Multifractality Nature of Microtubule Dynamic Instability Process

Vahid Rezania¹, Ferry C. Sudirga¹, and Jack A. Tuszynski^{3,4,5,*}

¹ Department of Physical Sciences, MacEwan University, Edmonton, Alberta, T5J 4S2, Canada.

³ Department of Physics, University of Alberta, Edmonton, AB, T6G 2E1

⁴ Department of Oncology, University of Alberta, Edmonton, Alberta, T6G 1Z2, Canada

⁵ DIMEAS, Politecnico di Torino, Turin, Italy

rezaniav@macewan.ca

sudirgaf@maymacewan.ca

jack.tuszynski@gmail.com (*corresponding author)

Abstract

The irregularity of growing and shortening patterns observed experimentally in microtubules reflects a dynamical system that fluctuates stochastically between assembly and disassembly phases. The observed time series of microtubule lengths have been extensively analyzed to shed light on structural and dynamical properties of microtubules. Here, for the first time, Multifractal Detrended Fluctuation analysis (MFDFA) has been employed to investigate the multifractal and topological properties of both experimental and simulated microtubule time series. We find that the time dependence of microtubule length possesses true multifractal characteristics and cannot be described by monofractal distributions. Based on the multifractal spectrum profile, a set of multifractal indices have been calculated that can be related to the level of dynamical activities of microtubules. We also show that the resulting multifractal spectra for the simulated data might not be comparable with experimental data.

Statement of Significance

Microtubules are some of the most important subcellular structures involved in a multitude of functions in all eukaryotic cells. In addition to their cylindrical geometry, their polymerization/depolymerization dynamics, termed dynamic instability, is unique among all protein polymers. In this paper we demonstrate that there is a very specific mathematical representation of microtubule growth and shrinkage time series in terms of multifractality. We further show that using this characteristic, one can distinguish real experimental data from synthetic time series generated from computer simulations.

Introduction

In a dynamical system, complexity arises when its subunits interact non-linearly with each other. Existence of a critical threshold, irregularity and sudden phase transitions are most common features of many complex systems. Among many biological systems with intrinsic complexities, dynamic instability of microtubules (MTs) can be ranked very high as it has been the subject of numerous investigations for decades [1-3].

Microtubules are cylindrical-shaped polymers of tubulin (a family of globular proteins) found in the cytoplasm of eukaryotic cells and are a component of the cytoskeleton [4]. Microtubules are crucially involved in many cellular activities, including cell divisions (meiosis and mitosis), trafficking, and motility. Not surprisingly, dysfunctionalities of microtubules are correlated with a number of diseases, including various cancers, Alzheimer disease, and Parkinson disease [5]. Tubulin proteins that make up a microtubule assemble and disassemble themselves spontaneously in a stochastic manner, leading to a non-equilibrium process commonly referred to as microtubule's dynamic instability [6].

Microtubules are formed by polymerization of its subunits, namely α and β tubulin heterodimers, into a cylindrical form with a hollow space inside which is called the lumen. In humans and most mammals, several distinct isotypes of α and β tubulins have been identified and their respective amino acid sequences mapped. Isotype variations within the β tubulin class are rather well understood in vertebrates where there are six main isotypes of β tubulins (I, II, III, IV, V, and VI) expressed by separate genes. These different β isotypes have been associated with different functional roles in eukaryotic cells [7]. Rezaei et al [8] listed some of the key biophysical properties of various human β -tubulin such as their electrostatic charge, dipole moment, volume, surface area as well as specifically the properties of their C-terminal tails which are crucially important for both protein binding and microtubule assembly processes. Interestingly, even though these physical characteristics of β tubulin isotypes are quite similar, microtubule stochastic behaviour seems to change notably depending on its β tubulin isotype composition [8].

Dynamic instability of microtubules is driven by a competition between assembly and disassembly states of tubulin dimers at the microtubule ends. The so-called plus-end (distal) is more dynamic than the minus-end (proximal). Assembly, or growth, requires the presence of two GTP (Guanine Tri-Phosphate) nucleotides bound to an $\alpha\beta$ heterodimer subunit. Tubulin-bound GTP is prone to hydrolyzation into GDP (Guanine Di-Phosphate) when bound to the β monomer while it does not hydrolyze when bound to the α -monomer. GDP-bound tubulin dimers are prone to disassemble. GTP-bound tubulins at the growing end of the microtubule appear to protect the entire structure from disassembly forming a so-called GTP-cap, as the microtubule's inner GDP-bound tubulins do not seem to spontaneously depolymerize. When GTP hydrolysis catches up to the tip of the microtubule, however, the entire structure is prone to a rapid disassembly called a catastrophe [10].

Earlier statistical studies on microtubule data have mainly focused on analyzing length histograms of microtubules in order to shed light on their probability distributions [8,11]. As shown in these studies, histograms of MT length distributions have an exponential long tail and a peak corresponding to relatively short microtubules that suggests a characteristic of Poissonian processes. Furthermore, the resulting inverse proportionality of the frequency of catastrophes

with respect to the growth velocity of a microtubule can be interpreted as an indication of the protective presence of a GTP cap at the growing end of a microtubule [12].

Recent studies employed different statistical methods for modeling stochastic signals such as compressed sensing, maximum likelihood, and hidden Markov models (HMMs) to capture the underlying signal variability [13]. Compressed sensing has been used to reconstruct signals using fewer samples than the Nyquist rate. The maximum likelihood technique has been applied to estimate parameters for a linear dynamic system from the observed data. Many wavelet-based hidden Markov processes have been also involved to capture real-world non-Gaussian signals [14].

Among the non-Gaussian variations, fractal behaviors have been found in biomedical signals from a wide range of physiological phenomena such as ECG, EEG, MR, and X-ray images [15]. Identifying scale-invariant variations not only can be used to differentiate between healthy and pathological conditions, but also to distinguish various types of unhealthy conditions. Biomedical time series such as interspike interval of neuron firing, stride interval of human walking, inter-breath interval of human respiration, and interbeat intervals of the human heart have all shown scale-invariant characteristics that can serve as a diagnostic tool [16-20]. In addition, scale invariant structures have also been found in spatial phenomena such as the branching of lungs, the vascular system and the nervous system [21-24], bone structure [25], and their properties such as the fractal dimension are able to differentiate between healthy and cancer tissues [26]. Furthermore, it has been proposed that changes in the scale-invariant structure of biomedical signals may result from the adaptability of physiological processes to new conditions [27]. As a result, a successful treatment of pathological conditions might change disturbed fractal structure and improve health. Therefore, fractal analysis of data structures can be used as a promising prognostic and diagnostic tool in biomedical signal processing.

Furthermore, a common theoretical approach to study stochastic properties of MTs is based on either the two-state model with growing-shrinking, or the three-state model with growing-pause-shrinking [1-3]. Effectively, such models are analogous to a 1D drift-diffusion random walk, where the drift velocity, v_{drift} , and diffusion constant, D , depend on the four main microscopic parameters: growth (v_+), shrinkage (v_-), catastrophe (f_c) and rescue (f_r) rates:

$$P(x, t) = \frac{P_0}{\sqrt{2\pi Dt}} e^{-\frac{(x - v_{\text{drift}} t)^2}{4Dt}}, \quad x \in (-\infty, \infty) \quad (1)$$

where $P(x, t)$ is the probability density function and

$$v_{\text{drift}} = \frac{f_c v_+ - f_r v_-}{f_c + f_r}, \quad D = \frac{v_+ v_-}{f_c + f_r}.$$

As a result, $\langle x \rangle = v_{\text{drift}} t$ and $\langle x^2 \rangle = v_{\text{drift}}^2 t^2 + 2Dt \sim t^2$ at large t [28]. Many physical and mathematical models have been proposed to shed light on the MT dynamic instability. Recently, Jonasson et al. [29] proposed a new model by introducing two different critical tubulin concentrations to provide a more unified explanation for many experimental observations of MT dynamic instability. Using their results, one can calculate the equivalent four parameters of the two-state model, i.e. (v_+ , v_- , f_c , f_r). In addition, Cassimeris et al. [30] have also simulated an

array of MTs using a Monte Carlo model to study the role of phase transitions and tubulin concentration on MT dynamic instability and MT length distribution. In a completely different approach, Rezanian and Tuszynski studied MT polymerization and depolymerization processes by considering a semi-classical method that is often used in quantum mechanics [31]. They showed that the dynamics of a MT can be mathematically expressed via a 3D cubic-quintic nonlinear Schrodinger equation.

In this paper, however, we aim to investigate whether dynamic instability processes involving microtubules can be modelled using a fractal stochastic approach. We implement this approach on both experimental data, using *in-vitro* raw data primarily obtained from [8] as well as simulated data generated by methods described in [29], [30] and [31]. This provides a new comparison methodology between experiments and simulations.

Here, we adopted the *multifractal detrended fluctuation analysis* (MFDFA) method described in [32] to perform a statistical analysis of MT length data to shed light on the dynamics occurring during polymerization/depolymerisation processes. In addition, we used MT data assembled from two different isotypically-purified β -tubulins, namely $\alpha\beta_{II}$ and $\alpha\beta_{III}$ in order to understand how different isotype compositions may affect MT dynamics.

Fractal signals are usually divided into two categories: monofractal and multifractal structures. In general, the monofractal structures are those whose variations can be scaled by a single power law exponent that is independent of time and space. However, spatial and temporal variations often appear in the scale-invariant structure of a biomedical signal. As a result, these types of data cannot be scaled by a single exponent. These spatial and temporal variations indicate a multifractal structure of the biomedical signal that is defined by a multifractal spectrum of power law exponents. The multifractal spectrum identifies the deviations in fractal structure within time periods with large and small fluctuations [32-34]. Such variabilities have been observed in heart rate signals where the width and shape of the spectrum can also differentiate between patients with heart diseases such as ventricular tachycardia, ventricular fibrillation and congestive heart failure [35-36]. The multifractal structure analysis of EEG and series of interspike intervals has also been able to differentiate between the neural activities of brain areas [18].

Figure 1 shows selected experimental time series data for the lengths of microtubule isotypes $\alpha\beta_{II}$ (top panel) and $\alpha\beta_{III}$ (middle panel). The lower panel demonstrates a simulated uniformly random data set over the same time sample as MT data for comparison purposes. By a simple observation from the following sample data one can conclude that:

- (i) microtubules usually spend most of their time in the growing phase rather than in the shortening phase [6,8].
- (ii) changes in length are trending, suggesting the possibility of a positive serial correlation [6].
- (iii) the shortening phase can be fast and drastic.
- (iv) volatility of length changes does not seem to be constant; i.e. volatility clustering and relatively frequent sharp jumps are observed.

Hence, modelling the stochastic behaviour of microtubules requires a different framework and might not be captured fully by the 1D drift-diffusion random walk models. One can explain the aforementioned behaviour by using a model that involves fractality, a pattern of self-affinity that shows up both at larger and smaller scales. Some fractals are deterministic, such as the Mandelbrot set and the Koch snowflake that are defined by iterative equations [37]. In Euclidian geometry, the length of an object can be defined as the limit of a sequence of approximations. For example, if we construct a polygon and increase the number of its sides, the length of the polygon converges to a limit, which is the circumference of the circle surrounding it. However, in fractal geometry, this might not be the case. Consider the Koch snowflake, which is made of an infinite number of self-affine structures. The perimeter length of the snowflake converges to an infinite number. Empirically, fractality can be quantified by the fractal dimension d_F , which is a measure of fractal “roughness of shape” [37].

Fractals can also be stochastic. For example, almost no two snowflakes are the same. One measure of stochasticity in fractals is to calculate the Hurst exponent H introduced by Hurst [38], see the Methods section. For noisy time series the Hurst exponent will be in the interval between 0 and 1 whereas for a random walk type time series, it will be above 1. A Hurst exponent of 0.5 indicates a purely random process, or white noise. That means that a corresponding time series has an independent or short-range dependent structure. A process with $H > 0.5$, however, implies a positive persistence or a correlation between the past trend and future behavior. This means that if a variable was increasing in the past, it should continue to increase. In such a process, time series differences have some tendency to trend together: positive movements are more likely to be followed by positive movements and *vice versa*. $H < 0.5$ indicates a negative persistence or anti-correlated behavior resulting in a dampening process. The closer the H value is to 0, the rougher or more jagged the trend looks like.

Hence, a fractal model with $H > 0.5$ can explain the trending patterns of microtubule growth and why microtubules tend to spend more time in the growing phase [6]. One relatively simple example of a fractal model is the Fractional Brownian Motion developed by Mandelbrot and van Ness [39].

Fractal analysis can capture some behaviors observed in microtubule data but not all. Specifically, the volatility clustering behavior observed in the microtubule data can only be explained if the stochastic process is multifractal. This is a more complex multi-layered fractal structure (i.e. fractals of fractals), in which different segments of the time series exhibit localized behaviors, driven by different H values for different statistical moments [40]. This is unlike the Fractional Brownian Motion process, in which H is constant for all statistical moments, i.e. a monofractal stochastic process. A model data set from a multifractal stochastic process is charted below. Note that the cumulative time series exhibits localized patterns that are similar to the longer series and that volatility ebbs and flows throughout the time series.

Methodology

By its very nature, empirical data that come from a (hypothesized) multifractal process would be non-stationary and many conventional statistical procedures cannot be applied to such data. Kantelhardt et al. [41] pioneered a rather ‘computationally friendly’ method called the Multifractal Detrended Fluctuation Analysis (MFDFA) that has been used extensively by

subsequent researchers [32, 42]. This research utilizes MFDFA algorithms as configured by [32]. The MFDFA method for a time series $\{x_i\}$, $i = 1, \dots, N$ proceeds as follows:

1. Integrate the time series to create the profile:

$$Y(i) = \sum_{k=1}^i (x_k - \langle x \rangle), \quad i = 1, 2, \dots, N \quad (2)$$

where $\langle x \rangle$ is the average value of $\{x_i\}$.

2. Divide the profile $Y(i)$ into $N_s = N/s$, non-overlapping segments of equal length s . Then, calculate the local trend for each segment by fitting the data using a least-squares fit, $y_v(i)$, with either linear, quadratic, or higher polynomial orders.
3. Detrend each segment $v = 1, 2, \dots, N_s$ by subtracting the best fit, $y_v(i)$, from the data. Then compute the mean-squared variance:

$$F^2(s, v) = \frac{1}{s} \sum_{i=1}^s \{Y[(v-1)s + i] - y_v(i)\}^2 \quad (3)$$

where $y_v(i)$ is the polynomial fitted to the segment $v(i)$.

4. Average all segments to obtain the q -order fluctuation function

$$F_q(s) = \left\{ \frac{1}{N_s} \sum_{v=1}^{N_s} [F^2(s, v)]^{\frac{q}{2}} \right\}^{\frac{1}{q}} \quad (4)$$

where q is a real number. It is interesting to note that for positive q , $F_q(s)$ measures large fluctuations, while for negative q , $F_q(s)$ measures small fluctuations in data.

5. Determine the localized generalized Hurst exponent H_q at each q -order by repeating the above calculation to obtain the fluctuation function $F_q(s)$ for a different box size s :

$$F_q(s) \sim s^{H_q} \Rightarrow \log(F_q(s)) \sim H_q \log(s) \quad (5)$$

Hence, the generalized Hurst exponent, H_q , is the fluctuation parameter that describes the correlation structures of the time series at different magnitudes and can be empirically quantified by running a regression of $\log F_q(s)$ vs $\log s$ for each q -order. In practice, quantification becomes more inaccurate as q becomes more positive or negative. Here, H_q is calculated for $-5 \leq q \leq 5$, as suggested in [32]. Note that H_2 (i.e. the Hurst exponent at $q = 2$) is identical to the basic Hurst exponent of a monofractal stochastic process [6, 43].

Note that the value of the Hurst exponent at the zeroth order, H_0 , cannot be determined by using Eq. (4) because of the diverging exponent. The logarithm averaging procedure should be used [41, 44]:

$$F_0(s) = \exp \left\{ \frac{1}{4N_s} \sum_{v=1}^{2N_s} \ln[F^2(s, v)] \right\} \sim s^{H_0} \quad (6)$$

Information contained in the relationship between H_q and q can be further transformed into useful statistical parameters as defined below [41]:

$$\tau_q = q H_q - 1, \quad (7)$$

$$\alpha_q = \frac{d\tau_q}{dq}, \quad (8)$$

$$D_q = q \alpha_q - \tau_q \quad (9)$$

Here, τ_q is called the q -order mass exponent and can be used to quantify the scaling properties of the data. α_q is the q -order singularity exponent that represents local Hurst exponents and D_q is the q -order dimension or the singularity spectrum. Each local Hurst exponent, α_q , quantifies the local singular behavior and thus relates it to the local scaling of the series. The plot of D_q as a function of α_q is also called the multifractal spectrum. For white noise time series, the mass exponent τ_q in Eq. (7) is a linear function of the q -order. This linear relationship of τ_q leads to a constant α_q , the first derivative of τ_q , as expected (see Figure 3). The nonlinear dependence of the mass exponent τ_q with the q -order, however, signifies the multifractality in the time series that can be decomposed into many subsets exhibiting different local Hurst exponents, α_q . The resulting multifractal spectrum will be usually a large arc where the difference between the maximum and minimum α_q is called the multifractal spectrum width, also called *the degree of multifractality*, $\Delta\alpha_q = \max(\alpha_q) - \min(\alpha_q)$, see Figure 3. The multifractal spectrum width is an important parameter; a wider spectrum implies more intense volatility clustering and extreme jumps in data values, something that Mandelbrot [45] referred to as wilder randomness. Conversely, in the random walk systems, the multifractal spectrum width is much smaller than in the multifractal systems, congruent with their properties of constant volatility [32].

In general, the multifractality in time series arises either from: (i) fat-tailed probability distributions of increments, compared with a Gaussian distribution, or from (ii) different long-range temporal correlations for small and large fluctuations [42, 46]. To identify the contribution of either of these sources, two procedures of shuffling and phase randomization have been used in the literature. These procedures are usually employed to study the degree of complexity of time series to distinguish different sources of multifractality in the time series. The type (i) can be tested by phase randomized surrogates that weaken the non-Gaussian effects in the data. The procedure preserves the amplitudes of the Fourier transform but randomizes the Fourier phases that eliminate nonlinearities while keeping only the linear properties of the original time series [46]. In the shuffling procedure, the values are put in random order. As a result, the type (ii) can be tested by shuffling as it erases the long-range temporal correlation for small and large fluctuations and conserves the overall distribution of the data. In other words, through the shuffling procedure the distribution function of the original data remains exactly the same but without memory. This leads to the q -order singularity spectrum, D_q , reaching a maximum at $\alpha_q \sim 0.5$. Furthermore, if multifractality in the time series is due to a broad distribution function, α_q obtained by the surrogate method will be independent of q . However, if both kinds of multifractality are present in time series, the shuffled and surrogate series will show a weaker multifractality strength than the original one [41]. Here, we apply these methods to MT time series data and compare with the original data.

Results and Discussions

Experimental data: Here, we studied experimental microtubule data based on tubulin samples extracted from bovine brains as described in [8]. There are 26 data sets of $\alpha\beta_{II}$ and 30 data sets of $\alpha\beta_{III}$ tubulin isotypes. It is necessary to note that the availability of samples was limited, which constrains our analysis. In general, the growing and shortening dynamics data for individual MTs at their plus ends were recorded by differential interference contrast video microscopy. Data points representing microtubule lengths were collected as a time series with 2-6 s intervals for as long as 10 minutes at a time. This means each microtubule data set only provided 120-250 data points. As suggested by Ihlen [32], however, ideally more than 1,000 points per sampling might be needed for robustness of the analysis.

Due to limited sample sizes, we increase the number of sampling rates by using Matlab resampling utility that implements an anti-aliasing lowpass filter. It should also be noted that sampling rates were uneven (i.e. time intervals between data points are not fixed), possibly due to some manual snap-shooting of microtubule lengths during the experiments.

As stated previously, Figure 1 demonstrates a selected raw experimental length data of microtubule isotypes $\alpha\beta_{II}$ (top panel) and $\alpha\beta_{III}$ (middle panel). For comparison purposes, uniformly random (Gaussian) data over the same time sample as microtubule data have been generated (lower panel). Resampled data have not been shown.

Figure 2 shows the difference data, $\Delta x_i = x_{i+1} - x_i$, $i = 2, \dots, N$, for each data set in Figure 1 after the resampling procedure. The results are plotted in terms of the sample number rather than time for better comparison. The root-mean-squared (RMS) value is only sensitive to differences in the amplitude of variation and not differences in the structure of variation. Notice the extreme amplitudes in both $\alpha\beta_{II}$ and $\alpha\beta_{III}$ results that are 3-4 times of their RMS values in comparison with the white noise with no such changes. As expected, these extreme changes are due to catastrophes in the corresponding data sets.

To utilize the multifractality of the data set, the multifractal spectrum is plotted and describes the singularity D_q as a function of α_q . The typical procedure to graph the multifractal spectrum in the MFDFA method is to first convert multifractal Hurst exponent H_q to the q -order mass exponent τ_q , and then convert τ_q to the q -order singularity exponent (or local Hurst exponent) α_q , and finally to q -order singularity dimension D_q [41], see Figure 3. To determine the multifractal Hurst exponents for different orders q , the minimum window length of 5 data points that varies in the range of 10–20 s was used. The maximum window length that we used here is $\text{int}(N/5)$ where N denotes the total number of data points in each data set. A total of 20 successive windows have been chosen that sets the length of each successive window increases by a factor of $\sim 2^{1/8}$ [32]. The multifractal Hurst exponents, H_q , were determined by varying q in the range, -5 to +5 for all three data sets.

Figure 3 shows the q -dependence of H_q , α_q , τ_q , and D_q for the selected $\alpha\beta_{II}$ and $\alpha\beta_{III}$ data sets as well as the simulated white noise for comparison purposes. As expected, the white noise time series depicts a linear relationship of the mass exponent τ_q with q -orders. The nonlinearity shape of τ_q reveals a multifractal behavior of the time series. For a monofractality feature that appears,

the function τ_q would be a linear function of q -orders with a constant slope. The linear q -dependence of τ_q leads to a constant α_q that reduces the multifractal spectrum to a small arc for the white noise time series (see Figure 3d). Similar behavior to that for white noise data is also expected for a monofractal data set. Furthermore, the average value of the Hurst exponent for the white noise is about 0.52, which is expected and implies a *short-range* dependence in the time series. The mass exponents τ_q for the $\alpha\beta_{II}$ and $\alpha\beta_{III}$ data sets, however, show a curved q -dependence and, consequently, a decreasing singularity exponent α_q . The resulting multifractal spectra represent a large arc where the difference between the maximum and minimum α_q is called the multifractal spectrum width, $\Delta\alpha_q$, see arrow in Figure 3(d). Figure 4 shows the generalized Hurst exponent and multifractal spectrum for all $\alpha\beta_{II}$ and $\alpha\beta_{III}$ data sets. The Hurst exponent for both $\alpha\beta_{II}$ and $\alpha\beta_{III}$ data sets converges to ~ 1.5 , which signifies the non-Gaussian behavior of the data sets. In general, the Hurst exponent indicates for a time series by how fast the overall RMS of local fluctuations RMS grows with an increasing segment sample size (i.e., scale). The fast changing fluctuations in the time series will influence the overall RMS, for segments with small sample sizes (i.e., small scale) whereas slow changing fluctuations will influence it for segments with large sample sizes (i.e., large scale). The larger Hurst exponent, H , is usually translated as more slowly-evolving variations or a more persistent structure in multifractal time series. Furthermore, the Hurst exponent greater than 0.5 describes a *long-range* dependence and memory effects on all time scales according to the level of persistence. As H approaches 1 or higher, the time series becomes increasingly periodic [42].

Figures 3 and 4 show that even though both $\alpha\beta_{II}$ and $\alpha\beta_{III}$ data sets show multifractality characteristics, there are still subtle differences between them. To measure the differences among these spectra, a set of multifractality indices has been calculated. These indices can also be related to the level of dynamical activities of microtubules. Following [42], the first index derived for a multifractal spectrum is the *degree of multifractality* of the spectrum, $\Delta\alpha_q$

$$\Delta\alpha_q = \max(\alpha_q) - \min(\alpha_q) \quad (11)$$

The wider $\Delta\alpha_q$ corresponds to the richer and stronger multifractality in data. On average, the degree of multifractality for $\alpha\beta_{II}$ and $\alpha\beta_{III}$ spectra is found to be 2.441 and 2.731, respectively.

Furthermore, in an ideal case, the shape of the multifractal spectrum, D_q , is symmetric. However, for real-world data, this is not necessarily symmetric. The asymmetry in a spectrum is related to the volatility of the data set and is likely to be a genuine result and not merely a data or methodology artifact [47]. The left side of the multifractal spectrum D_q is affected by the positive q values, which filter out larger fluctuations, and the vice versa. As a result, asymmetry in D_q reveals non-uniformity, indicating that the corresponding large and small-scale fluctuations are driven by different mechanisms [47]. If the spectrum is symmetric, it means there is a balance between the mechanisms that determine each side of the spectrum. The equilibrium will be broken when one of the fluctuations (small or large) is dominant.

The *degree of asymmetry*, A , or the skewness in the shape of the spectrum can be calculated as [42]:

$$A = \frac{\max(\alpha_q) - \alpha_{q_0}}{\alpha_{q_0} - \min(\alpha_q)} \quad (12)$$

where α_{q_0} is the value of α_q when D_q reaches its maximum. This spectrum is symmetric if $A = 1$, left-skewed for $0 < A < 1$, and right-skewed when $A > 1$. For the $\alpha\beta_{II}$ and $\alpha\beta_{III}$ spectra, the skewness values are found to be 2.114 and 2.749, respectively. On average, the $\alpha\beta_{III}$ isotype series appear to have a bolder right-skewed asymmetry (relatively more segments of low volatility than segments of high volatility) while the $\alpha\beta_{II}$ isotype series appear to have a modest left-skewed asymmetry, see Figure 4.

In addition, the multifractal spectrum can also have either a left or right truncation [32]. The truncations basically measure a leveling of the q -order Hurst exponent for negative or positive q values, respectively. The truncation is estimated by the *singularity ratio* C that is defined by the ratio between $\Delta D_q^{\text{left}}(q)$ of the left side of the spectrum and $\Delta D_q^{\text{right}}(q)$ of the right side of the spectrum calculated in relation to the maximum fractal dimension, $D_q^{\text{max}}(q)$:

$$C = \frac{\Delta D_q^{\text{left}}(q)}{\Delta D_q^{\text{right}}(q)} \quad (13)$$

where

$$\begin{aligned} \Delta D_q^{\text{left}}(q) &= 1 - \min \{D_q^{\text{left}}(q)\} \\ \Delta D_q^{\text{right}}(q) &= 1 - \min \{D_q^{\text{right}}(q)\} \end{aligned}$$

Here, $\min \{D_q^{\text{left}}(q)\}$ and $\min \{D_q^{\text{right}}(q)\}$ are minimum values of D_q in the left and right part of the spectrum, respectively. In this sense, for the singularity index $C < 1$, the right-hand side is truncated, while for $C > 1$ the truncation occurs on the left-hand side. Interestingly, a long left tail indicates that the time series has a multifractal structure that is insensitive to local fluctuations with small magnitudes. Conversely, a long right tail implies that the time series has a multifractal structure that is insensitive to local fluctuations with large magnitudes. For the $\alpha\beta_{II}$ and $\alpha\beta_{III}$ spectra presented here, the average singularity ratios are 0.8595 and 0.6736, respectively.

In order to identify the source of multifractality whether it is due to type (i) the broad (fat-tail) probability density function or type (ii) the long memory correlations in the small and large fluctuations, we calculated the multifractal indices for the original MT data sets and the resampled ones, the shuffled and surrogated data sets. Figure 5 demonstrates the averaged generalized q -order Hurst exponent and multifractal spectrum for the original (black curve), shuffled (red curve) and surrogated (blue curve) of $\alpha\beta_{II}$ and $\alpha\beta_{III}$ data sets. As shown, the multifractal spectrum D_q of $\alpha\beta_{II}$ is fairly symmetric in comparison with that of $\alpha\beta_{III}$. This suggests the large and small-scale fluctuations contribute more or less equally in $\alpha\beta_{II}$ data. Furthermore, the result illustrates that both shuffling and surrogate applications reduce the multifractality strength of the original time series that demonstrates that both the long-range correlation and the fat-tail distribution contribute to multifractality. The multifractality is, however, much less pronounced in the shuffled time series than original and surrogated time series. As a result, one can say that the long-range correlations among fluctuations have major contribution in the observed multifractality of the MTs' time series. This can also be seen by

calculating various multifractality parameters that are summarized in Table 1. As shown, except for the degree of asymmetry, A , the $(h_{q0}, H_2, \Delta\alpha_q, C)$ -quadruplets in the original time series are greater than their shuffled and surrogated partners, i.e. $(h_{q0}, H_2, \Delta\alpha_q, C)_{\text{original}} > (h_{q0}, H_2, \Delta\alpha_q, C)_{\text{surrogated}} > (h_{q0}, H_2, \Delta\alpha_q, C)_{\text{shuffled}}$. The nonzero spectrum width of the shuffled time series, $\Delta\alpha_q^{\text{shuffled}} > 0$, shows a weak multifractal effect in the shuffled time series. This implies that while we have removed the correlations in the original series, the multifractality still remains in the shuffled time series. This highlights that the multifractality of MT data arises from both types with the main contribution from type (i), as the shuffled time series shows much less a pronounced multifractality than the original and surrogated time series [48-49].

Simulated data: The MFDFA algorithm has been applied to three differently simulated data for comparison purpose. The first simulated data is based on the model proposed by Jonasson et al. [29] (hereafter model A) who studied MT polymerization/depolymerisation processes by considering two different critical tubulin concentrations to provide a better explanation for experimental observations of MT dynamic instability. The second set of data is generated based on Cassimeris et al. [30] (hereafter model M) who simulated an array of MTs using a Monte Carlo model and studied the effect of phase transitions and free tubulin concentrations on MT dynamics instability and MTs' length distribution. The computational codes were provided as a supplementary material. The third data set was generated based on the work of Rezania and Tuszynski [31] (hereafter model C) who studied MT polymerization and depolymerization processes by applying a semi-classical method that is often used in quantum mechanics. They showed that the dynamics of a MT can be mathematically expressed via a 3D cubic-quintic nonlinear Schrodinger equation.

In Figure 6, we plotted two time series of the normalized MT length over the sample number as an example for each models. The upper panels are constructed based on the model A, the middle panels are simulated based on model B and the lower panels are generated using the model C. For each data set, similar parameters are chosen as described in corresponding papers.

Similarly to experimental MT data sets, we performed the shuffling and the surrogated procedures on the simulated data from all three different models. Figure 7 shows the averaged generalized q -order Hurst exponent and multifractal spectrum for the original (black curve), shuffled (red curve) and surrogated (blue curve) of each simulated data set. Again, the upper, middle and lower panels are based on models A, B and C, respectively. Comparing Figures 5 and 7, the q -order Hurst exponent and multifractal spectrum, H_q and D_q , for the simulated data in model C is more or less comparable with the experimental time series. Although the simulated time series from models A and B show some degrees of multifractality, their multifractal spectrums are sharply truncated from the right side of the spectrum. This means the large-scale fluctuations have very little contributions in their data constructions. This is obviously not the case for the simulated data in the model C.

Conclusions

In this paper, we have used the MFDFA method to investigate the apparently stochastic nature of the experimental microtubule length data recorded in the course of 5–10 min intervals. These time series data describe the dynamic instability of microtubules that is driven by a competition

between assembly and disassembly processes. In particular, we examined the data obtained for two different MT isotypes, namely $\alpha\beta_{II}$ and $\alpha\beta_{III}$, to compare their similarities and differences. Our results indicate that the microtubule growth data sets exhibit a multifractal characteristic: a nonlinear curve of H_q as a function of q -orders, as opposed to a constant $H_q = 0.5$ for all q -orders, which is characteristic of a random walk process, or a linear function of q -orders for a monofractal process. Furthermore, by comparing the generalized Hurst exponents of the original time series with the shuffled and surrogated ones, we have found that multifractality due to the broadness of the probability density function has a greater contribution than the long-range correlations. This may be interpreted as indicating that the MT polymerization dynamics is governed by different physical mechanisms at different time scales. For example, MTs start polymerizing from nucleation, which is a highly nonlinear process. This is then followed by a largely linear assembly process, which is interspersed with stochastic catastrophes and non-uniformly distributed rescues [50]. Therefore, on different time scales, different specific mechanisms dominate the MT dynamic instability process.

In addition, we have examined three different simulated data sets based on the models A, B, and C proposed by Jonasson et al. [29], Cassimeris et al. [30], and Rezania and Tuszynski [31], respectively. Models A and B generally lead to the known two-state/three-state model for MT dynamic instability with some additional information. Even though the time series fluctuations in all three models considered here can represent MT dynamic instability, see Figure 6, the multifractal analysis shows that their resulting spectrums are quite different. Both models A and B have shown a sharply truncated spectrum from the right side suggesting that the structure of the data is mainly due to small-scale fluctuations. Model C, however, has shown a more symmetric spectrum, similar to experimental data, confirming that both small- and large-scale fluctuations contribute more or less equally. See Figure 7.

Available microtubule data used in this research suggest the possibility that the microtubule dynamic instability process is stochastically multifractal. Due to the fact that our analysis was limited to two tubulin isotypes and the data sets were not sufficiently large, further research should be performed with longer time-series data beyond 10 minutes of microtubule observations in order to achieve greater statistical robustness. The evolution patterns of the multifractality and the geometrical properties of the visibility graphs would have further clinical usefulness to investigate patterns of polymerization time series for drug-treated microtubules and to derive causality relationships between microtubule stochastic behaviours with mechanistic effects at a molecular level.

Acknowledgements

V.R. wishes to thank sincerely Dr. Goodson and her group as they provided the computational codes for this purpose and Philip Winter for his help in running some of the codes. V. R. also is grateful to MacEwan University for research grant for supporting this research. J.A.T. acknowledges with gratitude research support from NSERC (Canada) and the Allard Foundation.

Conflict of Interest

Authors declared that there is no conflict of interests.

References

- [1] M. Dogterom and S. Leibler, Physical Review Letters **70**(9), 1347 (1993).
- [2] A. Desai and T. J. Mitchison, Annual review of cell and developmental biology, **13**(1), 83 (1997).
- [3] T. Antal, P. L. Krapivsky, and S. Redner, Journal of Statistical Mechanics: Theory and Experiment, **5**, L05004 (2007).
- [4] J. Howard, *Mechanics of Motor Proteins and the Cytoskeleton* (Sinauer Associates is an imprint of Oxford University Press, 2001).
- [5] N. J. Woolf, A. Priel, and J. A. Tuszynski, *Nanoneuroscience: structural and functional roles of the neuronal cytoskeleton in health and disease* (Springer Science & Business Media, 2009).
- [6] J. A. Tuszyński, D. Sept, H. Bolterauer, and H. –J. Limbach, Advances in Structural Biology, **5**, 169 (1999). doi:10.1016/s1064-6000(98)80011-2
- [7] R. F. Luduena, International Review of Cell and Molecular Biology, **302**, 41 (2013). doi: 10.1016/B978-0-12-407699-0.00002-9.
- [8] V. Rezania, O. Azarenko, M. Jordan, H. Bolterauer, R. Luduena, J. T. Huzil, and J. A. Tuszynski, Biophysical Journal **95**,1993 (2008).
- [9] J. A. Tuszynski, E. Carpenter, J. T. Huzil, W. Malinski, T. Luchko, and R. F. Luduen., Int. J. Dev. Biol. **50**, 341 (2006).
- [10] T. Mitchison and M. Kirschner, Nature **312**, 237 (1984).
- [11] D. Panda, H. P. Miller, A. Banerjee, R. F. Ludueña, and L. Wilson, Proceedings of the National Academy of Sciences **91**(24), 11358 (1994).
- [12] M. Caplow and J. Shanks, Molecular biology of the cell **7**(4), 663 (1996).
- [13] S. Yarahmadian, V. Menon, and V. Rezania, Bioinformatics and Biomedicine (BIBM), 2015 IEEE International Conference On. IEEE; 417 (2015)
- [14] V. Menon, S. Yarahmadian, and V. Rezania, BMC Systems Biology **12**(Suppl 6), 112 (2018). <https://doi.org/10.1186/s12918-018-0631-5>
- [15] R. Lopes and N. Betrouni, Medical Image Analysis **13**, 634 (2009). <http://dx.doi.org/10.1016/j.media.2009.05.003>
- [16] P. C. Ivanov, L. A. Amaral, A. L. Goldberger, S. Havlin M. G. Rosenblum, Z. R. Struzik, and H. E. Stanley, Nature **399**(6735), 461 (1999).

- [17] C. K. Peng, J. E. Mietus, Y. Liu, C. Lee, J. M. Hausdorff, H. E. Stanley, A. L. Goldberger, L. A. Lipsitz, *Ann. Biomed. Eng.* **30**(5):683 (2002).
- [18] Y. Zheng, J. B. Gao, J. C. Sanchez, J. C. Principe, and M. S. Okun, *Phys. Lett. A* **344**, 253 (2005).
- [19] J. M. Hausdorff, *Hum. Mov. Sci.* **26**, 555 (2007).
- [20] G. Wang, H. Huang, H. Xie, Z. Wang, and X. Hu, *Med. Eng. Phys.* **29**, 375 (2007).
- [21] J. Bassingthwaite, J. VanBeek, and R. King, *Ann. N.Y. Acad. Sci.* **591**, 392 (1990).
- [22] S. Abboud, O. Berenfeld, and D. Sadeh, *Circ. Res.* **68**, 1751 (1991).
- [23] E. R. Weibel, *Am. J. Physiol.* **261**, 361 (1991).
- [24] G. Krenz, J. Linehan, and C. Dawson, *J. Appl. Physiol.* **72**, 2225 (1992).
- [25] I. Parkinson and N. Fazzalari, *Australas. Phys. Eng. Sci. Med.* **470**, 64 (1994).
- [26] C. Atupelage, H. Nagahashi, M. Yamaguchi, M. Sakamoto, and A. Hashiguchi, *Anal. Cell. Pathol. (Amst.)* **35**, 123 (2012).
- [27] A. L. Goldberger, L. A. Amaral, J. M. Hausdorff, P. C. Ivanov, C. K. Peng, and H. E. Stanley, *Proc. Natl. Acad. Sci. USA.* **99**, 2466 (2002).
- [28] E. A. Codling, M. J. Plank, and S. Benhamou, *J. R. Soc. Interface* **5**, 813 (2008).
- [29] E. M. Jonasson, A. J. Mauro, C. Li, E. C. Labuz, S. M. Mahserejian, J. P. Scripture, I. V. Gregoretti, M. Alber, and H. V. Goodson, *Molecular Biology of the Cell* **7**, 589 (2018)
<https://doi.org/10.1091/mbc.E19-02-0101>
- [30] L. Cassimeris, J. C. Leung, and D. J. Odde, *PLoS ONE* **13**(5): e0197538 (2018).
- [31] V. Rezaei and J. A. Tuszynski, *Physics Lett. A* **372**(47), 7051 (2008)
- [32] E. A. F. Ihlen, *Frontiers in Physiology* **3**. (2012). doi:10.3389/fphys.2012.00141
- [33] J. Suckling, A. M. Wink, F. A. Bernard, A. Barnes, and E. Bullmore, *J. Neurosci. Methods* **174**, 292 (2008).
- [34] E. A. F. Ihlen and B. Vereijken, *J. Exp. Psychol. Gen.* **139**, 436 (2010).
- [35] D. Makowiec, A. Rynkiewicz, J. Wdowczyk-Szulc, M. Zarczynska-Buchowiecka, R. Galaska, and S. Kryszewski, *Physiol. Meas.* **32**, 1681 (2011).
- [36] G. Wang, H. Huang, H. Xie, Z. Wang, and X. Hu, *Med. Eng. Phys.* **29**, 375 (2007).

- [37] B. B. Mandelbrot, *Physica Scripta* **32**, 257 (1985). doi:10.1088/0031-8949/32/4/001.
- [38] H. E. Hurst, *T. Am. Soc. Civ. Eng.* **116**, 770 (1951).
- [39] B. B. Mandelbrot and J. W. van Ness, *SIAM Review* **10**, 422 (1968).
- [40] R. H. Riedi, *An Introduction to Multifractals*, Technical Report, Rice University, Unpublished Version (1999).
- [41] J. W. Kantelhardt, S. A. Zschiegner, E. Koscielny-Bunde, S. Havlin, A. Bunde, and H. Stanley, *Physica A* **316**, 87 (2002). doi:10.1016/s0378-4371(02)01383-3.
- [42] D. B. de Freitas, M. M. Nepomuceno, M. G. Souza, I. C. Leão, M. L. Chagas, A. D. Costa, B. L. C. Martins, and J. R. Medeiros, *Astrophysical Journal* **843**(2), 103 (2017). doi:10.3847/1538-4357/aa78aa
- [43] J. Feder, *Fractals* (New York: Plenum Press 1988).
- [44] L. Zhao, W. Li, C. Yang, J. Han, Z. Su, and Y. Zou, *PLoS ONE* **12**(1): e0170467 (2017). doi:10.1371/journal.pone.0170467
- [45] B. B. Mandelbrot, B.B., *Fractals and Scaling in Finance Discontinuity, Concentration, Risk. Selecta Volume E*. New York: Springer (1997).
- [46] P. Norouzzadeha, W. Dullaertc, and B. Rahmanid, *Physica A* **380**:333 (2007).
- [47] S. Drożdż, and P. Oświęcimka, *Physical Review E* **91**(3) (2015). doi:10.1103/physreve.91.030902
- [48] R. Selmi, A. Tiwari, and S. Hammoudeh, *Economics Bulletin*. **38**, 2037 (2018).
- [49] E. Chandrasekhar, D. Subhakar, and Y. V. Vardhan, *Proceedings of the 17th annual conference of the International Association for Mathematical Geosciences, Editors: H. Schaeben, R. Tolosana Delgado, K. G. van den Boogaart, R. van den Boogaart, Freiberg (Saxony), Germany* (2015).
- [50] H. Bolterauer, H. J. Limbach, and J. A. Tuszyński, *Journal of Biological Physics* **25**(1), 1 (1999).

Table 1. Multifractal indices for the original, shuffled and surrogated data

	α_{q0}	H_2	Degree of multifractality, $\Delta\alpha_q$	Degree of asymmetry, A	Singularity ratio, C
$\alpha\beta_{II}$ – original	1.925	1.604	2.441	2.114	0.860
$\alpha\beta_{III}$ – original	2.016	1.617	2.731	2.749	0.674
$\alpha\beta_{II}$ – shuffled	0.605	0.534	1.083	3.152	0.297
$\alpha\beta_{III}$ – shuffled	0.560	0.476	0.990	3.763	0.270
$\alpha\beta_{II}$ – surrogated	1.567	1.376	1.132	1.630	0.654
$\alpha\beta_{III}$ – surrogated	1.503	1.307	1.645	2.623	0.526

Figure Captions:

Figure 1. Selected sample data of the growth pattern of microtubule isotypes $\alpha\beta_{II}$ (top panel) and $\alpha\beta_{III}$ (lower panel) from [8].

Figure 2. The difference time series, $\Delta x_i = x_{i+1} - x_i$, $i = 2, \dots, N$. $\alpha\beta_{II}$ (top panel), $\alpha\beta_{III}$ (middle panel), and white noise (lower panel) with zero average (red dashed lines) and $\pm$ RMS (red solid lines). The data are plotted in terms of the sample number rather than time for better comparison.

Figure 3. Multiple representations of the multifractal spectrum for the selected $\alpha\beta_{II}$, $\alpha\beta_{III}$ and the simulated white noise time series. (a) q -order Hurst exponent H_q (b) mass exponent τ_q (c) q -order multifractal spectrum of D_q and α_q (d) multifractal spectrum of D_q as a function of α_q . The arrow indicates the difference between the maximum and minimum α_q , which is called the multifractal spectrum width.

Figure 4. q -order Hurst exponent H_q (left panels) and multifractal spectrum of D_q as a function of α_q (right panels) for all available $\alpha\beta_{II}$ and $\alpha\beta_{III}$ data sets. The black line represents the average over corresponding data.

Figure 5. Averaged q -order Hurst exponent H_q (left panels) and multifractal spectrum of D_q as a function of α_q (right panels) for original, shuffled, and surrogated $\alpha\beta_{II}$ and $\alpha\beta_{III}$ data sets.

Figure 6. Simulated microtubule length data sets using three different methods described in models A (top panel), B (middle panel), and C (lower panel).

Figure 7. Averaged q -order Hurst exponent H_q (left panels) and multifractal spectrum of D_q as a function of α_q (right panels) for original, shuffled, and surrogated simulated microtubule data sets.

Top panel: model A, middle panel: model B, and lower panel: model C.

Figure 1.

Figure 2.

Figure 3.

Figure 4.

Figure 5.

Figure 6.

Figure 7.