It is important for users to be aware that the values of inputted thresholds will affect the numerical values of results of STADIA (as well as any other DI analysis method; e.g., Odde *et al.*, 1996; Gierke *et al.*, 2010; Matov *et al.*, 2010; Smal *et al.*, 2010; Prahl *et al.*, 2014; Guo *et al.*, 2018). For examples of the effects of changing these values, see the analyses with varied values of the Minimum Segment Duration and Maximum Error Tolerance in STADIA and the data acquisition rate of the length-history data as shown in Supplemental Sections S2 and S3. We recommend that users try at least a few different values of thresholds to test the strength of any conclusions they draw. In articles using STADIA, users should report the values

of the input parameters that they use in STADIA, in addition to reporting the resolution of their measurements, quantity of data, and values of any other relevant quantities.

At the segmentation stage, users should examine the piecewise linear approximation to ensure that the approximation is not overfitting or underfitting the raw data. The user's choices for the values of the Minimum Segment Duration and Maximum Error Tolerance determine how closely the piecewise linear approximation will fit the raw length-history data. When choosing the values of these thresholds, the user should take into account the resolution and noise level of their data as well as the timescale of the dynamics that the user wishes to study. For example, there are small-amplitude stochastic fluctuations that occur within growth, shortening, and stutter segments; if the user is studying phases at a scale similar to what we study in this article, which is a larger scale than the small-amplitude fluctuations, then the Minimum Segment Duration and Maximum Error Tolerance should not be so small as to pick up these fluctuations.

Note that for certain combinations of the Minimum Segment Duration and Maximum Error Tolerance, STADIA will produce "irreconcilable errors." These errors occur because it is not always possible to satisfy both the Minimum Segment Duration and the Maximum Error Tolerance. In such cases, STADIA outputs a warning to the user for each error. Such errors are most likely to occur if the user has chosen a long Minimum Segment Duration with a small Maximum Error Tolerance. The specific values of Minimum Segment Duration and Maximum Error Tolerance that result in irreconcilable errors will depend on the particular data set being analyzed. If such errors occur, the user should either change the parameter values or recognize that some segments of the piecewise linear approximation will not meet the input criteria.

If the user is aiming to identify one set of input parameter values or a small number of parameter sets that are ideal for their particular data set, then we recommend that the user choose input parameter values that minimize the number of irreconcilable errors. Our parameter sensitivity analysis (Section 2.6 and Supplemental Material Sections S2 and S3) indicates that the number of irreconcilable errors is more sensitive to the Maximum Error Tolerance than to the Minimum Segment Duration. For our in silico data set, the percentage of segments that have irreconcilable errors has a local minimum at Maximum Error Tolerance = 20 dimer lengths. The percentage of segments that have irreconcilable errors is also low for Maximum Error Tolerance > 40 dimer lengths but is very high for Maximum Error Tolerance < 15 dimer lengths. If the user is performing a parameter sensitivity analysis with a large range of parameter values (similar to the range used in Supplemental Material Sections S2 and S3), then the user should be aware that some parameter combinations may result in a large number of irreconcilable errors.

For further instructions on how to use the STADIA MATLAB code, we refer readers to Patel *et al*. (2020). Note that the input parameter called the "Minimum Segment Duration" here was referred to as the "minimum time step" in Patel *et al*. (2020).

### Software and data availability

STADIA software (MATLAB code) and tutorials can be downloaded from GitHub (https://github.com/GoodsonLab/STADIA/). Data analyzed in this paper are available from the authors upon request.

## ACKNOWLEDGMENTS

This work was supported by National Science Foundation (NSF) grants MCB-1244593 to H.V.G. and M.A., MCB-1817966 to H.V.G., and MCB-1817632 to E.M.J., National Institutes of Health (NIH) grant R35GM119552 to M.Z., and National Institutes of Health Integrated Biological Systems Training in Oncology training grant T32CA119925 to E.J.L. M.Z. also acknowledges the Searle Scholars Program. Portions of the work were also supported by funding from the University of Massachusetts Amherst (A.J.M.), NSF-GFRP DGE-1313583 (K.S.M.), and a fellowship from the Dolores Zohrab Liebmann Fund (S.M.M.). We thank the members of the Goodson laboratory and the Chicago Cytoskeleton community for their insightful discussions.

## REFERENCES

Aher A, Kok M, Sharma A, Steinmetz MO, Dogterom M, Akhmanova A (2018). CLASP suppresses microtubule catastrophes through a single TOG Domain. Dev Cell 46, 40–58.

Andrecka J, Arroyo JO, Lewis K, Cross RA, Kukura P (2016). Label-free imaging of microtubules with sub-Nm precision using interferometric scattering microscopy. Biophys J 110, 214–217.

Applegate KT, Besson S, Matov A, Bagonis MH, Jaqaman K, Danuser G (2011). PlusTipTracker: quantitative image analysis software for the measurement of microtubule dynamics. J Struct Biol 176, 168–184.

Blackwell R, Sweezy-Schindler O, Edelmaier C, Gergely ZR, Flynn PJ, Montes S, Crapo A, Doostan A, McIntosh JR, Glaser MA, *et al*. (2017). Contributions of microtubule dynamic instability and rotational diffusion to kinetochore capture. Biophys J 112, 552–563.

Castle BT, Odde DJ (2013). Brownian dynamics of subunit addition-loss kinetics and thermodynamics in linear polymer self-assembly. Biophys J 105, 2528–2540.

Castoldi M, Popov AV (2003). Purification of brain tubulin through two cycles of polymerization–depolymerization in a high-molarity buffer. Protein Expr Purif 32, 8388.

Coombes CE, Yamamoto A, Kenzie MR, Odde DJ, Gardner MK (2013). Evolving tip structures can explain age-dependent microtubule catastrophe. Curr Biol 23, 1342–1348.

Desai A, Mitchison TJ (1997). Microtubule polymerization dynamics. Annu Rev Cell Dev Biol 13, 83–117.

Dhamodharan R, Jordan MA, Thrower D, Wilson L, Wadsworth P (1995). Vinblastine suppresses dynamics of individual microtubules in living interphase cells. Mol Biol Cell 6, 1215–1229.

Dhamodharan R, Wadsworth P (1995). Modulation of microtubule dynamic materials and in vivo by brain microtubule associated proteins. J Cell Sci 108(Pt 4), 1679–1689.

Dillon GM, Tyler WA, Omuro KC, Kambouris J, Tyminski C, Henry S, Haydar TF, Beffert U, Ho A (2017). CLASP2 links reelin to the cytoskeleton during neocortical development. Neuron 93. 1344–1358.

Dogterom M, Leibler S (1993). Physical aspects of the growth and regulation of microtubule structures. Phys Rev Lett 70, 1347–1350.

Doodhi H, Prota AE, Rodríguez-García R, Xiao H, Custar DW, Bragsten K, Katrukha EA, Hilbert M, Hua S, Jiang K, *et al*. (2016). Termination of protofilament elongation by eribulin induces lattice defects that promote microtubule catastrophes. Curr Biol 26, 1713–1721.

Duan AR, Jonasson EM, Alberico EO, Li C, Scripture JP, Miller RA, Alber MS, Goodson HV (2017). Interactions between Tau and different conformations of tubulin: implications for Tau function and mechanism. J Mol Biol 429, 1424–1438.

Duellberg C, Cade NI, Holmes D, Surrey T (2016a). The size of the EB cap determines instantaneous microtubule stability. eLife 5, e13470.

Duellberg C, Cade NI, Surrey T (2016b). Microtubule aging probed by microfluidics-assisted Tubulin washout. Mol Biol Cell 27, 3563–3573.

Fees CP, Estrem C, Moore JK (2017). High-resolution imaging and analysis of individual astral microtubule dynamics in budding yeast. J Vis Exp 2017, 55610.

Fu TC (2011). A review on time series data mining. Engineering Applications of Artificial Intelligence 24, 164–181.

Gardner MK, Charlebois BD, Jánosi IM, Howard J, Hunt AJ, Odde DJ (2011). Rapid microtubule self-assembly kinetics. Cell 146, 582–592.

Gell C, Bormuth V, Brouhard GJ, Cohen DN, Diez S, Friel CT, Helenius J, Nitzsche B, Petzold H, Ribbe J, *et al*. (2010). Microtubule dynamics reconstituted in vitro and imaged by single-molecule fluorescence microscopy. Methods Cell Biol 95, 221–245.

Gierke S, Kumar P, Wittmann T (2010). Analysis of microtubule polymerization dynamics in live cells. Methods Cell Biol 97, 15–33.

Gildersleeve RF, Cross AR, Cullen KE, Fagen AP, Williams RC (1992) Microtubules grow and shorten at intrinsically variable rates. J Biol Chem 267, 7995–8006.

Gillespie DT (1976). A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. J Comput Phys 22, 403–434.
Gillespie DT (1977). Exact stochastic simulation of coupled chemical reactions. J Phys Chem 81, 2340–2361.
Girão H, Okada N, Rodrigues TA, Silva AO, Figueiredo AC, Garcia Z, Moutinho-Santos T, Hayashi I, Azevedo JE, Macedo-Ribeiro S, *et al*. (2020). CLASP2 binding to curved microtubule tips promotes flux and stabilizes kinetochore attachments. J Cell Biol 219, e201905080.
Goodson HV, Jonasson EM (2018). Microtubules and microtubule-associated proteins. Cold Spring Harb Perspect Biol 10, a022608.
Guo Y, Li Di, Zhang S, Yang Y, Liu J-J, Wang X, Liu C, Milkie DE, Moore RP, Tulu US, *et al*. (2018). Visualizing intracellular organelle and cytoskeletal interactions at nanoscale resolution on millisecond timescales. Cell 175, 1430–1442.
Gupta KK, Li C, Duan A, Alberico EO, Kim OV, Alber MS, Goodson HV (2013). Mechanism for the catastrophe-promoting activity of the microtubule destabilizer Op18/Stathmin. Proc Natl Acad Sci USA 110, 20449–20454.
Hastie T, Tibshirani R, Friedman J (2009). Springer Series in Statistics: The Elements of Statistical Learning, 2nd ed., New York: Springer Science+Business Media, LLC.
Howard J, Hyman AA (2009). Growth, fluctuation and switching at microtubule plus ends. Nat Rev Mol Cell Biol 10, 569–574.
Hyman A, Drechsel D, Kellogg D, Salser S, Sawin K, Steffen P, Wordeman L, Mitchison T (1991). Preparation of modified tubulins. Methods Enzymol 196, 478–485.
Hyman AA, Salser S, Drechsel DN, Unwin N, Mitchison TJ (1992). Role of GTP hydrolysis in microtubule dynamics: information from a slowly hydrolyzable analogue, GMPCPP. Mol Biol Cell 3, 1155–1167.
Jánosi IM, Chrétien D, Flyvbjerg H (2002). Structural microtubule cap:s, catastrophe, rescue, and third state. Biophys J 83, 1317–1330.
Janson ME, Dogterom M (2004). Scaling of microtubule force-velocity curves obtained at different tubulin concentrations. Phys Rev Lett 92, 248101.
Jonasson EM, Mauro AJ, Li C, Labuz EC, Mahserejian SM, Scripture JP, Gregoretti IV, Alber M, Goodson HV (2020). Behaviors of individual microtubules and microtubule populations relative to critical concentrations: dynamic instability occurs when critical concentrations are driven apart by nucleotide hydrolysis. Mol Biol Cell 31, 589–618.
Kamath K, Oroudjev E, Jordan MA (2010). Determination of microtubule dynamic instability in living cells. Methods Cell Biol 97, 1–14.
Kapoor V, Hirst WG, Hentschel C, Preibisch S, Reber S (2019). MTrack: automated detection, tracking, and analysis of dynamic microtubules. Sci Rep 9, 3794.
Karaca-Mandic P, Norton EC, Dowd B (2012). Interaction terms in nonlinear models. Health Serv Res 47, 255–274.
Keller PJ, Pampaloni F, Lattanzi G, Stelzer EHK (2008). Three-dimensional microtubule behavior in xenopus egg extracts reveals four dynamic states and state-dependent elastic properties. Biophys J 95, 1474–1486.
Keogh E, Chu S, Hart D, Pazzani M (2001). An online algorithm for segmenting time series. In: Proceedings—IEEE International Conference on Data Mining, ICDM, 289–296.
Kiris E, Ventimiglia D, Feinstein SC (2010). Quantitative analysis of MAP-mediated regulation of microtubule dynamic instability in vitro-focus on tau. Methods Cell Biol 95, 481–503.
Komarova YA, Vorobjev IA, Borisy GG (2002). Life cycle of MTs: persistent growth in the cell interior, asymmetric transition frequencies and effects of the cell boundary. J Cell Sci 115, 3527–3539.
Kruse R, Krantz J, Barker N, Coletta RL, Rafikov R, Luo M, Højlund K, Mandarino LJ, Langlais PR (2017). Characterization of the CLASP2 protein interaction network identifies SOGA1 as a microtubule-associated protein. Mol Cell Proteomics 16, 1718–1735.
Lawrence EJ, Arpag G, Norris SR, Zanic M (2018). Human CLASP2 specifically regulates microtubule catastrophe and rescue. Mol Biol Cell 29, 1168–1177.
Lawrence EJ, Zanic M (2019). Rescuing microtubules from the brink of catastrophe: CLASPs lead the way. Curr Opin Cell Biol 56, 94–101.
Li C, Li J, Goodson HV, Alber MS (2014). Microtubule dynamic instability: the role of cracks between protofilaments. Soft Matter 10, 2069–2080.
Lloyd SP (1982) Least squares quantization in PCM. IEEE Trans Inf Theory 28, 129–136.
Macqueen J (1967). Some methods for classification and analysis of multivariate observations. In: Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, Volume 1: Theory of Statistics, 5.1, 281–297. Berkeley and Los Angeles, CA: University of California Press.
Maechler M (2021). R: gap statistic for estimating the number of clusters, 2021. https://stat.ethz.ch/R-manual/R-devel/library/cluster/html/clusGap.html.
Mahrooghy M, Yarahmadian S, Menon V, Rezania V, Tuszynski JA (2015). The use of compressive sensing and peak detection in the reconstruction of microtubules length time series in the process of dynamic instability. Comput Biol Med 65, 25–33.
Majumdar S, Kim T, Chen Z, Munyoki S, Tso S-C, Brautigam CA, Rice LM. (2018) An isolated CLASP TOG domain suppresses microtubule catastrophe and promotes rescue. Mol Biol Cell 29, 1359–1375.
Maly IV (2002). Diffusion approximation of the stochastic process of microtubule assembly. Bull Math Biol 6, 213–238.
Mangeol P, Prevo B, Peterman EJG (2016). KymographClear and KymographDirect: two tools for the automated quantitative analysis of molecular and cellular dynamics using kymographs. Mol Biol Cell 27, 1948–1957.
Margolin G, Goodson HV, Alber MS (2011). Mean-field study of the role of lateral cracks in microtubule dynamics. Phys Rev E 83, 041905.
Margolin G, Gregoretti IV, Cickovski TM, Li C, Shi W, Alber MS, Goodson HV (2012). The mechanisms of microtubule catastrophe and rescue: implications from analysis of a dimer-scale computational model. Mol Biol Cell 23, 642–656.
Matov A, Applegate K, Kumar P, Thoma C, Krek W, Danuser G, Wittmann T (2010). Analysis of microtubule dynamic instability using a plus-end growth marker. Nat Methods 7, 761–768.
Matuschek H, Kliegl R (2018). On the ambiguity of interaction and nonlinear main effects in a regime of dependent covariates. Behav Res Methods 50, 1882–1894.
Maurer SP, Cade NI, Bohner G, Gustafsson N, Boutant E, Surrey T (2014). EB1 accelerates two conformational transitions important for microtubule maturation and dynamics. Curr Biol 24, 372–384.
Mauro AJ, Jonasson EM, Goodson HV (2019). Relationship between dynamic instability of individual microtubules and flux of subunits into and out of polymer. Cytoskeleton 76, 21557.
McIntosh RJ, O'Toole E, Morgan G, Austin J, Ulyanov E, Ataullakhanov F, Gudimchuk N (2018). Microtubules grow by the addition of bent guanosine triphosphate tubulin to the tips of curved protofilaments. J Cell Biol 217, 2691–2708.
Mickolajczyk KJ, Geyer EA, Kim T, Rice LM, Hancock WO (2019). Direct observation of individual tubulin dimers binding to growing microtubules. Proc Natl Acad Sci USA 116, 7314–7322.
Mitchison T, Kirschner M (1984). Dynamic instability of microtubule growth. Nature 312, 237–242.
Moriwaki T, Goshima G (2016). Five factors can reconstitute all three phases of microtubule polymerization dynamics. J Cell Biol 215, 357–368.
Odde DJ, Buettner HM, Cassimeris L (1996). Spectral analysis of microtubule assembly dynamics. AIChE J 42, 1434–1442.
Odde DJ, Cassimeris L, Buettner HM (1995). Kinetics of microtubule catastrophe assessed by probabilistic analysis. Biophys J 69, 796–802.
Panda D, Jordan MA, Chu KC, Wilson L (1996). Differential effects of vinblastine on polymerization and dynamics at opposite microtubule ends. J Biol Chem 271, 29807–29812.
Patel RJ, Murray KS, Martin PO, Sinclair M, Scripture JP, Goodson HV, Mahserejian SM (2020). Using STADIA to quantify dynamic instability in microtubules. Methods Cell Biol 158, 117–143.
Pedigo S, Williams RC (2002). Concentration dependence of variability in growth rates of microtubules. Biophys J 83, 1809–1819.
Pham DT, Dimov SS, Nguyen CD (2005). Selection of K in K-means clustering. Proc Inst Mech Eng Part C J Mech Eng Sci 219, 103–119.
Portran D, Schaedel L, Xu Z, Théry M, Nachury MV (2017). Tubulin acetylation protects long-lived microtubules against mechanical ageing. Nat Cell Biol 19, 391–398.
Prahl LS, Castle BT, Gardner MK, Odde DJ (2014). Quantitative analysis of microtubule self-assembly kinetics and tip structure. Methods Enzymol 540, 35–52.
Rawlings J, Pantula S, Dickey D (eds.) (1998). Applied Regression Analysis. Applied Regression Analysis (Second Edition), New York, NY: Springer-Verlag.
Rickman J, Duellberg C, Cade NI, Griffin LD, Surrey T (2017). Steady-state EB cap size fluctuations are determined by stochastic microtubule growth and maturation. Proc Nat Acad Sci USA 114, 3427–3432.
Rusan NM, Fagerstrom CJ, Yvon AMC, Wadsworth P (2001). Cell cycle-dependent changes in microtubule dynamics in living cells expressing green fluorescent protein-α tubulin. Mol Biol Cell 12, 971–980.
Sammak PJ, Borisy GG (1988). Direct observation of microtubule dynamics in living cells. Nature 332, 724–726.

Schek HT, Gardner MK, Cheng J, Odde DJ, Hunt AJ (2007). Microtubule assembly dynamics at the nanoscale. Curr Biol 17, 1445–1455.
Schulze E, Kirschner M (1988). New features of microtubule behaviour observed in vivo. Nature 334, 356–359.
Smal I, Grigoriev I, Akhmanova A, Niessen WJ, Meijering E (2010). Microtubule dynamics analysis using kymographs and variable-rate particle filters. IEEE Trans Image Process 19, 1861–1876.
Sousa A, Reis R, Sampaio P, Sunkel CE (2007). The *Drosophila* CLASP homologue, mast/orbit regulates the dynamic behaviour of interphase microtubules by promoting the pause state. Cell Motil Cytoskeleton 64, 605–620.
Steinley D (2006). K-means clustering: a half-century synthesis. Br J Math Stat Psychol 59, 1–34.
Tibshirani R, Walther G, Hastie T (2001). Estimating the number of clusters in a data set via the gap statistic. J R Statist Soc Series B Methodol 63, 411423.
Toso RJ, Jordan MA, Farrell KW, Matsumoto B, Wilson L (1993) Kinetic stabilization of microtubule dynamic instability in vitro by vinblastine. Biochemistry 32, 1285–1293.
Tran PT, Walker RA, Salmon ED (1997). A metastable intermediate state of microtubule dynamic instability that differs significantly between plus and minus ends. J Cell Biol 138, 105–117.
VanBuren V, Cassimeris L, Odde DJ (2005). Mechanochemical model of microtubule structure and self-assembly kinetics. Biophys J 89, 2911–2926.
Verde F, Dogterom M, Stelzer E, Karsenti E, Leibler S (1992). Control of microtubule dynamics and length by cyclin a- and cyclin B-dependent kinases in xenopus egg extracts. J Cell Biol 118, 1097–1108.
Walker RA, O'Brien ET, Pryer NK, Soboeiro MF, Voter WA, Erickson HP, Salmon ED (1988). Dynamic instability of individual microtubules analyzed by video light microscopy: rate constants and transition frequencies. J Cell Biol 107, 1437–1448.
Waterman-Storer CM, Salmon ED (1997). Actomyosin-based retrograde flow of microtubules in the lamella of migrating epithelial cells influences microtubule dynamic instability and turnover and is associated with microtubule breakage and treadmilling. J Cell Biol 139, 417–434.
Yenjerla M, Lopus M, Wilson L (2010). Analysis of dynamic instability of steady-state microtubules in vitro by video-enhanced differential interference contrast microscopy with an appendix by emin oroudjev. Methods Cell Biol 95, 189–206.
Zakharov P, Gudimchuk N, Voevodin V, Tikhonravov A, Ataullakhanov FI, Grishchuk EL (2015). Molecular and mechanical causes of microtubule catastrophe and aging. Biophys J 109, 2574–2591.
Zaliapin I, Gabrielov A, Keilis-Borok V (2003). Multiscale trend analysis. Fractals 12, 275–292.
Zaliapin I, Semenova I, Kashina A, Rodionov V (2005). Multiscale trend analysis of microtubule transport in melanophores. Biophys J 88, 4008–4016.
Zanic M (2016). Measuring the effects of microtubule-associated proteins on microtubule dynamics in vitro. In: The Mitotic Spindle, New York: Humana Press, 47–61.
Zanic M, Widlund PO, Hyman AA, Howard J (2013). Synergy between XMAP215 and EB1 increases microtubule growth rates to physiological levels. Nat Cell Biol 15, 688–693.
Zwetsloot AJ, Tut G, Straube A (2018). Measuring microtubule dynamics. Essays Biochem 62, 725–735.