



GEORG-AUGUST-UNIVERSITÄT
GÖTTINGEN

Master's Thesis

Measuring the Mean Back Relaxation via Darkfield Microscopy

Messung von Mean Back Relaxation mittels Dunkelfeldmikroskopie

prepared by

Julian Artur Schulz

from Göttingen

at the Third Institute of Physics Biophysics
Georg-August University Göttingen

Thesis period: 1st April 2023 until 30th September 2024

Supervisor: Dr. Till Münker

First Referee: Prof. Dr.Timo Betz

Second referee: Prof. Dr.Marcus Müller

Abstract

This thesis presents a new method for measuring Mean Back Relaxation (MBR) using darkfield microscopy, offering a more accessible alternative to laser-based optical tweezers. MBR can extract information about cellular activity from passive observations of microbead motion within living cells.

We develop an automated pipeline for tracking microbeads in darkfield microscope images, denoising trajectories, calculating MBR, and fitting microrheological models. Each step of the process is presented and tested, with areas for potential improvement identified.

By utilizing widely available darkfield microscopy, this approach aims to make MBR measurements more accessible to researchers studying cell mechanics. To further facilitate adoption, all data analysis methods are provided in an open-source Python package. This method could enable broader application of MBR as a tool for investigating cellular properties and dynamics.

Keywords: Physics, Master thesis

Contents

1. Introduction	1
2. Theory	3
2.1. Cell Mechanics	3
2.1.1. Models of Viscoelasticity	5
2.1.2. Connection between Power Spectrum and mechanical properties	7
2.1.3. Predicting Mechanical properties from active measurement . .	10
2.1.4. Predicting Mechanical properties from the Power Spectrum . .	10
2.2. Mean Back Relaxation	12
2.3. Filtering and Fitting the Power Spectrum	14
2.3.1. Fitting Exponentially Distributed Data	15
2.3.2. Finding Peaks	17
2.3.3. Model Selection using Bayesian Information Criterion	19
3. Materials and Methods	21
3.1. Dark-field Microscopy	21
3.1.1. Principle of Dark-field Microscopy	21
3.1.2. Microscope Setup	21
3.2. Hydrogel Sample preparation	23
3.3. Cell Sample preparation	23
3.3.1. Materials and Equipment	23
3.3.2. Procedure	23
3.4. Algorithms to Determine Trajectories	24
3.4.1. Mean Pixel Activation	24
3.4.2. Fitted Convolution	25
3.4.3. Fitted Particle Shape	29
3.5. Algorithm to Compute Mean Back Relaxation	30
3.6. Filtering and Fitting the Power Spectrum	31
3.7. Simulating Trajectories from Power Spectral Density	34

Contents

4. Results	37
4.1. Imaging	37
4.1.1. Imaging Hydrogel	37
4.1.2. Imaging HeLa Cells	38
4.1.3. Image cleaning	40
4.2. Tracking Beads	40
4.2.1. Tracking simulated trajectories	40
4.2.2. Tracking Piezostage Movement	43
4.2.3. Calibration via statistical measures	45
4.3. Filtering and Fitting	46
4.3.1. Validating the Fitting Method with Synthetic Data	46
4.3.2. Filtering and Fitting the Trajectories	50
4.4. Calculating Mean Back Relaxation	53
4.4.1. Validating Mean Back Relaxation Calculation on a simulation	53
4.4.2. Calaculating Mean Back Relaxation on experimental data	54
5. Discussion	57
5.1. Imaging	57
5.2. Tracking	58
5.3. Fitting and Filtering	59
5.4. Calculating Mean Back Relaxation	60
5.5. Determining Viscoelastic Properties	61
5.6. Outlook	61
6. Appendix	63
A. Power Spectral Density and Mean Back Relaxation Data	65
B. Entropy of the prior over length	71
C. Choosing $n_{typical}$	73
D. Distribution of the Mean Back Relaxation	75

1. Introduction

The mechanical properties of materials, characterized by their response to applied forces, play a crucial role in understanding both inanimate objects and living systems. These properties, encompassing elasticity (the tendency to deform reversibly in response to a force) and viscosity (the tendency to deform irreversibly in response to a force), are particularly significant in biological contexts. Living cells often optimize their mechanical properties to suit specific biological functions, making the study of cellular mechanics essential for a comprehensive understanding of cellular behavior and function [7, 11, 25, 26].

Traditionally, measuring the mechanical properties of materials in thermodynamic equilibrium has relied on the Fluctuation-Dissipation Theorem, which relates the system's response to external perturbations to its spontaneous fluctuations [16]. However, living cells operate far from equilibrium, actively consuming energy and breaking detailed balance. This non-equilibrium state necessitates alternative approaches to probe cellular mechanics.

One established method involves directly manipulating cellular components using specialized equipment such as optical tweezers. These instruments use focused laser beams to exert precise forces on micrometer-sized probe particles (e.g., silicone beads) within living cells while simultaneously tracking their positions [9]. While effective, this approach requires expensive, specialized equipment and considerable expertise.

Recent developments suggest that similar information about cellular mechanics might be obtainable through simpler, passive observation techniques. The key lies in analyzing how particles naturally move within the cellular environment in response to thermal fluctuations and active cellular processes. A promising statistical measure for this purpose is the Mean Back Relaxation (MBR), which quantifies how a particle reacts to displacement within a material. The MBR can reveal deviations from thermal equilibrium and has shown empirical links to specific ways biological systems break detailed balance [19]. If these relationships hold, it may be possible to

1. Introduction

infer a system’s mechanical properties and active behavior solely through passive observation. While previous studies have demonstrated the feasibility of this approach using laser tracking in cells [19], the goal of this master’s thesis is to develop a more accessible pipeline for measuring MBR using widely available darkfield microscopy equipment. This aims to make micromechanical investigations more broadly accessible to researchers across various fields.

Our approach emphasizes sophisticated data analysis to compensate for the lower precision of simpler measurement instruments. We implement and validate three different particle tracking methods, develop a novel technique for filtering noise from Power Spectra of measured trajectories, and create a comprehensive pipeline for calculating MBR and fitting viscoelastic models to the data.

We apply this pipeline to analyze simulated data, as well as experimental measurements of microbeads in hydrogels and HeLa cells. The thesis provides a detailed overview of the entire process, from image acquisition to data analysis and interpretation. To promote wider adoption and further development, we have made much of our data analysis code available as an open-source Python package.

While identifying areas for further refinement, this work establishes a foundation for fast, accessible measurements of MBR and power spectra in living cells. By lowering the barrier to entry for such investigations, we aim to facilitate novel research into cellular mechanics across a broader range of biological systems and research settings.

2. Theory

2.1. Cell Mechanics

Cells are not just biologically fascinating entities; they're also physical objects with measurable properties like stiffness. By examining cells through a physics lens and studying their mechanical characteristics, we can uncover important biological insights. The physical properties of cells play crucial roles in various biological functions.

For example, in stem cell differentiation, the physical environment, especially the elasticity of the surrounding matrix, strongly influences which cell type a stem cell will become [7]. During cell migration, such as in wound healing, cells interact mechanically with their surroundings. This allows them to navigate through tissues and respond to the stiffness and structure of the extracellular matrix [11].

Immune cells showcase their ability to dynamically adjust shape and stiffness, squeezing through tight spaces in tissues, which is essential for their function in reaching infection or inflammation sites [26]. Studies on cancer cell mechanics reveal that cancer cells exhibit distinct mechanical properties, including changes in elasticity and adhesiveness, which facilitate cancer dissemination [25].

While the complete mechanical behaviour of a cell might be very complex, a linear approximation of its reaction to forces suffices to describe its basic mechanical properties. When a material is subjected to a pulling force, the exerted force per unit area is referred to as stress, denoted by σ . Strain, symbolized as ϵ , measures the material's deformation as a proportion of its original length, indicating how much longer the material becomes relative to its initial state. For now we assume, that the relationship between stress and strain is linear and isotropic. This leaves us with the degrees of freedom, how much the material strains for a given stress, and whether this deformation is permanent, or reverts back when the stress vanishes. Mathematical models of these behaviours might look like this:

$$\sigma = k\epsilon \quad (2.1) \qquad \qquad \qquad \sigma = \mu\dot{\epsilon} \quad (2.2)$$

2. Theory

Here equation 2.1 describes a material that will only uphold the strain as long as the stress is still present like a spring with spring constant k . Meanwhile equation 2.2 describes a material which is permanently altered by any force, like a dashpot with friction μ . In both cases we assume overdamped dynamics. We can combine these notions by considering the linear dependency of the Fourier transform of stress $\hat{\sigma}$ and strain $\hat{\epsilon}$.

$$i\omega \hat{G} = \frac{\hat{\sigma}}{\hat{\epsilon}} \quad (2.3)$$

Here, G is the complex modulus. To see how we can describe viscous and elastic properties with this one quantity, we can solve equation 2.3 for the stress σ :

$$\begin{aligned} \hat{\sigma} &= \hat{G} \cdot \hat{\epsilon} i\omega \\ \rightarrow \sigma(t) &= \int_{-\infty}^{\infty} G(t - \Delta t) \dot{\epsilon}(\Delta t) d(\Delta t) \end{aligned}$$

We can see that due to causality, the G must vanish for all $t > 0$. From here we can reconstruct equations 2.1 and 2.2 by specific choices of \hat{G} .

Elastic Case	Viscous Case
$\hat{G}(\omega) = \frac{k}{i\omega}$	$\hat{G}(\omega) = \mu$
$\rightarrow \hat{\sigma} = k\hat{\epsilon}$	$\rightarrow G(t) = \delta(t)\mu$
$\rightarrow \sigma = k\epsilon$	$\rightarrow \sigma(t) = \int_{-\infty}^{\infty} \delta(t - \Delta t) \mu \dot{\epsilon}(\Delta t) d(\Delta t)$
	$\rightarrow \sigma = \mu \dot{\epsilon}(t)$

We see here that if \hat{G} is purely real, it describes viscous behavior, and when it is purely imaginary, it describes elastic behavior. By choosing \hat{G} with nonzero real and imaginary parts, we obtain a way to describe viscoelastic behavior.

So far we have assumed, that the deformation in the object is a uniform stretching along one dimension. However, we see in figure 2.1, that this is not the only possible deformation. In addition to the extension and compression of the material, moving a single point leads to the shearing of its surrounding. The shear strain γ is defined as the derivative of the displacement along one axis, derived by the position along an orthogonal axis. The relationship between shear stress and shear strain is analogous

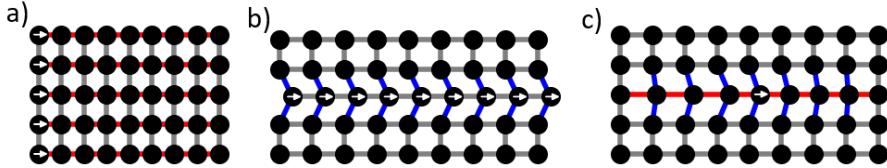


Figure 2.1.: Illustration of various grid deformation types: (a) Uniform Compression - the grid is uniformly compressed, reducing distances between all points equally. (b) Shear Deformation - a single layer is sheared, altering the alignment of points within that layer while maintaining the overall structure of adjacent layers. (c) Localized Displacement - a specific point is displaced, resulting in a combination of compression and shear deformation in its vicinity.

to our previous calculation:

$$\hat{\sigma}_{\text{shear}} = \hat{G}_{\text{shear}} \hat{\gamma} i\omega$$

The complex shear modulus is not the same as the complex modulus for compression. The relative factor between these two measures is called the Poisson ratio, and it typically lies between 0 and 0.5 [13]. In the rest of this thesis, we refer only to the complex shear modulus, denoted by G , since the movement of objects in cells that we are interested in mainly incurs shear stress [9].

2.1.1. Models of Viscoelasticity

Viscoelastic materials exhibit both viscous and elastic characteristics when subjected to deformation. This dual nature is modeled through various theoretical frameworks to predict their response under mechanical stresses. Among the most significant models are the Maxwell model, the Kelvin-Voigt model, and the Fractional Kelvin-Voigt model[23]. These models do not determine the value of the viscous and elastic part of the complex shear modulus, but describe how it depends on the frequency of movement.

Maxwell Model

The Maxwell model represents viscoelastic materials as a combination of a purely elastic element (spring) and a purely viscous element (dashpot) arranged in series. This model is particularly effective in describing materials that exhibit fluid-like

2. Theory

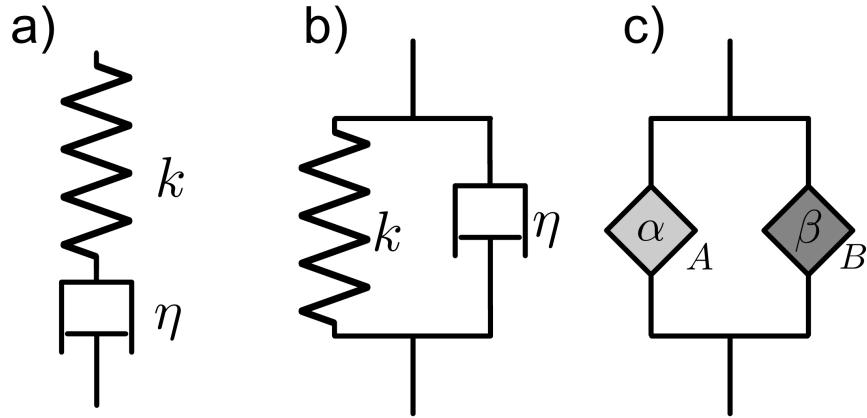


Figure 2.2.: Illustration of three viscoelastic models as a contraption of viscous and elastic objects: a) The Maxwell model: a spring and a dashpot in series. b) The Kelvin Voigt model: A spring and a dashpot in parallel. c) The fractional Kelvin Voigt Model: two elements of mixed viscous and elastic properties in parallel.

behavior over a long time scale, such as relaxation phenomena.

Because both objects are set in series, they are under the same stress, and the strain of the whole system is the added up strain of its component:

$$\begin{aligned}
 \epsilon &= \epsilon_{\text{spring}} + \epsilon_{\text{pot}} \\
 \sigma &= \sigma_{\text{spring}} = \sigma_{\text{pot}} \\
 &= k\epsilon_{\text{spring}} = \mu\dot{\epsilon}_{\text{spring}} \\
 \rightarrow \hat{\epsilon} &= \hat{\sigma}\left(\frac{1}{k} + \frac{1}{i\omega\mu}\right) \\
 \rightarrow \hat{G}_{\text{Maxwell}} &= \frac{1}{\frac{i\omega}{k} + \frac{1}{\mu}}
 \end{aligned}$$

Kelvin-Voigt Model

In contrast to the Maxwell model, the Kelvin-Voigt model arranges the elastic and viscous elements in parallel. This configuration allows the model to describe materials that exhibit immediate elastic response followed by a time-dependent viscous response, characteristic of solid-like behavior.

The parallel contraption leads to the strain being the same across both elements,

while the stress is separated out.

$$\begin{aligned}\epsilon &= \epsilon_{\text{spring}} = \epsilon_{\text{pot}} \\ \sigma &= \sigma_{\text{spring}} + \sigma_{\text{pot}} \\ &= k\epsilon_{\text{spring}} + \mu\dot{\epsilon}_{\text{spring}} \\ \rightarrow \hat{\sigma} &= \hat{\epsilon}(k + i\omega\mu) \\ \rightarrow \hat{G}_{\text{Kelvin-Voigt}} &= \frac{k}{i\omega} + \mu\end{aligned}$$

Fractional Kelvin-Voigt

The Fractional Kelvin-Voigt model does not derive from any particular model system of spring and dashpots. It is a generalisation of the Kelvin-Voigt model with elements, whose stress relates to their strain via a fractional derivative. This model is particularly useful in capturing the complex, non-linear behavior of polymers and biological tissues. The fractional derivative introduces an element of memory into the model, reflecting the material's history-dependent response.

$$\begin{aligned}\epsilon &= \epsilon_A = \epsilon_B \\ \sigma &= \sigma_A + \sigma_B \\ &= A' \left(\frac{d}{dt} \right)^\alpha \epsilon_A + B' \left(\frac{d}{dt} \right)^\beta \epsilon_B \\ \rightarrow \hat{G}_{\text{Fraktional Kelvin-Voigt}} &= A(i\omega)^{\alpha-1} + B(i\omega)^{\beta-1}\end{aligned}$$

2.1.2. Connection between Power Spectrum and mechanical properties

A Power Spectrum is a representation of a signal's power distribution across different frequencies. It provides information about the relative strength of various frequency components within the signal. The Power Spectrum is commonly used in signal processing and data analysis to identify dominant frequencies, noise characteristics, and periodic patterns [21, 24].

To calculate the Power Spectrum from a series of real-space data with a constant time step (Δt) between each point pair, the following steps can be followed:

2. Theory

1. Start with a time series of N data points, denoted as $x(t)$, where t represents the time index.
2. Apply the Discrete Fourier Transform (DFT) to the time series data. The DFT converts the data from the time domain to the frequency domain. The DFT of $x(t)$ is given by:

$$X(f) = \sum_{t=0}^{N-1} x(t) \cdot \exp(-i \cdot 2\pi \cdot f \cdot t \cdot \Delta t) \quad (2.4)$$

where f represents the frequency index and i is the imaginary unit.

3. Calculate the Power Spectrum by taking the absolute square of the DFT:

$$P(f) = |X(f)|^2 \quad (2.5)$$

4. Normalize the Power Spectrum by dividing it by the number of data points (N) to obtain the power spectral density (PSD):

$$\text{PSD}(f) = \frac{P(f)}{N} \quad (2.6)$$

5. The resulting PSD represents the power distribution across different frequencies. The frequency resolution (Δf) of the Power Spectrum is given by:

$$\Delta f = \frac{1}{N \cdot \Delta t} \quad (2.7)$$

where N is the number of data points and Δt is the time step between each point pair.

Passive Systems

In equilibrium systems, where no external forces or energy sources are present, the dynamics obey a fundamental principle known as detailed balance. This principle states that for any pair of microscopic states, the transition rates between these states are balanced, ensuring that the overall probability distribution remains stationary.

The detailed balance condition can be expressed mathematically as:

$$P(x)W(x \rightarrow y) = P(y)W(y \rightarrow x) \quad (2.8)$$

Here, $P(x)$ and $P(y)$ represent the probabilities of being in states x and y , respectively, while $W(x \rightarrow y)$ and $W(y \rightarrow x)$ denote the transition rates from state x to state y , and vice versa.

In such passive systems, the Fluctuation-Dissipation Theorem (FDT) establishes a fundamental relationship between the dissipative properties of the system, characterized by the response function, and the fluctuations present in equilibrium. The FDT provides a direct link between the imaginary part of the response function, also known as the dissipative part, and the equilibrium fluctuations of the observable being considered.

One form of the FDT relates the power spectral density (PSD) of equilibrium fluctuations to the imaginary part of the response function $\chi(\omega)$ [16].

$$S(\omega) = \frac{2k_B T}{\omega} \Im(\chi(\omega)) \quad (2.9)$$

Here, k_B is the Boltzmann constant, and T is the absolute temperature.

Active Systems

Living systems, on the other hand, represent a class of systems that are not in detailed balance, as they are actively driven by energy-consuming processes. In such non-equilibrium systems, the Fluctuation-Dissipation Theorem (FDT) can be violated. Empirically, some studies found a modified relationship to hold between the power spectral density (PSD) of fluctuations and the imaginary part of the response function [19, 20]. Those observations were made for objects moving within the cytoplasm of living cells, and do not hold for any systems out of equilibrium. This new equation takes the form:

$$S(\omega) = \frac{2(k_B T + E_0 \frac{\omega_0}{\omega})}{\omega} \Im(\chi(\omega)) \quad (2.10)$$

where E_0 represents an energy scale associated with the active processes driving the system out of equilibrium, and ω_0 is a characteristic frequency scale. By convention, without loss of generality, ω_0 is set to 1 Hz. This empirical law has been observed to hold within the frequency range of $10^0 - 10^3$ Hz, but has to break down

2. Theory

for lower frequencies, as long as the particle under consideration is confined to any finite space. Nevertheless, for the purpose of this thesis, we will tentatively assume that this modified relationship holds in the considered cases.

2.1.3. Predicting Mechanical properties from active measurement

Previous work by Münker et al.[20] has employed optical tweezers-based active microrheology to measure the frequency-dependent viscoelastic properties of the cytoplasm of living cells. In this technique, a phagocytosed probe particle is manipulated using an optical trap, applying oscillatory forces at varying frequencies. By monitoring the particle's displacement in response to these forces, the complex shear modulus $G^*(f)$ can be determined. To model the viscoelastic behavior, the authors used a double power law approach from the fractional Kelvin-Voigt model. In addition to the viscoelastic properties, the authors quantified the active forces within the cell by measuring the violation of the Fluctuation-Dissipation Theorem, as described in equation 2.10. The combination of all fitted parameters forms the mechanical fingerprint of a cell. Münker et al. applied this approach to various cell types and conditions, revealing several key findings. They demonstrated that the mechanical fingerprint can distinguish between different cell types, with the most discriminatory parameters being the scale of the detailed balance breaking part of the Power Spectrum E_0 and the prefactor to the more elastic part of the response function A .

Furthermore, the authors showed that the mechanical fingerprint is functionally tuned to match the biological roles of cells. For example, muscle cells (C2C12) exhibited higher stiffness (higher A), while immune cells (macrophages) were more fluid-like with higher activity (higher E_0). This suggests that cells adjust their mechanical properties in accordance with their function.

2.1.4. Predicting Mechanical properties from the Power Spectrum

By establishing connections between different theoretical frameworks, we can predict the mechanical properties of a system from its observed Power Spectrum of fluctuations. The generalisation of the Fluctuation-Dissipation Theorem (Eq. 2.10)

relates the power spectral density (PSD) to the imaginary part of the response function and an active energy scale that characterizes the departure from equilibrium. Considering a round small object with radius r , submerged in a material with a given shear modulus G (as defined in Eq. 2.3), we can deduce its response function χ , using Stokes' law[18]:

$$\chi(f) = \frac{1}{6\pi RG(f)} \quad (2.11)$$

Substituting this relationship into the Fluctuation-Dissipation Theorem yields an expression linking the PSD directly to the complex shear modulus:

$$S(\omega) = \frac{2(k_B T + E_0 \frac{\omega_0}{\omega})}{\omega} \Im \left(\frac{1}{6\pi R G} \right) \quad (2.12)$$

We can then substitute different models of viscoelasticity, such as the Maxwell, Kelvin-Voigt, and Fractional Kelvin-Voigt models, to predict the PSD's functional form. Additionally, if the trajectory measurement includes Gaussian noise with standard deviation σ , a constant noise term σ^2 is added to the PSD expressions:

$$S_{\text{Mx}}(\omega) = \frac{k_B T}{3\pi R \mu} \frac{1}{\omega^2} + \frac{E_0 \omega_0}{3\pi R \mu} \frac{1}{\omega^3} + \sigma^2 \quad (2.13)$$

$$S_{\text{KV}}(\omega) = \frac{k_B T \mu + E_0 \mu \frac{\omega_0}{\omega}}{3\pi R (k^2 + \mu^2 \omega^2)} + \sigma^2 \quad (2.14)$$

$$S_{\text{fKV}}(\omega) = \frac{k_B T + E_0 \frac{\omega_0}{\omega}}{3\pi R} \frac{A \omega^{\alpha-1} \sin\left(\frac{\pi\alpha}{2}\right) + B \omega^{\beta-1} \sin\left(\frac{\pi\beta}{2}\right)}{A^2 \omega^{2\alpha} + 2AB \omega^{\alpha+\beta} \cos\left(\frac{\pi}{2}(\alpha-\beta)\right) + B^2 \omega^{2\beta}} + \sigma^2 \quad (2.15)$$

Given a measured trajectory, these models can be fit to the data, providing insights into the underlying viscoelastic model and its parameters. One challenge lies in determining the active energy scale E_0 , as it only manifests at very low frequencies. The validity of the fluctuation-dissipation relation (Eq. 2.10) has not been extensively tested at such low frequencies, and discrete Fourier transforms yield fewer data points in this regime when frequency is considered on a log scale. An alternative approach to estimate E_0 involves the use of Mean Back Relaxation.

2. Theory

2.2. Mean Back Relaxation

The analysis of particle trajectories plays a fundamental role in modern statistical physics, enabling insights into microrheological and non-equilibrium properties of complex media. While correlation functions like the mean squared displacement (MSD) are powerful tools in equilibrium systems, they are insensitive to the breaking of time-reversal symmetry. To detect such non-equilibrium behavior, multi-point correlation functions that depend on multiple time points can serve as intriguing observables.

One such observable is the mean back relaxation (MBR), which correlates particle positions at three distinct time points. The MBR quantifies the average future displacement of a particle under the condition of a displacement in the past. Mathematically, the MBR in one dimension is defined as:

$$M(\tau, t) = \left\langle \frac{[x(t) - x(0)]}{[x(0) - x(-\tau)]} \right\rangle \quad (2.16)$$

Here, $x(t)$ denotes the particle's position at time t , and the angular brackets represent an ensemble average. The numerator represents the particle's displacement between times 0 and t , while the denominator corresponds to the displacement between times $-\tau$ and 0. By convention, a minus sign is included to ensure that the MBR is positive in equilibrium, as the particle typically travels in the opposite direction or relaxes back compared to its earlier movement.

The MBR captures how conditions affect the particle's subsequent motion, probing the interplay between past and future displacements. It can be interpreted as the conditional average of the future displacement given the past displacement, divided by the past displacement itself to render the quantity dimensionless.

Notably, for confined systems in equilibrium satisfying detailed balance, the MBR exhibits a universal long-time limiting value [15]. Under the assumptions of ergodicity, finite correlation times, and the factorization of the three-point probability distribution for large times, the long-time limit of the MBR takes the value:

$$\lim_{t \rightarrow \infty} M(\tau, t) = \frac{1}{2} \quad (2.17)$$

This result holds irrespective of the specific dynamics or the time period τ , as long as detailed balance is fulfilled. The proof leverages the fact that, in equilibrium, any observable anti-symmetric under time reversal must have a vanishing average. Con-

2.2. Mean Back Relaxation

sequently, the long-time MBR, which can be expressed as the sum of a symmetric and an anti-symmetric part, reduces to the constant value of 1/2 in the presence of detailed balance. This universal long-time limiting value serves as a hallmark of equilibrium conditions and provides a powerful tool for detecting deviations from detailed balance in confined systems.

While the MBR's behavior in equilibrium systems is well understood, we only know its analytical form for few non-equilibrium systems. However, one system where we know the MBR is the "horse and cart" model[19]. The horse and cart model consists of a particle (the cart) confined in a harmonic potential that itself undergoes random motion (the horse). This system mimics a particle trapped in a fluctuating environment, such as those found in active biological systems. In this model, the MBR in the limit of $\tau \rightarrow 0$ has been analytically derived:

$$MBR(t) = \frac{1}{2} \left(1 - \frac{D_q}{D} \right) \left(1 - e^{-\kappa\beta D t} \right) \quad (2.18)$$

Here, D is the diffusion coefficient of the particle, D_q is the diffusion coefficient of the potential's center, κ is the spring constant of the harmonic potential, and β is the inverse temperature.

Thus far, the MBR has been experimentally measured and reported in published work using optical tweezers by Münker et al. in 2023 [19]. In their study, micrometer-sized silica beads were phagocytized (internalized) by different types of living cells. Optical tweezers, also known as optical traps, are a powerful tool that allows for the precise manipulation and tracking of microscopic objects using highly focused laser beams. The underlying principle is that a tightly focused laser beam exerts an attractive or repulsive force on dielectric particles, enabling their controlled movement and positioning.

By employing optical tweezers, Münker et al. could measure particle trajectories with high temporal and spatial resolution, achieving a temporal resolution of 65 kHz and nanometer-scale precision. These high-precision measurements revealed that when the silica beads were placed inside living cells, the long-term limit of the MBR deviated from the expected equilibrium value of 1/2. Additionally, their data suggested a linear dependency between the active energy scale E_0 from Eq. 2.10 and the long-term limit of the MBR for short time periods τ .

2. Theory

This study highlighted the crucial importance of achieving high precision in measuring particle trajectories, both in time and space. To access short time scales τ , a high frame rate is essential, while resolving the particle's steps at such short times necessitates nanometer-scale spatial resolution. These stringent requirements underscore the valuable role of optical tweezers in enabling precise MBR measurements.

The ability to measure MBR with such high accuracy opens up several useful applications. Not only can systems that break detailed balance be identified, but passive imaging techniques can also provide insights into their viscoelastic properties. By independently measuring E_0 , Eq. 2.12 can be leveraged to determine other free parameters of viscoelastic models, which are more reliably accessible at higher frequencies.

However, a significant drawback of this technique is that optical tweezers are expensive and time-consuming to operate, limiting their widespread adoption. To make the MBR a more broadly applicable and relevant observable, alternative measurement methods that are easier to implement, such as microscopy, would be highly desirable. Nevertheless, achieving the required high precision in both time and space resolution for particle trajectories remains a crucial challenge for such alternative approaches.

2.3. Filtering and Fitting the Power Spectrum

When analyzing the Power Spectrum of a bead embedded in a viscoelastic material to infer its mechanical properties, several challenges arise.

Firstly, the elements of a Power Spectrum are not normally distributed but follow an exponential distribution[4]. This precludes the use of least-square error fitting, which assumes a normal distribution of errors.

Another significant challenge lies in the presence of peaks in the measured Power Spectrum that do not originate from the movement of the bead within the material[24]. These peaks can arise from resonance frequencies of the microscope or other external factors, such as vibrations from the table on which the experiment is conducted. To obtain a reliable representation of the material's viscoelastic properties, it is crucial to identify and filter out these extraneous peaks.

Furthermore, the choice of an appropriate viscoelastic model is essential for accurately describing the material's response, making a trade off between accuracy of

2.3. Filtering and Fitting the Power Spectrum

the fit and simplicity of the model

2.3.1. Fitting Exponentially Distributed Data

In physics, when fitting measurement data, the least squares method is commonly used. This approach is valid under the assumption that the data points are well described by a model in the considered model space, with measurement errors drawn from a Gaussian distribution[5].

Let's see how this assumption leads to this practice. Assuming the data points y_i are generated from a true function $f(x_i)$ with added Gaussian noise ϵ_i , the model can be expressed as:

$$y_i = f(x_i) + \epsilon_i, \quad \epsilon_i \sim \mathcal{N}(0, \sigma^2) \quad (2.19)$$

The probability density function (PDF) of observing a specific y_i given x_i is:

$$p(y_i|x_i) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(y_i - f(x_i))^2}{2\sigma^2}\right) \quad (2.20)$$

Let $\mathcal{D} = (x_i, y_i)_{i=1}^N$ be our dataset, where each (x_i, y_i) represents a pair of observed values, and N is the total number of data points. The likelihood of observing this dataset given the model parameters is the product of individual probabilities:

$$L = \prod_i p(y_i|x_i) \quad (2.21)$$

Taking the negative log-likelihood (NLL) and ignoring constant terms, minimizing NLL is equivalent to minimizing the sum of squared errors (SSE):

$$\text{SSE} = \sum_{i=1}^N (y_i - f(x_i))^2 \quad (2.22)$$

This leads to the least squares error criterion.

While the assumption of Gaussian errors is well justified in most contexts, it does not hold for power spectra. Consider a simple example of a stationary particle measured with Gaussian measurement noise. The particle's position over time is normally distributed, as shown in Figure 2.3a/b. However, when calculating the Power Spectrum of this trajectory using the method described in Section 2.1.2, the resulting Power Spectrum follows an exponential distribution, as illustrated in Figure 2.3c/d.

2. Theory

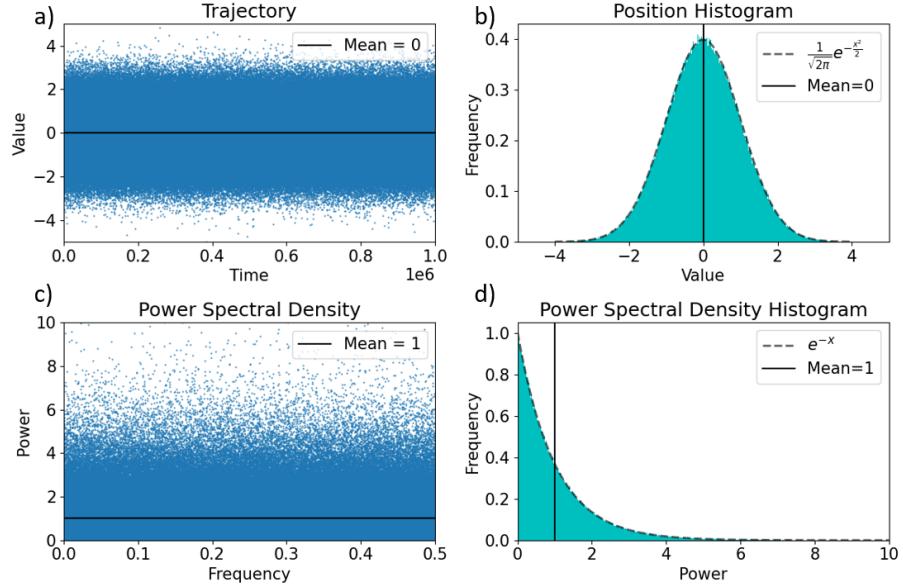


Figure 2.3.: Simulation exemplifying the properties of Gaussian random numbers and their Power Spectrum. (a) Trajectory of one million points drawn from a normal distribution with mean 0. (b) Histogram of the same million points, superposed with a normal distribution curve. (c) Power spectral density (PSD) of the trajectory, showing the frequency distribution of the power. (d) Histogram of the PSD, superposed with an exponential distribution curve with mean 1.

In general, if the expected value of the power spectral density is S , a point x in the Power Spectrum at that location follows the distribution:

$$p_S(x) = \frac{1}{S} e^{-\frac{x}{S}} \quad (2.23)$$

A common approach in microrheology is to average multiple points of the Power Spectrum into a single point for fitting [24]. If the number of points in each bucket is sufficiently large, the central limit theorem leads to a Gaussian distribution of these averages. While this method allows for the use of standard fitting techniques, it results in a loss of information, particularly at low frequencies where the power spectral density can vary significantly for small frequency differences.

An alternative approach is to calculate the negative log-likelihood (NLL) that needs to be minimized to fit exponentially distributed data. Given a dataset x_i, y_i

2.3. Filtering and Fitting the Power Spectrum

and a function $f(x, \theta)$ to be fitted, the NLL is:

$$NLL(y_i|x_i, \theta) = \sum_i \frac{y_i}{f(x_i, \theta)} + \log(f(x_i, \theta))$$

Here θ are the parameters of the function f , that we want to determine. Minimizing this NLL allows for fitting the power spectral data without the need for averaging, preserving more information from the original data.

2.3.2. Finding Peaks

When measuring the trajectory of a particle within a material to investigate its properties, a significant challenge arises: the observed trajectory is a combination of the particle's movement within the material and the material's movement through space. For example, when observing a bead moving inside a cell, the cell's movement is also captured. Separating these two components is crucial for accurate analysis.

Fortunately, if the material is fixed to a macroscopic object (e.g., a cell attached to a coverslip or pinned on a table), the two types of movement exhibit distinct characteristics in Fourier space. As shown in Equation 2.12, the movement of a bead within a viscoelastic material has a smooth Power Spectrum that decreases with frequency. In contrast, the movement of a lab table or microscope is dominated by resonance frequencies, leading to peaks in the Power Spectrum.

In microrheology literature, these peaks are often manually removed from the data or without explicit statistical consideration [24]. While this approach is effective for small datasets where a human can easily review and identify peaks, more rigorous methods are needed for larger datasets or ambiguous cases.

One approach to model datasets containing points x_i drawn from different distributions $p_\alpha(x|\theta)$ is Probabilistic Mixture Modeling for Outlier Analysis [1]. This method involves optimizing model parameters θ to maximize the likelihood of the data while simultaneously assigning each point i to a specific model α to maximize the overall likelihood. However, this method is not directly applicable to the power spectra under consideration, as we lack an explicit model for the appearance of noise peaks. Instead, we expect noise peaks to be rare and continuous sections of the Power Spectrum.

To address this, we propose a modified approach that still involves finding model parameters and assigning points to either the viscoelastic model or noise peaks, but without explicitly modeling the peaks. Instead, we treat the number and lengths of

2. Theory

peaks as additional model parameters to be optimized.

Given the datapoints (x_i, y_i) and the parametrized probability distribution $p(y|x, \theta)$ for non-noise peak points, where y is the power spectral density (PSD), x is the frequency, and θ are the model parameters, the probability distribution p should spread around the predicted Power Spectrum f as follows:

$$p(y|x, \theta) = \frac{1}{f(x, \theta)} e^{-\frac{y}{f(x, \theta)}} \quad (2.24)$$

We define a set of noise peaks $\rho := \{i_{peak\ start}, i_{peak\ end}\}_j$, such that the set of points in noise peaks is $\Sigma_\rho := i : \exists(i_{peak\ start}, i_{peak\ end}) \in \rho : i_{peak\ start} < i < i_{peak\ end}$. We assume a prior over the number of peaks $p_{number\ peaks}(n)$ and a prior over the length of a peak $p_{length\ peak}(l)$.

To quantify the evidence for a set of points being drawn from a specific probability distribution, we make use of the Bayesian notion of surprise.[3] Given a probability distribution $p(x)$ and a datapoint x , the surprise S is the amount by which the negative log-likelihood of x given p exceeds the expected negative log-likelihood of a datapoint randomly sampled from p :

$$\begin{aligned} S(x, p) &= -\log(p(x)) - \langle -\log(p(x)) \rangle \\ &= NLL(x|p) - H(x) \end{aligned}$$

Here, $H(p)$ is the entropy of the probability distribution p . The surprise of multiple points is additive, and the expected value of the surprise is 0. If the surprise is greater than 0, it provides evidence that the points x are not drawn from the probability distribution p . In a Bayesian framework, the surprise directly measures the amount of evidence that a set of points x_i is drawn from the probability distribution p , given by $e^{-\sum_i S(x_i, p)} \cdot p_{prior}$, where p_{prior} is the prior probability of x_i being drawn from p .

When analyzing a Power Spectrum containing a given number of datapoints and noise peaks, we can calculate the total surprise as the sum of the surprise from the datapoints and the surprise of the peaks:

2.3. Filtering and Fitting the Power Spectrum

$$\begin{aligned}
S_{\text{spectrum}} &= \sum_{i \notin \Sigma_p} \underbrace{\left(\frac{y_i}{f(x_i, \theta)} + \log(f(x_i, \theta)) - \underbrace{\log(f(x_i, \theta)) - 1}_{H(p(y_i|x_i, \theta))} \right)}_{NLL(y_i|x_i, \theta)} \\
&= \sum_{i \notin \Sigma_p} \left(\frac{y_i}{f(x_i, \theta)} - 1 \right) \\
S_{\text{peaks}} &= -\log(p_{\text{numberpeaks}}(n_{\text{peaks}})) - H(p_{\text{numberpeaks}}) \\
&\quad + \sum_{i=0}^{n_{\text{peaks}}} (-\log(p_{\text{lengthpeak}}(l_i)) - H(p_{\text{lengthpeak}}))
\end{aligned}$$

The total surprise, which we aim to minimize, is given by:

$$S_{\text{total}} = S_{\text{spectrum}} + S_{\text{peaks}}$$

The computational implementation of this minimization process is described in Section 3.6.

By minimizing the total surprise, we can simultaneously find the optimal model parameters for the viscoelastic material and identify the noise peaks in the Power Spectrum. This approach provides a principled and statistically rigorous method for separating the particle's movement within the material from the material's movement through space, enabling more accurate analysis of the material's properties.

The proposed method offers several advantages over manual peak removal or ad hoc approaches. It incorporates prior knowledge about the expected rarity and continuity of noise peaks while allowing for flexibility in the number and lengths of peaks. By considering the surprise of both the viscoelastic model and the noise peaks, the method balances the fit to the data with the complexity of the peak structure.

2.3.3. Model Selection using Bayesian Information Criterion

When fitting a viscoelastic model to the Power Spectrum data, it is crucial to select a model that accurately describes the material's properties while avoiding overfitting. One widely used approach for model selection is the Bayesian Information Criterion (BIC) [6].

BIC is a model selection criterion that balances the goodness of fit with the

2. Theory

complexity of the model. It is derived from a Bayesian perspective, where the goal is to find the model that maximizes the posterior probability given the observed data. The BIC is defined as:

$$\text{BIC} = -2 \cdot \log(\hat{L}) + k \cdot \log(n) \quad (2.25)$$

where \hat{L} is the maximum likelihood estimate of the model, k is the number of parameters in the model, and n is the number of data points. The first term, $-2 \cdot \log(\hat{L})$, measures the goodness of fit, while the second term, $k \cdot \log(n)$, penalizes the model complexity.

The BIC has several desirable properties for model selection:

It is consistent, meaning that as the sample size increases, the probability of selecting the true model (if it is in the candidate set) approaches 1. It penalizes model complexity more heavily than other criteria, such as the Akaike Information Criterion (AIC), which helps to prevent overfitting. It has a clear interpretation in terms of the posterior probability of the models, as the difference in BIC between two models is approximately equal to the log of the Bayes factor [6].

It is important to note that the BIC is not without limitations. It assumes that the true model is among the candidate models, and it may not perform well if the models are misspecified or if the sample size is small relative to the number of parameters.

By calculating the BIC for each candidate model, we can identify the model that best describes the material's properties while avoiding overfitting. In cases where multiple models fit the data well, softmaxing the negative BIC values can be used to obtain a posterior distribution over the models, incorporating uncertainties in the model selection process.

The softmax function assigns higher probabilities to models with lower BIC values while still considering the uncertainties associated with the model selection process[5]:

$$P(M_i|D) = \frac{\exp(-\text{BIC}_i/2)}{\sum_{j=1}^m \exp(-\text{BIC}_j/2)} \quad (2.26)$$

where $P(M_i|D)$ is the posterior probability of model M_i given the data D , BIC_i is the BIC value for model M_i , m is the total number of candidate models.

3. Materials and Methods

3.1. Dark-field Microscopy

3.1.1. Principle of Dark-field Microscopy

Dark-field microscopy is an optical microscopy technique that enhances contrast in unstained, transparent specimens. Unlike brightfield microscopy, where the specimen is illuminated directly, dark-field microscopy uses an illumination technique that prevents direct light from entering the objective lens. In dark-field microscopy, the condenser is designed to form a hollow cone of light with a numerical aperture (NA) that exceeds the NA of the objective lens. This oblique illumination is focused onto the specimen, and only the light scattered, refracted, or diffracted by the specimen enters the objective. As a result, the specimen appears bright against a dark background, providing enhanced contrast and visibility of fine details, especially for unstained or low-contrast samples.

Darkfield microscopes often employ specialized condensers with high NA, such as cardioid or paraboloid condensers, to achieve the necessary oblique illumination. These condensers have a larger NA than the objective lens to ensure that no direct light enters the objective. The condenser is typically immersed in oil to maintain the high NA illumination and to prevent light loss due to total internal reflection at the glass-air interface.

3.1.2. Microscope Setup

In this study, a Nikon ECLIPSE Ti2 equipped with a dark-field condenser was used for imaging. The microscope was fitted with a NIKON CFI super Fluor 100X objective lens, which had a numerical aperture of 0.5-1.3. The dark-field condenser used was a Nikon Dark Field Oil condenser, with a numerical aperture of 1.43-1.20, ensuring proper dark-field illumination conditions.

3. Materials and Methods

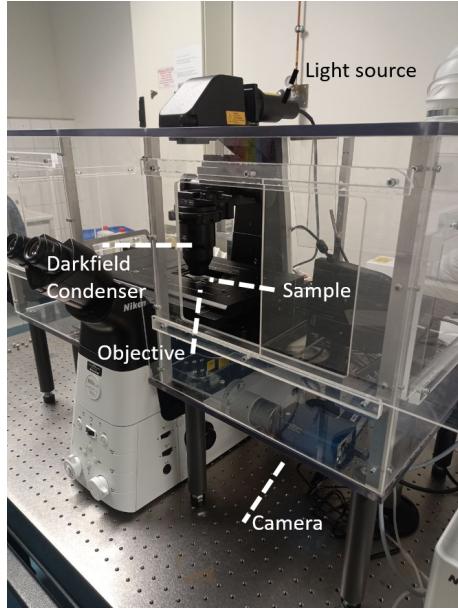


Figure 3.1.: Picture of the experimental setup of the Darkfield microscope with labeled parts.

Illumination was provided by a SOLA Light Engine light source, which offered high brightness and uniform illumination necessary for dark-field microscopy. The lamp was adjusted to provide maximum brightness without introducing glare or reflections.

For experiments requiring precise sample positioning, the microscope can be equipped with a high-precision piezo stage. This stage is based on the P-611.3S NanoCube XYZ Nanopositioner from Physik Instrumente (PI), controlled by an E-727 controller [22]. The piezo cube is fitted with a custom-made stage platform. This setup allows for extremely precise movements in three dimensions. The P-611.3S offers a travel range of up to $120\mu\text{m} \times 120\mu\text{m} \times 120\mu\text{m}$, with a resolution as fine as 0.2 nm.

Videos were captured using a Mikrotron EoSens Quad 1.1S digital camera attached to the microscope. The resulting videos had a pixel size of 36.7nm and the camera can make videos with 2.500 frames per second. By cropping the field of few, the frame rate can be increased beyond 100.000 frames per second. The imaging was controlled by VisualIMARC2 imaging software. The gain settings were optimized for each video to achieve the best image quality for the given framerate.

Before imaging, the microscope was properly aligned and the dark-field condenser was centered according to the manufacturer's instructions. Zeiss Immersion oil 518

3.2. Hydrogel Sample preparation

F with a refractive index of 1.518 was used between the condenser and the specimen slide to maintain the high NA illumination.

3.2. Hydrogel Sample preparation

The viscoelastic polyacrylamide hydrogels were prepared following a detailed protocol [17]. The process involves creating a premix of acrylamide, bisacrylamide, and acrylic acid, which is then combined with a dextran solution, APS (ammonium persulfate), and TEMED (tetramethylethylenediamine) to initiate polymerization. Importantly, the gel mixture also incorporates micrometer-sized polystyrene particles, referred to as beads.

3.3. Cell Sample preparation

This section describes the preparation of HeLa cell samples with embedded microbeads for dark-field microscopy observations.

3.3.1. Materials and Equipment

Item	Description
Cells	Wild Type HeLa cells
Culture medium	Dulbecco's Modified Eagle Medium (DMEM), HEPES-supplemented DMEM
Other solutions	PBS, Trypsin, 1% Fibronectin in PBS, Trypan Blue
Microbeads	1 μ L solution of 1 μ m Polybead® Microspheres silicone beads per sample
Labware	Culture flasks, coverslips, object holders
Equipment	Incubator, heating plate, centrifuge, hemocytometer

Table 3.1.: Materials and equipment for cell sample preparation

3.3.2. Procedure

The procedure can be divided into four main steps:

Cell Culture and Harvesting

HeLa cells were cultured using standard techniques. Cells were detached using trypsin, neutralized with medium, and centrifuged to obtain a cell pellet. A portion

3. Materials and Methods

of the cells was used to maintain the culture line.

Sample Preparation

Glass coverslips were coated with fibronectin to promote cell adhesion. A cell suspension was prepared with a concentration of 2000 cells per sample, mixed with the microbead solution, and applied to the fibronectin-coated coverslips.

Incubation

The prepared samples were incubated in a cell culture chamber for 4 hours, allowing cells to adhere and internalize the microbeads.

Final Sample Assembly

After incubation, the samples were washed with PBS. HEPES-supplemented medium was added to maintain cell viability during microscopy. Next, a double-sided adhesive tape with a square cutout in the center is taken. This tape is used to attach the sample to a microscope slide. The cutout area serves two purposes: it keeps the medium and cells visible for microscopic observation, and it creates a well to contain the medium, preventing it from flowing out.

3.4. Algorithms to Determine Trajectories

We have implemented multiple methods for tracking particle trajectories from the movies captured with the microscope. All of the methods track trajectories in a 2D space and rely on only one particle being visible at a time in the camera's field of view. While this assumption might seem to restrict the density of beads in the sample, it turns out not to be a problem because when making high-speed videos, we have to crop the field of view to allow fast enough readout of the pixels for frame rates beyond 2.500 and minimize the amount of data that needs to be transferred from the camera to the computer.

3.4.1. Mean Pixel Activation

The first method is among the simplest methods to track objects in a movie. It is mentioned here mostly as a sanity check to ensure that the more complicated and

potentially error-prone methods roughly agree with this one and are able to provide a more fine-grained resolution.

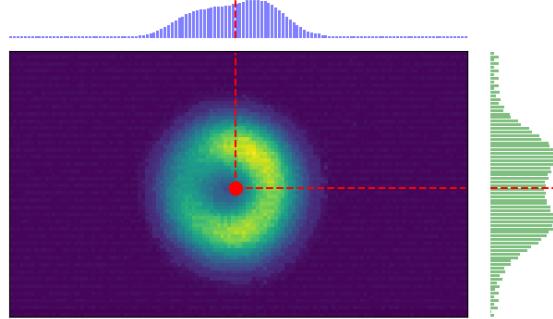


Figure 3.2.: Illustration of the Mean Pixel Activation method for detecting the bead position in an image. Example image of a bead in hydrogel from the experiment in section 4.1.1.

As shown in Figure 3.2, the idea behind the Mean Pixel Activation method is to calculate the mean x and y values of pixels, weighted by their brightness. This only works when we set the brightness value of the background to zero. This is done by subtracting the most common brightness value of all pixels from each pixel's brightness.

The advantages of this method are that we not only get the relative movement of the particle through space but also obtain the position of the particle's center for each image. This makes it easy to check that we are tracking the particle roughly correctly, as we can plot the measured particle center over the image, as shown in Figure 3.2. Another desirable property of this method is that the position of the particle in each picture is measured independently of the other pictures. We will see later that we sometimes have to make this assumption when estimating the error of a tracking method.

3.4.2. Fitted Convolution

In this method, we measure the shift that an object has undergone between two pictures by finding the maximum of the convolution of those two pictures.

Given two functions f and g , the convolution of f and g , denoted as $f * g$, is defined as:

$$(f * g)(t) = \int_{-\infty}^{\infty} f(\tau)g(t - \tau)d\tau \quad (3.1)$$

3. Materials and Methods

In the discrete case, such as when dealing with digital images, the convolution can be expressed as:

$$(f * g)[n] = \sum_{m=-\infty}^{\infty} f[m]g[n-m] \quad (3.2)$$

In our case, we compute the convolution of two images, I_1 and I_2 , which represent the particle at two different time points. The maximum value of the convolution indicates the shift that best aligns the particle between the two images.

However, because we are in a discrete setting, simply taking the point that maximizes the convolution would yield a maximal precision of one pixel. If we want to detect steps with a finer resolution, we need to interpolate between the points of the convolution. We have observed that the maximum of the convolution can be locally well approximated by a Gaussian.

As we see in Figure 3.3, the discrete convolution has the shape of a Gaussian for a radius of about 10 pixels around the maximum. We can obtain the maximum of the convolution by taking the maximum of this fitted two-dimensional Gaussian. As we will see later, this method not only yields numbers of fractional pixels but also provides step sizes that are precise to a sub-pixel level.

By taking into account more points around the maximum, we achieve a more precise detection of the maximum of the convolution. However, we also see in Figure 3.3 that as we incorporate a larger radius around the maximum, the convolution deviates more from a Gaussian shape, resulting in a worse fit. We can assess how well a given set of points allows us to determine the center of the fitted function by calculating the error of the fit variables that encode the center of the Gaussian. The error of the fit is calculated by taking the inverse of the Fisher information matrix. By finding the radius that minimizes that error for a small subset of successive pictures, we can estimate the ideal radius for particle tracking in the whole movie.

We end up making one fit of 4 parameters for each set of successive points (x and y position of the center of the Gaussian, its variance, and an additive term accounting for the fact that the convolution does not level off to zero). As successive pictures typically do not differ much, we can initialize the fit of one step by the variance and additive term of the previous fit and initialize the x and y position of the Gaussian fit in the center of the discrete grid, which would correspond to a step of zero. This initialization of the fit parameters speeds up the convergence of the fit procedure but has the consequence that particle positions are not determined independently.

3.4. Algorithms to Determine Trajectories

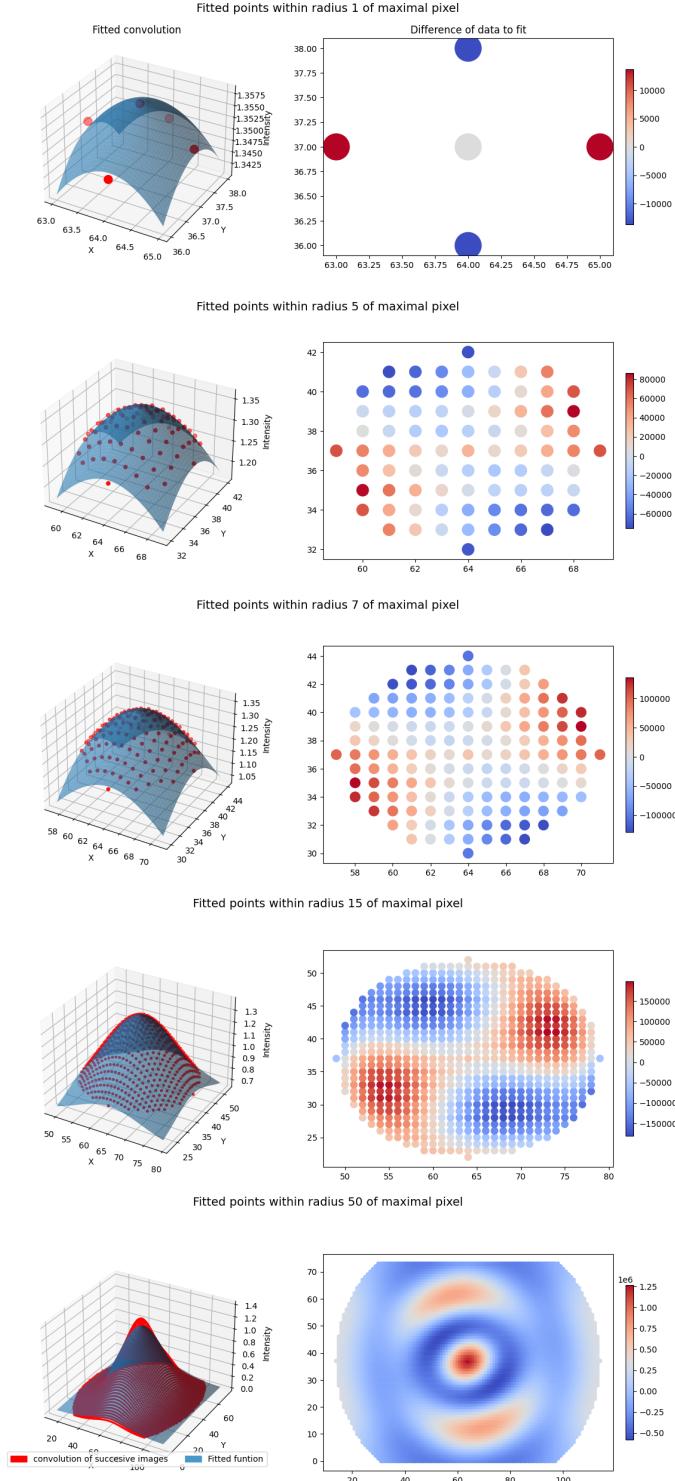


Figure 3.3.: Illustration of the Fitted Convolution method for detecting the bead position in an image. Example successive images of a bead in hydrogel from the experiment in section 4.1.1. On the left side, the convolution of two images is plotted around its maximum with a Gaussian fit of those points. On the right side, the error of the fit on each point is plotted. From top to bottom, we see a widening of the radius of the points around the center that we still consider for the fit.

3. Materials and Methods

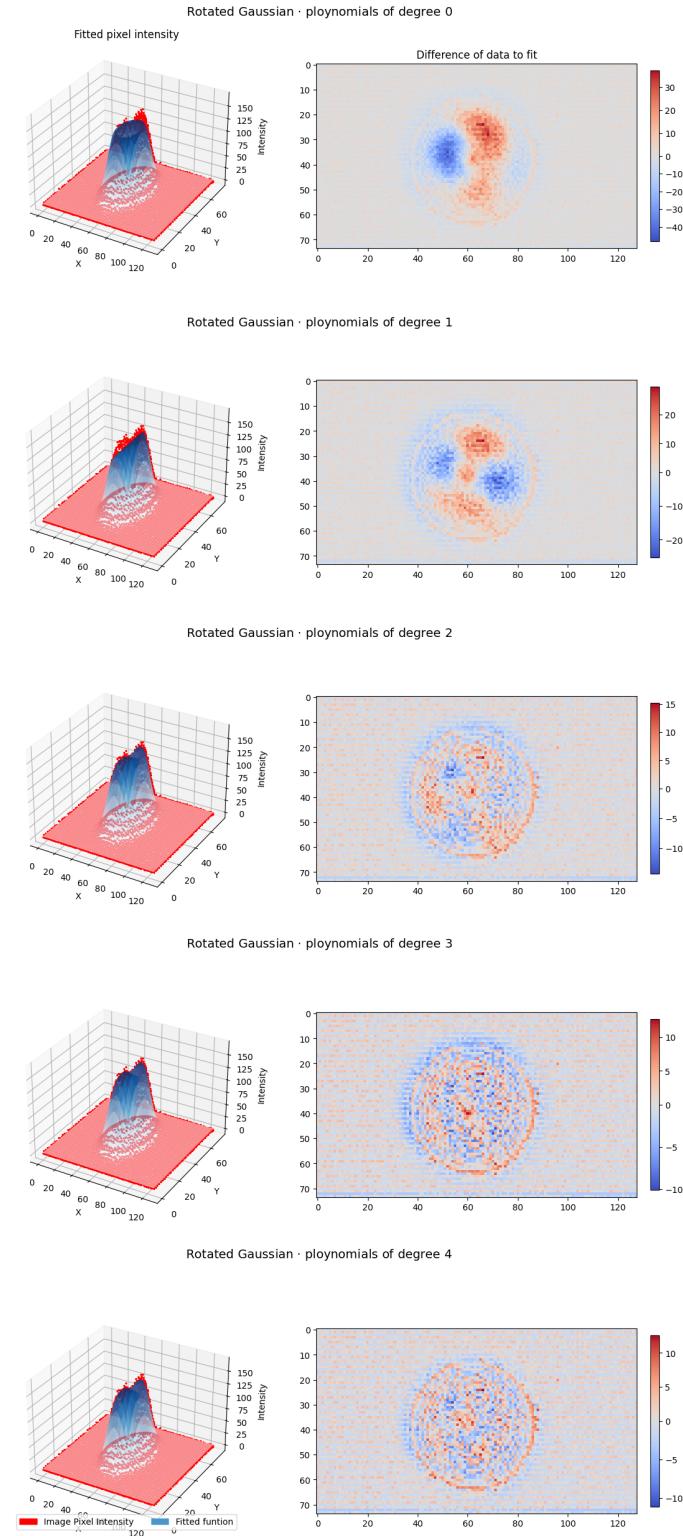


Figure 3.4.: Illustration of the Fitted Particle Shape method for detecting the bead position in an image. Example image of a bead in hydrogel from the experiment in section 4.1.1. On the left side, we see the image represented in a 3D plot where the pixel brightness is the z-dimension. Fitted on top of the image is a rotated Gaussian multiplied with a polynomial of n-th degree. The maximal degree of the polynomial increases from top to bottom. On the right, we see the error of the fit.

3.4.3. Fitted Particle Shape

In this method, we directly try to find the position of the particle in each frame by fitting the shape of the particle's image on the frame. For this, we need to have a mathematical model of the shape of the image of a silicone particle. We could already see in Figure 3.2 that the bead appears to be doughnut-shaped.

As shown in Figure 3.4, we can approximate the particle shape using a rotated Gaussian multiplied with a polynomial term:

$$f(x, y) = A \exp \left(-\frac{(\sqrt{(x - x_0)^2 + (y - y_0)^2} - r_0)^2}{2\sigma^2} \right) \left(c_0 + \sum_{i=1}^n \sum_{j=0}^i c_{ij} x^j y^{i-j} \right) + B \quad (3.3)$$

where:

- A is the amplitude of the Gaussian profile
- x_0 and y_0 are the coordinates of the center of the Gaussian profile
- r_0 is the radius offset of the Gaussian profile
- σ is the standard deviation of the Gaussian profile
- c_0 is the constant term of the polynomial background
- c_{ij} are the coefficients of the polynomial terms, with i representing the total degree of the term and j representing the degree of x
- n is the degree of the polynomial background
- B is the constant background level

We see in Figure 3.4 that as we increase the degree of the polynomial, we obtain a closer fit to the actual shape of the bead's image. However, we face a trade-off between the accuracy of the fit and the time it takes to fit those parameters, as higher-order polynomials greatly increase the number of parameters that need to be fitted. In practice, we never used a polynomial beyond the first degree, as it takes a prohibitively long time to fit.

Here, we also make use of the fact that the image changes little from frame to frame when initializing the fit parameters. However, if one were to initialize the fit parameters using the fitted values from the previous frame, one would have to

3. Materials and Methods

compute all fits serially. To make greater use of parallelization, we first fit every 100th frame without an informed initialization and then use those fitted parameters as initial parameters for the fits of the frames 50 time steps before and after that frame. While speeding up the fit, this also means that the fit of each frame is not independent of the others.

3.5. Algorithm to Compute Mean Back Relaxation

Efficiently computing the Mean Back Relaxation (MBR) is crucial for its practical application to large datasets. Here, we present a high-level overview of the algorithmic approach employed to calculate the MBR efficiently.

As we can see in Equation 2.16, the Mean Back Relaxation is the average fraction of the displacement of a particle within two different time intervals. If we look at the denominator of that fraction, we see that its most likely values lie around zero, as long as the particle is not on a ballistic trajectory. So when we empirically measure the MBR by averaging over a finite number of data points, a data point with an extremely small denominator might dominate the mean. That would be despite the fact that those data points are the least informative. After all, if the negligible motion in the denominator were in the opposite direction (only a small change in the actual particle position), the data point dominating the empirical MBR would have a different sign. For this reason, we restrict the data points we consider for calculating the Mean Back Relaxation to those where the absolute value of the denominator is above some threshold value. This technique was also used by Münker et al. [19].

The algorithm for computing the MBR can be summarized in the following steps:

1. For each data point, identify the data point that lies the time distance τ in the past. If the sampling is done regularly and all points have the same time difference dt , this is choosing the data point τ/dt time steps ago. But when the sampling is irregular, this step is done by a Two Pointers Algorithm.
2. Calculate the displacement $x(0) - x(-\tau)$ for each data point. Estimate the standard deviation of these displacements and establish a threshold, typically one standard deviation. Retain only the data points whose displacement exceeds this threshold. If the uncertainty in step size detection is larger than this

3.6. Filtering and Fitting the Power Spectrum

threshold, use the step size uncertainty as the minimum displacement criterion instead. This ensures that we only consider displacements that are actual particle movements instead of measurement noise.

3. For each data point, identify the data points that are up to T time in the future. Here, T is the maximal value of t from Equation 2.16. This can again be done by counting or the Two Pointers algorithm. From these data points, we calculate the forward displacements $x(t) - x(0)$.
4. Now, we divide each forward displacement by its respective backward displacement and average all forward displacements over different data points. In the case of irregular sampling, we have to bin forward time distances here. If we have different trajectories and we want to calculate their combined MBR, we can also average over trajectories here.

To further optimize the computation, the algorithm can be parallelized by processing multiple trajectories or overlapping parts of the same trajectory simultaneously using multi-threading or distributed computing techniques. This allows for efficient utilization of computational resources and enables the analysis of large datasets in a reasonable amount of time.

These algorithms are part of the Python package accompanying this master thesis for use in other projects involving Mean Back Relaxation. The package is available on GitHub.

3.6. Filtering and Fitting the Power Spectrum

For finding noise peaks in the Power Spectrum, we have described a surprise minimization criterion in section 2.3.2. While this method tracks noise peaks more precisely than simpler methods used in similar works [24], it is more computationally intensive and requires some tweaking to be feasible in polynomial time.

3. Materials and Methods

Let us first unpack the formula for the total surprise that we want to minimize:

$$\begin{aligned}
S_{total} &= S_{spectrum} + S_{peaks} \\
&= \sum_{i \notin \Sigma_\rho} \left(\frac{y_i}{f(x_i, \omega)} - 1 \right) \\
&\quad - \log(p_{numberpeaks}(n_{peaks})) - H(p_{numberpeaks}) \\
&\quad + \sum_{i=0}^{n_{peaks}} (-\log(p_{lengthpeak}(l_i)) - H(p_{lengthpeak}))
\end{aligned}$$

Now, let us make some assumptions about our priors for the number of peaks and the length of each peak. These priors are primarily motivated to simplify the above expression. Since there is no obvious way to choose a prior in this context, we might as well opt for one that simplifies the computation:

$$\begin{aligned}
p_{lengthpeak}(l) &\propto 2^{-\frac{l}{l_{typical}}} \\
\rightarrow p_{lengthpeak}(l) &= \frac{1}{2^{1/l} - 1} 2^{-\frac{l}{l_{typical}}} \tag{3.4}
\end{aligned}$$

$$p_{numberpeaks}(n_{peaks}) \propto 2^{-\frac{1}{2} \left(\frac{n_{peaks}}{n_{typical}} \right)^2} \tag{3.5}$$

Finding the noise peaks in the data now reduces to finding the set of noise peaks Σ_ρ , such that the following expression is minimized:

$$\begin{aligned}
&\sum_{i \notin \Sigma_\rho} \left(\frac{y_i}{f(x_i, \omega)} - 1 \right) + \log(2) \left(\frac{n_{peaks}}{n_{typical}} \right)^2 + \sum_{i=0}^{n_{peaks}} \log(\sqrt[1]{2} - 1) + \log(2) \frac{l}{l_{typical}} - H(p_{lengthpeak}) \\
&= \sum_{i \notin \Sigma_\rho} \left(\frac{y_i}{f(x_i, \omega)} - 1 \right) + \sum_{i \in \Sigma_\rho} \left(\log(2) \frac{1}{l_{typical}} \right) + \log(2) \left(\frac{n_{peaks}}{n_{typical}} \right)^2 - n_{peaks} (H(p_{lengthpeak}) - \log(\sqrt[1]{2} - 1)) \\
&\propto \underbrace{\sum_{i \notin \Sigma_\rho} \left(\frac{y_i}{f(x_i, \omega)} - 1 - \log(2) \frac{1}{l_{typical}} \right)}_{L_i} + \log(2) \left(\frac{n_{peaks}}{n_{typical}} \right)^2 - n_{peaks} (H(p_{lengthpeak}) - \log(\sqrt[1]{2} - 1)) \tag{3.6}
\end{aligned}$$

Let us call the final term the Loss L , as it is the quantity we want to minimize. We have omitted all terms that do not depend on the choice of Σ_ρ . In particular, we obtain the last line by subtracting the term $\log(2) \frac{n_{data}}{l_{typical}}$. $H(p_{lengthpeak})$ is just a constant once we choose $l_{typical}$. See appendix B for the calculation. The Loss is

3.6. Filtering and Fitting the Power Spectrum

now decomposed into a term for every data point L_i . We search for the Σ_ρ that minimizes this expression using the following algorithm:

1. We begin by assuming there are no noise peaks ($\Sigma_\rho = \emptyset$).
2. We look for the n_{start} and n_{end} such that $L_\Sigma = \sum_{i=n_{start}}^{n_{end}} L_i$ is minimized. This can be done with Kadane's Algorithm in linear time [12].
3. Now we compare L_Σ with $L_n = \log(2)^{\frac{n_{peaks}}{n_{typical}}} - H(p_{lengthpeak})$. If L_Σ is smaller than L_n , we can decrease the Loss L by adding $n \in \mathbb{N} | n_{start} \leq n \leq n_{end}$ to Σ_ρ .
4. Next, we return to step two to try to find a new peak. To make it possible to unlabel data points as noise peaks, we invert the sign of the L_i of each data point currently labeled as a noise peak. Since unlabeled data point i increases L by L_i , we can add a noise peak by labeling some of the data points as non-noise peaks.
5. We repeat this process until no further peak is found. When this condition is met, we refit the PSD function to the data that was not classified as a noise peak, and restart the entire process. This refitting of the PSD function is performed a predefined number of times (arbitrarily chosen as 10 in this case).

This algorithm has two free parameters that are not obvious how to choose: $n_{typical}$ and $l_{typical}$. In Bayesian data analysis, it is common for priors to be free parameters, as they do not tell us something about the system under consideration, but about our state of expectation about the system. However, we can still try to choose these free parameters in a principled way:

For $n_{typical}$, we choose the number by default in such a way that there is only a 5% chance of finding a peak when there is none. This way, this Bayesian approach conforms to the frequentist conception of statistical significance[8]. See appendix C for the calculation.

For $l_{typical}$, we choose the default value of 1. Examining equation 3.5, this implies that we require one bit of evidence to classify an additional data point currently bordering a peak as part of that peak. This means there can be up to a 50% chance that we obtain such strong evidence even if this point is not actually part of the noise spectrum. Consequently, when we converge on a peak of length l , this represents our best estimate of the peak's length, with the understanding that the marginal data point at the peak's edge has an approximately equal likelihood of belonging

3. Materials and Methods

to the peak or not. This approach contrasts with our more conservative estimation of the number of peaks: we only add another peak if we can be reasonably certain that at least that many peaks exist.

It's important to note that this algorithm's success heavily depends on the initial fit, which takes all data points into account, being approximately correct. We begin by identifying noise peaks as those points that seem unlikely given that initial fit. Therefore, if there is excessive noise at the outset, we may not achieve a sufficiently accurate first fit to correctly identify noise peaks, potentially causing the process to fail. Some particularly noisy datasets can be salvaged by adjusting the hyperparameters. Specifically, choosing a lower value for $l_{typical}$ makes the algorithm more hesitant to declare large sections of the Power Spectrum as noise peaks.

3.7. Simulating Trajectories from Power Spectral Density

To simulate trajectories from a given Power Spectral Density (PSD), we can leverage the relationship between the PSD and the Fourier transform of a trajectory. While the process of calculating the PSD from a trajectory (as described in equations 2.4, 2.5, and 2.6) is not directly reversible due to the loss of phase information, we can generate trajectories with the desired PSD characteristics.

The approach involves creating a Fourier series with magnitudes derived from the target PSD and random phases. Specifically, we draw each phase from a uniform distribution between 0 and 2π . This method allows us to generate trajectories that, on average, exhibit the desired PSD shape.

To account for the inherent variability in real-world power spectra, we can introduce stochasticity into the process. As observed in Figure 2.3, elements of the Power Spectrum follow an exponential distribution. Therefore, we can draw each element of the simulated Power Spectrum from an exponential distribution with the target PSD value as its mean. This approach produces trajectories that not only match the desired PSD on average but also capture the natural variability observed in experimental data.

This method can also be adapted to recreate trajectories without noise peaks that were present in the original data. Using the noise peak detection method discussed earlier, we can identify elements of the Power Spectrum associated with noise peaks. These elements can then be replaced by drawing new values from an exponential

3.7. Simulating Trajectories from Power Spectral Density

distribution, with the mean determined by our best-fit PSD curve at those frequencies. This process effectively removes the influence of noise peaks while maintaining the overall characteristics of the trajectory.

By applying this technique to all trajectories in which we identify noise peaks, we can generate a set of filtered trajectories that retain the essential features of the original data while eliminating unwanted noise artifacts. We can also simulate trajectories that adhere to a any model from section 2.1.1 with a given set of model parameters.

This algorithm is also published on the accompanying GitHub page, which also demonstrates how to use it on new datasets.

4. Results

4.1. Imaging

The first step towards measuring the Mean Back Relaxation of an object is to measure its trajectory with high resolution in time and space. Here we used Dark field Microscopy, to image samples of silicone beads in living and non living materials.

4.1.1. Imaging Hydrogel

Using the microscope setup described in Section 3.1.2, we imaged silicone beads infused into hydrogel samples. The samples were prepared following the protocol outlined in Section 3.2.

Our imaging experiments revealed several interesting observations and technical

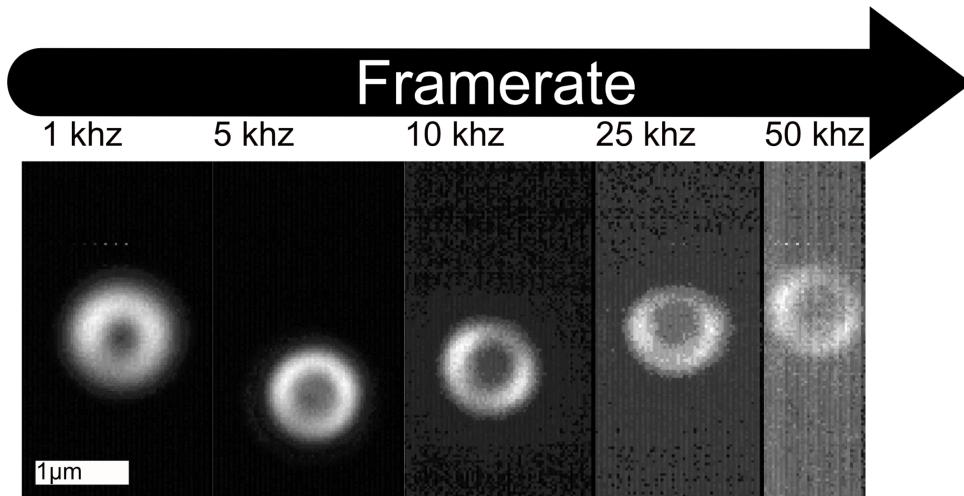


Figure 4.1.: Pictures of silicone beads in hydrogel made with darkfield microscopy.

The different pictures are frames of movies taken with increasing frame rate, such that the pictures from left to right have decreasing illumination time.

limitations. One of the most notable aspects of our setup was the ability to achieve

4. Results

remarkably high frame rates, up to 50 kHz. It's worth noting that the upper limit of our time resolution was primarily constrained by the camera's capabilities rather than the microscope setup itself. Figure 4.1 presents a series of darkfield microscopy images of silicone beads in hydrogel, captured at progressively higher frame rates from left to right. As the frame rate increases, the illumination time per frame decreases, resulting in visibly dimmer images. This reduction in light collection leads to several observable effects. Most notably, the contrast between the particle and the background diminishes, and the overall image becomes noisier. At the highest frame rate of 50 kHz, we observed artifacts in the form of bright stripes across the image.

Interestingly, artifacts were not limited to high frame rates. At lower frame rates, such as 1 kHz, we noticed persistent bright pixels in the images. These anomalies remained fixed relative to the camera sensor, even when the field of view was shifted, indicating that they were caused by defects in the camera's detector rather than features of the sample or optical setup. While these artifacts were infrequent enough to avoid directly overlapping with the bead images, they posed challenges for image cropping and analysis.

Another consequence of increasing the frame rate was a gradual degradation of the bead's image quality. At lower frame rates, the bead appeared as a well-defined donut shape, characteristic of darkfield imaging. However, this distinct shape became increasingly distorted at higher frame rates.

It's important to note that achieving such high frame rates required a trade-off in the field of view. The camera could only handle these rapid acquisitions by significantly cropping the image along one dimension. At 50 kHz, the field of view was barely large enough to encompass a single bead. This limitation suggests that even higher time resolutions might be achievable by using smaller particles, which would allow for further cropping of the field of view while still capturing the entire particle.

4.1.2. Imaging HeLa Cells

Following our experiments with hydrogel, we proceeded to image silicone beads infused in HeLa cells, prepared according to the protocol described in section 3.3. This presented additional challenges compared to the hydrogel measurements.

The environment in which the cells were maintained during the measurement was not ideal, as it lacked temperature control and a CO₂-independent medium. However, the most significant issue turned out to be light exposure. The cells did not

4.1. Imaging

exhibit any visible differences after being under the microscope for 20 minutes when not illuminated by the microscope lamp. However, after capturing the first video of beads within a cell with full illumination, we observed that the cell's morphology became irregular and distorted, indicative of necrosis or apoptosis [10, 14]. This necessitated a careful approach to minimize light exposure and preserve cell viability.

To address this challenge, we implemented the following strategy:

1. We only turned on the full illumination during actual image capture.
2. When searching for new samples or adjusting the microscope, we used dimmed illumination.
3. We turned off the light completely during data transfer from the camera to the computer.

This approach allowed us to maintain cell viability for multiple sample readings, preventing signs of necrosis in the probe for an extended period. However, it's important to note that during the actual recording, the cells were still receiving a UV dose that would be lethal if sustained for more than a few hundred seconds.

The need to limit light exposure made imaging at high frame rates particularly challenging. Each measurement required extensive calibration of all microscope settings immediately before capture to ensure optimal imaging of the bead. This was necessary because we couldn't rely on continuous illumination for adjustments.

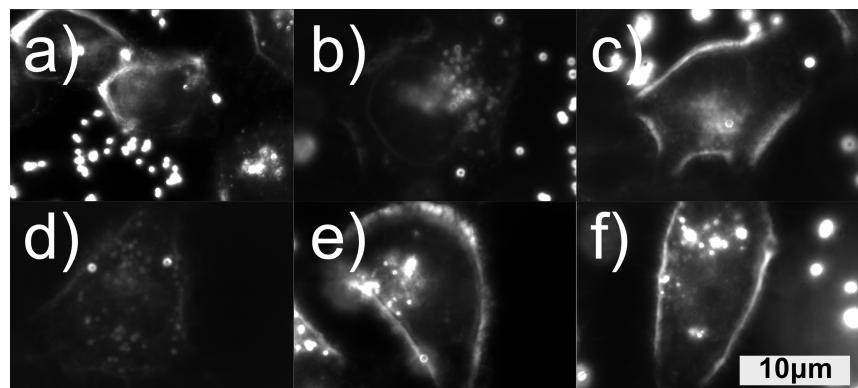


Figure 4.2.: Pictures of the cells in which the beads where measured. Table A.2 tracks which measurement is taken from which cell.

4. Results

Despite these challenges, we were able to image beads in a variety of HeLa cells. Figure 4.2 shows the pictures of some imaged cells.

4.1.3. Image cleaning

4.2. Tracking Beads

To calculate the Mean Back Relaxation (MBR) from a video of a particle, we first need to extract its trajectory from the video. We employ the three tracking methods introduced in section 3.4. However, obtaining a high-resolution trajectory in time and space is not sufficient for calculating the MBR. We also need to determine the error associated with that trajectory to establish the minimum value of τ in equation 2.16 that can be used without measuring the Mean Back Relaxation of measurement error rather than actual particle movement. Understanding the precision of our measurement setup is crucial for evaluating the potential of Dark Field Microscopy for these types of measurements. To assess our accuracy as reliably as possible, we employ three independent approaches.

4.2.1. Tracking simulated trajectories

In this section, we analyze the performance of our tracking algorithms using simulated particle trajectories. We begin by taking a single frame from the videos discussed in section 4.1.1 and create a simulated movie by concatenating copies of this image, each shifted by a specific amount along the x and y directions. To enhance realism, we add Gaussian noise to each pixel, with the noise magnitude determined by the variance of non-bead pixels in the original image.

This method allows us to simulate sub-pixel movements. For example, to shift a particle 10% of a pixel to the right, we replace each pixel's brightness with 90% of its original value plus 10% of its left neighbor's brightness.

With these simulated movies, we can apply the tracking algorithms described in section 3.4 and compare the reconstructed trajectories against the known true paths. Our primary interest lies in determining the range of step sizes for which the error in the reconstructed step is significantly smaller than the actual step size.

While it might seem more intuitive to focus on the absolute error of particle position detection, our two most promising tracking methods work differently. The Fitted Convolution method directly estimates step sizes, and the Fitted Particle

4.2. Tracking Beads

Shape method uses the previous frame's predicted position as its initial guess. Consequently, the accuracy of these methods in measuring inter-frame steps likely depends on the step size itself: smaller steps may lead to more accurate fits as the initial parameters are closer to the true values, reducing the risk of getting stuck in local minima.

Figure 4.3 demonstrates that the tracking algorithms generally reconstruct the original trajectories quite well. Interestingly, the reconstruction error does not appear to depend significantly on the step size. For the 1 kHz frame rate movie, all three tracking algorithms qualitatively track the true trajectory, with the Mean Pixel Activation method showing a higher reconstruction error of 2 nm, while the other methods perform similarly with errors around 0.5 nm.

As we increase the frame rate to 10 kHz, the Mean Pixel Activation method becomes unreliable, and the accuracy of other methods decreases, resulting in an error of about 3 nm. At 25 kHz, the Mean Pixel Activation method and the Fitted Particle Shape methods have an average error between 4 and 5 nm. However, when examining the plotted trajectories, we observe that the Fitted Particle Shape method tracks the true trajectory much more accurately. The Fitted Convolution method fails completely, consistently returning a step size of zero. At these noise levels, the Gaussian fit in the convolution, as described in section 3.4.2, may fail and return its initial guess value, corresponding to a step size of 0. This then determines the initial guess for the next fit, which becomes even more inaccurate and also fails. While this doesn't always occur with the Fitted Convolution method at these frame rates, we chose not to cherry-pick the data. This example showcases the relative brittleness of this method in the presence of significant noise.

4. Results

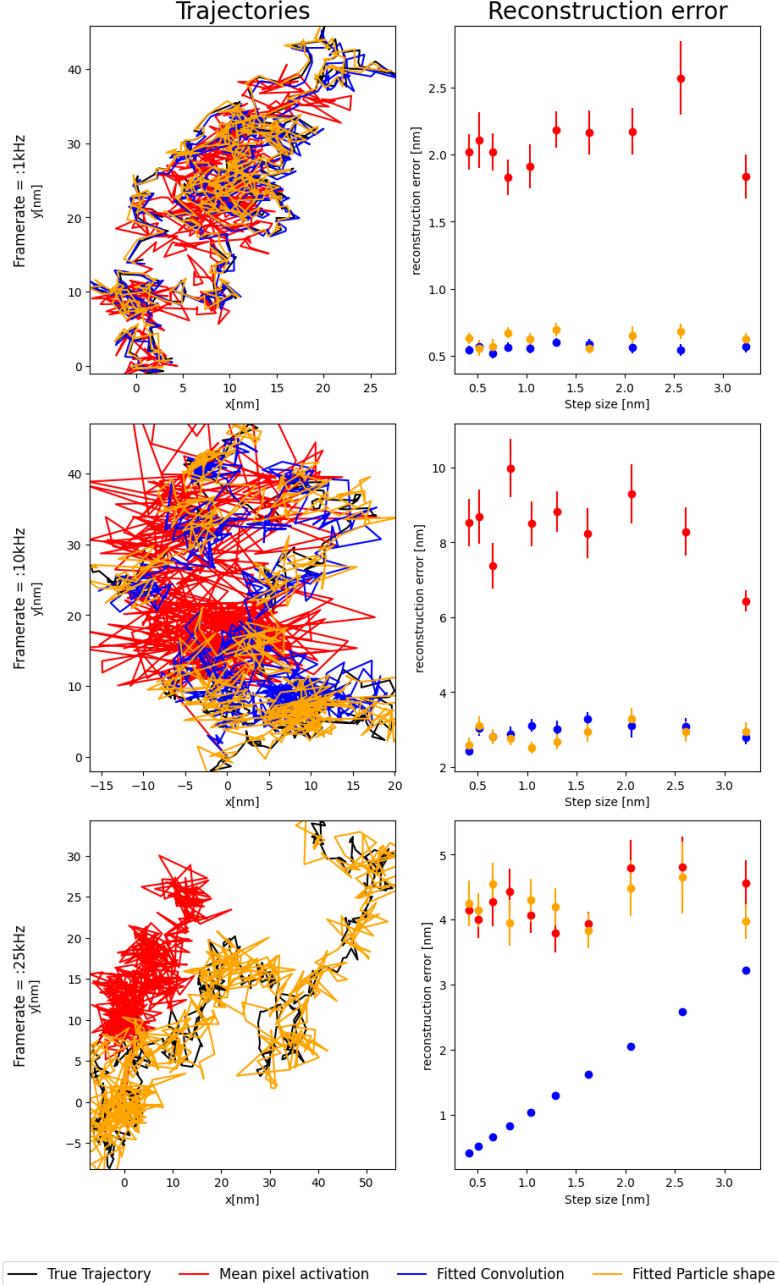


Figure 4.3.: Reconstruction of simulated particle trajectories using methods described in section 3.4. Trajectories are created by shifting images from figure 4.1 by fractional pixel values. The images represent frames from movies with increasing frame rates from top to bottom. Each trajectory consists of 1000 steps, with step sizes varying from 0.3 to 10 nm. The left panel shows the first steps of the original trajectory (black) overlaid with reconstructed trajectories (colored lines), providing a qualitative comparison. The right panel presents a quantitative error analysis, plotting average reconstruction errors for steps grouped by size. In the final panel, the Fitted Convolution method failed to detect any movement, and the plotted trajectory is therefore not visible.

4.2.2. Tracking Piezostage Movement

To further analyze the precision of our tracking algorithms, we created a video of a bead trajectory with controlled movement. We embedded beads in a two-component glue and inserted the piezo-controlled stage into the microscope as described in section 3.1.2. We then imaged the immobilized beads while moving them with the piezo stage.

This approach provides a known trajectory (the piezo stage movement) against which we can compare the measured trajectories from our detection methods. However, we cannot directly compare these trajectories because the x-y axes along which the piezo stage moves do not align perfectly with the camera's x-y axes. To address this misalignment, we used the first 100 steps of the trajectory as "training data." Given the first 100 detected particle positions \vec{d}_i and piezo stage positions \vec{s}_i , we optimized a translation matrix M to minimize $\sum_i(\vec{d}_i - M\vec{s}_i)^2$. This matrix converts the piezo stage positions into the camera's frame of reference. We then used this matrix to convert the remaining piezo stage positions and compared only those with their corresponding detected positions. This approach prevents data leakage from the translation matrix fit from artificially inflating the apparent accuracy of the detection.

As shown in Figure 4.4, the reconstruction error increases with step size. The Fitted Particle Shape method consistently outperforms or matches the other methods. The Fitted Convolution method is a close second for low frame rates but falls behind beyond 10 kHz. The Mean Pixel Activation method consistently performs worst, though qualitatively, all three methods generally follow the true trajectory. The Fitted Particle Shape method visibly tracks the true trajectory most accurately. Errors range from approximately 4 nm for high frame rates (40 kHz) to 2 nm for lower ones (5 kHz).

4. Results

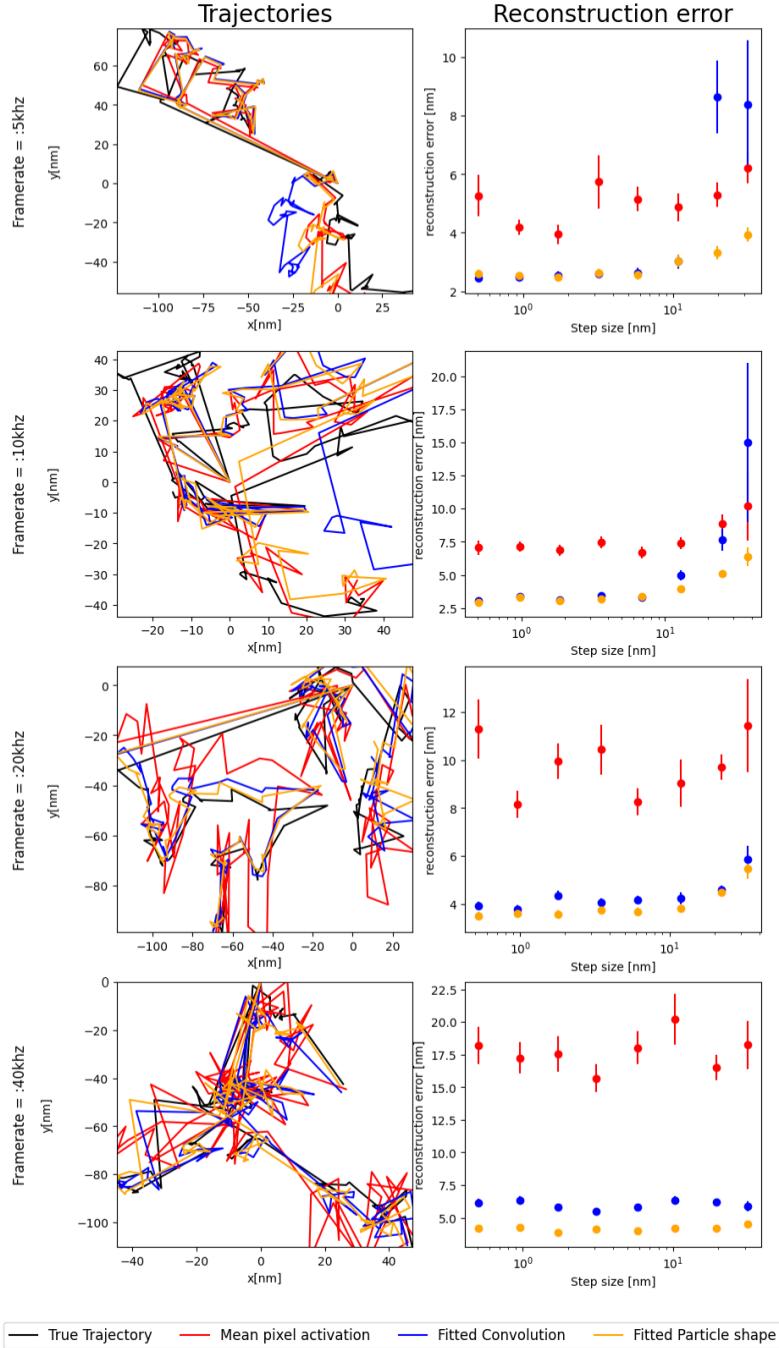


Figure 4.4.: Reconstruction of particle trajectories created with a piezo stage using methods described in section 3.4. The left panel shows the first steps of the original trajectory (black) overlaid with reconstructed trajectories (colored lines), providing a qualitative comparison. The right panel presents a quantitative error analysis, plotting average reconstruction errors for steps grouped by size. From bottom to top, we see an increase in the frame rate at which the videos were made.

4.2.3. Calibration via statistical measures

The final method we use to estimate the error of our detection methods utilizes our model of the Power Spectrum laid out in formulas 2.13-2.15. We observe that all terms in the Power Spectrum vanish for high frequencies, except for the noise term. This means that by examining the convergence of a Power Spectrum, we can estimate the measurement noise. This approach does not rely heavily on our specific models for the Power Spectrum of the particles. As long as the particle moves continuously, the Power Spectrum of the actual particle motion must approach zero at high frequencies. Therefore, any constant component in the high-frequency range of the Power Spectrum must originate from noise in the particle detection.

However, it's important to note that some types of detection noise may not manifest in this way in the Power Spectrum. For example, imagine a tracking algorithm that correctly determines the step size between two points in time but randomly guesses the direction of the step. This would be an inaccurate tracking method, but its noise would appear as zero in the Power Spectrum. As we see in Figure 4.4, there is evidence that some of the tracking methods under consideration do make errors that scale with step size.

As shown in Figure 4.5, the high-frequency limit for the Power Spectrum of the trajectory extracted via the Mean Pixel Activation method is higher than those corresponding to the other methods. The trajectory plot on the left also qualitatively confirms this observation, as the Mean Pixel Activation method's trajectory exhibits more erratic jumps, while the other methods track each other relatively closely. The measured noise levels are approximately 1 nm for both the Fitted Particle Shape method and the Fitted Convolution method, and around 4 nm for the Mean Pixel Activation method.

4. Results

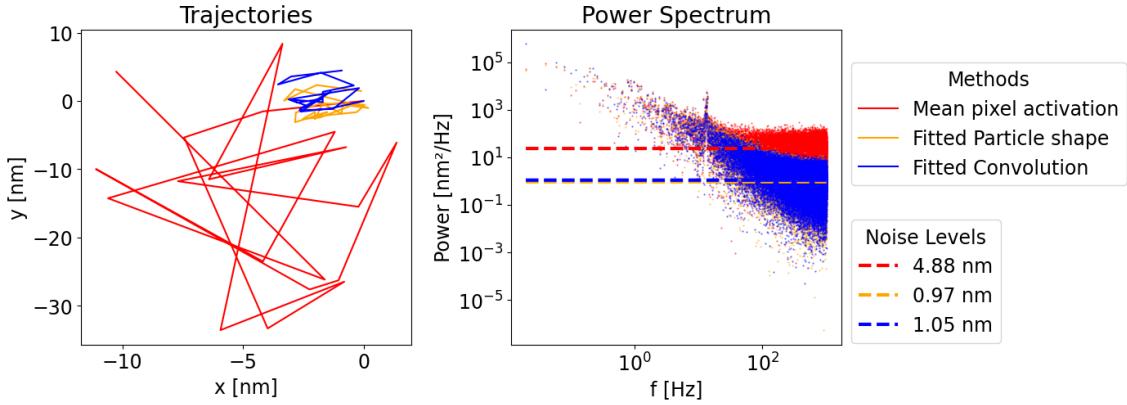


Figure 4.5.: Estimation of tracking method errors from section 3.4 using high-frequency behavior in the Power Spectrum. Beads in movies generated as described in section 4.1.1 with a frequency of 10 kHz are tracked with the algorithms under consideration. The Power Spectra of these trajectories are shown. The final term from equation 2.15 is fitted to the high-frequency parts of the Power Spectrum. The value of the fit parameter σ , which can be interpreted as the error of the tracking method, is marked. The left panel shows the first steps of the trajectory, while the right panel displays the Power Spectra with their high-frequency asymptotes marked.

4.3. Filtering and Fitting

Given that the process for fitting data and filtering out noise peaks described in Section 2.3 differs from previously implemented algorithms, we begin by validating our method against simulated data. Subsequently, we apply this method to filter and fit the trajectories extracted from our experimental data.

4.3.1. Validating the Fitting Method with Synthetic Data

To test our Power Spectrum fitting method, we use the Maxwell Model, the Kelvin-Voigt Model, and the Fractional Kelvin-Voigt Model to generate different Power Spectra with the simulation method described in section 2.3. This is accomplished using equations 2.13, 2.14, and 2.15, respectively, with arbitrarily chosen model parameters. The selected parameters are listed in Table 4.2, and the resulting Power Spectra are illustrated in Figure 4.6a.

We then simulate trajectories based on these Power Spectra, as shown in Figure 4.6b. Next, we calculate the Power Spectrum of each simulated trajectory and fit equations 2.13, 2.14, and 2.15 to these spectra using the method described in Section

4.3. Filtering and Fitting

2.3. The results of these fits are presented in Figures 4.6c-e.

If our fitting algorithm is effective, we should be able to reconstruct from the Power Spectrum data both the model used to create the underlying trajectory and its parameters. We employ the Bayesian Information Criterion, as described in Section 2.3.3, to obtain a posterior distribution over the different models for each dataset. The results of this analysis are presented in Table 4.1.

As evident from Table 4.1, we can identify the underlying model with high confidence. We then compare the parameters of the best-fitting model to the original simulation parameters. This comparison is shown in Table 4.2. While the fitted parameters are generally close to the original values, some discrepancies remain. In particular, the parameters A and α show errors approximately one order of magnitude smaller than their values.

	Maxwell	K-V	Frac. K-V
Maxwell	0.997	0.00315	3.18e-8
K-V	0	1	0
Frac. K-V	0	3.44e-53	1

Table 4.1.: Posterior probabilities of best-fitting models for Power Spectra generated from simulations shown in Figure 4.6. Rows represent the model used to generate each Power Spectrum, while columns show the fitted models. Values are calculated using the Bayesian Information Criterion as described in Section 2.3.3. The high values along the diagonal indicate that our fitting method correctly identifies each model. K-V: Kelvin-Voigt; Frac. K-V: Fractional Kelvin-Voigt.

4. Results

	A	B	α	β	noise
Simulation Parameters	0.5	1.0	0.1	1	0.01
Maxwell data fit	0.488				0.010
K-V data fit	0.521	0.991			0.010
Frac. K-V data fit	0.470	1.012	0.089	0.982	0.010

Table 4.2.: Comparison of simulation parameters and recovered fit parameters from Figure 4.6. Parameters correspond to equations 2.13, 2.14, and 2.15. For readability we here just name the prefactors in front of the ω terms A and B , even though they have more complicated forms in the Maxwell and Kelvin-Voigt model. Fit parameters are taken from the model with the highest posterior probability according to the Bayesian Information Criterion (see Table 4.1). Note that α and β are only applicable to the Fractional Kelvin-Voigt model and thus cannot be recovered when fitting data generated by other models. K-V: Kelvin-Voigt; Frac. K-V: Fractional Kelvin-Voigt.

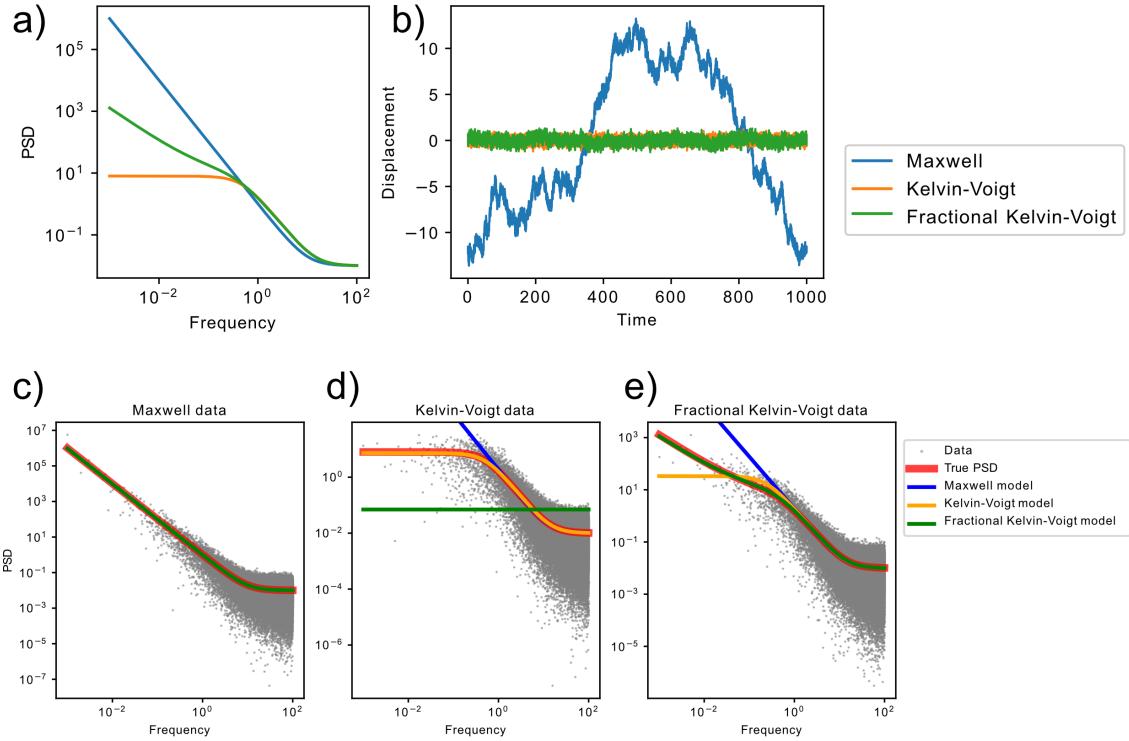


Figure 4.6.: Demonstration of fitting methods recovering information from simulated trajectories. Three trajectories with 200,000 data points in arbitrary units of time and space are created. (a) Average Power Spectral Density calculated from equations 2.13, 2.14, and 2.15. (b) The three trajectories in real space. (c-e) All points of the Power Spectrum with fits from the three models overlaid on each dataset. Red lines show the actual expected Power Spectral density from (a). The models used to create the Power Spectra are: (c) Maxwell model, (d) Kelvin-Voigt model, and (e) Fractional Kelvin-Voigt model. Model comparison using the Bayesian Information Criterion is shown in Table 4.1. Fit parameters of the best fit compared to actual simulation parameters are presented in Table 4.2.

4. Results

4.3.2. Filtering and Fitting the Trajectories

We proceeded to filter all trajectories extracted from the experiments performed in sections 4.1.1 and 4.1.2. As an example, Figure 4.7 shows the fit and filter of a trajectory taken from the hydrogel sample. As evident in Figure 4.7, there is

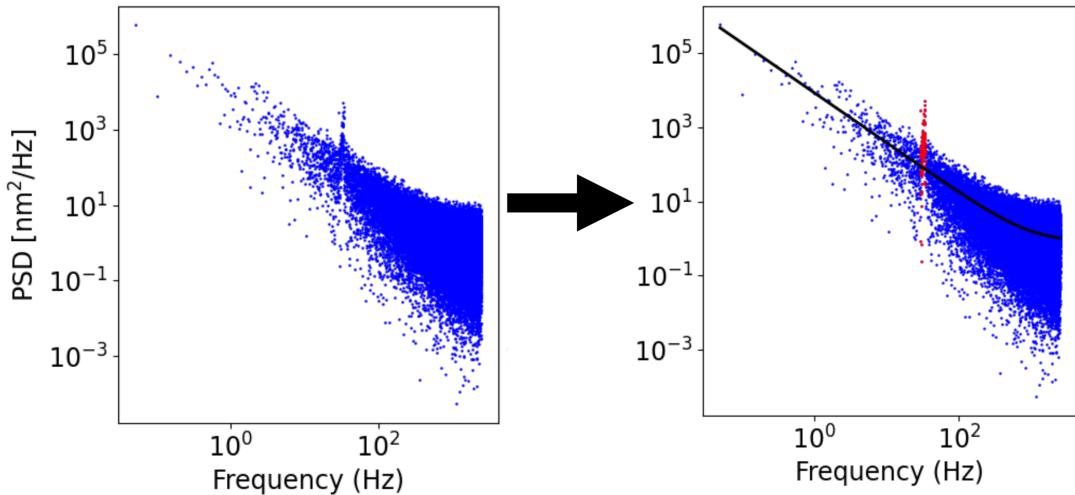


Figure 4.7.: Power Spectrum filtering and fitting for tracked beads. The depicted Power Spectrum is from a bead in hydrogel filmed at a frame rate of 5 kHz. The left panel shows the raw Power Spectrum. The right panel displays the same spectrum with the best fit (black line) and identified noise peaks (red points).

a clear noise peak at a specific frequency, which is successfully identified by our noise peak detection algorithm. The fit of the Power Spectrum appears to match the data points well. We observe the characteristic shape of a decreasing Power Spectrum for higher frequencies that levels off into a plateau, which is typical of noisy measurements. All Power Spectra plots and their corresponding fits can be found in Appendix A. Although each trajectory is two-dimensional, our previous data analysis methods (e.g., Power Spectral Density and Mean Back Relaxation calculations) assume one-dimensional trajectories. To address this, we decompose each particle trajectory into its x and y components, which are treated separately. This decomposition allows for an interesting cross-check: since each noise peak in the Power Spectrum represents a vibration of the experimental setup, it is likely not perfectly aligned with the x and y axes we use to decompose our trajectories. Consequently, we expect each noise peak to appear in both the x and y components

of the same trajectory, albeit potentially with different intensities.

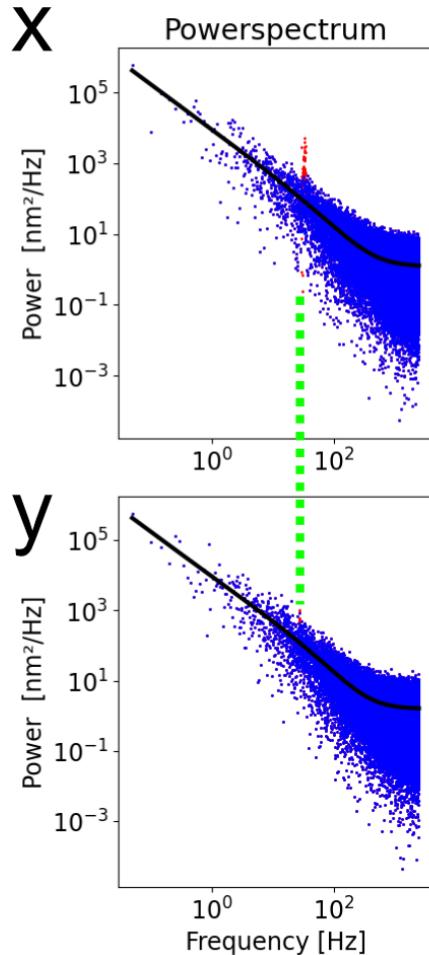


Figure 4.8.: Noise filtering can identify peaks not perceptible to the naked eye. The figure shows filtered and fitted Power Spectra of the x (top) and y (bottom) components of a bead in hydrogel filmed at 5 kHz. The top panel displays a clearly visible and successfully filtered noise peak. At the same frequency in the bottom panel (marked by the green line), a few red dots indicate a filtered noise peak that is not immediately apparent.

This property allows us to validate our peak-finding method by checking whether noise peaks that mirror those in the other component are detected, even when they are not visually apparent. Upon examining the Power Spectra, we indeed find such cases, as showcased in Figure 4.8. The red dots in the lower plot, while not easily visible, can be seen upon close inspection. In addition to the distinct bumps in the Power Spectrum expected from resonance frequencies in the experimental setup, we observed some identified noise peaks with a different appearance, particularly in the

4. Results

cell measurements. Figure 4.9 illustrates an example of this phenomenon. In Figure

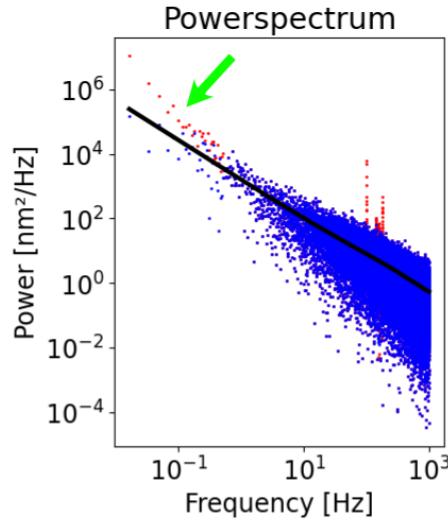


Figure 4.9.: Cell activity can also be classified as a noise peak. The figure shows the filtered and fitted Power Spectrum of a bead in a HeLa cell filmed at 10 kHz (same dataset as in Figure A.4 c). In addition to the characteristic noise peaks from resonance frequencies, we observe higher Power Spectrum values at low frequencies (indicated by the green arrow). This can be explained by the activity term in equation 2.10, corresponding to the cell's activity.

4.9, we observe an upward deviation from the expected Power Spectrum in the low-frequency regime. As these deviations only occur in trajectories taken from beads in cells, they likely originate from the activity term in equation 2.10. In our fits, we always assume E_0 to be 0, but we know this is not the case in living organisms. Directly incorporating E_0 as a free parameter in the fit proved challenging due to the limited number of data points in the low-frequency spectrum. Nevertheless, this observation suggests that information about the active driving of these beads is encoded in the Power Spectrum.

4.4. Calculating Mean Back Relaxation

4.4.1. Validating Mean Back Relaxation Calculation on a simulation

To ensure the accuracy and versatility of our MBR calculation methods, we tested them against a well-understood model system: the "horse and cart" model. This model, described in detail in section 2.3.2, provides an analytical solution for the MBR (equation 2.18), allowing us to compare our numerical results directly with the theoretical prediction.

We implemented several MBR calculation methods, as discussed in section 3.5, to handle various data scenarios:

1. Single dataset with uniform time steps
2. Single dataset with varying time steps
3. Multiple datasets with uniform time steps
4. Multiple datasets with varying time steps

To test these methods, we generated four different datasets based on the horse and cart model:

1. A single trajectory of one million time steps
2. The same trajectory with half of the data points randomly removed, simulating varying time steps
3. The original trajectory divided into 12 equal parts, simulating multiple shorter trajectories
4. The divided trajectory with half of the data points randomly removed from each part

These datasets allow us to evaluate our MBR calculation methods under different conditions that might be encountered in real experimental data.

Figure 4.10 presents the results of our MBR calculations alongside the analytical solution. Remarkably, all four datasets yield results that closely match the theoretical prediction. This demonstrates that our MBR calculation methods are robust and accurate across different data structures and sampling schemes.

4. Results

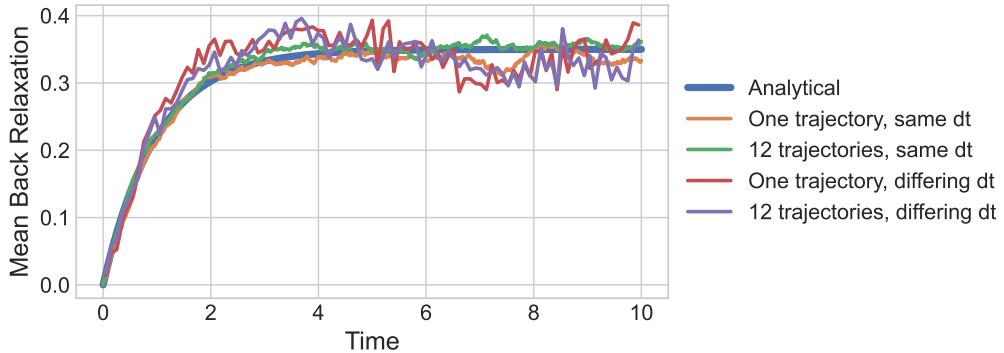


Figure 4.10.: Mean Back Relaxation of the Horse and Cart model. The analytical solution (blue line) is compared with empirical MBR calculations from various simulated datasets. We vary whether the time step between two points in the trajectory is constant or variable, and whether we have a single long trajectory or multiple shorter trajectories. Model parameters: $D_q = 0.5$, $D = 1$, and $\beta = 1$.

The empirical MBR curves consistently follow the analytical solution, indicating that our methods can handle both uniform and varying time steps, as well as single long trajectories and multiple shorter ones. This accuracy is maintained across different timescales, with the empirical curves capturing both the short-time behavior and the long-time limit of the analytical solution. These results validate our implementation of the MBR calculation methods and provide confidence in their application to real experimental data.

4.4.2. Calculating Mean Back Relaxation on experimental data

Following the extraction and filtering of particle trajectories, we proceeded to calculate the Mean Back Relaxation (MBR) for all previously discussed trajectories. The complete set of results can be found in Appendix A.

The calculated MBRs exhibited the expected general shape, starting at 0 and approaching a seemingly stable value over a time scale of approximately 0.5 seconds. For beads in hydrogel, we observed the anticipated behavior of relaxation towards the 0.5 value expected for passive systems. Interestingly, the relaxation time increased for larger values of τ (compare with figure A.1 a/b).

The MBR measurements in HeLa cells showed mostly similar patterns to those in hydrogel. The majority of these measurements also relaxed to the passive value of 0.5. However, some measurements, particularly in the 2 kHz range with high τ

4.4. Calculating Mean Back Relaxation

values, tended towards values different from 0.5.

Figure 4.11 presents a direct comparison between the MBR of a bead in hydrogel and a bead in a cell with a notably small long-time MBR value.

This figure clearly illustrates the stark difference in shape that can be attributed

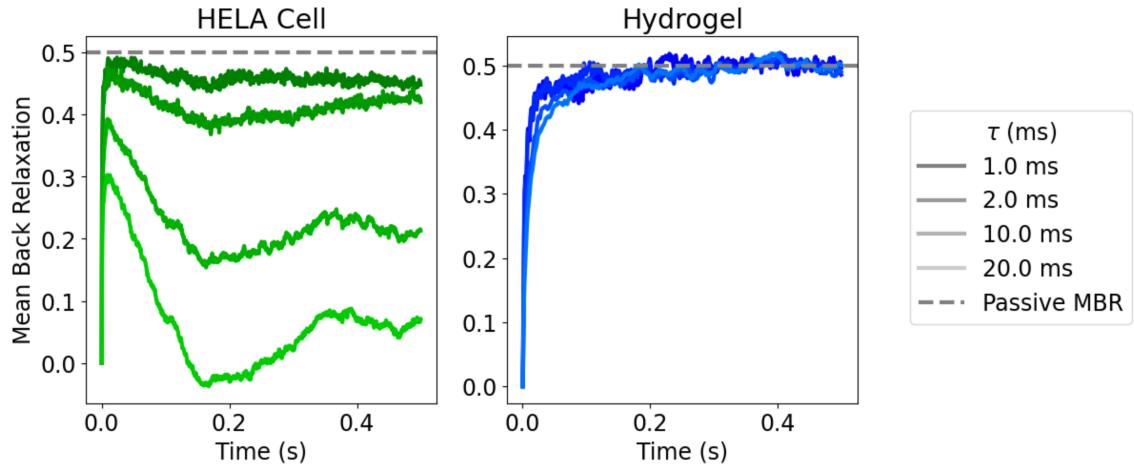


Figure 4.11.: Mean Back Relaxation (MBR) distinguishes between active and passive systems. The MBR, defined in equation 2.16, is plotted with the t parameter along the x-axis and the τ parameter represented by lines of varying darkness. Left panel: MBR values for a bead in a HeLa cell (corresponding to Fig. A.2 c). Right panel: MBR for a bead in hydrogel (corresponding to Fig. A.1 a). The grey horizontal line indicates the expected long-term MBR value (0.5) for passive systems. Note the distinct behaviors: the hydrogel (passive) consistently approaches 0.5, while the HeLa cell (active) shows deviations, particularly for larger τ values.

to the distinction between active and passive motion. The hydrogel sample (passive system) shows a consistent relaxation to 0.5 in almost all measurements, while the HeLa cell sample (active system) exhibits a deviation from this value, particularly for larger τ values. The differences in how frequently the Mean Back Relaxation significantly deviates from .5 can be seen in tables A.1 and A.2.

However, it is important to note that these measurements were not consistently reproducible across all samples. The MBR shape and long-term value varied widely between cells of the same type and sometimes even between the x and y components of the same trajectory. While some MBR measurements, such as the one shown in Figure 4.11, aligned with expectations, we were unable to consistently reproduce the MBR measurements as made with an optical tweezer[19].

4. Results

The most likely explanation for this inconsistency lies in the condition of the cells during measurement. As mentioned in Section 4.1.2, we encountered difficulties in maintaining cell viability throughout the measurement process. Although all cells used in the presented datasets showed no visible signs of necrosis before and after imaging (as seen in Figure 4.2), it is possible that the cells were in a non-standard state or even dying during the actual imaging process. This cellular stress could account for the variability in our MBR measurements and the deviation from expected results.

5. Discussion

This study aimed to develop an experimental and data analysis pipeline for measuring Mean Back Relaxation (MBR) of beads in living and non-living materials using darkfield microscopy. We successfully imaged silicone beads in hydrogels and HeLa cells at frame rates ranging from one to fifty kilohertz. We implemented and evaluated different tracking methods, estimating their precision to be on the order of one to ten nanometers. We developed a method for detecting noise peaks in Power Spectra that can label exact points of noise and identify peaks not visible to the human eye. We also fitted models of viscoelasticity to the Power Spectrum of test particles in materials, verifying our approach with simulated data. Finally, we applied these methods to our experimental data, reproducing expected results for particles in hydrogels and partially reproducing expected results for beads in HeLa cells.

5.1. Imaging

Our goal in imaging beads using a darkfield microscope was to perform measurements similar to previous MBR studies [19] but with quicker experiments and with more widely available equipment. We successfully imaged beads in both hydrogels and cells using the darkfield microscope. An initial concern was that imaging beads in cells might be challenging due to background interference from cellular structures. However, this did not prove to be a significant problem, as we could adjust the lighting to make the beads clearly visible while minimizing background noise from cells.

We were able to achieve the microsecond-scale frame rates of previous measurements [19] with our darkfield setup. However, increasing the frame rate required progressively restricting the field of view, which diminishes one of the biggest advantages of darkfield microscopy: its speed. A larger field of view would allow for capturing multiple beads simultaneously by segmenting the movie into different

5. Discussion

parts, each centered around a bead. While this is not possible at the highest frame rates due to the cropped field of view, it remains feasible for frame rates around 1 kHz, which is comparable to previous studies.

The most significant bottleneck in the imaging process was the time required to transfer image files from the camera to the controlling computer. Transferring a 2-minute movie at 50 kHz took approximately 40 minutes, which presents an opportunity for future improvement in data handling and transfer protocols.

Imaging HeLa cells posed a significant challenge due to their sensitivity to light exposure. We implemented a measurement protocol that minimizes continuous light exposure, allowing us to make multiple measurements on the same sample without observing visible effects of necrosis. However, the MBR results later suggested that most beads in HeLa cells had a long-term MBR value of 0.5, indicative of confined systems in thermodynamic equilibrium. Since living systems operate out of thermodynamic equilibrium, this would suggest that the cells were dead. This discrepancy between the apparent health of cells during imaging and the MBR results highlights an area for future improvement. Future work should focus on refining the cell imaging protocol, possibly incorporating fluorescent viability dyes to continuously monitor cell health during measurements.

5.2. Tracking

Our tracking methods aimed to determine bead positions as precisely as possible without needing to deal with complex backgrounds or multiple objects. This focus allowed us to implement three analytical and understandable approaches to particle tracking. We also developed three different ways to estimate the precision of our tracking algorithms, which is crucial for accurate MBR calculation.

Simulated trajectory reconstruction showed that the Fitted Convolution Method and the Fitted Particle Shape method achieved precisions of a few nanometers, dropping to half a nanometer for 1 kHz frame rates. This precision is surprisingly high, considering that one pixel is 36 nm wide. While this setup is idealized, it provides a lower bound for resolution and demonstrates the robustness of the tracking methods to additive pixel noise.

The piezo stage measurements provided a more robust real-world test, yielding precisions of 2.5 nm at lower frame rates, increasing to 4 nm at 20-40 kHz. This suggests we can track beads to a precision of about 10% of the camera's pixel size.

However, it remains an open question how well these results transfer to different materials (glue vs. hydrogel vs. cell).

The noise level estimated from the high-frequency plateau of the Power Spectrum yielded errors of approximately one nanometer, which is about 2-3 times smaller than expected from the previous methods. This discrepancy might be due to differences in the experimental setup or because only about half of the tracking algorithm imprecision manifests as Gaussian noise, with the rest appearing as biases.

The high precision independently suggested by different estimation methods indicates that this part of the pipeline is ready for integration into future research. While machine learning methods could potentially be applied to particle tracking in the future, the methods developed in this thesis may generalize better to other setups, as they rely on basic assumptions that can be easily verified for new experimental conditions.

5.3. Fitting and Filtering

We developed a principled algorithm to fit the Power Spectrum of a trajectory to a viscoelastic model while identifying noise peaks. Our method successfully retrieved the type of viscoelastic model from simulated data and recovered model parameters with reasonable precision. However, a more rigorous test would involve using a material with precisely known viscoelastic properties and embedded microbeads to see how well the properties can be reconstructed.

Our noise peak detection method demonstrated the ability to find peaks more precisely than visual inspection, as evidenced by identifying the same peaks in both x and y data even when they were only visible in one dimension. However, further validation could involve artificially introducing oscillations into trajectories and assessing the algorithm's ability to filter them out.

The fitting and filtering method developed in this project is potentially the most broadly applicable outcome, as it can be used for noise filtering of trajectories of objects in viscoelastic materials beyond the study of mean back relaxation. However, it also rests on the least theoretically grounded foundation. The Bayesian approach allows for direct calculation of which points are noise peaks, but requires defining priors somewhat arbitrarily. The method's parameters were chosen to make the algorithm computationally feasible while attempting to maintain a principled approach. As a result, the method may require parameter adjustments when applied

5. Discussion

to significantly different datasets.

The process is computationally intensive, with the time required to filter and fit a dataset scaling linearly with the number of data points. Future work could focus on optimizing the method, such as utilizing GPU acceleration for fitting functions to the data and summarizing multiple high-frequency data points to reduce computational load.

When applied to our experimental data, the fitting and filtering method generally performed well, successfully identifying visible noise peaks in most cases. The only exceptions were in cases of extreme noise where the initial fit was too far off to allow for proper peak filtering. The method also identified some peaks not visible to the naked eye, particularly in the low-frequency range, which may correspond to cell activity.

5.4. Calculating Mean Back Relaxation

We developed a method for calculating Mean Back Relaxation (MBR) that closely follows the approach presented by Münker et al. [19], discarding data points with small denominators to avoid dominating the empirical mean. We generalized the method for multiple datasets and datasets with irregular time steps, which could be relevant for future applications with different microscopy techniques.

Our MBR calculation method was validated using simulated data from the 'horse and cart' model, showing excellent agreement with the analytical solution. However, this validation used trajectories with a million data points, which may present an overly optimistic picture of how close we typically get to the actual MBR of a system with real data.

The MBR results from our experimental data showed smooth curves without bigger oscillations, indicating successful elimination of noise peaks from the Power Spectrum. As expected, all MBR long-term values for hydrogel samples approached 0.5. However, most MBR long-term values for HeLa cells also approached 0.5, which was unexpected and suggests potential issues with the cell measurements.

The most likely explanation for this discrepancy is that the HeLa cells were dead or dying at the moment of imaging, or that the beads were not correctly labeled as being within the cells. This highlights the need for further refinement of the measurement process and collection of more data to reliably measure MBR in living cells.

One cell measurement did show the expected MBR behavior consistent with previous results from Münker et al. [19], where the MBR value immediately rises to a value of around .4 within milliseconds, and then relaxes to a value around .3 over the timescale of one second. However, we observed a strong dependence on the value of τ , making it difficult to confidently determine the correct asymptotic value for short τ .

5.5. Determining Viscoelastic Properties

The next step in this research would be to use the long-term value of the MBR to subtract out the active term from the Power Spectrum according to equation 2.10, and then refit the Power Spectrum to determine the mechanical properties of cells. While a function for this purpose is implemented in the Python package for fitting Power Spectra, our results for the long-term effect of Power Spectra were too varied and unreliable to apply this step in the current study.

Future work building on this research could compare the measured mechanical properties of cells derived from this passive method with those obtained from active measurements using optical tweezers.

5.6. Outlook

This study has established a pipeline for measuring mean back relaxation via dark-field microscopy. The most exciting next steps would be to apply these techniques and data analysis tools to analyze large quantities of microscope data, reliably measure Mean Back Relaxation in living cells, and analyze their mechanical properties.

The most pressing open problem that merits further work is the handling of living cells in the imaging process. Additionally, there are several ways the entire pipeline can be sped up and automated:

- Imaging multiple beads simultaneously
- Accelerating data transfer from camera to computer
- Utilizing GPU acceleration for computationally intensive data analysis steps
- Clustering high-frequency data for Power Spectrum fitting

5. Discussion

Future studies could also explore varying the probe particle, such as using gold particles. Preliminary experiments with smaller gold particles showed they appeared much brighter under the darkfield microscope, potentially enabling higher frame rates and investigation of mechanical properties at smaller scales.

Once this methodology is optimized to work at scale, it could facilitate the creation of large datasets on the mechanical fingerprint and MBR of various cell types. This could enable investigations into how the behavior of a bead or gold particle varies depending on its location within a cell.

In conclusion, while this study has made significant progress in developing a dark-field microscopy-based approach to measuring Mean Back Relaxation, there remain several challenges and opportunities for refinement. The methods and tools developed here provide a foundation for future research into cellular mechanics using passive observation techniques, potentially offering a more accessible alternative to active measurement methods like optical tweezers.

6. Appendix

A. Power Spectral Density and Mean Back Relaxation Data

This section presents comprehensive data on the Power Spectral Density (PSD) and Mean Back Relaxation (MBR) for all measurements discussed in the results section. For each trajectory, we analyze and plot the x and y components separately.

In the PSD plots, we overlay the best-fit model as a black line and highlight data points identified as noise peaks in red. For the MBR plots, we vary the time step τ from equation 2.16 and report the long-term MBR value. This value is calculated by averaging the last 100 MBR data points up to 0.5 seconds. The associated error is represented by the standard deviation within these 100 data points.

For measurements involving beads in cells, we also provide a wider field of view image of the cell to provide context.

Panel	Frame Rate	Orientation	Long-time MBR (separated by value for τ)						
			1ms	2ms	5ms	10ms	20ms	50ms	100ms
a)	1kHz	x	.50±.01	.50±.01	.50±.01	.50±.01	.50±.01	.49±.01	.49±.01
b)	1kHz	y	.49±.01	.50±.01	.50±.01	.50±.01	.49±.01	.48±.01	.46±.02
			0.2ms	0.4ms	1ms	2ms	4ms	10ms	20ms
c)	5kHz	x	.49±.01	.50±.01	.49±.01	.50±.01	.49±.01	.49±.01	.49±.02
d)	5kHz	y	.51±.00	.49±.01	.50±.01	.49±.01	.49±.01	.49±.01	.48±.01
			0.1ms	0.2ms	0.5ms	1.0ms	2.0ms	5.0ms	10.0ms
e)	10kHz	x	.49±.01	.49±.01	.50±.01	.50±.01	.50±.01	.50±.01	.50±.01
f)	10kHz	y	.52±.01	.49±.01	.50±.01	.50±.01	.50±.01	.50±.01	.49±.01
			0.04ms	0.08ms	0.2ms	0.4ms	0.8ms	2.0ms	4.0ms
g)	25kHz	x	.49±.01	.50±.01	.49±.01	.50±.01	.49±.01	.49±.01	.51±.01
h)	25kHz	y	.52±.01	.51±.01	.50±.01	.50±.01	.49±.01	.49±.02	.49±.02

Table A.1.: Frame rates, orientations, and long-time MBR values corresponding to each panel in Figure A.1. Long-time MBR values significantly smaller than .5 are marked in **bold**.

A. Power Spectral Density and Mean Back Relaxation Data

Panel	Frame Rate	Orientation	Long-time MBR (separated by value for τ)						
			.5ms	1.0ms	2.5ms	5ms	10ms	25ms	50ms
A.2a)	2kHz	x	.50±.01	.50±.00	.49±.00	.50±.00	.50±.00	.50±.01	.49±.01
A.2b)	2kHz	y	.50±.00	.49±.00	.50±.00	.50±.00	.49±.00	.49±.00	.48±.00
A.2c)	2kHz	x	.48±.01	.47±.01	.45±.01	.40±.01	.35±.01	.24±.01	.09±.01
A.2d)	2kHz	y	.49±.00	.49±.00	.49±.00	.49±.00	.47±.00	.45±.00	.43±.00
A.2e)	2kHz	x	.49±.01	.50±.01	.50±.01	.49±.01	.48±.01	.49±.01	.47±.01
A.2f)	2kHz	y	.51±.01	.49±.01	.49±.01	.50±.01	.49±.01	.48±.01	.46±.01
A.3a)	2kHz	x	.51±.01	.50±.01	.49±.01	.50±.01	.50±.01	.50±.01	.48±.01
A.3b)	2kHz	y	.51±.01	.49±.01	.50±.01	.50±.01	.50±.01	.50±.01	.49±.01
A.3c)	2kHz	x	.48±.01	.50±.01	.49±.00	.50±.01	.49±.01	.49±.01	.48±.01
A.3d)	2kHz	y	.49±.01	.51±.01	.50±.01	.50±.01	.49±.01	.48±.01	.47±.01
A.3e)	2kHz	x	.48±.01	.50±.01	.50±.00	.49±.01	.49±.01	.48±.01	.46±.02
A.3f)	2kHz	y	.47±.01	.50±.01	.50±.01	.49±.01	.49±.01	.48±.01	.48±.01
			.1ms	.2ms	.5ms	1ms	2ms	5ms	10ms
A.4a)	10kHz	x	.50±.00	.50±.00	.50±.00	.50±.00	.50±.01	.50±.01	.50±.01
A.4b)	10kHz	y	.50±.00	.50±.00	.50±.00	.50±.00	.50±.00	.49±.01	.49±.01
A.4c)	10kHz	x	.50±.00	49±.00	.50±.00	.50±.00	.50±.00	.49±.01	.48±.01
A.4d)	10kHz	y	.49±.00	.50±.00	.50±.00	.49±.00	.50±.00	.49±.01	.48±.01
A.4e)	10kHz	x	.50±.00	.49±.00	.50±.00	.50±.00	.50±.00	.50±.01	.50±.01
A.4f)	10kHz	y	.50±.00	.50±.00	.50±.00	.50±.00	.49±.00	.50±.00	.49±.01

Table A.2.: Frame rates, orientations, cell, and long-time MBR values corresponding to each panel in Figures A.2, A.3, and A.4. Long-time MBR values significantly smaller than .5 are marked in **bold**.

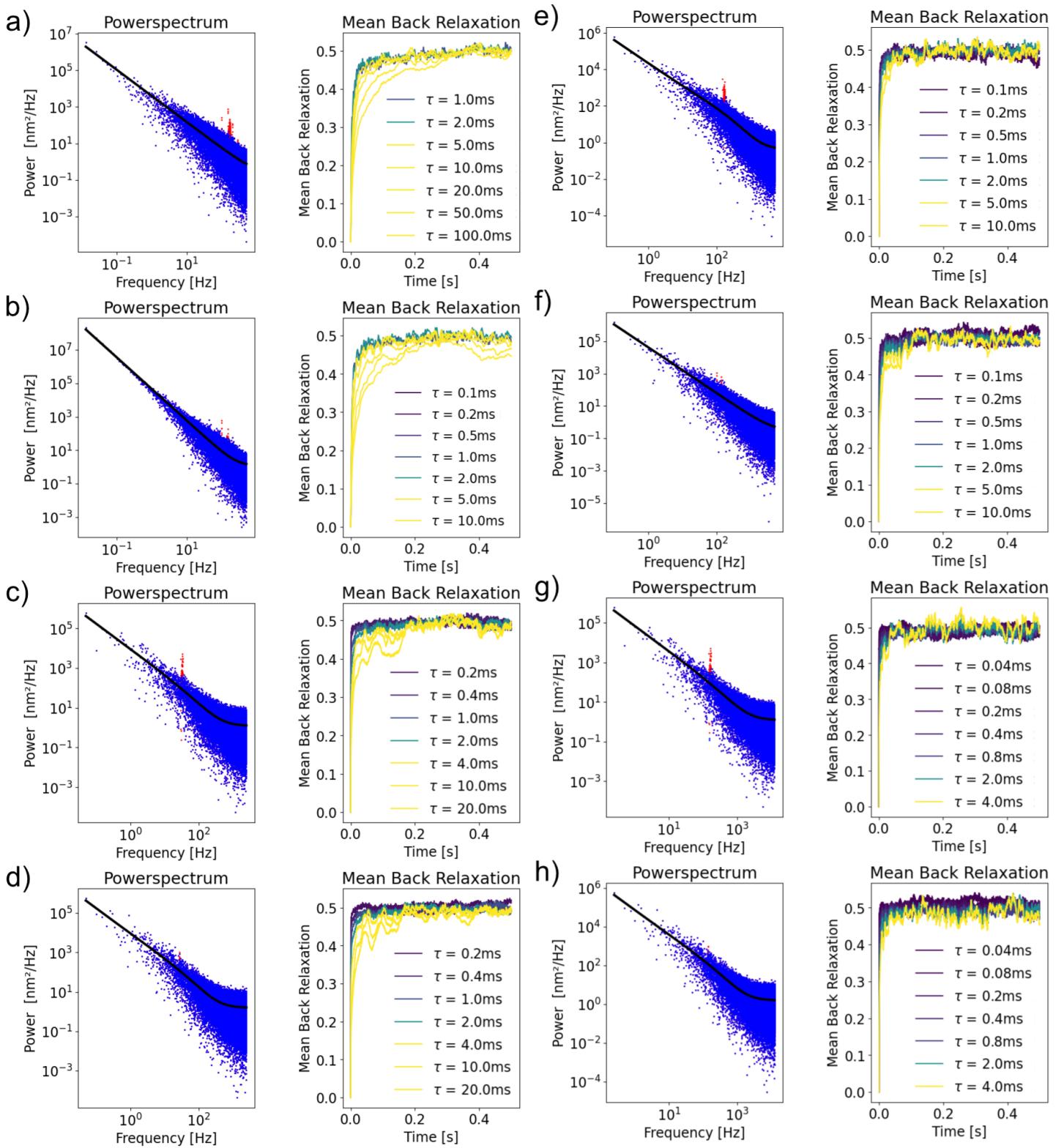


Figure A.1.: Power Spectra and Mean Back Relaxations for Beads in Hydrogel. These measurements are described in section 4.1.1. Table A.1 provides information about the frame rate, orientation and long term average value of the Mean Back Relaxation for each panel.

A. Power Spectral Density and Mean Back Relaxation Data

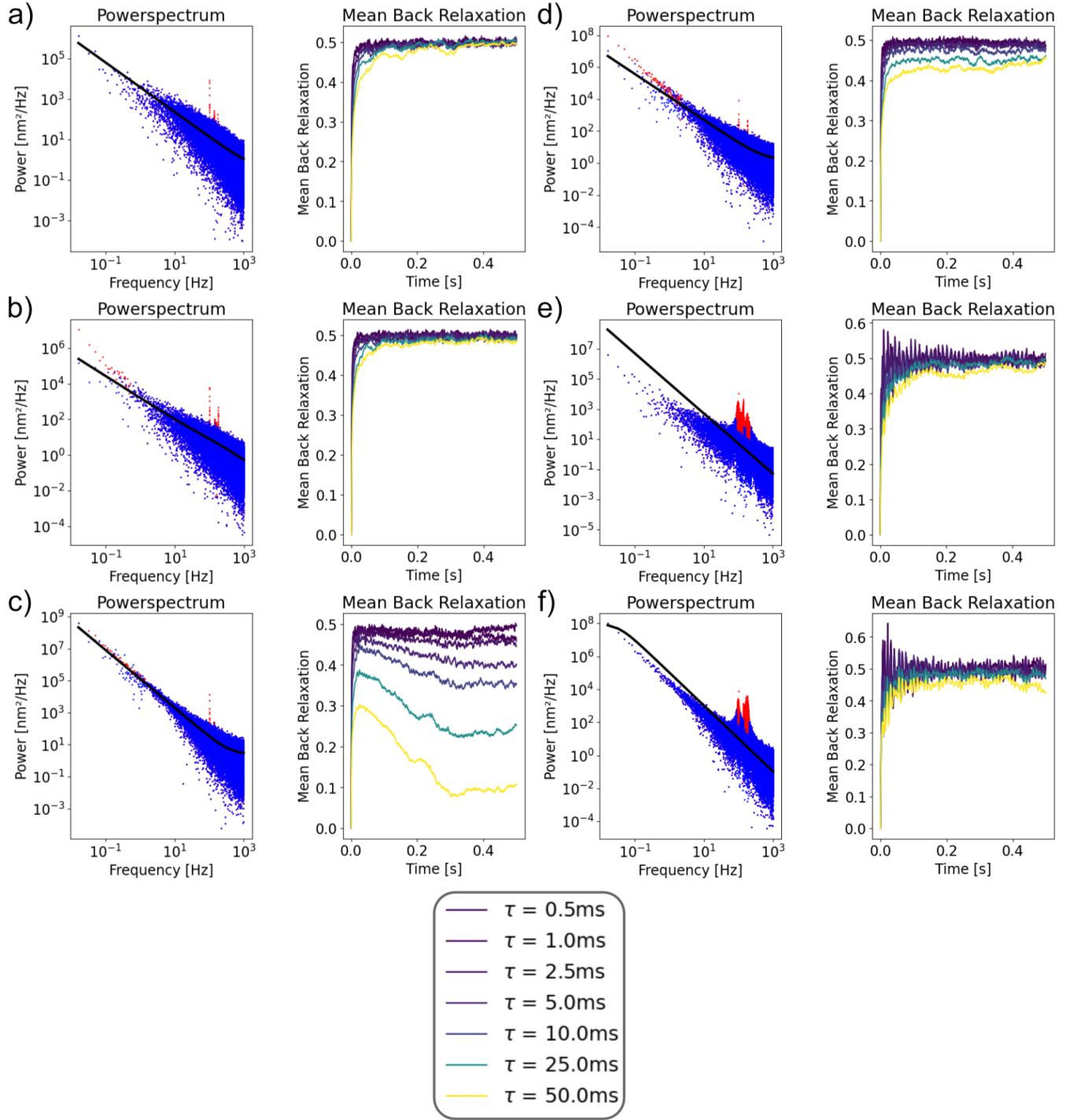


Figure A.2.: Power Spectra and Mean Back Relaxations for Beads in HeLa Cells at 2 kHz frame rate. These measurements are described in section 4.1.2. Table A.2 provides information about the cell orientation and long term average value of the Mean Back Relaxation for each panel.

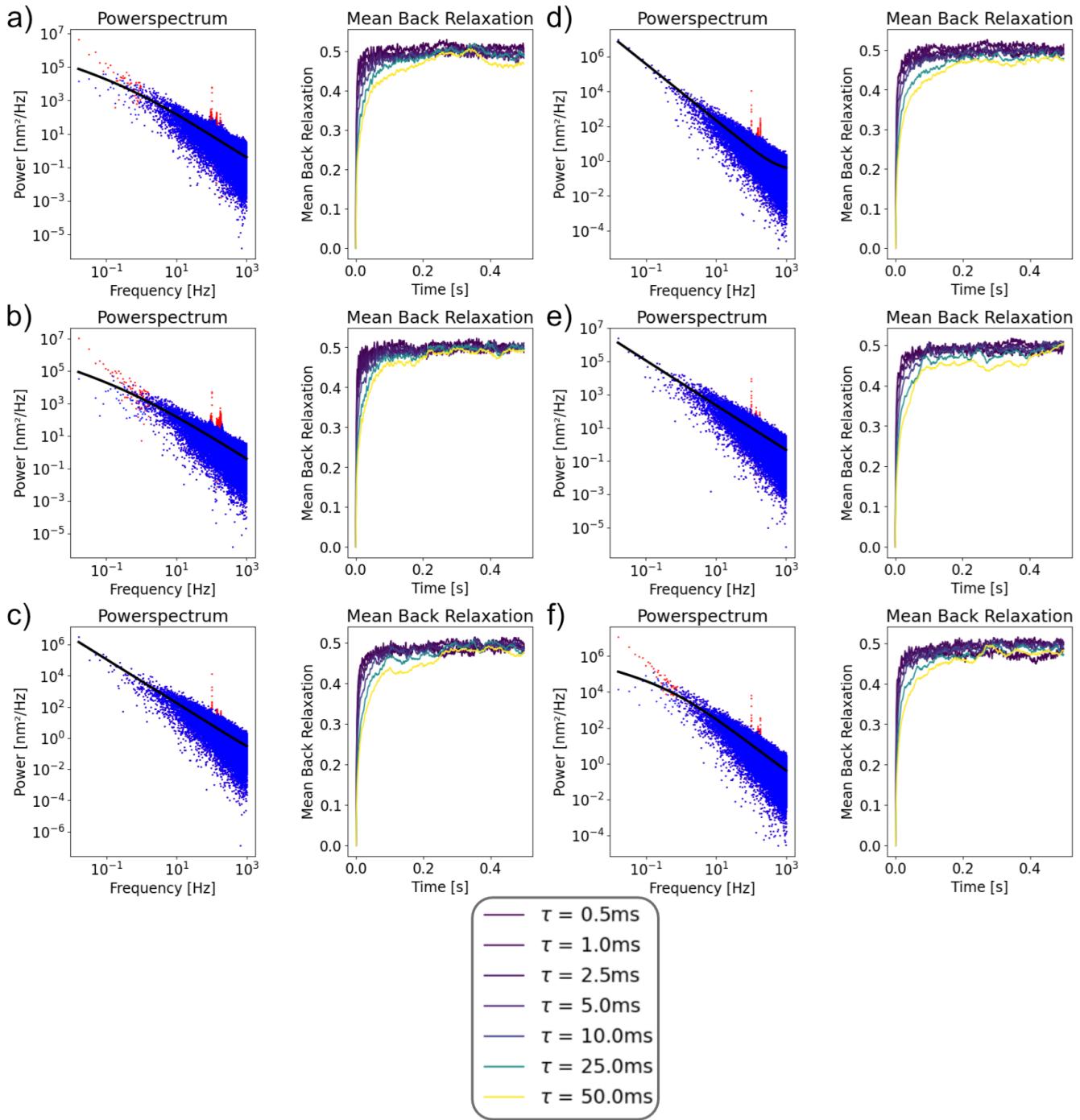


Figure A.3.: Power Spectra and Mean Back Relaxations for Beads in HeLa Cells at 2 kHz frame rate. These measurements are described in section 4.1.2. Table A.2 provides information about the cell orientation and long term average value of the Mean Back Relaxation for each panel.

A. Power Spectral Density and Mean Back Relaxation Data

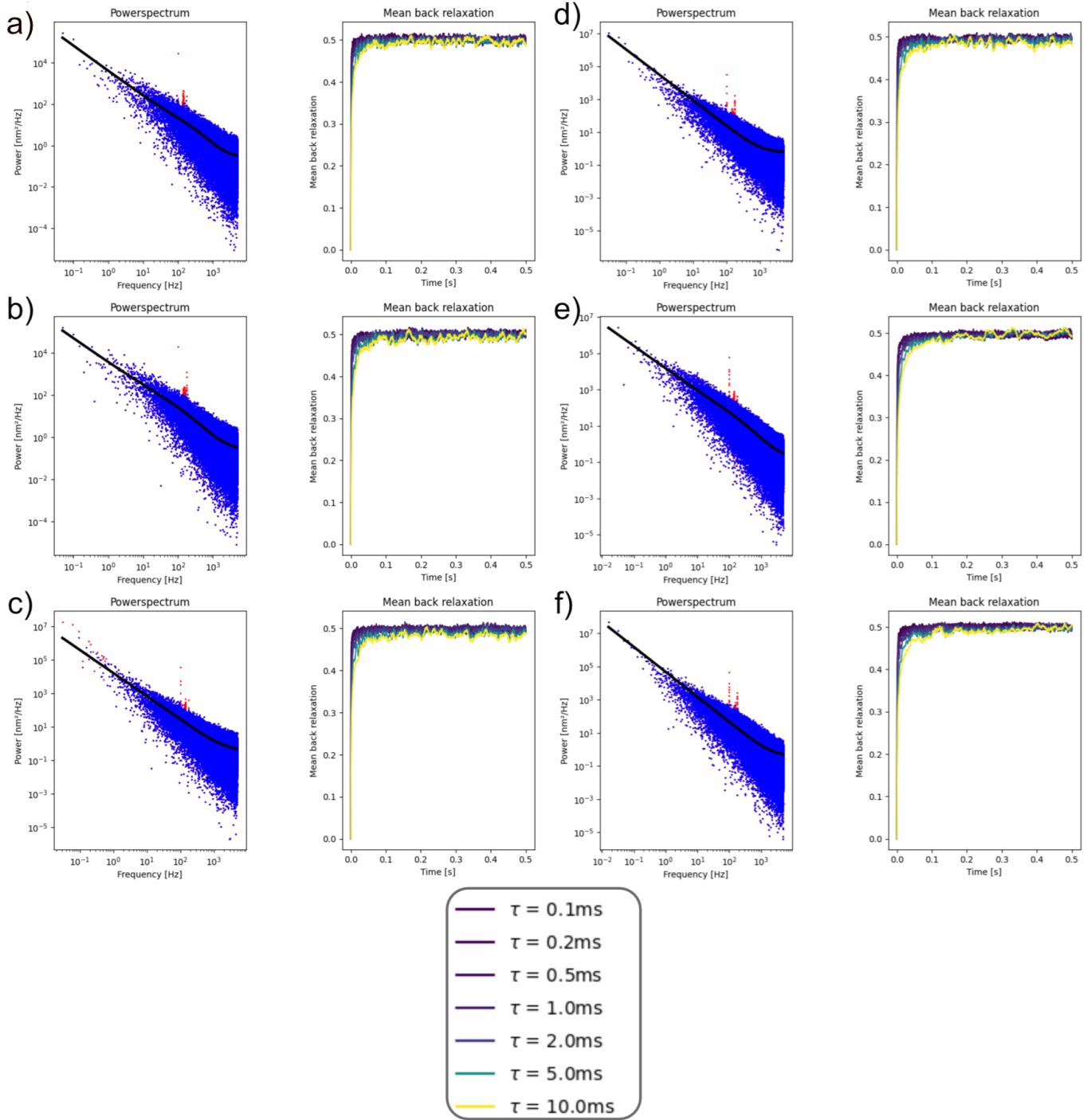


Figure A.4.: Power Spectra and Mean Back Relaxations for Beads in HeLa Cells at 10 kHz frame rate. These measurements are described in section 4.1.2. Table A.2 provides information about the cell orientation and long term average value of the Mean Back Relaxation for each panel.

B. Entropy of the prior over length

In equation 3.6, we need the entropy of the our prior over peak lengths $H(p_{lengthpeak})$. Our prior as mentioned in equation 3.5:

$$\begin{aligned}
 p_{lengthpeak}(l) &= \frac{1}{2^{1/L} - 1} 2^{-\frac{l}{l_{typical}}} \\
 H(p_{lengthpeak}) &= \frac{1}{2^{1/L} - 1} \sum_{l=1}^{\infty} \frac{l}{l_{typical}} 2^{-\frac{l}{l_{typical}}} \\
 &= \frac{1}{2^{1/L} - 1} \frac{2^{\frac{1}{L}}}{\left(-1 + 2^{\frac{1}{L}}\right)^2 L} \\
 &= \frac{2^{\frac{1}{L}}}{\left(-1 + 2^{\frac{1}{L}}\right)^3 L}
 \end{aligned} \tag{B.1}$$

C. Choosing $n_{typical}$

We want to choose $n_{typical}$ in equation 3.4 in such a way, that a data set only is labeled as having noise peaks, when those noise peaks are statistically significant. That means that any evidence that makes there being one peak more likely than there being no peak, must only have a 5% chance of occurring, if there is actually no peak. Let us call the prior of there being n peaks $P(n)$, and let E be evidence about the number of peaks. Now let $P(n|E)$ be the posterior after updating on this evidence and $p(E|n)$ be the probability of the evidence E , when there are indeed n peaks.

We want to choose the priors in such a way that:

$$P(1|E) = P(0|E) \iff p(E|0) = 0.05$$

Now the ratio of the posteriors can be calculated using Bayes Theorem:

$$\begin{aligned} \frac{P(1|E)}{P(0|E)} &= \frac{p(E|1)}{p(E|0)} \frac{P(1)}{P(0)} \\ \rightarrow 1 &= \frac{0.95}{0.05} \frac{P(1)}{P(0)} \\ \rightarrow 1 &= 20 \frac{P(1)}{P(0)} \\ \rightarrow \frac{1}{20} &= \frac{2^{-\frac{1}{2} \left(\frac{1}{n_{typical}} \right)^2}}{2^{-\frac{1}{2} \left(\frac{0}{n_{typical}} \right)^2}} \\ \rightarrow \log(20) &= \frac{1}{2} \frac{1}{n_{typical}^2} \\ n_{typical} &= \frac{1}{\sqrt{2 \log(20)}} \approx 0.408... \end{aligned}$$

Here we assume, that the evidence E only has probabilities on there being 0 or 1 peaks. This is because when we follow the algorithm described in section 3.6, at the

C. Choosing $n_{typical}$

moment, where we have to decide whether to classify the first noise peak, these are the only scenarios under consideration. We only consider the possibility of multiple noise peaks, once the first one was classified.

D. Distribution of the Mean Back Relaxation

While the Mean Back Relaxation is defined in 2.16, as the average of the expression $\frac{[x(t) - x(0)]}{[x(0) - x(-\tau)]}$ (from here on referred to as pointwise MBR) over a trajectory, the actual empirical calculation of this average, as discussed in section 3.5, is not as straightforward. As mentioned before, this is due to outliers of the pointwise MBR where the denominator is small. A visualization of this phenomenon is given in figure D.1. As we can see here, the values which have a relatively high denominator look like

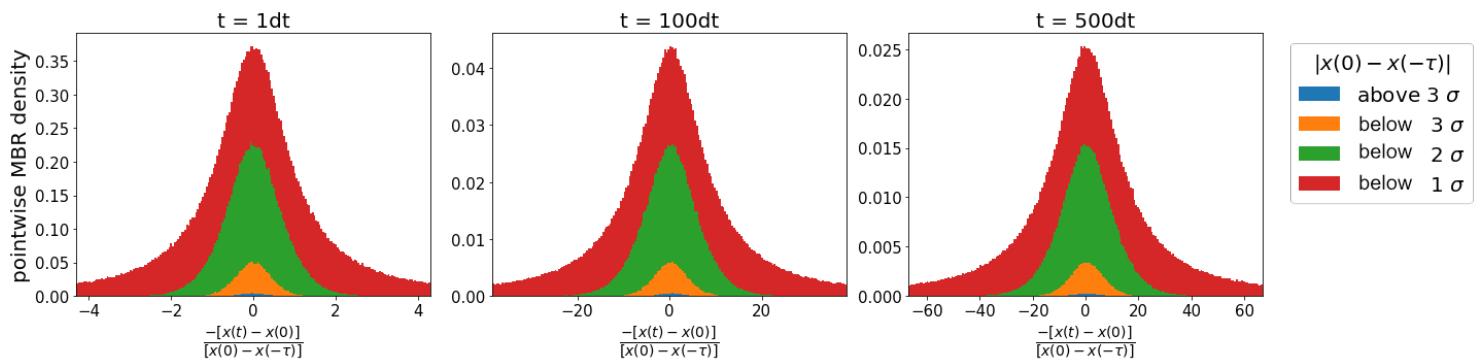


Figure D.1.: Distribution of pointwise values of the Mean Back Relaxation for different times t , the density of pointwise values is color-coded for size of the denominator $|x(0) - x(-\tau)|$ normalized by the distribution of denominator values. The density function is non-Gaussian with tails getting thicker for a more out-of-distribution denominator. The trajectory these plots are based on is from a Horse and Cart model as discussed in section 2.2.

value distributions that cluster around a mean. When excluding pointwise MBR values where the denominator is within the one sigma bound of its own distribution, we exclude the values here plotted in red. We can see that those values are not distributed in a long-tailed distribution.

While we can just throw the data away and take the mean over the rest, this is not

D. Distribution of the Mean Back Relaxation

ideal. For one, there is some information in the data we do not consider. While our estimate does not improve by just considering them as much as any other datapoint, there must be some weighting where we can learn something about the mean of the distribution of pointwise MBR samples, since they are drawn independently of that mean. It also is arbitrary to choose the standard deviation of the denominator distribution as the cutoff point. Lastly, the choice of weighing each other data point equally does also not seem justified, since different sizes of the denominator seem to spread with different variances around the actual mean of the whole distribution, which might justify different weightings.

To find a more principled way to use all the data we have to estimate the value of the Mean Back Relaxation as well as possible, we would benefit from knowing the distribution of the pointwise MBR value explicitly. In the following, we have made an attempt to derive this distribution analytically for the simplest case, where the particle is trapped in a harmonic potential. We did succeed in deriving a solution, but the form of the solution is too difficult to simplify to derive a better MBR estimation method from there. The derivation is reprinted here for completeness and the hope that someone might be able to derive a simpler or more generalized solution for it.

To be precise with the distribution we want to calculate, we define it here:

$$p \left(\frac{-[x(t) - x(0)]}{[x(0) - x(-\tau)]} = y \right) (t, \tau) = \int W_3(x(t), t; x(0), 0; x(-\tau), -\tau) \delta(\frac{-[x(t) - x(0)]}{[x(0) - x(-\tau)]} - y) d(x(\tau)) d(x(0)) d(x(t))$$

Where here W_3 is the three-point correlation function. We have to integrate over all values of $(x(t), x(0)$ and $x(-\tau)$ that lead to the value of $\frac{-[x(t) - x(0)]}{[x(0) - x(-\tau)]}$ being constant. First, let us substitute the denominator $[x(0) - x(-\tau)]$ with the variable d , such that each variable only appears once. Since for each value of $x(0)$, this is just a change of variables by a constant, we can do this without any factor from the Jacobian of the variable transform.

Now we do a more involved variable transform:

$$y = \frac{[x(t) - x(0)]}{d}$$

$$\alpha = d$$

$$\beta = \frac{x(0)}{d}$$

$$J = \begin{bmatrix} \frac{\partial y}{\partial x(t)} & \frac{\partial y}{\partial x(0)} & \frac{\partial y}{\partial d} \\ \frac{\partial \alpha}{\partial x(t)} & \frac{\partial \alpha}{\partial x(0)} & \frac{\partial \alpha}{\partial d} \\ \frac{\partial \beta}{\partial x(t)} & \frac{\partial \beta}{\partial x(0)} & \frac{\partial \beta}{\partial d} \end{bmatrix}$$

$$J = \begin{bmatrix} -\frac{1}{d} & \frac{1}{d} & \frac{x(t)-x(0)}{d^2} \\ 0 & 0 & 1 \\ 0 & \frac{1}{d} & -\frac{x(0)}{d^2} \end{bmatrix}$$

$$\det(J) = \frac{1}{d^2} = \frac{1}{\alpha^2}$$

$$x(t) = \alpha(\beta - y)$$

$$d = \alpha$$

$$x(0) = \alpha\beta$$

So we can now explicitly write down the integration steps:

$$p(-\frac{x(t) - x(0)}{d} = y)(t, \tau) = \int d\alpha \int d\beta W_3(\alpha(\beta - y), t; \alpha\beta, 0; \alpha(\beta - 1), -\tau) \alpha^2 \quad (\text{D.1})$$

Now for a particle in equilibrium in a harmonic oscillator, we can separate W_3 into the probability of being at place x_1 at time t_1 and the probabilities of transitioning to x_2 by time t_2 and to x_3 by t_3 :

$$W_3(x_3, t_3; x_2, t_2; x_1, t_1) = P_{\text{eq}}(x_1)P(x_2 - x_1, t_2 - t_1 | x_1)P(x_3 - x_2, t_3 - t_2 | x_2) \quad (\text{D.2})$$

$$(\text{D.3})$$

Here $P_{\text{eq}}(x_1)$ is the equilibrium probability distribution of a diffusing particle in a

D. Distribution of the Mean Back Relaxation

harmonic oscillator. The time derivative of probability distribution can be calculated using the Fokker-Planck equation (assuming for now, that $\gamma = 1$):

$$\frac{\partial P(x, t)}{\partial t} = D \frac{\partial^2 P(x, t)}{\partial x^2} - \frac{\partial(-kxP(x, t))}{\partial x} \quad (\text{D.4})$$

For the equilibrium solution, we can set the time derivative to be zero.

$$\begin{aligned} 0 &= D \frac{\partial^2 P_{\text{eq}}(x)}{\partial x^2} + \frac{\partial(kxP_{\text{eq}}(x))}{\partial x} \\ \rightarrow \frac{\partial(P_{\text{eq}}(x))}{\partial x} &= c - x \frac{k}{D} P_{\text{eq}}(x) \\ \rightarrow P_{\text{eq}}(x) &= \frac{1}{\sqrt{\frac{2\pi D}{k}}} e^{-\frac{kx^2}{2D}} \end{aligned}$$

$P(x, t|x_0)$ denotes the probability of the particle being in position x at time t , if it was at position x_0 at time 0. To get to an expression for this probability, let us take another look at the Fokker-Planck equation, and try to solve it with a separation Ansatz: $P(x, t) = P_x(x) \cdot P_t(t)$.

$$\begin{aligned} \frac{\partial P_x(x) \cdot P_t(t)}{\partial t} &= D \frac{\partial^2 P_x(x) \cdot P_t(t)}{\partial x^2} + \frac{\partial(kxP_x(x) \cdot P_t(t))}{\partial x} \\ \rightarrow \frac{\dot{P}_t}{P_t} &= D \frac{P''_x}{P_x} + \frac{kP_x + kxP'_x}{P_x} = c \\ \rightarrow P_t &= A \cdot e^{ct} \\ \rightarrow DP''_x + kxP'_x &= (c - k)P_x \\ \rightarrow P''_x &= -\frac{k}{D}xP'_x + \frac{(c - k)}{D}P_x \end{aligned}$$

This final equation can be solved as a Sturm-Liouville Equation. The class of solutions contains this function[2]:

$$P_x(x) = e^{-\frac{kx^2}{2D}} H_{-\frac{c}{k}}\left(\sqrt{\frac{k}{2D}}x\right)$$

Here H_n denotes the n th Hermite polynomial. To get the degree of the Hermite polynomial to an integer, we have to choose c as $-nk$. So we know now that this is

a solution for our Fokker-Planck equation:

$$P(x, t) = e^{-\frac{kx^2}{2D}} H_n(\sqrt{\frac{k}{2D}}x) e^{-nkt}$$

Notice that this solution is not normalized and can have negative values. But since the Fokker-Planck equation is a linear differential equation, we know that any linear combination of solutions is again a solution. If we now know that the particle was originally at x_0 , we can construct a solution of the Fokker-Planck equation:

$$\begin{aligned} P(x, t|x_0) &= e^{\frac{kx_0^2}{2D}} \sum_n c_{n,x_0} e^{-\frac{kx^2}{2D}} H_n(\sqrt{\frac{k}{2D}}x) e^{-nkt} \\ c_{n,P_0} &= \int_{-\infty}^{\infty} \delta(x - x_0) e^{-\frac{kx^2}{2D}} H_n(\sqrt{\frac{k}{2D}}x) d(\sqrt{\frac{k}{2D}}x) \\ &= \sqrt{\frac{k}{2D}} e^{-\frac{kx_0^2}{2D}} H_n(\sqrt{\frac{k}{2D}}x_0) \\ \longrightarrow P(x, t|x_0) &= \sqrt{\frac{k}{2D}} e^{-\frac{kx^2}{2D}} \sum_n H_n(\sqrt{\frac{k}{2D}}x_0) H_n(\sqrt{\frac{k}{2D}}x) e^{-nkt} \end{aligned}$$

We know that $P(x, 0|x_0) = \delta(x - x_0)$, because the Hermite polynomials are an orthonormal basis of the space of functions $f(x)$, which satisfy $\int |f(x)|^2 e^{-x^2} dx < \infty$. We can now use these results to find an analytical expression of W_3 :

$$\begin{aligned} W_3(x_3, t_3; x_2, t_2; x_1, t_1) &= P_{\text{eq}}(x_1) \cdot P(x_2, t_2 - t_1|x_1) \cdot P(x_3, t_3 - t_2|x_2) \\ &= \frac{1}{\sqrt{\frac{2\pi D}{k}}} e^{-\frac{kx_1^2}{2D}} \\ &\quad \cdot \sqrt{\frac{k}{2D}} e^{-\frac{kx_2^2}{2D}} \sum_n H_n(\sqrt{\frac{k}{2D}}x_1) H_n(\sqrt{\frac{k}{2D}}x_2) e^{-nk(t_2-t_1)} \\ &\quad \cdot \sqrt{\frac{k}{2D}} e^{-\frac{kx_3^2}{2D}} \sum_m H_m(\sqrt{\frac{k}{2D}}x_2) H_m(\sqrt{\frac{k}{2D}}x_3) e^{-mk(t_3-t_2)} \\ &= \sqrt{\frac{k}{2D}} \frac{e^{-\frac{k}{2D}(x_1^2+x_2^2+x_3^2)}}{\sqrt{\pi}} \sum_{n,m} H_n(\sqrt{\frac{k}{2D}}x_1) H_n(\sqrt{\frac{k}{2D}}x_2) \\ &\quad \cdot H_m(\sqrt{\frac{k}{2D}}x_2) H_m(\sqrt{\frac{k}{2D}}x_3) e^{-k(n(t_2-t_1)+m(t_3-t_2))} \end{aligned}$$

D. Distribution of the Mean Back Relaxation

Now we can set this explicit equation back into D.1:

$$\begin{aligned}
p\left(-\frac{x-x_0}{d} = y\right)(t, \tau) &= \int d\alpha \int d\beta W_3(\alpha(\beta - y), t; \alpha\beta, 0; \alpha(\beta - 1), -\tau) \alpha^2 \\
&= \int d\alpha \int d\beta \sqrt{\frac{k}{2D}} \frac{e^{-\frac{k}{2D}\alpha^2(3\beta^2 - 2\beta(y+1) + y^2 + 1)}}{\sqrt{\pi}} \\
&\quad \cdot \sum_{n,m} H_n\left(\sqrt{\frac{k}{2D}}\alpha(\beta - 1)\right) H_n\left(\sqrt{\frac{k}{2D}}\alpha\beta\right) \\
&\quad \cdot H_m\left(\sqrt{\frac{k}{2D}}\alpha\beta\right) H_m\left(\sqrt{\frac{k}{2D}}\alpha(\beta - y)\right) e^{-k(n\tau + mt)} \alpha^2
\end{aligned}$$

Numerically simulating this solution turned out to be hard, since even high-order terms are relevant, and the solution when truncating the sum does not look realistic. Further analytical simplification also was not successful, which is why this line of inquiry has been closed for now. We remained with the technique of estimating the MBR by throwing away some values and weighing the rest equally.

Bibliography

- [1] Charu C Aggarwal and Charu C Aggarwal. *An introduction to outlier analysis*. Springer, 2017.
- [2] Mohammed A Al-Gwaiz. *Sturm-Liouville theory and its applications*, volume 264. Springer, 2008.
- [3] Pierre Baldi and Laurent Itti. Of bits and wows: A bayesian theory of surprise with applications to attention. *Neural Networks*, 23(5):649–666, 2010.
- [4] Kirstine Berg-Sørensen and Henrik Flyvbjerg. Power spectrum analysis for optical tweezers. *Review of Scientific Instruments*, 75(3):594–612, 2004.
- [5] Christopher M Bishop. Pattern recognition and machine learning. *Springer google schola*, 2:5–43, 2006.
- [6] Kenneth P Burnham and David R Anderson. *Model selection and multimodel inference: a practical information-theoretic approach*. Springer, 2002.
- [7] Adam J Engler, Shamik Sen, H Lee Sweeney, and Dennis E Discher. Matrix elasticity directs stem cell lineage specification. *Cell*, 126(4):677–689, 2006.
- [8] Ronald Aylmer Fisher. Statistical methods for research workers. In *Breakthroughs in statistics: Methodology and distribution*, pages 66–70. Springer, 1970.
- [9] Jan Gieseler, Juan Ruben Gomez-Solano, Alessandro Magazzù, Isaac Pérez Castillo, Laura Pérez García, Marta Gironella-Torrent, Xavier Viader-Godoy, Felix Ritort, Giuseppe Pesce, Alejandro V Arzola, et al. Optical tweezers—from calibration to applications: a tutorial. *Advances in Optics and Photonics*, 13(1):74–241, 2021.
- [10] Georg Häcker. The morphology of apoptosis. *Cell and tissue research*, 301:5–17, 2000.

Bibliography

- [11] Brenton D Hoffman and John C Crocker. Cell mechanics: dissecting the physical responses of cells to force. *Annual review of biomedical engineering*, 11: 259–288, 2009.
- [12] Wirawan Istiono. Analyzing time and space complexity: Kadane vs. divide and conquer algorithms for maximum sub-array problem. *Journal of Applied Computer Science & Mathematics*, (2), 2023.
- [13] Zbigniew D Jastrzebski. Nature and properties of engineering materials. 1976.
- [14] Nabeeha Khalid and Mahzad Azimpouran. Necrosis. 2020.
- [15] Gabriel Knotz and Matthias Krüger. Mean back relaxation for position and densities. *arXiv preprint arXiv:2311.17477*, 2023.
- [16] Rep Kubo. The fluctuation-dissipation theorem. *Reports on progress in physics*, 29(1):255, 1966.
- [17] Dorian Marx. Manual: Viscoelastic polyacrylamide hydrogels. Technical manual, Drittes Physikalisches Institut, Georg-August-Universit"at G"ottingen, 2023.
- [18] Thomas G Mason and David A Weitz. Optical measurements of frequency-dependent linear viscoelastic moduli of complex fluids. *Physical review letters*, 74(7):1250, 1995.
- [19] Till M Muenker, Gabriel Knotz, Matthias Krüger, and Timo Betz. Onsager regression characterizes living systems in passive measurements. *bioRxiv*, pages 2022–05, 2022.
- [20] Till M Muenker, Bart E Vos, and Timo Betz. Intracellular mechanical fingerprint reveals cell type specific mechanical tuning. *bioRxiv*, pages 2023–08, 2023.
- [21] Henry Pfister. Discrete-time signal processing. *Lecture Note, pfister. ee. duke.edu/courses/ece485/dtsp. pdf*, 2017.
- [22] Physik Instrumente (PI) GmbH Co. KG. P-611.3 NanoCube® XYZ-Nanopositionierer, 2024. URL <https://www.physikinstrumente.de/de/produkte/piezo-nanopositioniertische/mehrachsen-piezotische/p-6113-nanocube-xyz-system-201700>. Accessed: 2024-07-28.

Bibliography

- [23] A Serra-Aguila, JM Puigoriol-Forcada, G Reyes, and J Menacho. Viscoelastic models revisited: characteristics and interconversion formulas for generalized kelvin–voigt and maxwell models. *Acta Mechanica Sinica*, 35:1191–1209, 2019.
- [24] Astrid van der Horst and Nancy R Forde. Power spectral analysis for optical trap stiffness calibration from high-speed camera position detection with limited bandwidth. *Optics express*, 18(8):7670–7677, 2010.
- [25] Joanna Zemła, Joanna Danilkiewicz, Barbara Orzechowska, Joanna Pabian, Sara Seweryn, and Małgorzata Lekka. Atomic force microscopy as a tool for assessing the cellular elasticity and adhesiveness to identify cancer cells and tissues. In *Seminars in cell & developmental biology*, volume 73, pages 115–124. Elsevier, 2018.
- [26] Xuexiang Zhang, Tae-Hyung Kim, Timothy J Thauland, Hongjun Li, Fate-meh Sadat Majedi, Chau Ly, Zhen Gu, Manish J Butte, Amy C Rowat, and Song Li. Unraveling the mechanobiology of immune cells. *Current Opinion in Biotechnology*, 66:236–245, 2020.

Acknowledgments

I would like to express my thanks to Till Münker and Timo Betz for supervising my Master's thesis, and for their patience during its completion. I also thank Marcus Müller for spontaneously stepping in as the second referee.

I am also grateful to the Betz Group for their help. Special thanks go to Lukas Abegg for providing the hydrogel, Dorian Marx for helping me understand microrheology, Ulrike Schulz for patiently teaching me how to work with cell cultures, and Sarah Lädke for continuing the project.

Statement on Usage of AI Tools

I hereby declare that I utilized the following artificial intelligence tools in the process of writing this master's thesis:

1. **Claude 3.5 Sonnet** for improving text formulation, grammar, and spelling; text summarization; and formatting of LaTeX
2. **Elicit** for finding relevant academic sources
3. **GitHub Copilot** to assist in the creation of the code presented in this thesis

All substantive ideas, analyses, and conclusions presented in this thesis are my own original work unless otherwise indicated.

Erklärung Ich versichere hiermit, dass ich die vorliegende Arbeit ohne fremde Hilfe selbstständig verfasst und nur die von mir angegebenen Quellen und Hilfsmittel verwendet habe. Wörtlich oder sinngemäß aus anderen Werken entnommene Stellen habe ich unter Angabe der Quellen kenntlich gemacht. Die Richtlinien zur Sicherung der guten wissenschaftlichen Praxis an der Universität Göttingen wurden von mir beachtet. Eine gegebenenfalls eingereichte digitale Version stimmt mit der schriftlichen Fassung überein. Mir ist bewusst, dass bei Verstoß gegen diese Grundsätze die Prüfung mit nicht bestanden bewertet wird.

Göttingen, den January 21, 2025

(Julian Artur Schulz)