POLITÉCNICA

UNIVERSIDAD POLITÉCNICA DE MADRID

ESCUELA TÉCNICA SUPERIOR DE INGENIERÍA AGRONÓMICA, ALIMENTARIA Y DE BIOSISTEMAS

MÁSTER EN BIOLOGÍA COMPUTACIONAL

DEPARTMENT OF UPM Tutor: INTELIGENCIA ARTIFICIAL/ESCUELA TÉCNICA SUPERIOR DE INGENIEROS INFORMÁTICOS/UPM

DEPARTMENT OF TFM Tutor: UNIVERSIDAD COMPLUTENSE DE MADRID / HOSPITAL DOCE DE OCTUBRE

KALMAN FILTERS FOR BIOLOGICAL SIGNAL ANALYSIS

MASTER FINAL THESIS

Author: Miguel Barroso de Frutos

Professional tutor: Francisco Monroy Muñoz

Academic tutor: Esteban García Cuesta

June, 2025

CONTENTS:

FIGURES INDEX:

LIST OF ABBREVIATIONS:

1. KF: Kalman Filter.
2. EKF: Extended Kalman Filter.
3. PSD: Power Spectral Density.
4. pN: PicoNewtons.
5. DNA: Deoxyribonucleic Acid.
6. TADs: Topologically Associating Domains.
7. MSD: Mean Squared Displacement.
8. ATP: Adenosine Triphosphate.
9. FRAP Fluorescence Recovery After Photobleaching.
10. DIC: Differential Interference Contrast

ABSTRACT:

In this project, work with both the Kalman Filter (KF) and the Extended Kalman Filter (EKF) was made to make sense of noisy spatiotemporal data from optical traps, using them in both lab-controlled and live biological settings. It started with control experiments, tracking polystyrene microspheres in water and in polyacrylamide gel. These setups gave me a stable baseline to test the filters. Both KF and EKF did a solid job cleaning up high-frequency noise, all while keeping the key oscillatory signals intact. This was especially clear when looking at the power spectral density (PSD). The EKF stood out by better handling the non-linear aspects of the data.

With those results in hand, the next step was to focus to a more complex biological environment: live HeLa cells. These cells showed different nuclear morphologies; some normal, others more granular, or with ramified chromatin structures. Out representative optical traps were picked for each type and looked at the filtered force trajectories. Even with the noise and variability that come with live cells, both filters (and especially the EKF) were able to tease out meaningful patterns and highlight shifts in dynamic behaviour. It was also created correlation heatmaps to see how traps across the nucleus interacted with one another.

Finally, a comparation was made between bead experiments with the messier but more interesting cellular data. Bead signals were stable and repeatable, piconewton(pN)-range forces that didn’t change much. In contrast, the signals from live cells were more unpredictable, smaller in magnitude, and clearly shaped by things like chromatin changes, molecular crowding, and active forces within the cell. This comparison really emphasized the importance of using the right filter settings, because being too aggressive, there is a risk wiping out the very signals you're trying to study.

1. INTRODUCTION:

In the eukaryotic nucleus which is a complex and dynamic structure genetic functions are regulated. The core of all these processes is the chromatin. It is a DNA-complex which has two main functions: it packages the genome and simultaneously it regulates the access to the genetic code. [1] The chromatin is organised following a hierarchical order, it starts with nucleosomes (DNA wrapped around histone octamers) [2] and then it is further folded into higher order domains such as chromatin loops, topologically associating domains (TADs), and chromosome territories that help establish functional nuclear compartments. [3]

Apart from the structural role, chromatin has a key part as a regulatory interface talking about transcriptional accessibility. When the chromatin is remodelled with diverse mechanisms such as repositioning or histone modifications, the cells regulate which regions of the genome must open to transcriptional machinery. [1][4] There are some processes like acetylation or methylation of the histone tails that can promote or inhibit the transcriptional factor binding. Both usually coincide with changes in the chromatin mobility and organisation. [3][5]

As provided in many studies the chromatin undergoes a continuous movement with the nuclear landscape, and there are both active reorganizations and passive fluctuations related to this activity. For example, there is a trend in the domains of active genes in which they tend to be more mobile and spatially segregated from heterochromatin. [3] These movements have been seen in real time thanks to super-resolution microscopy and single-particle tracking, showing that the chromatin which is active exhibits distinctive diffusion patterns and spatial reorganization. [5][6]

To study the dynamic behaviour of chromatin many approaches and tools have emerged, especially live-cell imaging. There are some tools like fluorescent histone probes, CRISPR-based genomic labelling, and high-speed microscopy that allow to monitor specific loci or entire chromatin domains over time. [7] These technologies yield time-resolved spatial datasets that can be leveraged to infer underlying biological states, including chromatin compaction, locus accessibility, and transcriptional activity. One of the main lines of studies focuses on the interplay between transcription factors and chromatin motion. The transcriptional activators tend to be localized in dynamic chromatin regions and there is also a correlation between gene activation and their residence time and mobility. [8] Moreover, the process of nucleosome remodelling is closely linked to the recruitment of

transcription factors, establishing a feedback loop in which chromatin structure and transcriptional activity exert mutual influence. [4]

Knowing that chromatin movements reflect transcriptional regulation a new window discerning meaningful biological signals from the pervasive physical noise intrinsic to the cellular environment appears. In the eukaryotic nucleus, thermal energy produces constant stochastic perturbations that affect the chromatin and its protein complexes. These random fluctuations, usually represented as Brownian motion, stem from collisions with the surrounding molecules and they are one of the fundamental sources of noise in biological systems. [10][11]

The Brownian motion is a physical phenomenon which describes the random movement of particles which are suspended in a fluid, produced by the thermal collisions with the surrounding solvent molecules. Related with the nuclear context, this means that chromatin fibres have unpredictable movements because of the interactions within a densely packed and heterogeneous environment. These fluctuations produce the passive diffusion of chromatin loci independently of active biological mechanisms. A new challenge from biologically meaningful chromatin dynamics appears because of thermal driven motion. [11][12]

Some empirical studies have shown that chromatin presents subdiffusive behaviour, in which the mean squared displacement (MSD) increases sublinearly with time. This suggest that the chromatin motion may be constricted probably due to the viscoelastic nature of the nuclear environment and because of the intricate structural organisation of chromatin itself. [12][5] It has also been seen that chromatin mobility is not uniform. Nucleosomes located in euchromatin-rich interior show greater mobility than those in the heterochromatin-dense periphery, highlighting variations in chromatin compaction and transcriptional accessibility. [5][3]

The main challenge is to distinguish passive (thermal driven chromatin movements) from the active ones (biologically significant processes). Some procedures such as transcription, chromatin remodelling and loop extrusion generate directed motion of chromatin loci, but all these active movements are superimposed on a background of thermal fluctuations. Because of this it is hard to attribute observed movements to specific molecular mechanisms. [13][14].

Also, the nuclear environment is not static, there are some factors such as metabolic activity and cytoskeletal interactions that influence these dynamic behaviours. It has been

seen that ATP-dependent processes contribute to chromatin mobility, showing that chromatin dynamics beyond passive diffusion can be regulated by energy-consuming mechanisms. [10][15]

To obtain meaningful information from chromatin movement data, it is essential to separate the signal from the noise. There are some techniques used to provide high resolution data on chromatin dynamics such as single particle tracking and interferometric scattering correlation. [12][16] But to interpret this data, models that account for stochastic nature of Brownian motion and the directed movement produced by biological processes are required. Some advanced statistical and computational methods, including some filtering processes such as the KF, offer promising ways to untangle these complex datasets, helping to uncover the hidden biological signals buried in all the noise.

The KF states as an optimal recursive algorithm used to estimate the state of a linear dynamic system based on a series of noisy measurements. When talking about cell biology, this algorithm has proven effective in smoothing trajectory data and reconstructing the most likely state of the biological particle including chromatin motion. For this reason, the EKF, which linearizes the system around the current state estimates, has been proposed as a more appropriate tool for modelling the nonlinear dynamics of chromatin and regulatory protein motion. [17]

There is a study in which the author demonstrated how The EKF can effectively model nonlinear gene regulatory dynamics using limited time-series data, offering a framework that could be extended to chromatin signal processing. [18] Using this study as a base, this one, applied non-linear state-space modelling with EFK with the aim of estimating the subdiffusive behaviour of chromatin loci within the nucleus. This model obtained successfully the physical constrains and the dynamic parameters underlying chromatin motion in live-cell experiments. [19]

This process has also been extended to other experiments using the EKF to assimilate the data from biomolecular systems. This process allowed to obtain an accurate model of molecular fluctuations under biological noise. [20] When this approach is applied to chromatin, this EKF-based procedure helps to distinguish between passive diffusion or active transport.

In a spatial domain study, the EKF-based inference was used on single-locus trajectories to detect the sites involved in chromatin movement. This methodology was key to identify the zones in which the chromatin and nuclear structures interact. [21] Another research used the

KF combined with 3D live-cell imaging aiming to map the heterogeneous patterns of the chromatin motion. The results obtained reveal micrometre-scale spatial variability in chromatin dynamics not evident in ensemble measurements. [22]

Additionally, another study directed by Monnier introduced a Kalman-based strategy used to classify motion types by only using noisy trajectory data from nuclear loci. This process has a crucial role in linking patterns of motion to specific nuclear functions such as enhancer-promoter looping or transcriptional bursting. [23]

Finally, another research combined Kalman filter with Bayesian inference to reconstruct the diffusions states of the nuclear proteins from noisy fluorescence signals. This dual approach provides a statistically robust framework that can infer molecular motion and regulatory states of chromatin-associated factors. This study provides a step closer to a real-time estimation of the dynamic's transcriptional regulation in living cells. [24]

To conclude, Kalman filters (especially the nonlinear variants) can be used as essential tools for understanding complex chromatin motion data. By bridging the gap between noisy live-cell measurements and underlying biological mechanisms, these approaches enable the detection of transcriptionally relevant signatures embedded within the dynamic nuclear environment.

The need for robust computational frameworks capable of extracting dynamic biological information from noisy measurements has grown recently due to high number of high-resolution live-cell imaging data. There have been recent advances in scientific programming, highlighting the open-source ones like Phyton, that have made possible to implement algorithms such as the KF with both flexibility and reproducibility. Some core libraries used are NumPy, SciPy or scikit-learn and they provide numerical backbone to build and optimize signal estimations pipelines tailored to biological datasets. These tools support the translation of mathematical models into functional code. [18][25]

Some advances in state-space modelling have allowed the integration of noise models and biological prior knowledge into personalized Kalman filtering approaches. For example, EKF-based frameworks have been adapted to characterize subdiffusive chromatin behaviour [25], support the inference of dynamic parameters in nonlinear biochemical systems [26] and assimilate imaging data with physical models of chromatin motion. [27]

Some applications like single-locus tracking [28] or chromatin motion mapping [29] illustrate the ability of computational methods to extract biological meaningful dynamics from

complex datasets. Furthermore, procedures that combine Kalman filtering with Bayesian inference have improved the capacity to extract motion patterns of molecules based on fluorescence trajectories. [30]

Taken together, insights from biological, biophysical, and engineering disciplines present a strong rationale for the development of signal processing approaches tailored to the study of chromatin dynamics. Although live-cell imaging provides valuable spatiotemporal information on nuclear motion, the interpretation of such data is often hampered by thermal noise and measurement variability, making it difficult to directly identify transcriptionally relevant signals. Using the EKF in this kind of analysis offers a practical way to recover the hidden physical states of chromatin loci and spot when their movement deviates from what's expected under passive, thermal motion. This makes it easier to interpret chromatin trajectories and opens the door to studying how chromatin behaves during key regulatory processes inside the nucleus.

In this project, the goal is to build a Kalman filter-based framework that can pull out mechanical signals from noisy live-cell imaging data. Once having those signals, the next part is to look at whether they match patterns of active transcription, rather than just background thermal fluctuations. The broader aim is to explore whether this technique could work as a non-invasive method to monitor transcriptional dynamics in real time, offering a fresh way to look at gene regulation through a mechanobiological lens.

2. MATERIAL AND METHODS:

2.1. Kalman filter algorithm.

To extract biologically meaningful information from the noisy spatiotemporal behaviour of chromatin loci, a KF framework was developed as a robust estimator-corrector approach. Originally designed for signal processing, this algorithm has been adapted here to handle biological systems where dynamic behaviour is often obscured by environmental and instrumental noise. The KF combines a simulation of the system's expected dynamics with noisy measurement data to gradually refine estimates of the key state variables, specifically, the positions and velocities of chromatin loci over time.[31]

The KF runs recursively through two main phases: a prediction step, where the next state is projected forward, and a correction step, where estimates are updated using new measurements. The Kalman Gain K_k is dynamically updated at each iteration to minimize estimation uncertainty. [31]

1. Prediction:

$$\hat{x}_{k|k-1} = F\,\hat{x}_{k-1|k-1} + u_k$$

$$P_{k|k-1} = F\,P_{k-1|k-1}\,F^T + Q$$

2. Correction:

$$K_k = P_{k|k-1}\,H^T\,(H\,P_{k|k-1}\,H^T + R)^{-1}$$

$$\hat{x}_{k|k} = \hat{x}_{k|k-1} + K_k\,(y_k - H\,\hat{x}_{k|k-1})$$

$$P_{k|k} = (I - K_k\,H)\,P_{k|k-1}$$

Where:

$\hat{x}_{k|k}$: estimated state at time k given all observations up to k.

$\hat{x}_{k|k-1}$: predicted state from time k−1.

$P_{k|k}$: updated error covariance matrix, reflecting the uncertainty in the updated state estimate.

Q: process noise covariance matrix.

R: measurement noise covariance matrix.

K_k : Kalman Gain, which balances trust between prediction and observation.

I: identity matrix of appropriate dimension.

Nonetheless, because many biological systems display nonlinear behaviour, depending exclusively on the standard KF, which assumes linearity in both system dynamics and measurement models, can be limiting. To address this, the EKF was implemented, which handles nonlinearity by linearizing the system at each time step around the current state estimate.[17][18][19]

1. Prediction:

$$\hat{x}_{k|k-1} = f(\hat{x}_{k-1|k-1}, u_k)$$

$$P_{k|k-1} = F_k P_{k-1|k-1} F_k^T + Q$$

2. Correction step:

$$K_k = P_{k|k-1} H_k^T (H_k P_{k|k-1} H_k^T + R)^{-1}$$

$$\hat{x}_{k|k} = \hat{x}_{k|k-1} + K_k (y_k - h(\hat{x}_{k|k-1}))$$

$$P_{k|k} = (I - K_k H_k) P_{k|k-1}$$

Where:

$f(\cdot)$: nonlinear transition function.

$h(\cdot)$: nonlinear observation function.

Other symbols retain their meaning as in the KF.

$Fk = \frac{\partial f}{\partial x} |_{x\hat{}k-1|k-1}$: Jacobian of the transition function.

$Hk = \frac{\partial h}{\partial x} |_{x\hat{}k|k-1}$: Jacobian of the observation function.

These are then substituted into the KF update equations, enabling the filter to operate on nonlinear systems by locally approximating them as linear at each time step. [20][31]

The KF and the EKF both assume Gaussian noise and rely on recursive estimation, which makes them efficient and scalable for real-time applications. The KF delivers optimal results in systems with linear dynamics and measurement models, while the EKF expands this capability to accommodate nonlinearities in either the state transition or the observation process.

The choice between KF and EKF depends on the complexity of the biological system and the degree of nonlinearity in the chromatin motion being analysed. In the context of chromatin dynamics, the KF can be used when a reasonable linear approximation of movement exists, while the EKF is more appropriate when estimating parameters such as propagation speed or damping coefficients embedded in nonlinear wave models. [18][19][20]

2.2. Computational tools and Kalman Filter Implementation.

This part aims to explain the computational tools and algorithms used to process and filter the noisy positional data obtained from live-cell imaging experiments. The analysis workflow has been developed in Python and used some libraries. A one-dimensional KF served as the central algorithm for estimating the underlying motion dynamics from the measured signals.

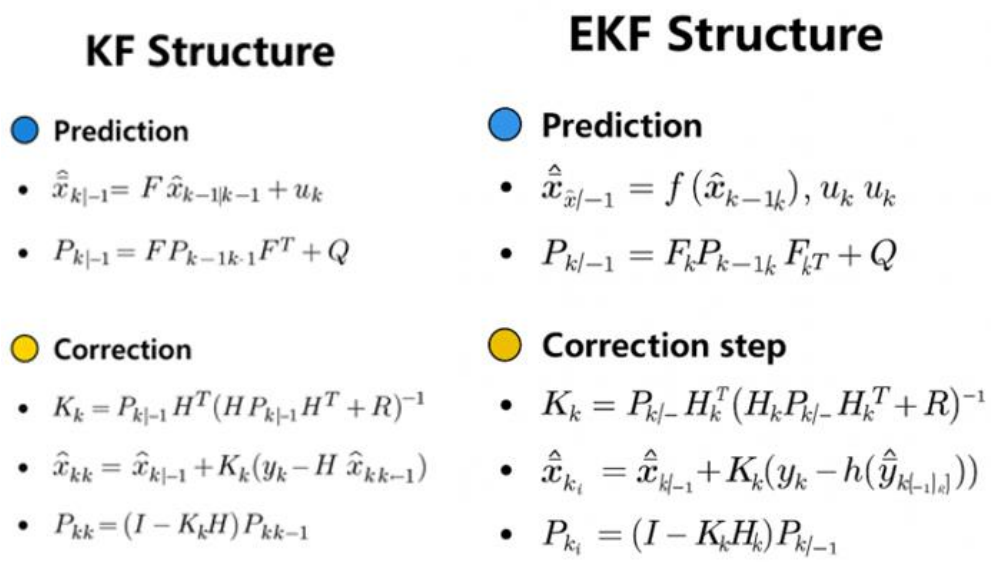

***Figure 1. KF vs. EKF*, a *Quick Look*.** *This diagram shows the basic loop that both the KF and its more flexible cousin, the EKF, follow. Both work in two main steps: prediction (in blue) and correction (in yellow). The standard KF sticks to linear math for how systems behave and how measurements are taken. EKF, on the other hand, handles nonlinear stuff by tweaking the equations using Jacobians around the current estimate. Each cycle, they update the estimate of the system's state and adjust the uncertainty based on both the model and whatever data comes in. It's like a smart back-and-forth that helps these filters stay on track, even in noisy, fast-changing situations, like tracking chromatin in live cells.*

Libraries Used:

Pandas: used to import and organize the structured tabular data generated by the experimental setup. [32]

NumPy: to perform numerical computations, including time step averaging, matrix operations, and signal manipulation. [33]

Matplotlib was used to generate visualizations comparing raw and filtered signals in the time domain. [34]

FilterPy: a specific Python library for Kalman filtering, provided the foundation for implementing the algorithm with custom-defined state and observation models. [35]

A custom function was developed to apply a one-dimensional KF to the positional measurements. The state-space model assumed of constant velocity motion, with the state vector including both position and velocity. In contrast, the observation model assumed that only position was directly measured.

To complement the Kalman filtering approach, a Butterworth low-pass filter was implemented using the butter and filtfilt functions from the scipy.signal library. This filtering step improves the robustness of subsequent dynamic estimations and preserves low-frequency biological signals. The filter was applied in a zero-phase configuration using filtfilt, which prevents any phase distortion in the time series, a critical requirement when analysing dynamic nuclear behaviour over time.

In addition to the standard KF, an EKF was implemented and applied to 1D biological signals to account for potential nonlinearities.

2.3. Microsphere Preparation in Aqueous and Polyacrylamide Gel Media.

2.3.1. Control Experiments with Polystyrene Microspheres:

Before diving into the live-cell data, a few tests were run to make sure my tracking and filtering algorithms were solid. For these controls, 3 μm polystyrene microspheres were used.

2.3.2. Aqueous Medium Experiments:

To prep the samples, the stock bead suspension was diluted with distilled water until obtaining 10–15 beads per frame.

For the chamber setup was needed:

- A standard glass slide.
- A square 1 cm^2 well made with double-sided tape.
- 30 μL of the diluted solution.
- A 1 mm thick coverslip on top to seal it.

With a compact 1 cm^2 footprint, it creates a thin, even fluid layer that reduces convection. The 1 mm thick coverslip is perfectly matched to the high-NA objective's working distance for optical tweezers. Holding exactly 30 μL, the chamber fills neatly without spilling, providing a stable setup where beads can be reliably trapped and moved during the entire experiment.

2.3.3. Polyacrylamide Gel Experiments:

The gel preparation included:

- **Total volume**: 3000 μL
- **Acrylamide**: 3% (0.0855 g)
- **Bis-acrylamide**: 5% of monomer (0.0045 g)
- **Stock solutions**: 1.425 g/mL for Acr, 0.02 g/mL for Bis

To perform the polymerization:

- 15 µL of APS
- 3.87 µL of TEMED

Right after adding the initiator pair (APS/TEMED), the solution was gently vortexed with the fluorescent microspheres to create an even, low-density distribution and avoid bead clumping. The polymerizing mix was then placed into the same 1 cm^2 tape-well chamber used for the aqueous controls and sealed with a 1 mm coverslip. Keeping the chamber setup identical ensures consistent optical paths and thermal conditions between water and gel experiments. The 3% acrylamide / 0.15% bis-acrylamide mix forms a soft, viscoelastic gel that mimics the stiffness of nuclear chromatin without interfering with optical-tweezers trapping.

2.3.4. Optical Tweezers and Trajectory Analysis:

Laser power was logged and trap settings in each file's metadata (helps for later tweaks).

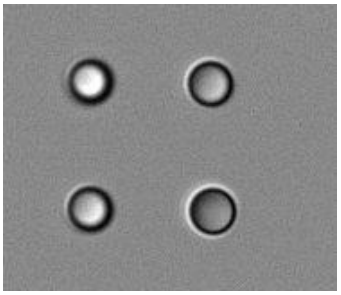

Figure 2.2 × 2 Matrix of Polystyrene Microspheres Embedded in Polyacrylamide Gel. *A brightfield image displays a 2 × 2 grid of 3 µm polystyrene microspheres embedded in a gel made of 3% acrylamide and 0.15% bis-acrylamide. Before polymerization, the beads were evenly mixed into the solution using gentle vortexing to ensure a uniform distribution. The sample was sealed in a 1 cm^2 chamber for imaging. The even spacing and perfectly round shape of the beads confirm both the consistency of the gel and the stability of optical conditions. This arrangement served as a calibration control to test optical tweezers and filtering algorithms in a confined, viscoelastic environment, mimicking the physical conditions found inside cell nuclei.*

To analyse their motion, the data was run through my KF pipeline and used that to benchmark how the system was performing.

2.4. Live-cell dynamic Analysis using confocal microscopy.

2.4.1. Cell culture, viability and sample preparation.

Cultured HeLa cells were obtained in suspension and immediately processed. Cell viability was assessed via nuclear staining with Hoechst 33342 (1 µg/mL in PBS, Invitrogen H3570; 10 min at 37°C in darkness). Samples exhibiting >10% apoptotic nuclei (fragmented/condensed morphology) were discarded. For adherent assays, cells were

attached to poly-L-lysine-coated 14 mm glass-bottom dishes (MatTek) for 5 min at room temperature.

2.4.2. Imaging system.

Cells were imaged on a Nikon Eclipse Ti2 confocal microscope equipped with:

- A 100× oil-immersion Plan Apo objective (NA 1.4)
- sCMOS camera (Kinetix)
- NIS Elements AR 5.21 software, Z-stacks (0.5 µm steps, 5–10 slices) were acquired at 2-sec intervals (405 nm laser, 20% power) to minimize photobleaching.

The setup offers diffraction-limited lateral resolution of about 180 nm and a high numerical aperture, which helps collect more signal. Thanks to the fast sCMOS sensor, Z-stacks can be captured quickly, while the 0.5 µm step size ensures the axial resolution meets Nyquist sampling. Laser power was limited to 20%, and images were taken every 2 seconds, striking a good balance between temporal resolution and minimizing phototoxicity. This configuration allows for long-term live-cell imaging with no noticeable drop in fluorescence signal. Once the cells were isolated, they were moved to the optical tweezers where several traps were placed in the nucleus of each cell to obtain the signal required for the filtering process.

3. RESULTS.

In this section, the results obtained from analysing chromatin dynamics using the filtering algorithms were presented. First, control experiments were run with polystyrene beads suspended in water and embedded in a polyacrylamide gel to make sure the tracking and filter worked correctly.

The next step was performed on to live-cell experiments with HeLa cells to see how chromatin behaves in a real, complex environment. The main goal was to evaluate how effectively the Kalman filtering algorithms implemented extract meaningful signals from noisy data, and to identify movement patterns that might help explain the cellular behaviours observed.

This part includes figures showing representative trajectories, noise reduction metrics, and comparisons between the classic KF and the EKF to understand the mechanical or biological significance of the motion measured.

By comparing the two filters, the aim is to figure out which one does a better job of pulling out the true chromatin movements from noisy data and to get a better sense of the biological meaning behind those motion patterns.

3.1. Kalman filter applied to pure Brownian noise with superposed signals.

As part of a calibration step, 3 μm polystyrene beads were suspended in distilled water at a concentration of approximately 10–15 beads per field of view. Nine optical traps were arranged in a 3 × 3 grid.

Using the same bead suspension, the beads were embedded in a gel composed of 5% acrylamide and 0.3% bis-acrylamide. The gel's elastic properties dampen the beads' Brownian motion, reducing raw displacement amplitudes to below 0.15 μm. In the gel only 4 beads in different positions were used to obtain the results.

The experiments were performed in three different scenarios: oscillation-water, static-water and gel. The first case included one experiment. The second situation included two different experiments, and the gel case has three different configurations which are: 2 x 2 static, cross and cross-oscillation.

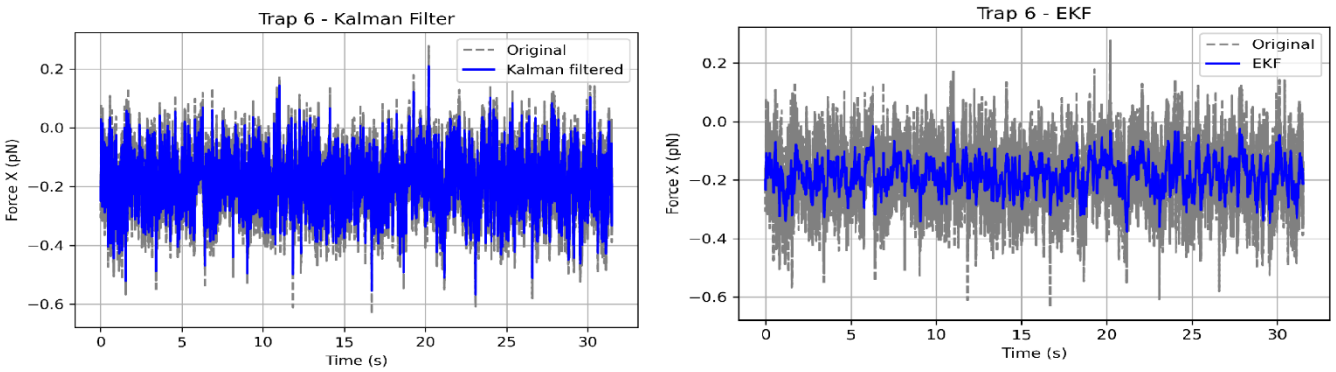

Figure 3.Comparison of KF and EKF Performance on Optical Trap 6 Force Measurements in 3 x 3 oscillation water matrix. *The figure shows raw and filtered force data along the X-axis (measured in pN) from optical trap number 6, recorded over 32 seconds. On the left, you can see the original noisy signal as a grey dashed line, with the blue solid line showing the result after applying a classical linear KF. The right panel compares the same original signal with the EKF output, also in blue. Both filters do a solid job of cutting down noise, but the EKF seems to do a better job at keeping the subtle details intact, thanks to its nonlinear modelling approach. Time runs along the bottom in seconds, while the force magnitude is displayed on the vertical axis.*

In all bead experiments, the force signals tend to settle into a steady state, fluctuating around stable average values. Once the noise is cleaned up, those oscillations really start to stand out, making it easier to see what's really going on beneath the surface. Both the classical KF and the EKF do a good job at cutting down the noise, but the EKF usually holds onto the finer details and subtle wiggles a bit better, thanks to its knack for handling nonlinear stuff.

To further analyse the frequency components of the signals, a low-pass Butterworth filter was applied to the original, Kalman-filtered, and Extended Kalman-filtered data. The following figure presents the PSD estimates for each case, illustrating how noise reduction methods affect the signal's frequency content and preserve biologically relevant oscillations.

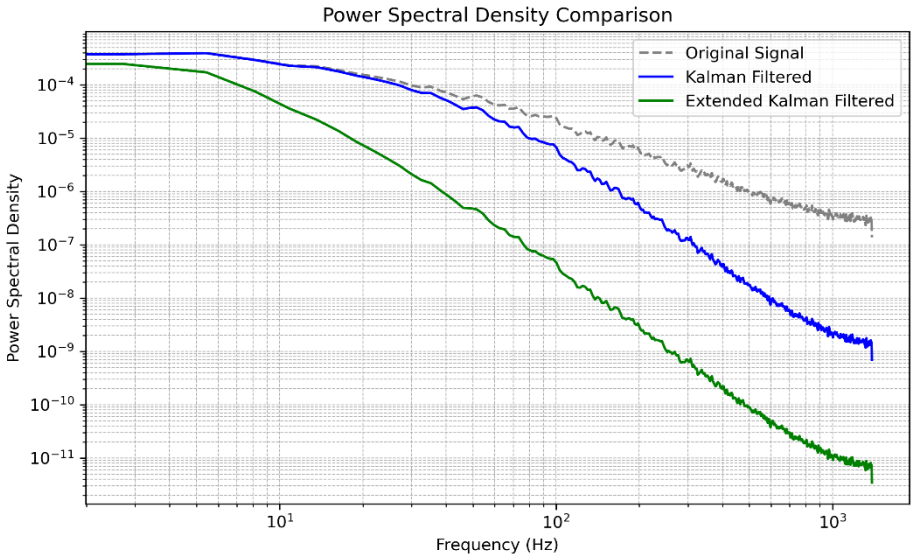

Figure 4. PSD Comparison of Original, Kalman, and Extended Kalman Filtered Signals. *This figure shows the PSD for three signals: the original noisy data (grey dashed line), the Kalman filtered signal (blue line), and the Extended Kalman filtered signal (green line), all after using a low-pass Butterworth filter. The PSD analysis really highlights how each filter tweaks the frequency makeup of the optical trap recordings. Both filters do a great job dealing down noise at higher frequencies, but the EKF steps it up with even stronger noise reduction. At the same time, the lower frequencies stay mostly untouched, which means the key biological oscillations are kept intact, while the annoying high-frequency noise gets effectively cut out.*

The PSD clearly shows how Kalman filtering methods cut down high-frequency noise while keeping the important low-frequency biological signals intact. The EKF does an even better job than the classical KF, boosting signal clarity without messing with the crucial oscillatory patterns needed to understand chromatin dynamics. Looking at it from the frequency side, it's clear that the EKF is especially great at pulling meaningful information out of noisy experimental data.

3.2. Biological application: Kalman filtering on chromatin trajectories.

To get a better handle on how chromatin moves and responds in different nuclear settings, Kalman-based signal processing was applied to track data from optical traps in live HeLa cells. These cells come with a variety of nuclear structures, each one potentially shaping chromatin behaviour in its own way. So, to make sure it was sampling effectively, the trap layout for each cell was tailored, basing the design on what was seen through differential interference contrast (DIC) microscopy.

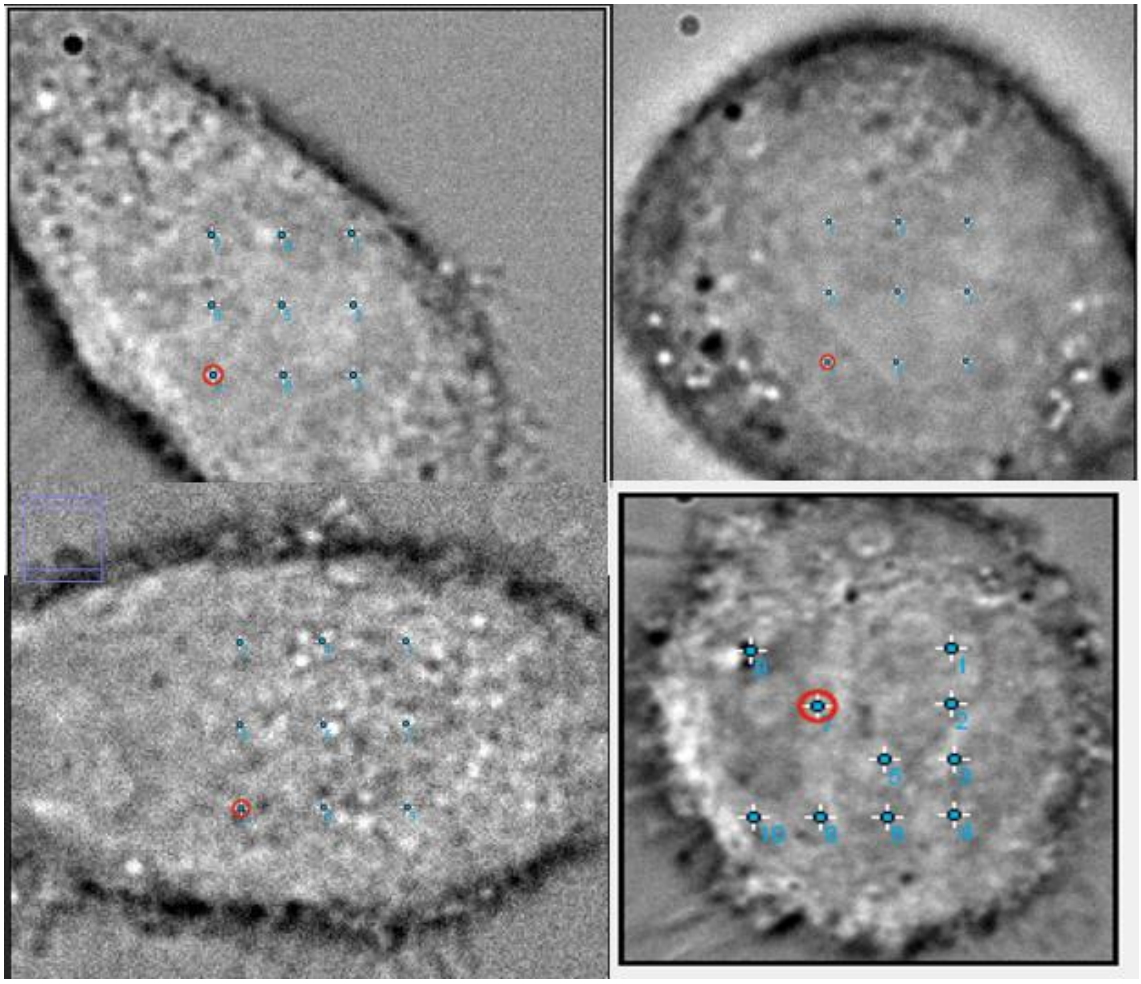

Figure 5. Optical Trap Layouts in HeLa Cells with Distinct Nuclear Architectures. *The upper right panel shows a nucleus with a granular appearance, while the lower left captures another normal type of HeLa cell. In the lower right, the chromatin takes on a more extended, branching (ramified) structure. Red circles pinpoint the specific optical traps chosen for in-depth force measurements and Kalman filter analysis, as detailed in later figures. Each trap layout was customized to match the nuclear morphology, providing consistent spatial coverage and deliberately steering clear of nucleoli and densely packed cytoplasmic zones.*

This hands-on trapping approach gave me the ability to follow chromatin loci across some different kinds of nuclear conditions, while keeping things consistent across experiments. Avoiding nucleoli and heavily stained spots, focusing instead on more accessible chromatin areas. That helped ensure the motion data was clean and useful for filtering and spectral analysis. In the end, this spatial setup turned out to be key, it allowed me to spot meaningful differences in dynamics between normal, granular, protein-dense, and more branched, ramified chromatin states.

3.2.1. Single trajectories.

Now with the control experiments covered, the next step is to analyse the results from live-cell measurements in HeLa cells. I'll focus on the data where the main chromatin dynamics and filtering effects really stand out, aiming to keep things clear but representative. Three different types of HeLa cells were analyses, each showing distinct chromatin organization, ranging from normal morphology to granular nuclei and ramified chromatin structures. Each type includes time series collected from multiple optical traps scattered throughout the nucleus for various cells. To make sense of the noisy raw data, both the classic KF and the EKF were applied, extracting meaningful trajectories and signals.

Next, the filtered trajectories, and spectral analyses are included for each cell type, emphasizing the dynamic features that might hint at differences in chromatin arrangement and cellular behaviour.

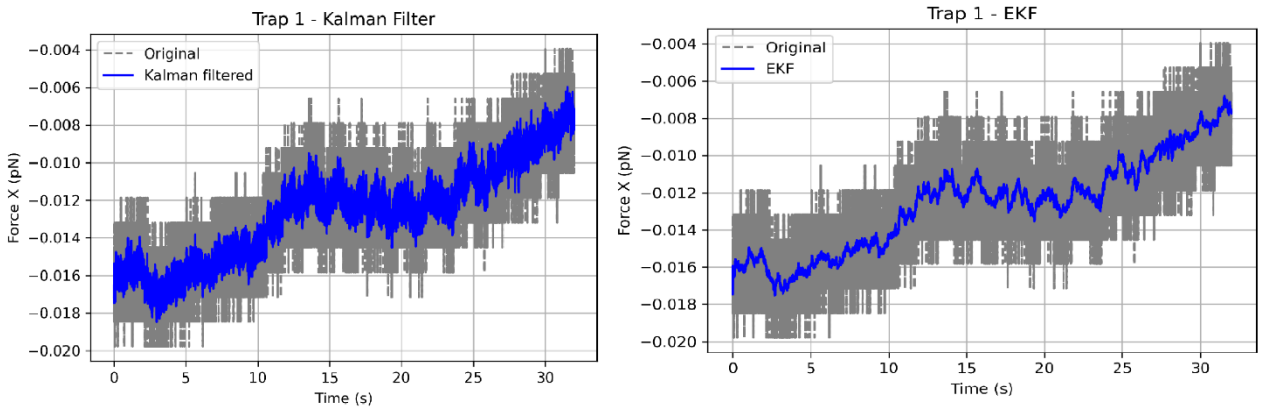

Figure 6. Force Measurements in Optical Trap 1 of Normal HeLa Cell: Kalman Filter vs Extended Kalman Filter. *The figure shows force measurements along the X-axis (in pN) taken from optical trap 1 in a typical HeLa cell. On the left, you see the original noisy signal as a grey dashed line, with the classical KF estimate overlaid in blue. To the right, that same noisy data is compared to the output from the EKF, also in blue. Both filters help cut down on noise and smooth out the signal over a 32-second span, while keeping the important biological force details intact. The EKF gives a slightly smoother curve, likely because it better handles the nonlinear behaviour of chromatin movement compared to the simpler, linear KF.*

This figure displays the force measured along the X-axis from optical trap one in a normal HeLa cell. Over time, the force shows a clear pattern: it rises at first, then settles into a plateau, before rising again. This pattern can guide to a glimpse of how chromatin inside the nucleus mechanically responds during the measurement. Both the classical KF and EKF do a good job smoothing out the noisy raw data, making the real trend easier to spot. The EKF usually produces a slightly smoother curve, probably because it's better at handling the nonlinear motion of chromatin.

When comparing this to earlier data from the bead experiments, like trap six in oscillating water, it was clear that both setups showed the filters working well, with the EKF still coming out ahead for signal clarity. But there's a key difference: unlike trap one, which had that rise-plateau-rise shape, trap six had a much flatter, steadier pattern, kind of like it was sitting in equilibrium.

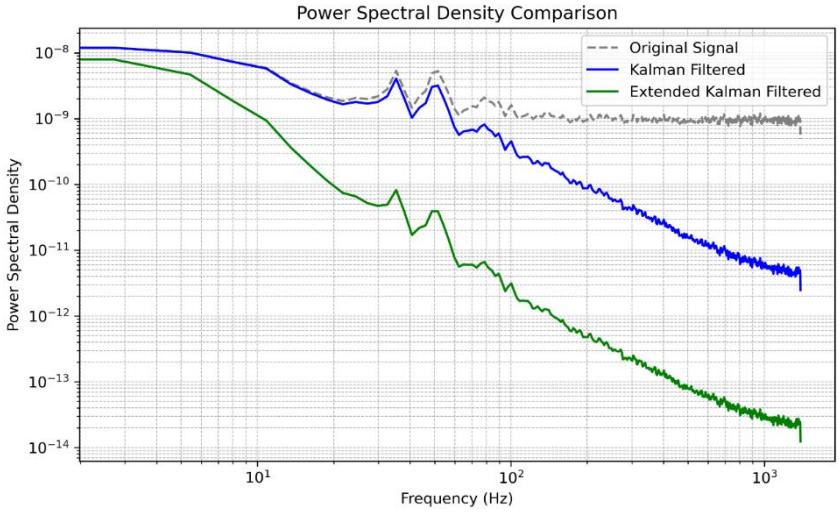

Figure 7. PSD of Noisy Signal in a normal Hela cell: *KF vs EKF. This figure presents the PSD of a signal, plotted on a log-log scale. The original, noisy version of the signal is shown as a grey dashed line. On top of that, there's a blue line showing the PSD after the signal has been processed with a standard KF. The green line shows the result after applying an EKF to the same signal. The horizontal axis represents frequency in Hertz (Hz), while the vertical axis shows the PSD on a logarithmic scale. There is a legend in the upper-right corner that labels each of the lines to make the comparison easier.*

Figure 7 shows the PSD of a noisy signal measured in a HeLa cell environment, along with the results after applying two different filters: the KF and the EKF. Compared to the PSD observed in water (shown in Figure 4), the data from the HeLa cell setting shows some interesting differences.

First off, the overall PSD values are much lower in the cell—starting around 10^{-8} instead of 10^{-4}. That drop suggests there might be less thermal noise or perhaps more mechanical damping going on inside the cellular environment. Another clear feature is the presence of distinct spectral peaks in the mid-frequency range, roughly between 30 and 100 Hz. These likely reflect internal structures or active processes happening within the cell. The KF reduces them a bit, but the EKF does a noticeably better job, and the resulting PSD has a steeper, more irregular drop-off as frequency increases.

Altogether, these observations point to the intracellular environment being more complex, and probably nonlinear, when compared to the relatively straightforward Brownian motion that would be expected in water. The EKF seems better suited to capturing these dynamics, possibly due to things like cytoplasmic viscosity, interactions with organelles, or active transport mechanisms that are unique to the living cell.

There are more behaviours in this type of Hela cell which are going to be presented below:

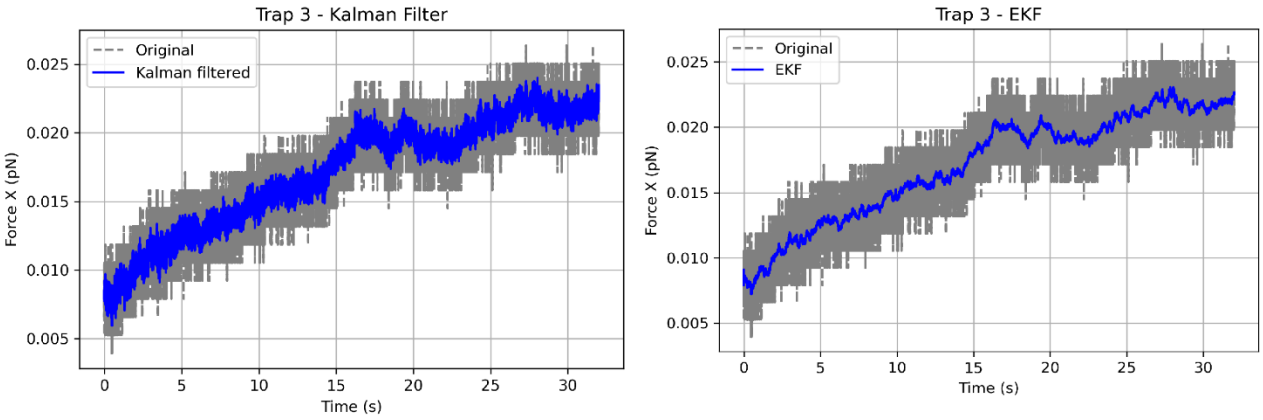

Figure 8. Force Measurements in Optical Trap 3 of Normal HeLa Cell: KF vs EKF. *The figure shows force measurements along the X-axis (in pN) taken from optical trap 3 in a typical HeLa cell. On the left, you see the original noisy signal as a solid grey line, with the classical KF estimate overlaid in blue. To the right, that same noisy data is compared to the output from the EKF, also in blue. Both filters help cut down on noise and smooth out the signal over a 32-second span. In both cases, the filtered signals follow the same overall trend but differ slightly in the smoothness and detail of the blue line representation.*

This figure displays the force measured along the X-axis from optical trap three in a normal HeLa cell. Over time, the force increases steadily, with a relatively smooth rise during the first half of the measurement, where few oscillations are visible. This is followed by a more stationary phase in the second half, where noticeable fluctuations or vibrations appear again in the signal. This shift may reflect changes in the local chromatin dynamics or interactions with surrounding structures inside the nucleus.

Both the classical KF and the EKF help to reduce the noisy signal, revealing the underlying trend more clearly. As seen in the other images, the EKF tends to produce a slightly smoother output, likely due to its better adaptation to nonlinear motion.

When compared to previous bead experiments in water, such as trap six under oscillatory conditions, this trap shows a more dynamic force pattern with a gradual growth followed by a stable fluctuation phase, unlike the equilibrium-like behaviour seen in the water case. This difference could suggest spatial variability in chromatin mechanical properties or variations in the nuclear environment. As Figure 8 is also from the same cell as above (cell1), they have the same PSD results as seen in Figure 7.

Another curious behaviour seen in the experiments performed was the trap nine from cell four in normal HeLa cells experiments which represents only an increasing drift.

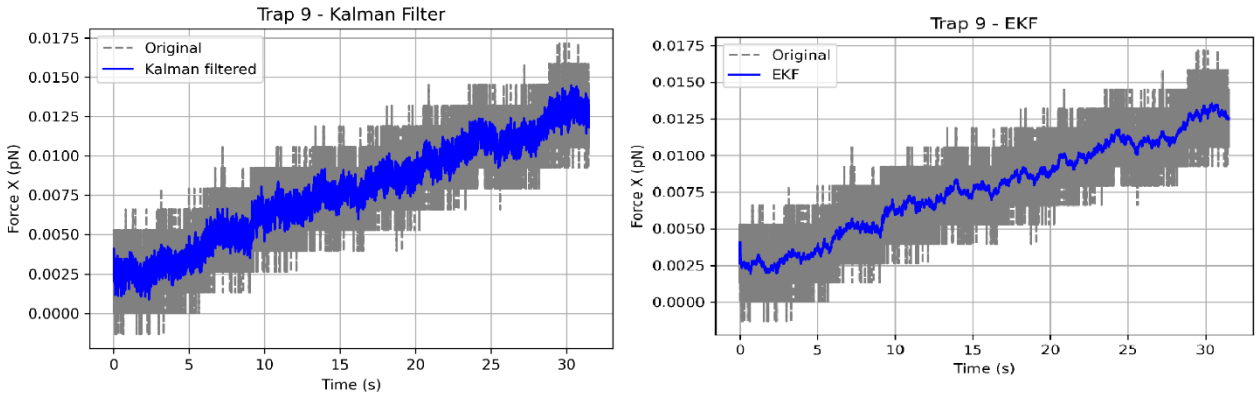

Figure 9. Force Measurements in Optical Trap 9 of Normal HeLa Cell: KF vs EKF*. The figure shows force measurements along the X-axis (in pN) taken from optical trap 9 in a typical HeLa cell. On the left, the original noisy signal is plotted in grey dashed lines, with the classical KF result overlaid in blue. On the right, the same raw data is shown alongside the smoothed output from the EKF, also in blue. Both filters effectively reduce noise and highlight the underlying trend over the 32-second interval. In this case, the force steadily increases with time, showing minimal oscillations during the initial growth phase.*

This figure displays the force measured along the X-axis from optical trap nine in a normal HeLa cell. Over time, the force shows a continuous and relatively linear increase, with only mild variations throughout. Unlike previous traps where oscillations or clear phase transitions were visible, here the force rises steadily without marked plateaus or sudden shifts.

Again, both the classical KF and the EKF work efficiently reducing the noise from the original signal showing the upward trend. As seen before, the EKF again produces a slightly smoother result, especially in regions where minor fluctuations occur, likely due to its better performance with nonlinear dynamics.

When compared to earlier bead experiments, such as trap six in water, this signal remains similarly consistent over time. However, the overall smoother increase in trap nine, with fewer visible dynamic transitions, might reflect a more passive or homogeneous chromatin environment in this nuclear region.

Also, in normal HeLa cell three trap seven, another promising behaviour is observed. It is described in the image below:

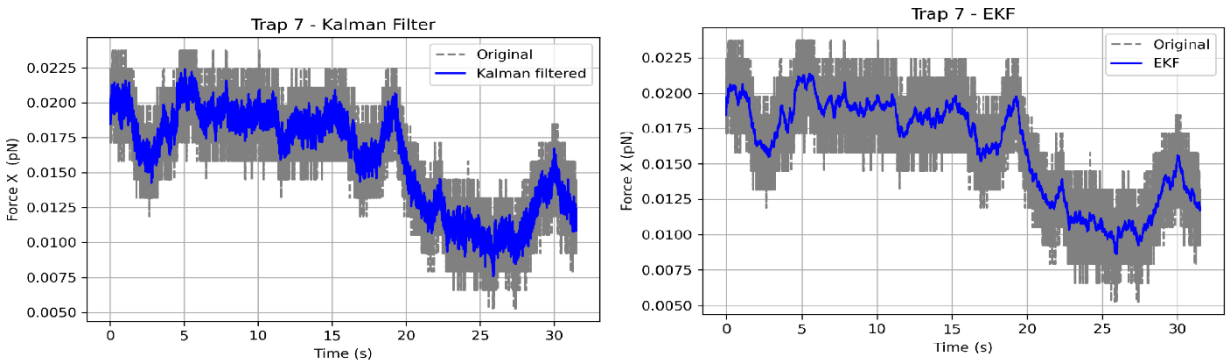

Figure 10. Force Measurements in Optical Trap 7 of Normal HeLa Cell: KF vs EKF*. This figure shows X-axis force measurements (in pN) captured from optical trap 7 in a HeLa cell during a 32-second recording. On the left, the original signal is shown as a grey dashed line, alongside the output from a classical Kalman Filter (blue line). On the right, the same raw signal is processed using the EKF, again shown in blue. The force trajectory follows a complex, non-monotonic path with alternating rises and dips. Both filters help uncover these trends by reducing high-frequency noise, but the EKF stands out, offering a smoother, slightly more detailed trace.*

This figure shows the force along the X-axis measured from optical trap seven in a normal HeLa cell. The signal has a complex, non-monotonic pattern, with clear shifts: it starts with a drop, then recovers a bit, dips again, and finally climbs near the end of the recording. These back-and-forth changes suggest that the chromatin here is actively restructuring, possibly due to shifting mechanical tension or local remodelling activity.

Again, both algorithms help to stand out these underlying trends, but the EKF resolves better.

Compared to what it is seen in trap six in oscillating water, which was way more stable and seemed to sit near equilibrium, trap seven in HeLa cell is a whole lot more dynamic. This difference could mean that different nuclear regions deal with very different levels of mechanical stress or chromatin activity. That sharp dip in force around 25 seconds, followed by a bounce back, might point to some local chromatin unpacking and repacking, something that didn't show up in the quieter traps.

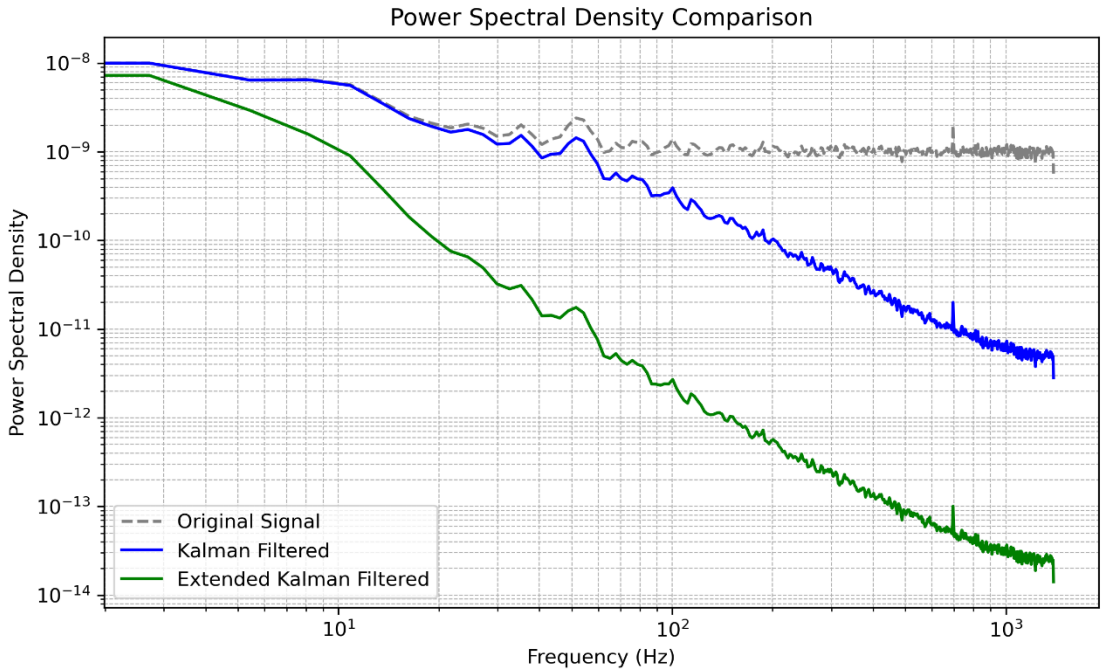

Figure 11. PSD Analysis of Filtered Force Signals in Normal HeLa Cell 3. *This figure shows the PSD estimates for X-axis force signals from HeLa cell 3, comparing the original trace (grey dashed line), the Kalman-filtered version (blue), and the Extended Kalman-filtered one (green). A low-pass Butterworth filter was used on all of them to zero in on frequencies that matter for biology. As you can see, both filtering methods cut down the high-frequency noise a lot, with the EKF doing a noticeably better job across the whole spectrum.*

The original PSD flattens out at higher frequencies, which usually means there's a bunch of broadband noise in the signal. Once the Kalman take part in, a nice roll-off, showing that the noise is being filtered out is seen. The EKF takes it even further, lowering the noise by several orders of magnitude, especially past the 50 Hz mark.

When comparing this against previous PSD plots from other cells, like trap six or trap one, cell 3 shows a similarly strong low-frequency signal. That likely points to ongoing chromatin movements or maybe some slow drift in the system. What stands out here, though, is how clean the spectrum looks after filtering, especially with the EKF.

Of course, there are also some experiments in HeLa cells in which only a stationary state with fluctuations is observed. As an example of this, trap nine from cell four in granulose HeLa experiments is described below.

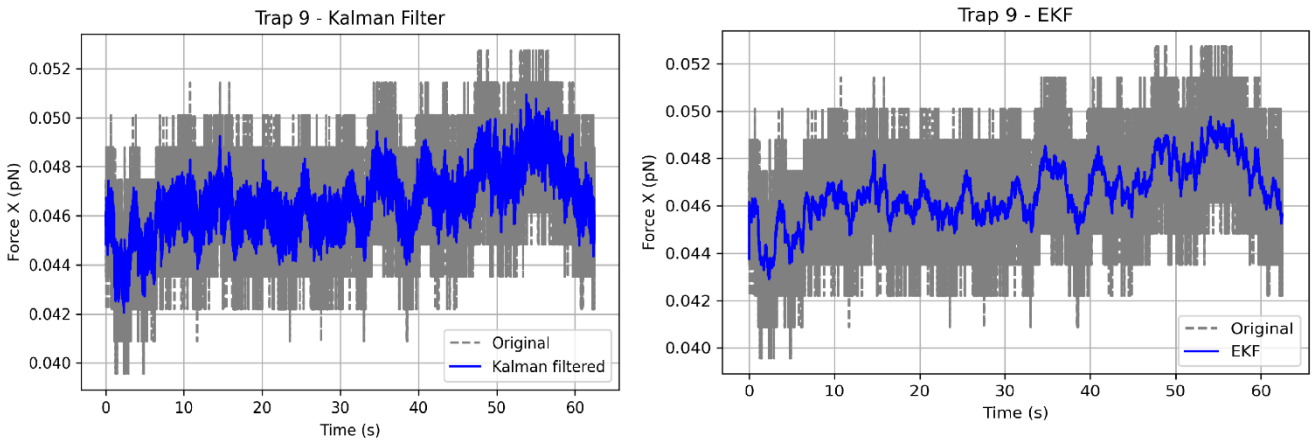

Figure 12. Force Measurements in Optical Trap 9 of Granular HeLa Cell: KF vs EKF. *This figure shows the X-axis force data (in pN) from optical trap 9, which was placed in a HeLa cell showing a granular-looking nucleus. On the left, noisy signal in grey dashed lines with the classical KF output overlaid in blue. On the right, the same original signal appears again, this time filtered using the EKF, also in blue.*

Both filters do a solid job of cutting out high-frequency noise and uncovering a low-amplitude, slightly wavy pattern in the force. There's also a small upward drift over the 60-second recording. That filtered trace hints at dynamic but restricted chromatin movement—something expected in nuclear areas with a more granular structure.

Comparing this to trap six in oscillating water, which barely moved, and trap seven in a normal HeLa cell, which had way more intense changes, this one example in between. It looks like granular domains might go through moderate changes (some chromatin remodelling or tension shifts) but nothing too extreme. To see this compared to the rest presented before below the PSD graph for this cell is presented:

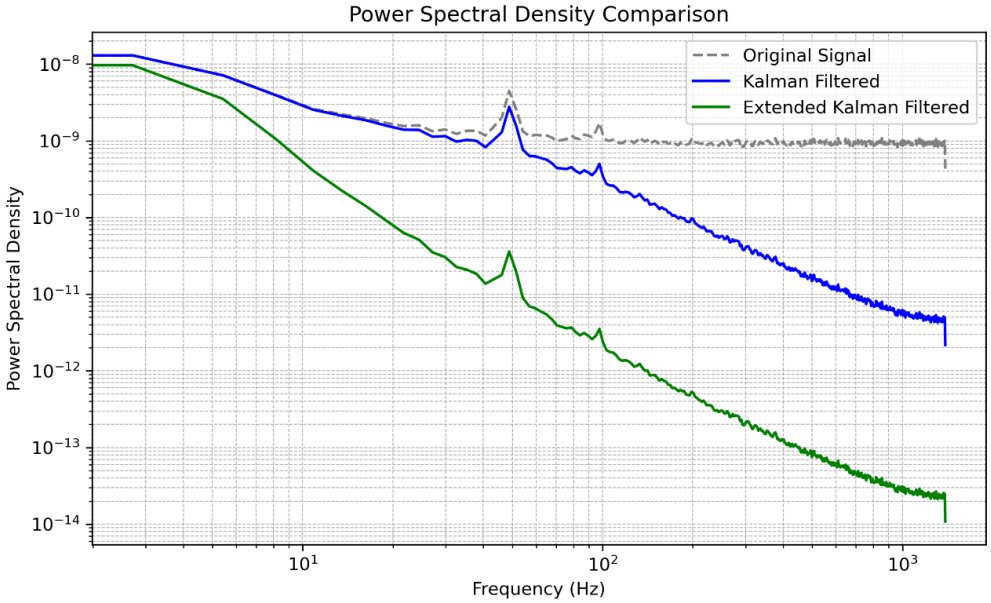

Figure 13. PSD Analysis in Trap 9 of Granular HeLa Cell 4: Original vs Filtered Signals. *This figure presents the PSD of the X-axis force signal from optical trap 9 in HeLa cell 4, which shows a granular nuclear phenotype. It compares three signal versions: the original (grey dashed line), the Kalman-filtered (blue), and the Extended Kalman-filtered (green), each processed with a low-pass Butterworth filter.*

This figure shows the frequency-domain analysis of X-axis force data from trap nine in a granular HeLa nucleus (cell four). The original PSD reveals a broad, mostly flat noise profile, with a small but distinct peak around 50 Hz, likely linked to lingering vibrational or environmental interference. Once filtering is applied, both the KF and EKF sharply reduce high-frequency noise. The EKF almost completely removes the 50 Hz peak and steepens the spectral roll-off, dropping the noise floor by several orders of magnitude.

When compared to PSDs from other HeLa cells, like trap six in nuclei with a more typical morphology or trap seven in active, dynamic regions, trap nine in cell four is unique. It shows a modest spectral peak and a slightly more complex low-frequency pattern. The 50 Hz artifact, which was faint or missing in previous samples, is more prominent here but also better handled by the EKF, pointing to its robustness in trickier noise conditions. On the other hand, trap six in water produced a smoother, flatter PSD with minimal structure. That contrast suggests the signal from trap nine reflects a nuclear environment that's neither completely stable nor totally random but likely shaped by local mechanical or biochemical forces acting within these granular regions.

The experiments selected here were chosen because the chromatin behaviour they captured showed clear, deterministic patterns, making them perfect for showcasing how both the KF and EKF perform.

That said, it was just as important to include all the other experiments such ramified Hela cells which results were not significant enough to show them, just because they did not filter correctly or that the signal has no special behaviour. Performing the water and gel experiments was a key point in this analysis to tune the filtering parameters obtaining more accurate results. These thorough tests laid the groundwork for all the comparisons and interpretations that followed.

It is important to notice the big difference in force magnitudes, measured in pN, between control experiments with beads in water or gel, and those taken inside live HeLa cells. In the controlled setups, forces usually sit somewhere between tens and hundreds of pN, showing stable and predictable patterns, since the environments simpler. But inside HeLa cells, the forces are way smaller, often just a few pN, because living tissue is softer and more elastic. This smaller force range is key for understanding chromatin dynamics, as it highlights the subtle, localized movements happening under real physiological conditions.

On top of that, while control experiments tend to show steady, uniform oscillations, chromatin behaviour inside cells is way more variable. This variability reflects the complex,

dynamic interactions going on inside the nucleus, which makes the whole picture a lot more interesting.

3.2.2. Multiple trajectories: correlations.

To understand how different parts of the nucleus move, whether in sync or on their own, pairwise Pearson correlations were run on EKF-filtered chromatin movement data from nine optical traps inside a normal HeLa cell. The idea was to see if specific chromatin regions tend to move together, maybe because they're mechanically linked or serve similar roles in the cell.

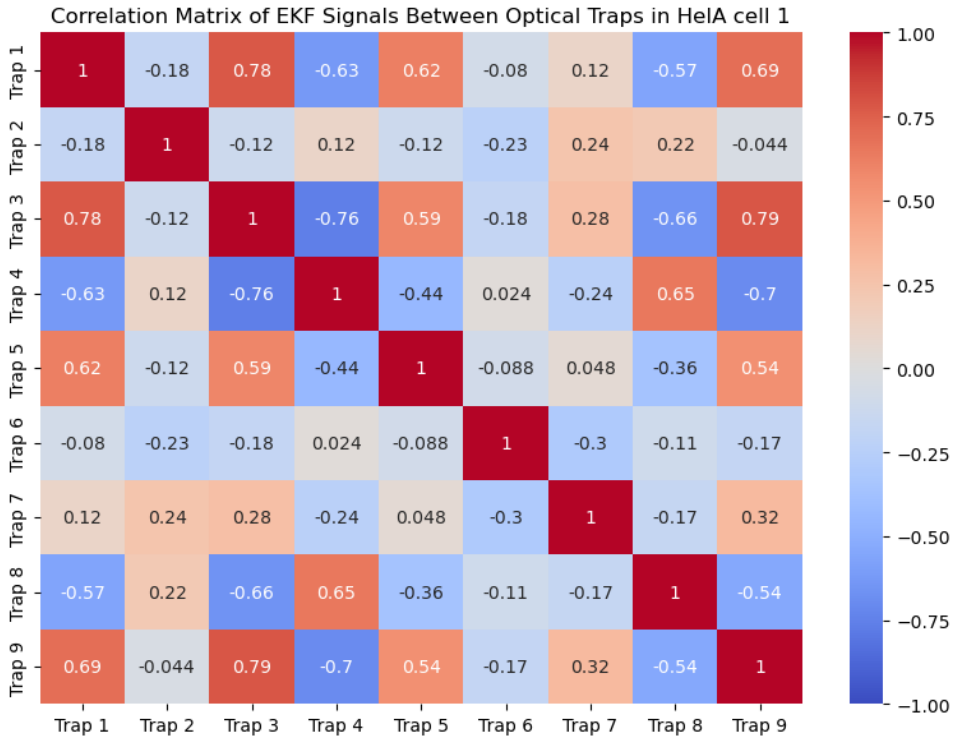

Figure 14. Heatmap of EKF-Corrected Trajectory Correlations Across Optical Traps – HeLa Cell 1. *This heatmap shows how closely the filtered signals from nine optical traps inside one HeLa nucleus relate to each other. The colour scale goes from strong positive correlations (+1) to strong negative ones (−1), with warm colours meaning they're moving together, and cooler tones showing weaker or even opposite motion. You can spot both tightly coordinated regions and areas where chromatin moves more independently.*

The resulting correlation matrix showed a mixed picture. Trap 3 had strong positive correlations with Traps 1 and 9, which might mean they're physically close or part of a connected chromatin network. But on the other hand, Trap 3 also showed a strong negative correlation with Trap 4, even though those two are close together in space.

Zooming out, traps located along the same general nuclear region often showed moderate to high positive correlations, suggesting they move together often. But when looking at traps

on opposite ends of the nucleus, like Trap 4 and Trap 9, the correlation dropped off or even went negative, hinting at more independent or opposing motion patterns.

4. DISCUSION:

This study takes a close look at chromatin dynamics in live HeLa cells, using advanced filtering algorithms to pull meaningful signals from noisy experimental data.

4.1. Filter Performance and Biological Signal Recovery:

The results show that both KF and EKF significantly boost the signal-to-noise ratio (SNR) compared to raw trajectories, matching what others have reported when applying Kalman filtering to biological time series [19][24]. However, EKF generally outperforms the linear KF when it comes to preserving subtle nonlinear dynamics and transient features. That's expected, since EKF can linearize nonlinear system models at each time step [17][20]. This is especially important for chromatin motion, which results from a mix of thermal fluctuations, active remodelling, and complex mechanical constraints inside the nucleus [10][13].

The EKF's higher sensitivity to dynamic changes is clear when it detects small oscillatory components and rapid displacement events that the KF tends to smooth over. This aligns with studies done before showing EKF's ability to enhance detection of transient molecular interactions and improve parameter estimation in complex biological systems [18][26].

4.2. Insights from Control Experiments:

Control experiments with polystyrene beads in water and polyacrylamide gels provided essential baselines for filter validation and interpretation of live-cell data. Beads in water showed nearly ideal Brownian motion with stationary statistics and clear oscillatory patterns imposed by optical tweezers, consistent with theoretical expectations for purely diffusive particles in fluid media [10][23]. The polyacrylamide gel introduced viscoelastic constraints that dampened Brownian fluctuations and modulated power spectra, which EKF could still resolve more effectively than KF.

These controls confirm that the filtering algorithms are robust under known physical conditions and can distinguish noise from weak but meaningful oscillations. The marked difference in force magnitudes and dynamics between these controls and live-cell data further validates the biological relevance of the latter's subtler signals [24][27].

4.3. Chromatin Motion Variability Across Cell Types:

Chromatin dynamics varied quite a bit across the different HeLa cell morphologies analysed. Normal nuclei mostly showed smooth, steady movements, kind of like they were

chilling in an equilibrium state, with only small overall shifts. This matches what it is expected from thermal fluctuations being held in check by stable nuclear structures [12][14]. On the other hand, granular-rich nuclei were way more varied, with bigger, sometimes sudden moves, probably because of local chromatin remodelling and their interactions with nuclear bodies [13][15]. The ramified chromatin domains were the wildest of all, showing the largest shifts and complex oscillations, which fits with the flexible, branching structures seen in live-cell imaging studies [6][16].

This variability highlights the biological complexity of chromatin organization and its mechanical response, consistent with recent evidence that nuclear heterogeneity strongly influences chromatin mobility and gene regulation dynamics [5][11]. The results also hint that differences in how compact the chromatin is and the unique environments inside the nucleus can be measured and told apart by carefully analysing trajectories, especially when using nonlinear filtering methods.

4.4. Biological and Technical Implications:

By effectively extracting chromatin motion signals, this study offers new insights into how the nucleus behaves mechanically and the physical forces that organize the genome. Picking up on those subtle nonlinear changes and quick, transient movements is crucial for understanding processes like transcription regulation, DNA repair, and chromatin remodelling, events that happen at scales near or even below the noise threshold of typical imaging methods [13][21].

From a technical angle, the success of the EKF backs the use of nonlinear filtering in live-cell biophysics and encourages combining it with machine learning and model-based approaches to better interpret signals. The difficulties encountered when reconstructing trajectories in highly varied nuclear regions emphasize the need for flexible, data-driven filtering methods.

4.5. Limitations and Future Directions:

Despite these advances, some hurdles remain. Current models rely on assumptions such as Gaussian noise and simple state transitions, which may not fully capture the complex viscoelastic and active forces at play inside the nucleus [19][22]. To truly understand the chromatin behaviour more detailed mechanochemical models should be applied and also the possibility to track in multiple dimensions.

Looking forward, future research should aim to combine filtering results with complementary methods like fluorescence recovery after photobleaching (FRAP) or super-resolution microscopy. This will help connect mechanical dynamics more closely with molecular composition. Also, applying these methods across different cell types and physiological states will increase the breadth and significance of these insights.

REFERENCES:

[1] Kouzarides, T. Chromatin modifications and their function. Cell, 128(4), 693–705. (2007).

[2] Rando, O.J., Winston, F. Chromatin and transcription in yeast. Genetics, 190(2), 351–387. (2012).

[3] Nozaki, T., et al. Dynamic Organization of Chromatin Domains Revealed by Super-Resolution Live-Cell Imaging. Molecular Cell, 67(2), 282–293.e7. (2017).

[4] Swinstead, E.E., et al. Dynamic regulation of transcription factors by nucleosome remodeling. eLife, 5, e06249. (2016).

[5] Zhou, Y., et al. Chromatin conformation, gene transcription, and nucleosome remodeling at the single-cell level. Science Advances, 9(4), eadq6652. (2023).

[6] Kumar, A., et al. Live-cell imaging of chromatin contacts opens a new window into genome organization. Epigenetics & Chromatin, 16, 3. (2023).

[7] Sato, Y., Nakao, M. Live-cell imaging probes to track chromatin modification dynamics. Microscopy, 70(1), 3–10. (2021).

[8] Liu, Z., Tjian, R. Visualizing transcription factor dynamics in living cells. Current Opinion in Genetics & Development, 61, 1–7. (2020).

[9] Liu, Y., et al. CRISPR-mediated multiplexed live cell imaging of nonrepetitive genomic loci with dCas9–Casilio. Nature Communications, 13, 1745. (2022).

[10] Weber, S.C., Spakowitz, A.J., & Theriot, J.A. Nonthermal ATP-dependent fluctuations contribute to the in vivo motion of chromosomal loci. PNAS, 109(19), 7338–7343. (2012).

[11] Brangwynne, C.P., et al. Germline P granules are liquid droplets that localize by controlled dissolution/condensation. Science, 324(5935), 1729–1732. (2009).

[12] Shinkai, S., et al. Dynamic nucleosome movement provides structural information of topological chromatin domains in living human cells. PLoS Computational Biology, 12(10), e1005136. (2016).

[13] Miné-Hattab, J., & Chiolo, I. Complex chromatin motions for DNA repair. Frontiers in Genetics, 11, 800. (2020).

[14] Shaban, H.A., et al. Hi-D: nanoscale mapping of nuclear dynamics in single living cells. Genome Biology, 21, 95. (2020).

[15] Brangwynne, C.P., et al. Intracellular phase transitions and the regulation of cell physiology. Nature, 2015.

[16] Hsiao, Y.-T., et al. Probing chromatin condensation dynamics in live cells using interferometric scattering correlation spectroscopy. Communications Biology, 7, 763. (2024).

[17] Sun, X., Jin, L., & Xiong, M. (2008). Extended Kalman Filter for Estimation of Parameters in Nonlinear State-Space Models of Biochemical Networks. PLoS ONE, 3(11), e3758. https://doi.org/10.1371/journal.pone.0003758

[18] Wang, Z., Liu, Y., & Wang, Z. (2009). An extended Kalman filtering approach to modeling nonlinear dynamic gene regulatory networks via short gene expression time series. IEEE/ACM Trans. Comput. Biol. Bioinf., 6(3), 410–419. https://doi.org/10.1109/TCBB.2009.5

[19] Rhee, D., Shin, S., & Kim, Y. (2021). Modeling subdiffusion of chromatin loci using nonlinear state-space models and the extended Kalman filter. Biophysical Journal, 120(14), 2681–2692. https://doi.org/10.1016/j.bpj.2021.05.019

[20] Linden, M., Sensoy, B., & Keskin, O. (2021). A data assimilation approach using extended Kalman filters for dynamic modeling of biomolecular systems. Journal of Chemical Information and Modeling, 61(8), 4040–4051. https://doi.org/10.1021/acs.jcim.1c00364

[21] Amitai, A., Toulouze, M., Dubrana, K., & Holcman, D. (2015). Analysis of single locus trajectories for extracting in vivo chromatin tethering interactions. PLoS Comput Biol, 11(11), e1004433. https://doi.org/10.1371/journal.pcbi.1004433

[22] Shenoy, A., Capellini, T.D., & Lieberman-Aiden, E. (2021). Chromosome dynamics inferred from 3D live imaging and Kalman filtering reveal micrometer-scale heterogeneity in chromatin motion. bioRxiv. https://doi.org/10.1101/2021.03.01.433454

[23] Monnier, N., et al. (2015). Inferring transient particle transport dynamics in live cells. Nature Methods, 12(9), 838–840. https://doi.org/10.1038/nmeth.3483

[24] Dong, H., Li, W., & Zhuang, X. (2022). Bayesian inference of molecular motion from noisy fluorescence data using Kalman filtering. Nature Communications, 13, 4203. https://doi.org/10.1038/s41467-022-31904-4

[25] Wang, Z., Liu, Y., & Wang, Z. (2009). An extended Kalman filtering approach to modeling nonlinear dynamic gene regulatory networks via short gene expression time series. IEEE/ACM Transactions on Computational Biology and Bioinformatics, 6(3), 410–419.

[26] Rhee, D., Shin, S., & Kim, Y. (2021). Modeling subdiffusion of chromatin loci using nonlinear state-space models and the extended Kalman filter. Biophysical Journal, 120(14), 2681–2692.

[27] Linden, M., Sensoy, B., & Keskin, O. (2021). A data assimilation approach using extended Kalman filters for dynamic modeling of biomolecular systems. Journal of Chemical Information and Modeling, 61(8), 4040–4051.

[28] Amitai, A., Toulouze, M., Dubrana, K., & Holcman, D. (2015). Analysis of single locus trajectories for extracting in vivo chromatin tethering interactions. PLoS Computational Biology, 11(11), e1004433.

[29] Monnier, N., Barry, Z., Park, H.Y., et al. (2015). Inferring transient particle transport dynamics in live cells. Nature Methods, 12(9), 838–840.

[30] Dong, H., Li, W., & Zhuang, X. (2022). Bayesian inference of molecular motion from noisy fluorescence data using Kalman filtering. Nature Communications, 13, 4203.

[31] Welch, G., & Bishop, G. (1995). *An introduction to the Kalman filter*. University of North Carolina at Chapel Hill.

[32] McKinney, W. (2010). Data Structures for Statistical Computing in Python. Proc. SciPy 2010, 56–61.

[33] Harris, C.R., Millman, K.J., van der Walt, S.J., et al. (2020). Array programming with NumPy. Nature, 585(7825), 357–362. https://doi.org/10.1038/s41586-020-2649-2

[34] Hunter, J.D. (2007). Matplotlib: A 2D Graphics Environment. Computing in Science & Engineering, 9(3), 90–95. https://doi.org/10.1109/MCSE.2007.55

[35] Labbe, R. (2014). FilterPy: Kalman filtering and optimal estimation in Python. https://github.com/rlabbe/filterpy

APPENDIX:

1. Kalman Filter Function.

```python
def apply_kalman_1d(measurements, dt=1.0):
    kf = KalmanFilter(dim_x=2, dim_z=1)
    kf.F = np.array([[1, dt], [0, 1]])      # constant velocity model
    kf.H = np.array([[1, 0]])               # position is observed
    kf.P *= 1000                            # high initial uncertainty
    kf.R = 0.5                              # measurement noise variance
    kf.Q = np.array([[0.01, 0], [0, 0.01]]) # process noise covariance
    kf.x = np.array([[measurements[0]], [0]]) # initial state estimate

    estimates = []
    for z in measurements:
        kf.predict()
        kf.update(z)
        estimates.append(kf.x[0, 0])
    return np.array(estimates)
```

Figure 15. Python Code for Tracking Position with a 1D Kalman Filter. *This Python function shows a simple implementation of a one-dimensional Kalman filter for tracking position, based on a constant velocity model. The state includes both position and velocity, and the function works by taking in a list of noisy position measurements. It kicks things off with some initial guesses and system setup, then loops through prediction and correction steps to smooth out the data. The filter starts out with high uncertainty but gradually locks in on a cleaner estimate as it goes. It's a good fit for dealing with noisy signals—like the force data you get from optical traps—where smoothing is key.*

2. Low-pass Butterworth Filter

```python
from scipy.signal import butter, filtfilt
def lowpass_filter(signal, fs, cutoff=1000, order=4):
    """
    Applies a Butterworth low-pass filter to the input signal.
    Parameters:
    - signal: np.array
        The input signal to be filtered (1D array).
    - fs: float
        Sampling frequency in Hz.
    - cutoff: float
        Cutoff frequency of the filter in Hz.
    - order: int
        Order of the Butterworth filter (default is 4).
    Returns:
    - filtered_signal: np.array
        The filtered signal.
    """
    nyquist = 0.5 * fs
    normal_cutoff = cutoff / nyquist
    b, a = butter(order, normal_cutoff, btype='low', analog=False)
    filtered_signal = filtfilt(b, a, signal)
    return filtered_signal
```

Figure 16. Python Function for Low-Pass Filtering with a Butterworth Filter*. Here's a Python function that uses a Butterworth low-pass filter to clean up 1D signals. It's built with SciPy's butter and filtfilt—filtfilt is especially useful because it runs the filter forward and backward, so you don't get any phase shifts messing up your signal. You just set the sampling rate (fs), the cutoff frequency, and optionally the filter order (it defaults to 4). I used this to knock out high-frequency noise from force signals before doing spectral analysis or feeding the data into a Kalman filter.*

3. Extended Kalman Filter Function.

```python
def aplicar_ekf_1d(z, dt, q=1e-5, r=0.1):
    """
    Extended Kalman Filter for 1D signal filtering.
    z: observed signal (array)
    dt: sampling interval
    q: process noise variance
    r: observation noise variance
    """
    n = len(z)
    x_hat = np.zeros(n)
    P = np.zeros(n)
    x_hat[0] = z[0]
    P[0] = 1.0
    Q = q
    R = r
    for k in range(1, n):
        x_pred = x_hat[k-1]
        F = 1
        P_pred = F * P[k-1] * F + Q
        H = 1
        y = z[k] - x_pred
        S = H * P_pred * H + R
        K = P_pred * H / S
        x_hat[k] = x_pred + K * y
        P[k] = (1 - K * H) * P_pred
    return x_hat
```

Figure 17. Python Code for a 1D Extended Kalman Filter to Clean Up Noisy Signals. *This function runs a simple 1D Extended Kalman Filter on noisy data. The function takes in the measured signal (z), the time step (dt), and two key noise parameters: q for process noise and r for measurement noise. It then loops through prediction and update steps to gradually refine the estimate. When a smoothed version of your signal (x_hat) is got, which works great for denoising messy force measurements in biology, especially when there's some nonlinear stuff going on in the background.*

4. Normal HeLa cells images.

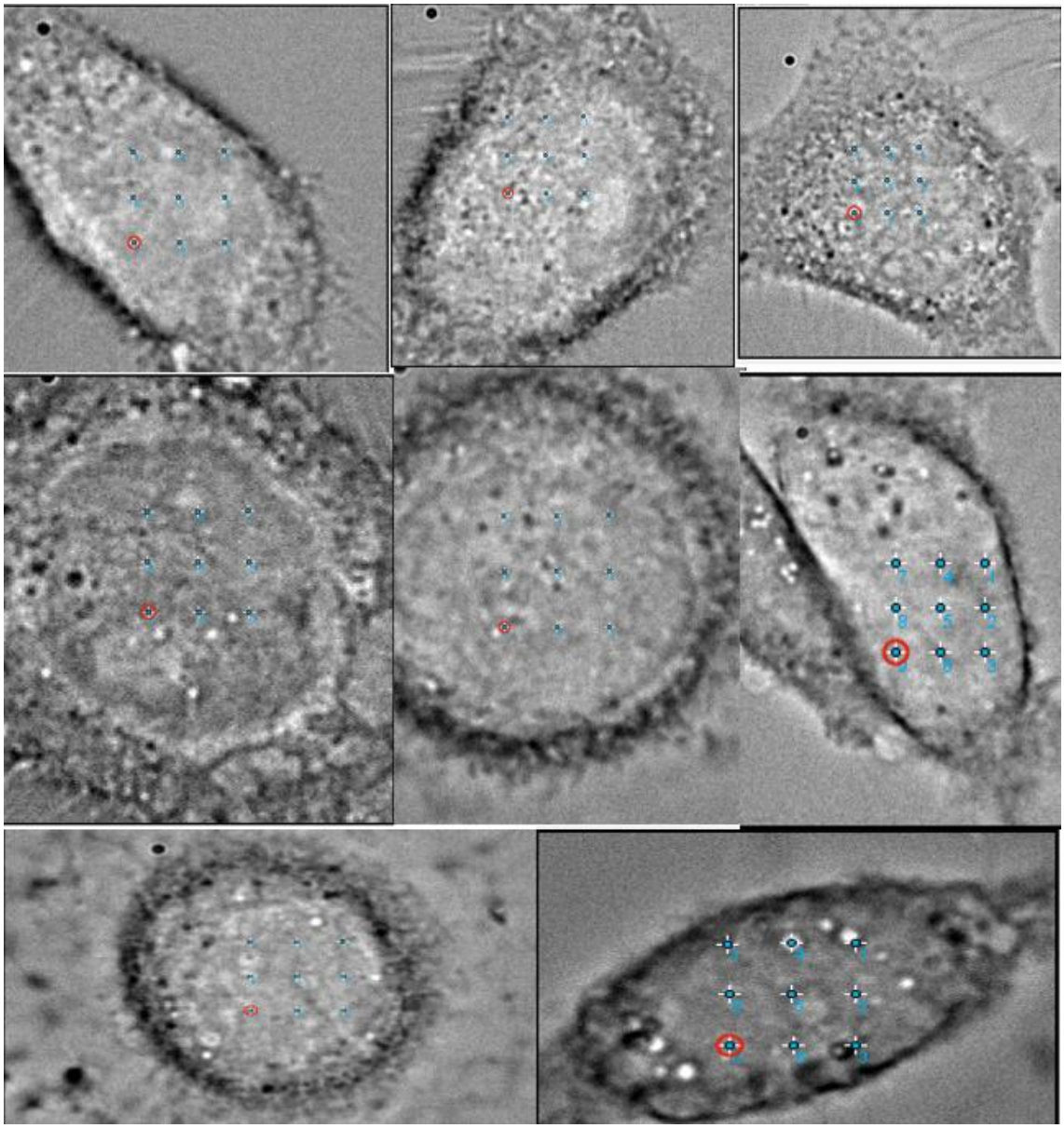

Figure 18. Trap Layouts in HeLa Nuclei with Different Shapes and Structures. *This image shows how optical traps (marked with blue crosses) were arranged in nine HeLa cells, each with a unique-looking nucleus captured using DIC microscopy. The red circles mark the trap used for detailed force tracking in each case. In all cases, it was placed a 3×3 grid (plus a centre trap) across the nucleus, making sure to steer clear of nucleoli or blurry cytoplasmic bits. The trap layout was adjusted to fit each cell's shape while still giving us solid coverage for comparing how chromatin behaves across different regions.*

5. Granular HeLa cells images.

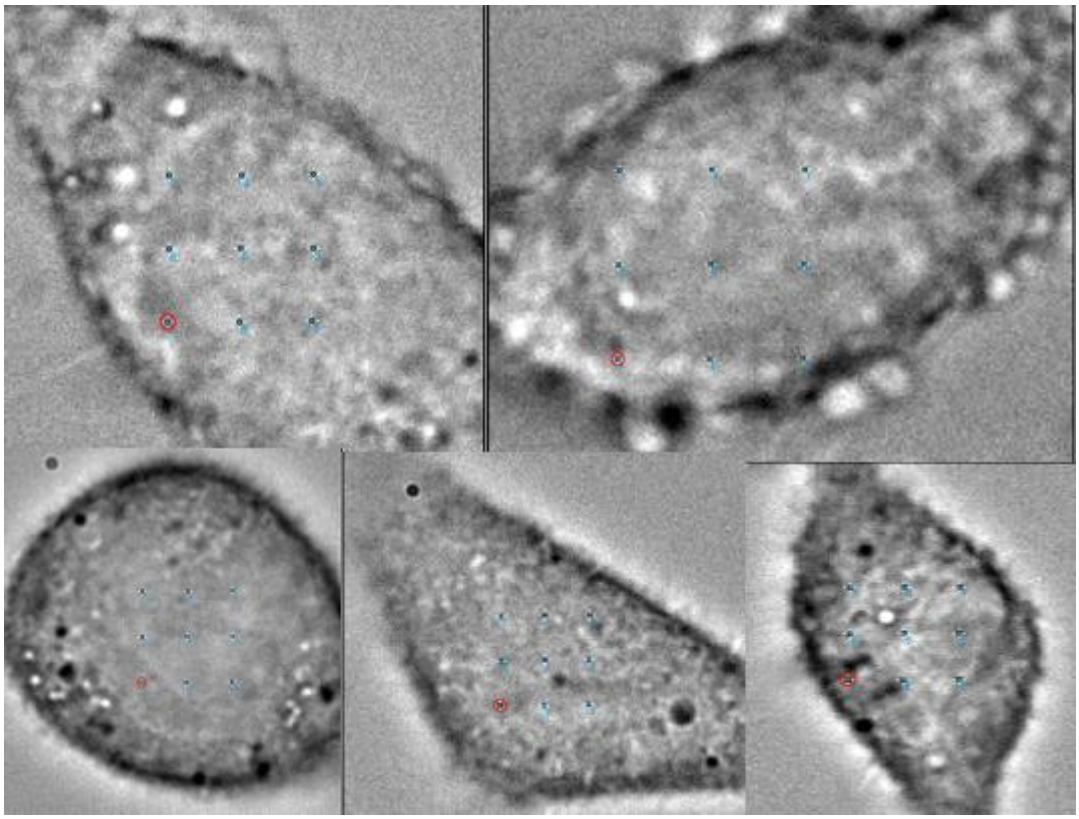

Figure 19. Optical Trap Layouts in HeLa Cells with Granular Nuclei. *Here are DIC images from five HeLa cells showing granular-looking nuclei, with trap spots marked by blue crosses. The red circle in each panel shows which trap used for deeper signal analysis. These granular nuclei tend to have uneven chromatin and dark spots that made it tricky to place traps, so I had to be careful to avoid nucleoli and dense patches. Again, there is a 3×3 grid (plus a centre trap) in each nucleus, adapting the layout just enough to match each cell's unique shape while keeping coverage even across the nuclear area.*

6. Ramified HeLa cells.

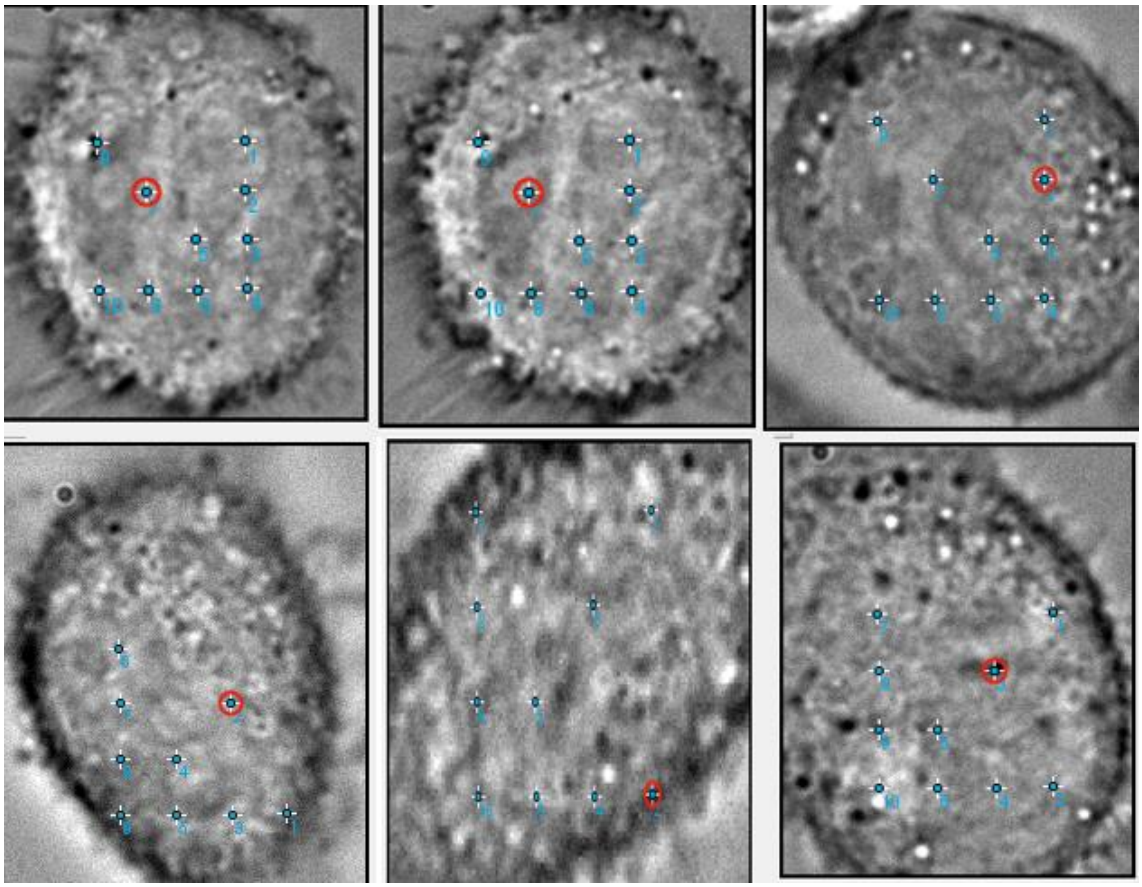

Figure 20. Trap Grids in HeLa Nuclei with Ramified Chromatin. *These DIC images show six HeLa cells where the chromatin takes on a ramified, branch-like pattern. Blue crosses mark where to place the optical traps, and red circles highlight the one picked for deeper signal analysis in each cell. These types of nuclei have a more spread-out, irregular shape, so the trap grid needed to be adjusted to reach across lobes and branches.*

7. Extra HeLa cells analysed images.

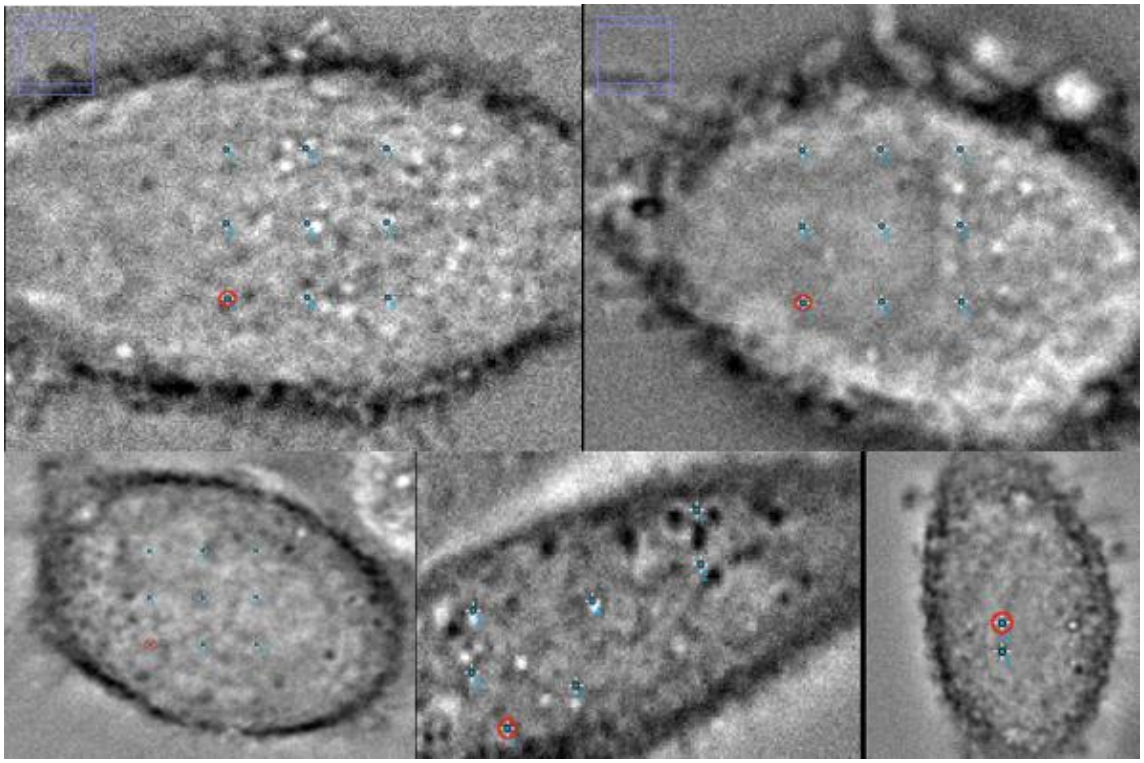

Figure 21. Additional HeLa Cells with Trap Layouts for Extended Analysis. *These DIC images show five more HeLa cells that were part of the extended dataset. Blue crosses mark the optical trap positions, and in each cell, the red circle points to the trap used for analysing force trajectories. These cells didn't fit neatly into the main categories like normal, granular, protein-rich, or ramified, but they still show interesting nuclear shapes or in-between types.*

8. Correlation matrix function.

```
import numpy as np

trampa_array = np.array(trampa)
```

```
ekf_signals = trampa_array[:, :, 2]

import pandas as pd
df = pd.DataFrame(ekf_signals.T, columns=[f"Trap {i+1}" for i in range(ekf_signals.shape[0])])

import seaborn as sns
import matplotlib.pyplot as plt

corr = df.corr(method='pearson')

plt.figure(figsize=(8, 6))
sns.heatmap(corr, annot=True, cmap='coolwarm', vmin=-1, vmax=1)
plt.title('Correlation Matrix of EKF Signals Between Optical Traps in HelA cell 1')
plt.tight_layout()
plt.show()
```

***Figure 22. Python Script for Correlation Mapping of EKF-Filtered Trap Signals**. This script builds and plots a Pearson correlation matrix using EKF-filtered force signals from several optical traps in a HeLa cell. The full dataset lives in a 3D array called trampa_array, and here out the X-axis signals were pulled and then were cleaned up by the EKF. They were included in a Pandas DataFrame, label each trap, and then use Seaborn to plot a heatmap of the correlations. It's a quick way to spot traps that might be moving together hinting at shared chromatin dynamics or mechanical coupling across different parts of the nucleus.*