Dissertation

submitted to the

Combined Faculty of Mathematics, Engineering and Natural Sciences

of Heidelberg University, Germany

for the degree of

Doctor of Natural Sciences

Put forward by

Jenna Elliott

born in: High Wycombe, UK

Oral examination: 2nd February 2026

Selective information transmission by particle distributions

Referees: Dr. Anna Erzberger

Prof. Dr. Ulrich Schwarz

Zusammenfassung

Lebende Zelle reagieren auf verrauschte Signale aus ihrer Umgebung. Sie eignen sich daher besonders gut, die physikalischen Prinzipien der Signalverarbeitung in weicher Materie zu untersuchen. Räumliche Heterogenität von Signalen ist ein dabei noch wenig erforschter Bestandteil zellulärer Regulierung. Motiviert durch die responsiven Eigenschaften biologischer Membranen untersuche ich, wie die räumliche Verteilung von oberflächengebundenen Teilchen eine physikalische Form von Informationsverarbeitung ermöglicht.

Aufbauend auf dem Prinzip der Pfad-Entropie Maximierung, leite ich her, wie mikroskopische Nebenbedingungen die Nichtgleichgewichtsdynamik von Teilchen beeinflussen. Ich untersuche sodann die räumliche Verteilung von Teilchen, die mit heterogenen Strukturen in der Umgebung interagieren. Teilcheninteraktionen, die zu einer nichtlineare Abhängigkeit der Teilchendichte führen, ermöglichen hierbei einen Mechanismus für Informationsübertragung. Ich charakterisiere, wie hieraus eine Form der Mustererkennung entsteht, bei der die Wechselwirkungen zwischen den Teilchen die Effizienz der Kodierung von Information beeinflussen, die für nachgelagerte Aufgaben relevant ist. Ich identifiziere Parameterbereiche, in denen die Teilchenverteilung als Schwellenwertfilter oder Kantendetektor fungiert, und quantifiziere die Informationsübertragung. Interessanterweise fallen viele biophysikalische Systeme in Parameterbereiche mit optimalen Übertragungseigenschaften – darunter die Verteilung von Kernporenkomplexen in Sphaeroforma arctica, die Anzeichen von teilchenvermittelter Schwellenwertbildung aufweisen. Weiterhin zeige ich, dass Rückkopplungen von Teilchenverteilungen auf die Wechselwirkungsenergien die selektive Informationsübertragung verbessern.

Meine Ergebnisse deuten darauf hin, dass die räumliche Verteilung makromolekularer Komplexe in biologischen Systemen eine Rolle bei der selektiven Informationsverarbeitung von Umwelteinflüssen spielen. Die zugrundeliegenden Prinzipien zeigen, wie sich mithilfe physikalischer Wechselwirkungen in weicher Materie Eigenschaften wie Mustererkennung in nicht-lebenden Systemen realisieren lassen.

Abstract

Living cells respond to environmental cues under noisy conditions, making them ideal platforms for uncovering physical principles of computation in soft materials. Signal processing by spatial heterogeneities remains a yet underexplored component of such cellular regulation. Motivated by the relaying of signals across membranes, I investigate information transmission by surface-bound particle distributions.

Extending maximum caliber methods, I address how microscopic constraints influence non-equilibrium particle dynamics, and identify a mechanism for information transmission arising from non-linear responses of equilibrium particle densities to spatial features in adjacent structures. This permits pattern recognition, with inter-particle interactions tuning the response function of noisy signal filters and resulting in the efficient encoding of information relevant to downstream tasks. I identify thresholding and edge-detecting regimes, and quantify how biophysical membrane properties affect information transmission by thresholding filters. I discover parameter regimes with optimal transmission where many biophysical systems fall – including nuclear pore complex distributions in Sphaeroforma arctica which show signatures of particle-mediated thresholding. Feedback from particle distributions to interaction energies is shown to improve selective information transmission.

These results indicate that spatial distributions of macromolecular complexes can selectively sense environmental cues, with fundamental implications for how physical interactions may encode computational logic in soft materials.

Contents

List of Figures

List of Tables

1 Introduction

Living matter has many material properties that set it apart from its non-living counterparts. Achieving such properties, including adaptability, intelligence, and responsiveness, is a prominent aim in the design of soft robots and smart materials [1–3]. Utilising inherent non-linearities in physical material properties, a collection of biology-inspired soft materials have been developed that can adapt and compute responses to changing environmental features [4, 5]. In particular, there has been investment into designing responsive nano and micrometer sized synthetic materials, for applications ranging from environmental remediation to drug delivery [6]. At these scales, typical electronics-based approaches for control and sensing are often unsuitable, requiring novel signal processing methods to achieve desired computational functions and to permit reliable signal transmission in the presence of noise [2, 7].

Single living cells have evolved to successfully respond to stimuli in noisy environments, and as such have become a source of inspiration for the design of these small-scale technologies [6, 8, 9]. Traditionally, signal processing in biology has been addressed on the level of genetic and biochemical signalling, whose implementation in bottom-up studies of synthetic life has already achieved many successes [6, 10, 11]. The quantification of information flow achieved in such signalling processes has helped uncover several principles behind the robustness and complexity of living matter, such as how chemical signals are able to reliably template the body plan of developing embryos [12, 13]. These information-theoretic approaches frequently neglect the important regulatory role of spatial organisation in biological signal processing, instead treating the inner cell as homogeneous. However, spatial variations in sub-cellular organisation are known to store mechanical memories; control cell migration; sense global cellular properties such as geometry; and compute responses to external stimuli [13–16].

Motivated by the role of biological membranes in forming a regulatory barrier surrounding otherwise isolated compartments, in this thesis I investigate the role interface-associated particle distributions play in the transmission of spatially resolved signals. Starting from a statistical description of particle dynamics (Chapter 2), I demonstrate that such distributions act as signal filters that selectively transmit information across these interfaces (Chapter 3), and detail how noise contributes to selective signal compression by quantifying the transmitted information (Chapter 4 and 5). I close the thesis by demonstrating that such filters

can accurately describe the distribution of macromolecular complexes within an example biological system (Chapter 6).

To provide context for these investigations, in the following section I first introduce the concept of signal filters in the context of traditional information technologies, before discussing the limitations of such technologies and the current biologically-inspired approaches aimed at addressing them. This is followed with a discussion of how information processing is achieved in living systems, introducing techniques for quantifying information. Focusing on subcellular scales, I then discuss the role of hierarchical protein patterning in processing sub-cellular spatial signals using two examples from my review article on pattern formation [17]. Motivated by the importance of particle interactions in forming these patterns, I also provide some examples of intra-particle interaction potentials. Finally, I describe current methods for modelling the stochastic dynamics that drive such processes, in particular maximum caliber methods which have been proposed as a non-equilibrium equivalent to equilibrium ensemble theories. I conclude this introduction, in Section 1.3, with an outline of the thesis aims and how I address them.

1.1 Signal processing in physical systems

The period spanning the 20th and 21st centuries has been heralded as an information-processing age, with household computers and mobile phones now performing computations at a scale unimaginable only 100 years ago. Progresses in silicon chip design, robotics, data storage, and communication theory have driven the development of rapid electronic communication technologies and computational tools for applications ranging from medical interventions to space travel. However, in many applications computational technologies are reaching the limits of what traditional, silicone-chip-based information processing methods are capable of achieving [4, 18]. Current work is pushing the frontier of these technologies away from such 'hard-matter' methodologies, developing soft materials capable of transmitting, storing, and computing signals using their innate physical properties to overcome traditional information processing limitations [1, 19–21]. These approaches require theoretical tools for describing and quantifying the performance of soft computational materials in processing information, collating signals from various sources, and using them to compute responses and outcomes.

1.1.1 Signal filters

Typically, the process of transforming an input signal into an output – called *signal processing* – is directional [22]. An input signal, or collection of signals, provides an initial source of information which 'flows' through various processing steps to produce the corresponding output. In analogy to the flow of water down a waterfall, such directional processes are sometimes called signal processing *cascades* [22, 23]. Although these cascades frequently entail complex networks of feedback loops, they are commonly broken into individual signal processing blocks which each perform some operation on the signal they receive [24]. The object or function comprising an individual processing step is called a *signal filter*, since it frequently enacts some selection on the input [25].

Each signal filter performs an operation on the signal it receives. The mathematical form of this operation is called the filter's *response function*, given as an input-output relation [25]. The form of these response functions determine the performance and application of the filter, and much of control theory is dedicated to designing the correct topology of signal processing networks to enact desired computational logic on signals [26]. Filters have been, for example, designed to aid in noise reduction, feature selection from images, and for data compression [25, 27, 28]. The successes of neural networks, such as in applications to computer vision, are also attributed to the learning of suitable signal filters [29].

One particularly interesting and useful response function is a sigmoidal function, providing a non-linear mapping that amplifies changes in input signals close to, and suppressing variations far from, a threshold value. Despite their apparent simplicity, such filters have proven invaluable in a variety of contexts, such as in band-pass filters; in computer vision; as non-linear activation functions in early neural networks; and in ligand-binding responses described by Hill functions [27, 30–33]. Taken to the extreme, these filters describe a discrete binarisation of the input signal. Binarising filters like these play a key role in converting analogue signals to digital, and are also heavily used in image analysis or manipulation in the form of thresholding filters [34].

1.1.2 Physical computation

Despite phenomenal advances in signal processing, current 'hard-matter' methods of computation have several undesirable features [1, 35, 36], including having centralised hubs of computation and suffering from poor resistance to extreme conditions. Furthermore, it can be especially energy and time intensive to compute complex, non-linear response

functions [37]. In contrast, many physical material properties naturally respond non-linearly to external perturbations, and offer the opportunity to enact non-localised computations. For example, disordered packings of grains, like sand, respond non-linearly to shear forces [38] and curvature-inducing proteins give rise to non-linear properties in biological membranes [39–41]. Non-linear elastic responses to perturbations are also being designed in articulated materials, such as chain-mail inspired jamming materials, for applications as smart, reconfigurable fabrics [42, 43].

It is therefore unsurprising that interest is growing in addressing the limitations of current computing methods using the innate physics of materials to perform signal processing. Examples of this include the use of the natural non-linear nature of physical systems to map inputs to a high-dimensional space of variables in reservoir computing [37, 44]; the use of inherent directionality of LED activation in optoelectronic neural networks [45]; and physical learning materials with tunable internal degrees of freedom [46]. Other approaches to develop non-centralised, energy efficient computations consider deformable 'soft matter' materials, where signal processing is achieved throughout the system [47], such as the design of complex materials with periodic internal structures – metamaterials – whose physical properties can be designed to perform certain tasks and have been shown to provide computational logic for a wide range of signal processing applications [19, 48–52].

One of the major appeals of these soft-matter approaches to signal processing is that they are expected to interface well with biological tissues and thus prove useful in medical applications [1]. Substantial effort is therefore being placed in designing signal processing materials that operate on a smaller scale [53, 54]. This presents a different set of challenges, in particular in overcoming system noise [55, 56]. Evolution appears to have solved these problems for living systems, where computations are reliably and robustly carried out at nano and micrometer scales [57]. Many research efforts therefore draw inspiration from how biological systems sense and respond to environmental signals [35, 58]. By uncovering how living systems achieve adaptive, responsive, and learning behaviours, it is hoped that synthetic materials may be constructed that exhibit these life-like properties [1, 2, 4]. In turn, such synthetic systems may provide a tractable means to test theories of how living systems process and respond to signals within the noisy constraints of their environment.

1.1.3 Biological computation

It is a defining property of living matter that it can sense and respond to conditions in its environment [59, 60], be it a plant changing its direction of

growth in response to lighting conditions, a bacteria swimming towards a food source, or indeed a human reading a textbook. Information stored in the environment – such as light, chemical concentrations, or words on a page – is received through one of various sensing mechanisms, processed, and used to compute an outcome. This permits the organism to adapt, respond, and learn. Uncovering the underlying principles that govern these behaviours is an active and fruitful area of research, with efforts typically focused on understanding the response functions involved in drug responses and cellular signalling [22, 61–63], search strategies of single cellular organisms [64], mechanisms for robust cell fate decisions in organism development [12, 13, 62, 65–67], and signal processing in neural systems [68]. Recent work has also addressed how information processing is achieved across scales, such as how signalling between cells can inform the formation of patterns in tissues [69–71]. Successes in these research directions have also lead to progress in synthetic signal processing, inspiring neural network architectures and methods for local contrastive learning in physical learning systems [33, 72, 73].

One of the great challenges facing signal processing in these systems is how the transmission of signals competes with innate system noise and variability [74–77]. Thermal and active fluctuations in sensors and information processing machinery act as a 'channel noise' garbling the input signal as it passes through the communication channel formed by the filter. The signal processing filters involved in life-like response behaviours exist in such noisy systems, and yet organisms must be able to reliably respond to stimuli. For decision making to be robust to fluctuations and random system variations, sufficient information must be transmitted through the signal processing cascade [13, 74, 78]. When studying information processing in such systems it is therefore important to quantify the capacity of the signal filters to transmit information [79, 80].

Although the notion of information stored in a signal is in many ways intuitive, it was not until 1948 that its treatment as a measurable quantity with an exact mathematical definition was first achieved by Shannon [81]. This treatment was based on the fact that, in real life scenarios, a signal will fluctuate due to stochastic processes. As such, signals can be treated as random variables, with any a particular realisation of the signal being sampled from the signal's probability distribution. Considering information as a measure of the uncertainty about a signal X – for example a chemical concentration – distributed with the probability density $p_X(x)$, Shannon defined information as the entropy $H(X) = -\int dx p_X(x) \ln(p_X(x))$ [82, Chapter 2]. This describes on average how much information a measurement, or 'realisation', x of the variable X would provide about its distribution.

In the context of the transmission of information during signal processing, or *information processing*, it is informative to ask how much information originally available in an input signal X will be mapped to the output signal Y of a filter. This question could alternatively be posed as how much information is gained about a signal X by taking a measurement of Y, or how much is the uncertainty about the signal X reduced by this measurement. Defined similarly to the information $H(X)$ introduced above, the Shannon entropy of X after a measurement of Y, $H(X|Y) = -\iint dxdy p_X(x|y)p_Y(y)\ln(p_X(x|y))$, is dependent on the conditional probability of a realisation x given a measurement of y, $p_X(x|y)$, and the marginal distribution $p_Y(y)$ of signal Y. The transmitted information is given by the difference between the information about the input X, i.e. $H(X)$, and the information about X given a measurement of Y, i.e. $H(X|Y)$, and called the *mutual information*, $I(X;Y) = H(X) - H(X|Y)$. The mutual information is more commonly written in an equivalent form as the Kullback-Leibler divergence [82, Chapter 2],

$$I(X;Y) = \int dxdy p_{X,Y}(x,y)\ln\left(\frac{p_{X,Y}(x,y)}{p_X(x)p_X(y)}\right), \tag{1.1}$$

where $p_{X,Y}(x,y) = p_X(x|y)p_X(x)$ is the joint probability distribution of the random variables x and y, and $p_X(x)$ and $p_Y(y) = \int dx p_{X,Y}(x,y)$ are the marginal distributions for x and y respectively. Since Equation (1.1) is symmetrical in x and y, the direction of information flow is irrelevant and the mutual information can be interpreted as the amount of information that is *shared* between the signals x and y, i.e. $I(X;Y) = I(Y;X)$. Information and entropy calculations are typically done in the discrete limit [83], in which case the mutual information becomes,

$$I(X;Y) = \sum_{x,y} P_{X,Y}(x,y)\ln\left(\frac{P_{X,Y}(x,y)}{P_X(x)P_Y(y)}\right). \tag{1.2}$$

In the limit where the filters are loss-less, a realisation y is directly mappable to a realisation of x such that $p_{X,Y}(x|y)$ is a delta function and the transmitted information is equal to the entropy of the input signal, $I(X;Y) = H(X)$. Equations (1.1) and (1.2) give the mutual information in units of 'Nats' (using a natural logarithm base of e), however, typically information is measured in 'Bits' (using a logarithm of base 2). One bit is the information required to distinguish between binary, equiprobable outcomes [13].

Mutual information quantifications have been used to assess the robustness of signal processing in a variety of biological systems [12, 59, 67, 77, 84]. Biological regulation at the scale of both genes and biochemical signals have been addressed from the viewpoint of information transmission, leading to productive theories of biological computation, as well

as estimates on the maximum number of inputs that a cell has sufficient information to distinguish [62, 75, 80, 85–88]. A particularly exciting example of this regards body plan templating during early fly embryo development, where stem cells make a commitment to a particular cell fate in a process called cellular decision making. This decision is informed by the concentration of different proteins, called morphogens, which vary in concentration along the embryo tissue. These morphogen gradients form a set of input signals that encode the position along the embryo body axis. Individual cells can readout the local concentration of these morphogen signals and use this to infer their position and thus determine their position-dependent fate. Calculations of the mutual information between these signal carrying morphogen concentrations and cell position, i.e. 'positional information', have been used to investigate how many noisy input signals are necessary to communicate the information required for reliable cellular decision making [13, 59]. It is important to note, however, that although the morphogen gradients are resolved in space, the information communication is local, meaning that an individual cell measures only a scalar value of each morphogen concentration. In addition to other developmental processes [89], similar information theoretic principles have also been applied in understanding immune memory and responses [90–92], as well as to develop models of evolutionary processes [93, 94]

Although the directionality of information flow is irrelevant for mutual information calculations applied to a single signal filter, it becomes important when considering information flow through a cascade. In particular, a two-step signal processing cascade involving signals X, Y, and Z can be denoted as a Markov chain, $X \rightarrow Y \rightarrow Z$, if the conditional distribution of Z depends only on Y and is independent of X [82, Chapter 2]. This is the case for two filters in series, where an initial signal is first transformed by a filter to produce Y, which is subsequently read out by a second filter to produce Z. For this cascade, the information that Z shares with X must be less than or equal to the information X shares with Y. This is a statement of the *data processing inequality*,

$$I(X;Z) \leq I(X;Y), \tag{1.3}$$

where equality occurs when the second filter is loss-less. Similarly, the amount of information Z has about X must be less than or equal to the amount it has about Y, i.e.

$$I(Z;X) \leq I(Z;Y). \tag{1.4}$$

Together these inequalities provide the limits of information transmission in a system.

Frequently signals contain details that are irrelevant to later processes within an information cascade. In these cases, instead of being a limitation, information loss can be beneficial in reducing the complexity of a signal. Compressive filters are integral to the function of, for example, radio communication, data storage, and photo sharing through social media, where compression is used to select important information and reduce signal complexity [95–99]. This feature may also be relevant in biological processes to discard redundant signal features to result in lower-complexity outputs, such as in binary cellular decision making [100]. Both the form of a filter's response function, and any channel noise in the filter mechanism, can lead to this controlled information loss and thus determine which features of a signal are amplified, and which are suppressed or removed.

Consider a compressive information processing cascade defined by the Markov chain $X \to Y \to Z$, where X is a signal relevant to some downstream task, Y is a complex signal containing features of this relevant signal, and Z is the final compressed output. A good signal compression would maximise the information shared between X and Z, i.e. $I(X;Z)$, while minimising the information shared between Z and Y, i.e. $I(Z;Y)$ [101]. In essence this means that the cascade would perform as close to the equality limit specified by both Equation (1.3) and Equation (1.4) as possible. For a noisy encoding of X into Y it is impossible for the compression to reach this equality. Instead the best performing compression has a less efficient limit, which may be calculated using the *information bottleneck* framework introduced by Tishby et al. in 2000 [101]. This framework aims to find an optimal compression algorithm and allows an assessment of compressive filter performance against this optimal. This has been applied – along with follow-up adjusted theories – with great success in a variety of systems [24, 102–105]. In particular, this framework has been suggested to explain some of the principles behind learning in deep neural networks and used to motivate neural network architectures for computer vision [106–109].

In a biological setting, the information bottleneck has been used to suggest that neuronal systems perform near-optimal compression [110], and also to infer optimal decoding schemes for the position-dependent cell fate decisions in the early fly embryo [83]. In the later of these, the cell fate position was treated as the relevant signal X, noisily encoded within the morphogen signals Y, and read out to produce a cell fate decision Z. The information bottleneck framework permitted the identification of optimal signal filters for the readout process $Y \to Z$ and suggested that such an optimal readout could be achieved through multiple noiseless binary readouts [16, 76], supporting other work indicating how multiple signals can be combined to combat information loss due to noise [80]. This work

also demonstrated that thresholded readouts of each morphogen, when used cooperatively across the different morphogen signals, could be sufficient for cell fate patterning, echoing the sentiment in Section 1.1.1 that such filters, although simple, can be useful in information processing cascades.

1.1.4 Processing of spatial signals in living matter

Spatial organisation plays an important role in sub-cellular signal processing. Indeed, subcellular physical properties and organisation are thought to be integral to how slime moulds compute [111–113]; pathogens are recognised by immune cells [114–117]; and migrating cells coordinate their motion [14, 15].

Of particular interest is the formation of inhomogeneous protein distributions [17], which have been shown to facilitate the sensing of system properties, such as boundary geometry, to inform cellular decisions [118, 119]. For example, surface contraction waves arising from mechanochemical couplings act as geometry sensors in starfish oocytes [120]. In these oocytes, information on the angle and distance of the cell membrane from the nucleus is encoded within a temporally decaying cytosolic protein concentration gradient and read out in a two-stage process [120]. In the first stage the initial protein gradient templates a second, sigmoidal protein distribution, compressing the input into an approximately binary signal demarcating surface regions above and below a threshold input value. As the initial signal decays, the boundary between these two regions sweeps along the cell surface at a geometry-dependent velocity, and the location of the boundary is read out via a local surface contraction that occurs as a membrane region is swapped between the binary regions. This second stage results in a contractile wave whose speed is modulated by the surface geometry such that it will always converge at the membrane location closest to the nucleus [121–123]. From the perspective of signal processing, this process can be viewed as a two-stage signal processing cascade. These cascades are termed *hierarchical patterning* processes, where one protein distribution templates another, which subsequentially acts as a guiding cue for further downstream patterning processes [124].

Hierarchical patterning can be used to both sense a signal and to compute a necessary response. For example, *E. coli* cells use a three-stage hierarchical mechanism to identify their midpoint ahead of cellular division [125–128]. Specifically, an initial geometry dependent process (the 'MinD' oscillator) gives rise to standing wave oscillations in protein concentrations at the cell membrane [129–133]. These directly template the distribution of a second membrane associated protein, which acts

as an inhibitor to the binding of a third, contraction-mediating protein. Eventually this third protein forms a contractile ring at the minimum of the time-averaged inhibitor distribution. This demonstrates explicitly how hierarchical patterning permits subcellular computation through particle distributions which select information from an input (in this case the cell geometry) and compute the desired output (in this case the location of the mid-cell plane).

Such sequential information processing systems are integral to biological function [65, 134], and there is a growing interest in quantifying the information stored in such particle distributions [66, 87, 135]. However, these approaches have so far addressed either single-valued time-varying signals [87], or the information content of self-organised patterns [66, 135]. In contrast, the information transmission through the hierarchical templating of protein distributions is under-researched. Quantifications of such information flows may elucidate principles of pattern-based computation and aid, for example, in the design of biocompatible nanomotors or programmable vesicles for drug delivery by providing a clearer understanding of how biological systems process spatial signals, and how best to interface with them [6].

Comprising interacting particle species, the above signal filters are examples of reaction diffusion systems, where spatial organisation arises due to feedback between system components. The 'MinD' oscillator, in particular, is a *mass-conserving* reaction diffusion system, which have been shown to lead to complex patterning dynamics [136–138]. Although originally described for biochemical reactions, theories describing the emergence of spatial organisation have been extended to consider a wide variety of feedback mechanisms, such as mechanical or mechanochemical couplings [17]. Indeed diverse self-organising patterns are observed across biological scales, from tissue-scale patterning of whole organisms [139] to symmetry breaking within single cells [126]. Both positive and negative feedback processes contribute to the emergence of such distributions. Positive feedback leads to the amplification of small perturbations, whereas negative feedback suppressing their unbounded growth and, when coupled to time delays, the emergence of oscillations and waves. Feedback is also known to play important roles in system control and in ensuring processes are robust to perturbations and thermal fluctuations, however it also introduces further system complexities in which noise can develop [140–142]. It remains unclear how the balance between signal robustness and additional noise constrains impacts the flow of information in such systems.

Traditionally, investigations into the role of spatial inhomogeneities in biological signal processing have been focused on large patterns at the scale of tissues, or easily visible subcellular protein distributions. With

recent improvements to high-resolution imaging, it is now possible to observe small-scale structures and patterns, opening up the possibility of understanding how spatial structures may store information and inform cellular decisions [143, 144]. For example, high resolution imaging techniques such as expansion microscopy imaging permit the observation of interesting and complex distributions of particles within subcellular membranes, such as nuclear pore complexes in the nuclear envelope [145]. Such observations provide an opportunity to investigate the role of protein organisation in the processing of spatially resolved signals in subcellular systems. Membrane-associated particle distributions, in particular, provide a basis for such investigations. In living systems, membranes form a barrier between the compartments they enclose and their environment, posing a challenge for inter-compartment communication. The coordination of compartment processes in response to changes in its surroundings requires the relaying of information across the membrane interface. Such a set-up aids in defining a direction of signalling cascades, and as such provides a starting point for modelling the information flow in hierarchical patterning processes.

1.1.5 Particle interactions

The response of a particle distribution to a guiding signal is controlled by interactions between the system's components. In reaction diffusion systems these interactions typically take the form of chemical reactions in which one biochemical species is converted to another. However, physical interactions between particles may also tune the formation of spatially-varying distributions. To characterise the ability of a distribution to process signals, it is therefore important to understand the relations between these physical interactions and the distributions they produce.

The most common form of particle-particle interaction is a short-range repulsive interaction, that accounts for the volume of the particle from which other particles are excluded. This is typically modelled as a hard-core repulsion, in which the interaction energy is infinite for particle separations shorter than the particle diameter, and zero elsewhere [146, Chapter 5]. Other short-range repulsive interactions can lead to effective particle sizes greater than is true diameter. For example, proteins embedded within biological membranes can induce a curvature into the surface [40]. Due to the resistance of the membrane to such curvature, two embedded particles at a separation distance d from each other are either attracted or repelled in order to best relax the curved membrane [147]. Typically, small induced curvatures lead to repulsion of other similar

particles, with an interaction potential described by [147]

$$u_{\text{curvature}}(d_{mk}) = 8\pi\kappa\theta^2\left(\frac{d_0}{2d}\right)^4, \tag{1.5}$$

where κ is the membrane bending rigidity, θ is the angle of the membrane incident to the embedded particle, d_0 is the true protein diameter. Such curvature-mediated interactions determine cell shape and facilitate geometry sensing in migrating cells, which has inspired approaches to control the movement of synthetic vesicles [148–150]. Alternatively, in some scenarios, charged particles may interact through electrostatic interactions. In a dielectric buffer, this results in a shielded electrostatic interaction potential, called the Yukawa potential [151, 152],

$$u_{\text{electrostatic}}(d) = \frac{Q^2}{4\pi\epsilon_0}\frac{e^{-d/\lambda_d}}{d}, \tag{1.6}$$

where Q is the charge per particle, λ_d is the Debye length, and ϵ_0 is the vacuum permittivity.

Extending beyond simple repulsive interactions opens the possibility of more complex proteins patterns. Indeed, particle systems with attractive interactions are known to phase separate [153]. For example, a mean-field description of particles that undergo long-range attraction, along with short-range volume exclusion, is the Van der Waals gas. In this description, particle interaction potentials are assumed to be attractive for particle pairs separated by more than a threshold distance, but infinite for separations lower than this critical value. In the mean-field limit with particle density ρ, the Helmholtz free energy density of this particle gas is given by [146, Chapter 5.5]

$$\beta f_{\text{VdW}} = \rho\left[\ln\left(\frac{C\rho}{1-\rho/\rho_{\max}}\right) - 1\right] - \beta a\rho^2, \tag{1.7}$$

where $\rho_{\max}$ is the maximum density that can be achieved by the particles, C is a constant arising from the discretisation of phase space, and a gives the strength of the attractive interaction. Throughout this thesis, I denote the inverse thermal (or effective thermal) energy scale by $\beta = 1/k_BT$.

Phase separation is known to contribute to signal processing in subcellular systems. Many works have indicated the role such processes can play in noise reduction and rapid sensing [154–156]. Furthermore, the spatially-variant nature of phase separating particle systems plays a key role in cellular regulation. In particular, phase separation is thought to influence nuclear processes by controlling transcription [157–159], identifying DNA double-stranded breaks [160], and identifying genetic crossover points during meiosis [161]. Outside of biological systems, such guided phase separating systems have also been implemented in the construction of

patterned materials [162–164], and in the design of controllable synthetic vesicles [165].

In summary, despite a wealth of evidence showing the importance of guided particle distributions in subcellular signalling, the effectiveness of these distributions to select and transmit information has been under-explored. Uncovering the capabilities and limitations of this signal processing modality, and how particle interactions tune such processes may extend the repertoire of biocompatible signalling modalities by which small-scale soft robots maybe be programmed, controlled, and interfaced with living matter.

1.2 Theoretical frameworks for particle dynamics

To understand how various signal-processing particle distributions occur and function, it is important to understand how particle interactions influence the resultant distributions. In the following, I outline the theoretical frameworks for particle dynamics and, in particular, how interactions influence the fluctuations or noise in these particle distributions.

1.2.1 Stochastic dynamics

One particularly important theory for modelling particle dynamics is *stochastic thermodynamics* – the study of non-equilibrium, mesoscopic systems that connects to macroscopic thermodynamics [166]. In this framework the evolution of a probability distribution $P(X)$ of system states X is described using stochastic jump rates $k_{X\to X'}$ which give the Poissonian rate of the system transitioning from state X to state X'. Typically such dynamics are written in the form of a master equation [166],

$$\frac{dP(X)}{dt} = \sum_{X'\neq X} [k_{X'\to X}P(X') - k_{X\to X'}P(X)], \tag{1.8}$$

where the mesostates X can represent any distinguishing mesoscopic property of the microscopic system, for example the set of occupation numbers n_i of different boxes in a discrete-lattice description. Equation (1.8) describes how the uncertainty of an initial system configuration, as determined by the distribution $P(X)$, is propagated via the exact rates $k_{X\to X'}$. As such, this provides a model for how noise propagates in time in stochastic systems, providing an important tool in understanding the limitations of information transmission using particle distributions.

Local detailed balance

In thermodynamically consistent theories, the jump rates are known to obey the *generalized detailed balance* – or *local detailed balance* – relation [166],

$$\frac{k_{X\to X'}}{k_{X'\to X}} = e^{\beta q_{X\to X'}}, \tag{1.9}$$

where $q_{X\to X'}$ is the amount of energy released to the environment due to the state transition. In the absence of external forces, this energy released is simply the difference in energy between the two states $q_{X\to X'} = E(X') - E(X)$. This relation allows stochastic thermodynamics to relate the system dynamics to the microscopic features of a physical system of interest. The justification of the local detailed balance relation stems from comparisons to equilibrium systems, were the rates can be shown to obey detailed balance through equilibrium ensemble approaches. Local detailed balance is then generalized to non-equilibrium systems through time-scale separation arguments [166].

Fluctuation theorems

On the scale of full trajectories, additional relations – called *fluctuation theorems* – are attainable [167]. These relations constrain fluctuations in the system according to thermodynamic quantities. One particularly famous form of these relations is the Crook's fluctuation theorem [168]. This states that, for out of equilibrium systems, the probability ratio of forward P_F and correspond time-reversed P_R transitions is equal to the exponential of the entropy production [166, Chapter 4][168],

$$\frac{P_F}{P_R} = e^{\frac{\Delta_F S}{k_B}}, \tag{1.10}$$

where $\Delta_F S = (W_0 - \Delta F)/T$ is the entropy production caused by the forward transition, and where F is the Helmholtz free energy and W_0 is the work done on the system. Such fluctuation relations are powerful theoretical tools for detailing how system fluctuations are constrained by thermodynamic quantities.

1.2.2 Fokker-Planck dynamics

Another important description of stochastic particle dynamics is given by the Fokker-Planck equation. This ubiquitous equation has been used in a diverse range of applications, from bacteria chemotaxis [169] to quantum electrodynamics [170], to describe how distributions of particles respond to external forces. The general Fokker-Planck equation takes the form [171,

Chapter 1.2.2],

$$\frac{\partial W}{\partial t} = \left[-\sum_{i=1}^{N} \frac{\partial}{\partial x_i} D_i^{(1)}(\{x\}) + \sum_{i,j=1}^{N} \frac{\partial^2}{\partial x_i \partial x_j} D_{ij}^{(2)}(\{x\}) \right] W, \tag{1.11}$$

where W is a probability distribution over the variables $x_i \in \{x\}$, indexed by $i \in 1, .., N$, and where $D_i^{(1)}$ is a drift vector and $D_{ij}^{(2)}$ is a diffusion tensor. The drift term accounts for the deterministic movement of particles, whereas the diffusion term accounts for fluctuations. These coefficients should be derived from microscopic equations however, in cases where this is intractable, they are commonly obtained using heuristic arguments [171]. Typically for continuous systems this is achieved using a Langevin or Markov formalism, using a master-equation similar to Equation (1.8) [172, Chapters 3, 4, and 13].

In many applications, where fluctuations in the particle distribution are irrelevant, it is sufficient to address particle dynamics in the mean-field approximation. In this case the Fokker-Planck equation is often combined with an additional term $\mathscr{R}_k$ accounting for reactions that add or remove particles from particle species k, resulting in the dynamical *reaction-diffusion equation* for particle density ρ_k of the form

$$\frac{\partial \rho_k(\mathbf{r})}{\partial t} = \nabla \cdot \left[D(\mathbf{r}) \beta \rho(\mathbf{r}) \mathbf{f}(\mathbf{r}) + \nabla (D(\mathbf{r}) \rho(\mathbf{r})) \right] + \mathscr{R}_k \tag{1.12}$$

where $\mathbf{f}(\mathbf{r}) = \nabla U$ is the external force applied to each particle by a potential U, and $\mathbf{r}$ is a position vector. In this case, it is typical to ignore cross-diffusion terms, which are frequently negligible compared to the rates of the reactions and self-diffusion.

There are several 'special cases' of the Fokker-Planck equation. One of which is the Smoluchowski equation, which was originally derived for particles with spatially invariant diffusion coefficients. In this equation, the diffusion coefficient is to the left of the gradient operation [171, Chapter 1.2.6]. However, as shown in [173], in the case of spatially varying diffusion coefficients, Equation (1.12) may also be written in Smoluchowski form by defining an adjusted force $\tilde{\mathbf{f}}(\mathbf{r}) = \mathbf{f}(\mathbf{r}) - k_B T \nabla [\ln(D(\mathbf{r})/D_0)]$, where D_0 is a reference diffusion constant.

1.2.3 Caliber methods

Although current non-equilibrium methods such as those described above are highly effective and have been successful in a broad range of applications, their derivations are often complex, and frequently arise from phenomenological arguments. Indeed, it has been noted that non-equilibrium methods lack a unifying statistically founded variational

method akin to the equilibrium ensemble methods first founded by Boltzmann and Gibbs [174, 175]. In such equilibrium methods, the probability distribution over an ensemble of possible system states is identified by maximising the system's state-entropy subject to constraints in known system parameters [176]. Since this variational method provides an elegant justification for how well-known phenomenological relations arise from system constraints (such as conservation laws) in equilibrium, there has been a substantial drive to develop an equivalent non-equilibrium variational theory [174, 175]. In particular, maximum path entropy – also called maximum *caliber* – formalisms have been suggested to fill this role [174]. In these approaches, probability distributions of system state *transitions* are identified by maximising the path entropy – a process called the *Principle of Maximum Caliber* – rather than finding the probability distributions of system states, as done in the equilibrium ensemble methods [177, 178].

In brief, maximum caliber methods work as follows. Inspired by the equilibrium methods, the probability of a system transitioning from an initial state to a final state is related to the number of different 'paths' or combinations of particle exchanges a system could take to move between the states. This number is called the *caliber* of the transition [174], and is directly related to the path entropy S_{path} associated with the change in mesoscopic states. The probability distribution of paths is then found as the extremum of the objective function

$$f = S_{\text{path}} + \sum_n \lambda_n B_n . \tag{1.13}$$

where the sum is over various system constraints B_n which are enforced by their corresponding Lagrange multipliers λ_n. This approach offers a means of deriving non-equilibrium dynamics motivated by a dynamical transition ensemble that mirrors the function of state ensembles in equilibrium approaches. The maximum caliber principle is particularly suited to memory-less systems, in which sequences of transitions form Markov chains described by the conditional probability relation [179, Chapter 5.4],

$$p(x_1, t_1; x_2, t_2; ...|y_1, \tau_1; y_2, \tau_2; ...) = p(x_1, t_1; x_2, t_2; ...|y_1, \tau_1), \tag{1.14}$$

where the systems history is denoted by the state variable y_i and time τ_i where $\tau_i > \tau_{i+1}$, and the future is denoted by the state variable x_i and time t_i where $t_i < t_{i+1}$. Indeed, it has been shown that Markov chains arise naturally from the maximum caliber principle [180].

In recent years, research efforts have focused on developing the general maximum caliber method, in particular showing how it recovers well-known non-equilibrium relations such as fluctuation theorems and

reaction diffusion equations [181–190]. In these approaches, the path entropy is typically given by [175]

$$S_{\text{path}} = \sum_{\Gamma} P_{\Gamma} \ln(P_{\Gamma}) \tag{1.15}$$

where P_{Γ} is the probability distribution of the trajectory Γ, and the sum is over all possible trajectory. In comparison, substantially less work has been undertaken connecting calibre approaches to discrete, microscopic toy models, where the path entropy can be directly related to microscopic model parameters. To the best of my knowledge, attempts at this have been limited to non-interacting particles in two-state systems, see [191]. One notable example has been made by Ghosh et al. [192], who implement a path entropy for non-interacting particles jumping between two degenerate particle states using a combinatorics argument. By nature of the path entropy construction, maximum caliber approaches are sensitive to the topology of the particle transition network, as briefly outlined in [193]. However, by limiting the treatments of microscopic toy models to only two particle states, works to date have been unable explore this feature of the dynamics. Similarly, microscopic models offer the opportunity investigate the impact of short-range particle interactions through constraints on arrangement statistics. Indeed, in equilibrium ensemble methods these arrangement statistics lead to, for example, the distinction between Boltzmann and Fermi-Dirac distributions [194, Chapters 4 and 5]. Yet by treating only non-interacting particles, caliber approaches so far have not addressed how such interactions influence particle dynamics.

1.3 Thesis objectives

In this thesis, I address the formation of membrane-associated protein distributions from an information-theoretic perspective, extending theoretical frameworks for biological signal processing to spatially-resolved signals and quantifying the information transfer achieved by such noisy pattering processes.

In Chapter 2, I present a framework for modelling particle dynamics that extends the maximum caliber principle to microscopic particle transition models on discrete lattices, with the aim of providing a clear statistical physics basis on which to address the formation of guided particle distributions. I demonstrate that this approach recovers known features of diffusion in both discrete graphs and continuous physical systems and close the chapter by showing that this method recovers the Fokker-Planck equation for mean-field particle dynamics, explicitly detailing how the

maximum caliber treatment of particle interactions maps to the drift dynamics in Equation (1.12).

Using this mean-field description, in Chapter 3 I address the impact of particle interactions on the formation of membrane-associated protein distributions. Considering particles diffusing within two-dimensional interfaces and interacting with structures in their environment, I derive analytical solutions for equilibrium particle distributions. Treating particle-environment interaction energy fields as input signals, the patterning of these particles is interpreted as a signal filter which permits the selective transmission of the input energy signals into resultant particle density fields. I show how inter-particle interactions give rise to non-linear filter response functions, which may be tuned by biophysical system parameters, and explicitly identify both binarising and edge-detecting regimes.

Accounting for noise in these guided particle patterning processes, in Chapter 4 I quantify the information transmission achieved by the particle-mediated binarising signal filters. Calculating the mutual information between entire input energy and output density fields, I identify an optimal information transmission regime for these filters, and demonstrate through comparisons with system parameter measurements from literature that these regimes are attainable by biological systems. I further find that input signals stored in filamentous, rather than surface-like, structures have optimal regimes with lower particle number requirements. With the aim of addressing how effective these filters are at binary signal compression, I compute the transmitted information relevant to desired binary input compressions and compare the results to the ideal filter limits defined by the data processing limit to find regimes of effective selective information transmission by filters where the tuning of physical system parameters allows selection of information for downstream tasks.

Extending these information theoretic approaches to systems with feedback on the input signal, in Chapter 5 I demonstrate a trade-off between noise arising from additional fluctuating degrees of freedom and improvements in filter response functions. I find that feedback leads to better selectivity of binarised input features but similar overall information transmission.

In Chapter 6, I conclude my investigations by demonstrating the applicability of these signal filters to information processing in living cells, showing that distributions of nuclear pore complexes in the nuclei of single cell protists *Sphearoforma arctica* exhibit signatures of the binarising filters, and that the these systems exist within the optimal information transmission parameter regime.

Overall these findings suggest a mechanism for sub-cellular pattern recognition through sequential noisy particle distributions, and provide a framework for quantifying the transmission of spatial information in sub-cellular systems. As such, these results have the potential to impact how physical interactions are used to encode computational logic in synthetic soft materials, as applicable to bottom-up nanorobot design and the development of synthetic living vesicles.

Statistical physics of interacting particles

2

To understand how the guided formation of particle distributions may be tuned to perform signal processing behaviours, it is important to understand how the dynamics is influenced by particle interactions. In this chapter, I implement a maximum caliber method to derive a non-equilibrium description of particle dynamics under such interactions. As introduced in Chapter 1, maximum caliber theories are variational approaches for calculating probable system trajectories by maximising the path entropy, and as such provide a statistical basis for modelling non-equilibrium processes [174]. These frameworks offer a *dynamical* equivalent to free energy minimization methods for predicting equilibrium states [192], however, in their current form they lack an explicit link to the microscopic structure of discrete modelled systems, and little progress has been made linking the caliber frameworks to microscopic models of interacting particles.

In this chapter, I provide a step towards overcoming these limitations in the form of a general microscopic caliber theory, through which system dynamics can be directly derived from state-transition statistics. Extending the approach of Ghosh et al. [192] to *general interacting particles*, I consider the dynamics of diffusive particles in a discrete limit, accounting for short range interactions though general constraints on particle arrangements. This results in a caliber formalism that highlights the role of network topology in particle dynamics and forms a bridge between graph theoretic models for diffusion on discrete networks and stochastic thermodynamics. Throughout this chapter, I illustrate key parts of the method through the specific example of particles that obey Fermi-Dirac state occupation statistics.

The general microcaliber theory presented in the beginning of Section 2.1 was developed in collaboration with Roman Belousov, who introduced me to caliber methods. I subsequently independently developed the contextualisation of this method to graph diffusion and stochastic mechanics (Sections 2.1.1 and 2.1.2), and the following recovery of mean-field reaction diffusion equations (Section 2.2).

2.1 A general microcaliber theory

I am grateful to Roman Belousov, who collaborated with me in the development of this general microcaliber method. My use of the pronouns 'we/our' indicates this collaboration.

In this section, a discrete, lattice-based model of particle diffusion is introduced, and the principle of maximum caliber is implemented to identify the most likely trajectory the model system will take through state-space, starting from a known initial state. Although caliber methods are typically used to describe the evolution of probability distributions, studying the evolution of an exact system state permits a clear demonstration of how constraints on microscopic systems influence the state trajectory.

Consider a model system in which particles are confined to degenerate particle states i and permitted to jump to states $l \in \mathcal{N}(i)$ within the neighbourhood $\mathcal{N}(i)$ of their original state. The mesoscopic state of the whole system, neglecting the microscopic details of particles within each degenerate state, is then described by the set of occupation numbers $\{n_i\}$ of each particle state.

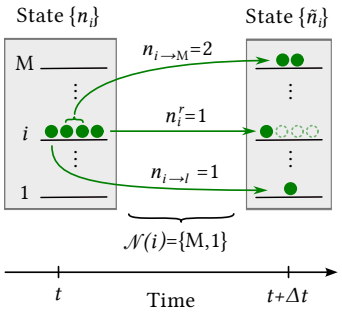


Figure 2.1: A mesoscopic particle system, with particles distributed over M particle states, evolves from state $\{n_i\}$ at time t to state $\{\tilde{n}_i\}$ at time $t + \Delta t$, through the movement of $n_{i\to l}$ particles from particle state i to particle state $l \in \mathcal{N}(i)$. The remaining n_i^{r} particle stay in their initial particle state i.

The time-evolution of these occupation numbers marks out the system's trajectory through configuration space. In the discrete-time limit, this trajectory can be decomposed into individual system state transitions, where the number of particles that move from state i to l in a time-step Δt is given by $n_{i\to l}$, and the number of particles that remain in the same state is denoted n_i^{r}, see Figure 2.1. A transition from one mesoscopic state $\{n_i\}$ to another $\{\tilde{n}_i\}$ is then fully described by the sets of numbers, $\{n_{i\to l}\}$ and $\{n_i^{\mathrm{r}}\}$, spanning all pairs of states and all states respectively.

As introduced in Chapter 1, the probability of a particular transition is related to the number of different 'paths' – or combinations of particle exchanges – a system could take to move from its initial state to its final state. Indeed, in most cases more than one path could produce the same outcome. The caliber of a system is then defined as

Definition 2.1.1 *The 'caliber' of a transition $\Omega(\{n_{i\to l}\}, \{n_i^{\mathrm{r}}\})$ is the number of paths a system may take to transfer $n_{i\to l}$ particles from state i to state l, whilst leaving n_i^{r} particles in state i, for all permitted pairs of states i, l.*

The principle of maximum caliber then states that the most probable trajectory is the one that maximises this caliber [174].

How would one construct a general caliber for this model system? Without specifying the underlying constraints on particle transitions, we first note that the total number of paths, Ω, is composed of the number of ways, $\omega(\{\tilde{n}_i\})$, to arrange distinguishable particles into their final configuration $\{\tilde{n}_i\}$, divided by the relevant Boltzmann factors to account for the indistinguishability of the particles that either did not move, or moved from the same initial state to the same final state [195, 196]. This

results in a caliber of

$$\Omega = \frac{\omega(\{\tilde{n}_i\})}{\prod_i n_i^{\mathrm{r}}! \prod_{il, l\neq i} n_{i\to l}!}, \tag{2.1}$$

where $\omega(\{\tilde{n}_i\})$ is a function of only the final states.

Why 'micro'-caliber? Models that treat particle dynamics using a coarse-grained construction of degenerate particle states, as presented here, are *mesoscopic* models. In these constructions, the microscopic details of the particle configurations are discarded, and instead the only degrees of freedom considered relevant to the dynamics are the state occupation numbers. Our construction is therefore a mesoscopic framework. However, by introducing a means of counting the different possible arrangements of the particles within the degenerate particle states (via ω) we are able to describe a level of detail that is typically lost to mesoscopic models. We have therefore chosen to call the approach 'microcaliber' to distinguish it from other approaches which neglect these smaller-scale degrees of freedom.

A system's dynamics depends on constraints to its particle occupation numbers, which lead to a reduction in the number of possible paths the system can traverse. In particular, we impose a global constraint on the total system energy E and local constraints on the particle movements, which account for particle conservation. These constraints take the form

$$E = U(\{\tilde{n}_i\}), \qquad n_i - n_i^{\mathrm{r}} = \sum_{l\in\mathcal{N}(i)} n_{i\to l}. \tag{2.2}$$

where $U(\{n_i\})$ is a generic energy function that may account for energy differences between the particle states, as well as interactions between the particles themselves, and depends on the particle distribution through the set of occupation numbers $\{\tilde{n}_i\}$. The particle arrangement statistics introduced using ω in Equation (2.1) could have alternatively been introduced here as a constraint to the particle arrangements, using an additional Lagrange multiplier.

Invoking the principle of maximum caliber, the most probable trajectory is found by extremising the caliber subject to these constraints, while treating the initial state $\{n_i\}$ and the number of remaining particles n_i^{r} for each state as fixed, independently determined system parameters. Fixing the initial state $\{n_i\}$ implies that the system starts in a known state, and we are only interested in the particle transitions from this state. In contrast, the number of particles n_i^{r} that remain in a state is determined by the size of the time-step Δt under consideration and, for small time-steps, defines the escape rate R_i and corresponding from state

i, such that

$$n_i^{\mathrm{r}} = (1 - \Delta t R_i) n_i. \tag{2.3}$$

Therefore, n_i^{r} is determined by the choice of time-discretisation and other microscopic modelling parameters, as discussed further in Sections 2.1.1 and 2.1.2, rather than through entropy maximisation. Alternative methods for introducing a similar time-step dependence, through either further constraints in the objective function or by taking small time-step limits, are discussed in [187, 191]. We therefore optimize the objective function

$$f = \ln \Omega + \beta(E - U(\{\tilde{n}_i\})) + \sum_i \theta_i \left(n_i - n_i^{\mathrm{r}} - \sum_{l \in \mathcal{N}(i)} n_{i \to l} \right), \tag{2.4}$$

in which we use the Lagrange multipliers β and θ_i to constrain the total energy and the initial occupation numbers respectively. Taking the extremum of f for $n_{i \to l}$, using equation (2.1) and applying Stirling's approximation, yields the most probable number of particles jumping from i to l

$$n_{i \to l} = e^{\mathcal{A}_l - \theta_i}, \tag{2.5}$$

where we have introduced the general statistical weight

$$\mathcal{A}_l = \frac{\partial \ln \omega}{\partial \tilde{n}_l} - \beta \frac{\partial U}{\partial \tilde{n}_l}. \tag{2.6}$$

Since $\mathcal{A}_l$ is dimensionless, we can identify the Lagrange multiplier β with the inverse energy scale, $\beta = 1/k_{\mathrm{B}}T$.

To fully evaluate equation (2.5) we must impose the constraints associated with the Lagrange multipliers θ_i. Evaluating the constraint on $n_{i \to l}$ from Equation (2.2), the most probable system trajectory is described by the particle jumps

$$n_{i \to l} = \frac{e^{\mathcal{A}_l}}{\xi_i} (n_i - n_i^{\mathrm{r}}), \tag{2.7}$$

with the dynamical partition function,

$$\xi_i = \sum_{l \in \mathcal{N}(i)} e^{\mathcal{A}_l}. \tag{2.8}$$

Interestingly, we find a dependence on the state i via ξ_i which arises due to the restriction of particle movements to those within the neighbourhood of the particle's initial state. Such neighbourhood dependence accounts for how network topology influences the particle dynamics, as briefly discussed in [193].

By construction, our result for $n_{i \to l}$ is provides an out-of-equilibrium description of the system dynamics in form of a weighted graph Markov

chain [197, Chapter 3]

$$n_l(t+dt) = \tilde{n}_l = n_l^{\mathrm{r}} + \sum_{i \in \mathcal{N}(l)} n_{i \to l}, \tag{2.9}$$

In fact, treating each state i as a vertex, we can frame the particle dynamics as a random walk on a weighted, undirected (i.e. both forward and back transitions exist at all edges) graph with asymmetric transition probabilities [197, Chapter 3]. The transition rate for a particle moving from vertex i to a neighbouring vertex l is then

$$k_{i \to l} = \frac{n_{i \to l}}{\Delta t n_i} = \frac{e^{\mathcal{A}_l} R_i}{\xi_i}. \tag{2.10}$$

Formulating the particle dynamics through the transition rates in Equation (2.10) highlights how the microcaliber approach bridges between microscopic system statistics and descriptions of random walks on graphs, as such providing a basis for system dynamics that depends on the fundamental properties of the chosen model system, such as the energy function H, arrangement statistics, ω, and escape rates R_i.

Example: Particles obeying Fermi-Dirac statistics

Although quantum systems are beyond the focus of this thesis, I illustrate the microcaliber framework with the example of particles that obey Fermi-Dirac statistics, such that each particle state can accommodate up to a maximum number of particles $g_i \geq n_i$, called the *degeneracy* of the state. These particles form a paradigmatic model for particles exhibiting volume exclusion – i.e via steric interparticle interactions – in classical thermodynamic systems. It is important to note that the degeneracy is system specific and in other systems it may be pertinent to introduce a degeneracy that combines several particle states (see the example in Section 2.2).

To construct Ω, I first count the number of ways distinguishable particles can be arranged into the lattices, ω. Starting with considering state l, there are $\omega_l = g_l!/(g_l - \tilde{n}_l)!$ ways to arrange $\tilde{n}_l$ particles within the g_l of the state. The total number of ways to arrange all the particles is then given by the product

$$\omega = \prod_l \omega_l = \prod_l \frac{g_l!}{(g_l - \tilde{n}_l)!}. \tag{2.11}$$

As discussed in the preceding section, the path count is corrected for indistinguishable particles by dividing by $\prod_l n_l^r! \prod_{li, i \neq l} n_{i \to l}!$, leading

to the caliber

$$\Omega = \prod_l \frac{g_l!}{(g_l - n_l^r - \sum_{i\in\mathcal{N}(l)} n_{i\to l})! n_l^r! \prod_{i\neq l} n_{i\to l}!}. \tag{2.12}$$

An alternative method of deriving Ω for volume-excluding particles is available in appendix A.1.

Considering the case where each particle state i contributes an energy ϵ_i per particle, the system energy becomes $U(\{\tilde{n}_i\}) = \sum_i \epsilon_i \tilde{n}_i$. Substituting the forms of Ω and U into equation (2.6), results in the general chemical potential $\mathcal{A}_l = \ln(g_l - \tilde{n}_l) - \beta\epsilon_l$. From which, using equation (2.5), the most probable number of particles jumping from i to l is

$$n_{i\to l} = (g_l - \tilde{n}_l) e^{-\beta\epsilon_l - \theta_i}.$$

Applying the constraints associated with the Lagrange multipliers θ_i result in

$$n_i - n_i^r = \sum_{l\in\mathcal{N}(i)} n_{i\to l} = e^{-\theta_i}\left(\zeta_i - \sum_{l\in\mathcal{N}(i)} \tilde{n}_l e^{-\beta\epsilon_l}\right)$$

in which $\zeta_i = \sum_{l\in\mathcal{N}(i)} g_l \exp(-\beta\epsilon_l)$ is a local partition function. Therefore, for the case of particles undergoing volume exclusion, $\xi_i = \zeta_i - \sum_{l\in\mathcal{N}(i)} \tilde{n}_l e^{-\beta\epsilon_l}$ accounts for the influence of the connectivity of the system through the dependence on box index i.

The most probable paths are then found to be

$$n_{i\to l} = \frac{(g_l - \tilde{n}_l)e^{-\beta\epsilon_l}}{\xi_i}(n_i - n_i^r). \tag{2.13}$$

Here it is interesting to note that the number of particles jumping from i to l depends not only on the number of particle in the initial state n_i but also on the number of particles in the *final* state of the system, $\tilde{n}_l$.

2.1.1 Connectivity-dependent steady-states

My use of the pronouns 'I/my' in the remainder of the chapter indicates that the remainder of this work was completed independently.

As previously mentioned, the microcaliber approach finds a dependence of the particle dynamics on particle state neighbourhoods. Such dependencies have previously been observed by Dixit et al. [193] on the scale of whole-system state topologies. In contrast, by detailing the possible transitions of individual particles, the microcaliber framework describes the dependence of the system's trajectory on the topology of permitted individual particle transitions. As such, this permits a direct link to

models of particle diffusion on graphs, where the network topology is a key feature of the physical system being modelled. For such systems free energy minimisation methods are unsuitable for finding the equilibrium particle distribution, since they cannot account for inhibited transitions between non-neighbouring particle states. In the following section, I illustrate this compatibility by recovering known steady-state distributions of particles undergoing random walks on weighted graphs.

The principle of detailed balance states that, at steady state, there exists a particle distribution $\{\bar{n}_i\}$ such that [172, Chapter 6.3.1]

$$\bar{n}_i \Delta t k_{i\to l} = \bar{n}_l \Delta t k_{l\to i}, \tag{2.14}$$

which, in the caliber-framework notation, can also be written as $n_{i\to l} = n_{l\to i}$. Upon substitution of equation (2.7), the principle of detailed balance becomes

$$\frac{e^{\alpha_l}}{\xi_i} R_i \bar{n}_i = \frac{e^{\alpha_i}}{\xi_l} R_l \bar{n}_l. \tag{2.15}$$

Which is satisfied by the steady-state distribution

$$\bar{n}_i \propto \frac{e^{\alpha_i}\xi_i}{R_i}. \tag{2.16}$$

This steady state depends on the network topology through ξ_i, and the choice of R_i sets the edge weight of self-loops within the network. In contrast, free energy minimization would result in the steady state solution $\bar{n}_i \propto e^{\alpha_i}$ which does not account for the network connectivity.

For example, diffusion on symmetric weighted graphs with no self-loops implies that in each time-step all particles must leave their initial state (i.e. $R_i = 1/\Delta t \; \forall \, i$). Considering a case where the particle states and statistics are state invariant such that $A_i = A$ = constant, the particle dynamics are known to resultant in steady state particle distributions where the occupation numbers at each vertex are proportional to the vertex degree, $\bar{n}_i \propto d_i$ [197, Chapter 3]. I indeed recover this steady state by first noting that, from equation (2.8), $\xi_i = \sum_{l\in\mathcal{N}_i} e^{\alpha_l} = e^{\alpha} d_i$ and further substituting into equation (2.15) to find the steady state condition

$$\frac{\bar{n}_i}{d_i} = \frac{\bar{n}_l}{d_l} \tag{2.17}$$

which is satisfied by $n_i \propto d_i$.

The microcaliber results for the equilibrium particle distribution therefore encompass dependencies on both particle jump statistics, through the system-dependent choice of R_i, and graph connectivity, through ξ_i. The principle of maximum caliber thus provides a microscopically-motivated method for constructing system trajectories via Markov chains, which encompasses a broader range of system dynamics than achievable

through free energy minimisation.

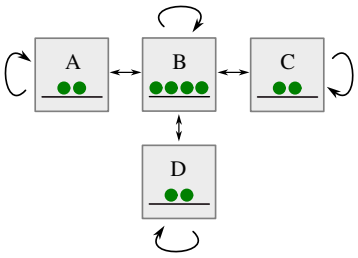


Figure 2.2: Diffusion on this example network of degenerate particle states $i \in \{\mathrm{A}, \mathrm{B}, \mathrm{C}, \mathrm{D}\}$ depends on the state escape rates. Arrows indicate possible state transitions to neighbouring particle states, and circles indicate the steady-state distribution for $R_i = R \,\forall\, i$.

Example: Explicit network topology To further illustrate the difference between random walks on networks and continuum systems, to which free energy minimisation is applicable, I consider a simple example of four vertices $\{\mathrm{A}, \mathrm{B}, \mathrm{C}, \mathrm{D}\}$ with one degenerate particle state each, and with neighbourhoods organised as represented by the graph in Figure 2.2. Assuming all the states are energetically identical, each configuration of particles within the four boxes has the same statistical weight (such that free energy minimisation would give a uniform distribution), however the true steady state would differ depending on the *accessibility* of these configurations. If all the states had equal escape rates $R_i = R$, the probability of a particle moving along an edge is inversely proportional to the degree of the vertex the particle is leaving (including self-loop edges). The transition rates would therefore be given by $\Delta t k_{i\to l} = 1/2 \,\forall\, i \in \{\mathrm{A}, \mathrm{C}, \mathrm{D}\}$, $j \in \{\mathrm{A}, \mathrm{B}, \mathrm{C}, \mathrm{D}\}$ and $\Delta t k_{\mathrm{B}\to j} = 1/4 \,\forall\, j \in \{\mathrm{A}, \mathrm{B}, \mathrm{C}, \mathrm{D}\}$, and the steady state distribution would satisfy $n_\mathrm{B} = 2n_\mathrm{A}$ and $n_\mathrm{A} = n_\mathrm{C} = n_\mathrm{D}$ rather than a uniform distribution[1].

1: In contrast, in continuous systems where network topology is a consequence of simplifying modelling assumptions, the escape rates are proportional to the vertex degree (as seen in the next section), such that $R_i \propto \sum_{l\in\mathcal{N}(i)} e^{\mathcal{A}} \propto d_i$, which recovers the uniform particle distribution as expected from the free energy minimisation, $n_i = n_l \,\forall\, il$.

2.1.2 Continuous physical systems

It is frequently informative to use discrete toy models, such as those described above, to model continuous physical systems. Under these modelling scenarios, the time discretisation and network topology are a modelling choice, and as such not relevant to the dynamics of the true physical system. Having demonstrated how microcaliber captures the network-topology dependence of particle dynamics in discrete systems, in this section I explore how such discrete modelling approaches may also be used to address continuous systems though consideration of how escape rates depend on the choice of model discretisation. Using this, I then recover ensemble-theoretic results for the steady-state distributions of continuum processes, and provide an independent derivation of the local detailed balance condition and a fluctuation theorem.

Returning to the markov chain description of the particle dynamics, described by Equation (2.9), and using Equation (2.3), the particle dynamics can be written in terms of the transition rates $k_{i\to l}$ and escape rates R_i as

$$\tilde{n}_i = (1 - \Delta t R_i) n_i + \sum_{l\in\mathcal{N}(i)} \Delta t k_{l\to i} n_l. \tag{2.18}$$

Through the constraint on particle number in Equation (2.2), the escape rate for particle state i is simply the sum of all the transition rates out of

state i,

$$R_i = \sum_{l \in \mathcal{N}(i)} k_{i \to l}. \tag{2.19}$$

Substituting this into Equation (2.18) and taking the small time-step limit, $\Delta t \to 0$, results in the deterministic dynamical equation for the occupation number of state i,

$$\frac{dn_i}{dt} = \sum_{l \in \mathcal{N}(i)} \left(k_{l \to i} n_l - k_{i \to l} n_i \right). \tag{2.20}$$

It has been noted that taking such continuum-time limits in caliber frameworks can lead to poorly-defined path entropies if not compared to a reference entropy due to explicit dependencies of the transition probabilities on the discrete time scale, see [187] for a brief discussion. Here, however, the limit is taken purely for the Markov chain construction with time-step-independent transition rates, where this discussion is irrelevant. Equation (2.20) describes how the occupation number of a particular particle state, given a known initial value, is expected to evolve over time according to the rates derived from microcaliber. For a mesoscopic system-state described by the set of occupation numbers $\{n_i\}$, the system dynamics can be described through a set of coupled equations of the form of Equation (2.20) for each of the particle state occupation numbers n_i.

Extending Equation (2.20) to consider a *distribution* of initial states, the dynamics can be described by a master equation of the form given by Equation (1.8) [166, Chapter 3],

$$\frac{d}{dt} P(n_i) = \sum_{l \in \mathcal{N}(i)} \left[k_{l \to i} P(n_l) - k_{i \to l} P(n_i) \right]. \tag{2.21}$$

In this master equation, the uncertainty in initial system configuration (as determined by the distribution $P(n_i)$) is propagated via the mesoscopic rates $k_{l \to i}$. By mapping this stochastic system to the microcaliber formalism, I am positing that these mesoscopic rates arise as the most likely transition rates of the underlying microscopic system, subject to energy and particle number constraints, and as such may be found using the microcaliber framework.

As previously mentioned, discrete toy models are frequently constructed to model continuous processes, where the choice of network topology is irrelevant to the true system dynamics. As such, a good toy model would permit the inference of likely system dynamics independent of the arbitrary choice of model topology. Since in the microcaliber framework the dynamical partition function ξ_i depends on the network topology,

imposing the requirement of network independence implies that

$$R_i \propto \xi_i, \tag{2.22}$$

as seen using the definition of the transition rates given in Equation (2.10). Imposing such proportionality factorises out any dependence on the network topology from the dynamics, and although justified heuristically, such arguments of network independence are valid far from equilibrium. In fact, such an equation would be expected to arise from rigorous coarse-graining of continuum systems into such discrete model systems, and it would be a pertinent extension of the work presented in this chapter to explicitly derive such a relation. Furthermore, assuming all relevant system constraints were applied in Equation (2.4), all other state-dependent contributions to the jump rates are expected be contained within the general statistical weights $\mathcal{A}_i$. As such, there is no further state dependence to the proportionality between R_i and ξ_i, resulting in the transition rates

$$k_{l\to i} \propto e^{\mathcal{A}_i}. \tag{2.23}$$

Recovering Gibbs-ensemble distributions

Incorporating the definition of $k_{l\to i}$ from equation (2.23) into the detailed balance relation (2.14), I recover the steady state particle distribution

$$\bar{n}_i \propto e^{\mathcal{A}_i} \tag{2.24}$$

as expected from minimization of the free energy. In particular, for particles obeying Boltzmann statistics and which each contribute an energy ϵ_i to the total system energy, the general chemical potential $\mathcal{A}_i = -\epsilon_i$ and Equation (2.24) recovers the Boltzmann distribution of particles.

Local detailed balance

On the scale of individual particle transitions, the ratio of the forward and backwards transition rates, in general using Equation (2.10), takes the form

$$\frac{k_{i\to l}}{k_{l\to i}} = \frac{e^{\mathcal{A}_l} R_i \xi_l}{e^{\mathcal{A}_i} R_l \xi_i} \tag{2.25}$$

which, in the case of continuum-time jump processes where $R_i \propto \xi_i$, becomes

$$\frac{k_{i\to l}}{k_{l\to i}} = e^{\mathcal{A}_l - \mathcal{A}_i}. \tag{2.26}$$

Here we have derived the *local detailed balance* condition for the individual rates [179, 198, Chapter 5.4.1]. Although commonly justified through

arguments that local particle exchanges must be independent of whether the global system has equilibrated [166, Chapter 3], I have shown that the microcaliber approach provides an independent justification of this relation using the maximisation of transition probabilities and without reference to equilibrium states.

Fluctuation Theorems

How does Equation (2.26) generalise to a system trajectory comprising a sequence of several particle transitions? Following a similar approach to the above, I address this explicitly for the discrete particle-transition model introduced above considering the ratio of forward and backward probabilities for the transition of a system from an initial state $\{n_i\}$ to a final state $\{\tilde{n}_i\}$.

Denoting the number of particles that move from state i to state l over the course of the transition as $\Delta n_{i\to l}$, the final state occupation numbers are given by $\tilde{n}_i = n_i + \sum_l(\Delta n_{l\to i} - \Delta n_{i\to l})$. For a single particle transition of duration t, the probability that one of the n_i particles initially in state i moves to state l is $P_{i\to l} = (1 - e^{-k_{i\to l}t})n_i \approx k_{i\to l}n_i t$. Decomposing the system transition into individual particle transitions indexed by k, and of duration t_k, and assuming that the number of exchanged particles is small compared to the occupation numbers of these states, $\Delta n_{i\to l} << n_i, \tilde{n}_l$, the probability of the full state transition is then given by $P_{\{n\}\to\{\tilde{n}\}} \approx \prod_{i,l\neq i}[(k_{i\to l}n_i)^{\Delta n_{i\to l}} \prod_k t_k]$. In the backwards transition, the exchanged particles return to their original state following the time-reversed order of the original transition sequence, resulting in the probability ratio

$$\frac{P_{\{n\}\to\{\tilde{n}\}}}{P_{\{\tilde{n}\}\to\{n\}}} \approx \prod_{i,l\neq i}\left(\frac{k_{i\to l}n_i}{k_{l\to i}\tilde{n}_l}\right)^{\Delta n_{i\to l}} = e^{-\sum_{i,l\neq i}[\beta\phi_l(\tilde{n}_l)-\beta\phi_i(n_i)]\Delta n_{i\to l}}. \quad (2.27)$$

where in the second step I have introduced $\beta\phi_i(n_i) = -\mathcal{A}_i + \ln(R_i) - \ln(\xi_i) + \ln(n_i)$.

Equation (2.27) resembles a fluctuation theorem which, valid far from equilibrium, relate the ratio of transition probabilities to the exponential of the change in a global quantity. Inspired by this I introduce a quantity, Φ, such that

$$\Delta_{\{n\}\to\{\tilde{n}\}}\Phi = \sum_{i,l\neq i}[\phi_l(\tilde{n}_l) - \phi_i(n_i)]\Delta n_{i\to l}. \quad (2.28)$$

Defining $\Phi(\{n_i\})$ as a functional of the set of occupation numbers, I approximate the change in this quantity by expanding for small numbers

of exchanged particles $\Delta n_{i\to l}$. This expansion gives

$$\Delta_{\{n\}\to\{\tilde{n}\}}\Phi \approx \sum_{i,l\neq i}\left(\frac{\partial\Phi}{\partial\tilde{n}_l}-\frac{\partial\Phi}{\partial n_i}\right)\Delta n_{i\to l}+\mathcal{O}(\frac{\Delta n^2}{n_i},\frac{\Delta n^2}{\tilde{n}_i}). \tag{2.29}$$

Therefore, through comparison of Equations (2.29) and (2.28), I identify $\partial\Phi/\partial n_i = \phi_i$. Taking the integral with respect to n_i, and substituting the Equation (2.6) for $\mathcal{A}_l$, results in the functional

$$\beta\Phi(\{n_l\}) = \left(-\ln\omega+\beta U-\sum_l n_l\left[\ln\left(\frac{\xi_l}{R_l}\right)-\ln(n_l)+1\right]\right)+\text{constant}. \tag{2.30}$$

Interestingly, the $\ln\omega$ and $n_l\ln n_l$ terms in Equation (2.30) combine to give the Boltzmann entropy of the system, $S_\mathrm{B} = k_\mathrm{B}\ln(\omega/\prod_l n_l!)$, under the Stirling approximation. Upon substitution into the above, this results in

$$\begin{aligned}\Phi(\{\tilde{n}_l\}) &= -TS_B + U - \sum_l k_\mathrm{B}Tn_l\ln\left(\frac{\xi_l}{R_l}\right)+\text{const.}\\ &= F - \sum_l k_\mathrm{B}Tn_l\ln\left(\frac{\xi_l}{R_l}\right)+\text{const.}.\end{aligned} \tag{2.31}$$

where F is the Helmholtz free energy. The quantity Φ is therefore identified as a grand potential, in which each particle state l contributes a *thermodynamic chemical potential*, $\mu_{\mathrm{th},l} = k_\mathrm{B}\ln(\xi_l/R_l)$, that accounts for the accessibility of the particle state l [146, Chapter 1]. Equation (2.27) therefore relates mesoscopic transition probabilities to changes in the constrained free energy, resulting from changes in the Helmholtz free energy and the additional contribution that takes into account the connectivity of the states (i.e. the spatial dependence on the neighbourhoods $\mathcal{N}(i)$).

In contrast to local detailed balance, which regards state changes on the scale of individual particle exchanges, fluctuation theorems provide relationships between the forward and backward probabilities of system-state transitions at the mesoscopic scale [199], as discussed in Chapter 1. General maximum caliber approaches have been shown to recover such relations [175, 184, 191]. To put Equation (2.27) into the form of a true fluctuation theorem, one must account for a probability of being initially in the $\{n\}$ state, $P(\{n\})$, for the forward transition, and in the $\tilde{n}$ state, $p(\{\tilde{n}\})$, for the backwards transition. Accounting for these terms I find

$$\frac{P_{\{n\}\to\{\tilde{n}\}}}{P_{\{\tilde{n}\}\to\{n\}}} = e^{-\beta\Delta\Phi+\Delta s}, \tag{2.32}$$

where Δs is the additional entropy change arising from changes in the probability distribution $\Delta s = \ln P(\{n\}) - \ln P(\{\tilde{n}\})$ [167]. For continuous systems, where $R_i \propto \xi_i$, the thermodynamic chemical potential is a

constant and I recover the Crooke's fluctuation theorem (see Chapter 1) for an isolated system (i.e. no work done, $W_0 = 0$), where the free energy, $F' = F - Ts$, also includes the stochastic entropy contribution from the particle distribution, $s = \ln(P(\{n\}))$ [167, 168].

Outside of maximum caliber methods, such fluctuation theorems are typically derived from solutions to dynamical equations such as the Fokker-Planck equation [199]. However, the construction of such equations commonly rely of equilibrium frameworks, or have a circular element by which a fluctuation theorem is required to fully define the dynamical equations (see discussion in [167]). As demonstrated here, and elsewhere, maximum caliber methods permit independent derivations of such fluctuation theorems that do not suffer such limitations [184]. Here, in particular, I have demonstrated how both underlying network topology and particle arrangement statistics contribute to the fluctuation relations, which may provide a basis by which to assess the influence of such microscopic properties on fluctuations in particle patterning systems.

2.2 Mean-field dynamics

In many modelling applications, it is sufficient to address purely mean-field particle dynamics. As introduced in Chapter 1, the mean-field Fokker-Planck equation provides a description of such systems, and it's derivation from a Markov-chain system description is widely available in physics textbooks, e.g. [166, 171]. Following such approaches, it has been confirmed that maximum caliber methods recover this particle description [178]. In this section, I present a brief derivation of the mean-field Fokker-Planck equation from the microcaliber framework, starting from the Markov chain for the system, Equation (2.9). In doing so I demonstrate explicitly how the particle statistics captured by the microcaliber approach are mapped into the mean-field description.

Since the generalisation to higher dimensions should be straightforward, I present this derivation for a one-dimensional system, where space is discretised into a lattice of boxes with length dx, indexed by i, such that box i is at position x_i (see the sketch in Figure 2.3). Each box contains several particle states, each of which may contain particles. The particle distribution can then be described by the set of occupation numbers $\{n_{ij}\}$ of each state j in each of the lattice sites i. Here an additional state index has been included to clearly distinguish between the particle states and positions. Assuming the particles can jump to all states in either its initial box i, or in neighbouring boxes $i-1, i+1$, and using the Markov chain in Equation (2.9), the change in state occupation number after a

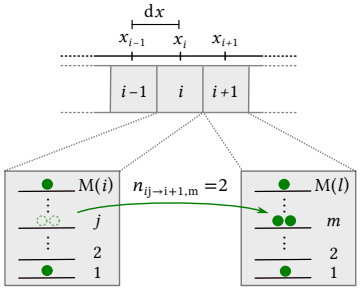


Figure 2.3: Simplified, one-dimensional model for interacting particles on a lattice of length-scale dx (top). Particles are distributed between the M degenerate particle states within each box (bottom), and – within a time-step dt – are permitted to jump to another state, either within their own box or in an adjacent box.

time-step dt is given by

$$\begin{aligned} n_{ij}(t+dt) - n_{ij}(t) = &\sum_{k\neq j}(n_{ik\to ij} - n_{ij\to ik}) \\ &+ \sum_{k} n_{i-1,k\to ij} + n_{i+1,k\to ij} - (n_{ij\to i-1,k} + n_{ij\to i+1,k}). \end{aligned} \tag{2.33}$$

In the limit of small dx and dt, I define continuous particle density fields $\rho_j(x) = \lim_{dx\to 0} n_{ij}/a$ and – following similar logic as used in Equation (2.10) – transition probabilities $P_{kj}(x) = \lim_{dx\to 0} n_{ik\to ij}/n_{ik}$ and $P_{\pm,kj}(x) = \lim_{dx\to 0} n_{ik\to i\pm 1j}/n_{ik}$.

Taking this limit of Equation (2.33), and expanding in small dx results in the Fokker-Planck equation [171]

$$\begin{aligned} dt\partial_t \rho_j(x) = &\sum_{k\neq j}\left[(P_{+,kj} + P_{kj} + P_{-,kj})\rho_k - (P_{-,jk} + P_{jk} + P_{+,jk})\rho_j\right] \\ &+ \sum_{k}\left[dx\partial_x([P_{-,kj} - P_{+,kj}]\rho_k) + \frac{dx^2}{2}\partial_x^2([P_{+,kj} + P_{-,kj}]\rho_k)\right], \end{aligned} \tag{2.34}$$

where the diffusion coefficient tensor can be identified as $D_{kj} = P_{+,kj} + P_{-,kj}$. The first two terms in this equation are reaction rates, whereas the latter two describe directed motion and diffusive motion respectively. Here I have included the possibility of transitions where particles change both position and state within the same jump, which give rise to cross diffusion terms. Although commonly discarded in typical Fokker-Planck derivations, these jumps have been shown to give rise to interesting phenomena in biological systems [200], and as such I have included them for completeness in the following derivation.

Expanding in terms of Poissonian rates

In Belousov et al. [173], the authors demonstrate the equivalence of the Smoluchowski and Fokker-Planck equations by formulating the transition rates as Poissonian rates using a directing function. Here, I extend this formalism to allow for multiple particle energy states.

As discussed in Section 2.1.2, continuous systems, when modelled in a lattice-gas framework, have an escape rate proportional to the connectivity of the imposed lattice, $R_{ij} \propto \xi_{ij}$. Under these conditions, the probability of a transition – given by equation (2.10) – takes the form $p_{ij\to lm} = dtR'_{ij}e^{\varpi_{lm}}$, where $dtR'_{ij} = dtR_{ij}/\xi_{ij}$ is a connectivity-independent probability of escaping state ij in time dt. Although this would typically be state-independent in systems where the free energy captures all non-network topology dynamics (as discussed in Chapter 1), including an additional state and position dependence here allows, for example, for the inclusion

of spatially varying diffusion coefficients arising from variations in particle jump rates. I therefore include such state dependence. Writing the connectivity-independent escape probability as an exponential of the form $dtR'_{ij} = e^{-\Gamma_{ij}}$, the continuum probabilities are found to be

$$P_{\pm,jk} = e^{-\Gamma_j(x)+\lambda_k(x\pm dx)} \quad \text{and} \quad P_{jk} = e^{-\Gamma_j(x)+\lambda_k(x)}, \tag{2.35}$$

where $\lambda_k(x \pm dx) = \lim_{dx\to 0} \mathcal{A}_{i\pm 1k}$, and $\Gamma_j(x) = \lim_{dx\to 0} \Gamma_{ij}$. Expanding these probabilities to first order in dx and substituting into the dynamical equation (2.34), the Fokker-Planck equation takes the form,

$$\begin{aligned}\partial_t \rho_j(x) = &\sum_{k\neq j} \Big[\frac{3}{dx^2}(D_{kj}\rho_k - D_{jk}\rho_j)\Big] \\ &+ \sum_k \Big[-\partial_x(D_{kj}\rho_k \partial_x(2\lambda_j(x))) + \partial_x^2(D_{kj}\rho_k)\Big],\end{aligned} \tag{2.36}$$

where

$$D_{kj} = e^{-\Gamma_k(x)+\lambda_j(x)}\frac{dx^2}{dt}. \tag{2.37}$$

Relating diffusive dynamics to fluctuation theorems

It now remains to relate the diffusion coefficients D_{jk} to global thermodynamic properties. Taking the ratio of the number of particles $P_{+,jk}\rho_j(x)$ jumping from state j at x to state k at $x + dx$ with the number jumping back $P_{-,jk}\rho_k(x + dx)$, using equation (2.35), and expanding in x to first order gives,

$$\frac{P_{+,jk}(x)\rho_j(x)}{P_{-,kj}(x+dx)\rho_k(x+dx)} \approx e^{\alpha_k - \alpha_j + dx\partial_x\alpha_k}, \tag{2.38}$$

where $\alpha_k = \Gamma_k(x)dt + \lambda_k(x) - \ln\rho_k(x)$ is defined up to a constant. This resembles Equation (2.27) in the case that only one particle transition from state ij to state $i+1, k$ which, in the limit of small boxes, can be written in the form

$$\begin{aligned}\frac{P_{+,jk}(x)\rho_j(x)}{P_{-,kj}(x+dx)\rho_k(x+dx)} &= e^{-\beta\mu_k(x+dx)+\beta\mu_j(x))} \\ &\approx e^{-\beta\mu_k(x)-\beta dx\partial_x(\mu_k(x))+\beta\mu_j(x))},\end{aligned} \tag{2.39}$$

where $\mu_j(x) = -\lim_{dx\to 0}\partial\Phi/\partial n_{ij}$ and $\mu_j(x+dx) = -\lim_{dx\to 0}\partial\Phi/\partial n_{i+1,j}$ are *generalised chemical potentials*. It is important to highlight that these chemical potentials are different from the thermodynamic chemical potential $\mu_{\text{th},l}$ defined in Section 2.1.2. In the case where there are no state-dependent contributions to the escape probabilities, such that $dtR'_{ij} =$ constant (see Section 2.1.2), the thermodynamic chemical potential is a constant, and the generalised chemical potentials become the gradient of the free energy (up to a constant). Comparing equations (2.38) and

(2.39), I find

$$-\beta\mu_k(x) \doteq \alpha_k \doteq \Gamma_k(x)dt + \lambda_k(x) - \ln\rho_k(x), \tag{2.40}$$

with $\doteq$ denoting equality up to a constant. This comparison relates the parameters of the Fokker-Planck description to the underlying microscopic dynamics encapsulated in the grand potential. Combining Equation (2.40) and (2.37) results in the relation,

$$-2\lambda_j(x) = \beta\mu_k(x) + \ln\frac{D_{jk}(x)}{D_{jj}(x)} - \ln D_{kj}(x)\rho_k(x) + \text{constant}, \tag{2.41}$$

which completes the link between the microcaliber results and the Fokker-Planck dynamical description. Substituting Equation (2.41) into Equation (2.36) and generalising to arbitrary spatial dimensions results in

$$\partial_t\rho_j(\mathbf{x}) = \mathcal{R}_j(\mathbf{x}) + \sum_k \nabla\cdot\left[D_{kj}(\mathbf{x})\rho_k(\mathbf{x})\nabla\left(\beta\mu_k(\mathbf{x}) + \ln\frac{D_{jk}(\mathbf{x})}{D_{jj}(\mathbf{x})}\right)\right], \tag{2.42}$$

where we have defined the reaction rates

$$\mathcal{R}_j = \sum_k 3D_{kj}(\rho_k - e^{-\mu_k+\mu_j-\ln\frac{\rho_j}{\rho_k}}\rho_j)/dx^2. \tag{2.43}$$

As previously discussed, it is common practise to neglect cross-diffusive transitions in derivations of the Fokker-Planck equation. Neglecting such contributions, I recover Fick's law, $\mathbf{J} = -D\rho\beta\nabla\mu$, and the dynamical equation for particle conservation,

$$\partial_t\rho_j = \mathcal{R}_j + \beta\nabla\cdot(D_{jj}\rho_j\nabla\mu_j). \tag{2.44}$$

Typically, Equation (2.44) is given in the form of the Smoluchowski equation of reaction-diffusion dynamics [173],

$$\partial_t\rho_j = \mathcal{R}_j + \nabla\cdot(D_{jj}\rho_j\beta\nabla\tilde{\mu}_j + D_{jj}\nabla(\rho_j)), \tag{2.45}$$

where

$$\tilde{\mu}_j = \mu_j - k_\mathrm{B}T\ln\rho_j + \text{constant} \tag{2.46}$$

is an *adjusted chemical potential* that accounts for corrections to the chemical potential of the ideal gas – which is described by the term $D_{jj}\nabla(\rho_j)$. The drift terms in the Fokker-Planck and Smoluchowski formalisms therefore describe how particle interactions with either their environment or with other particles influence particle dynamics. Using the microcaliber framework, I have provided a non-equilibrium derivation of how these interactions contribute to the dynamics through both density-dependent system energy U contributions and the microscopic particle arrangement

statistics included in ω. Overall, by deriving the mean-field Fokker-Planck equations using the microcaliber framework, I have demonstrated the consistency of the microcaliber method with well-known non-equilibrium modelling approaches and identified how microscopic properties of modelled systems may contribute to particle dynamics. Maximum caliber approaches have also been used to derive Langevin equations, and the dynamical equation (2.44) could have been derived from this starting point instead.

Particles obeying volume exclusion

For volume-excluding particles arranged on a lattice, such that the particle state degeneracy takes the form $g_i \geq \sum_k n_{ik}$ and the state-independent escape rate is constant $dtR'_{ik} = \text{constant}$, the grand potential is equal $\Phi \doteq F$ to the Helmholtz free energy,

$$F = U - TS = \sum_i [\sum_k \epsilon_{ik} n_{ik}] + k_B T \ln \frac{g_i!}{(g_i - \sum_k n_{ik})! \prod_k n_{ik}!}, \quad (2.47)$$

up to a constant. Applying the Stirling approximation, and taking the partial derivative with respect to n_{ik} gives the chemical potential $\mu_{ik} = \epsilon_{ik} - k_B T \ln(\frac{g_i - n_i}{n_{ik}})$ which, in the limit of vanishing boxes, becomes the chemical potential field

$$\mu_k(x) = \lim_{\Delta x \to 0} \mu_{ik} = \epsilon_k(x) - k_B T \ln(\frac{\rho_{\max}(x) - \rho(x)}{\rho_k(x)}). \quad (2.48)$$

Substitution into equation (2.45) results in the typical Smoluchowski reaction-diffusion equation

$$\partial_t \rho_j(x) = \partial_x \Big(D_{jj} \partial_x(\rho_j) + \beta D_{jj} \rho_j \partial_x(\epsilon_j(x) + E(x)) \Big) + \mathcal{R}_j, \quad (2.49)$$

where $E(x) = -k_B T \ln(\rho_{\max}(x) - \rho(x))$ is the adjustment to the chemical potential due to the volume exclusion interactions. In this example, short-range, repulsive particle-particle interactions introduced through constraints in particle rearrangement statistics are shown to result in a density-dependent contribution to the chemical potential.

2.3 Concluding remarks

Aiming to provide a foundation on which to investigate the role particle interactions play in out-of-equilibrium particle dynamics, in this chapter I have extended caliber methods to address microscopic system constraints. As discussed in Jaynes et al. [174], who first proposed caliber methods in 1980, it is desirable to have a non-equilibrium equivalent to Gibbs

ensemble methods, where conservation laws are imposed and from which correct phenomenological relations follow. Through my microcaliber framework, I recover the local detailed balance relation, fluctuation theorems, and expected mean-field dynamics, all of which lend support to the suggestion that caliber methods achieve this goal. Introduced either as energetic contributions to the system energy or via constraints on particle statistics, the microscopic caliber method directly addresses how particle interactions influence system state transitions and contribute to mean-field dynamics. Although the remainder of this thesis will focus on addressing equilibrium particle distributions, these detailed dynamics – and the unified, non-equilibrium view of their origins – may facilitate studies on particle pattering and organisation in living and soft robotic systems by providing a concrete foundation on which to build model systems.

Spatial signal transmission at cellular surfaces

3

Patterning – the emergence of spatial order – is ubiquitous in biology, and recent years have seen considerable progress in uncovering how such patterns contribute to sensing and signalling [17, 135, 201–204]. Many of these efforts are motivated by the expectation that understanding signal processing in living systems may help guide designs of synthetic soft signal processing materials. Key to these approaches is the idea that information relevant to some biological function may be encoded within the spatial organisation of a system, and subsequently 'read out' by particle distributions that use these heterogeneities as a template [124]. Most of the work in understanding these signal transmission cascades has focused on molecular and gene circuits, for which a wealth of theoretical tools have been developed [32, 61, 80, 84, 85, 205, 206]. Yet, the role that *physical* interactions and mechanical system properties play in how living materials process signals has been less explored, contrary to the vast array of observational evidence that biology does indeed use these properties [148, 207–211].

In this chapter, I investigate how physical interactions at interfaces shape the encoding and transmission of spatially-resolved mechanical signals. Focusing on the special case of particles diffusing within a two-dimensional membrane and interacting with their environment through binding to external structures, I derive analytical equations for the equilibrium distributions of particles, and identify how particle-particle interactions give rise to non-linear relations that control signal transmission across the interface. Membrane-enclosed systems provide a particularly interesting opportunity to study pattern-mediated signal transduction because the membrane forms a regulatory barrier to the material that it encloses, through which information must be relayed in order to coordinate responses to external stimuli [212].

A paper containing part of the results presented in this chapter is accepted at PRL [213]. The notebooks and scripts used to produce the results of this chapter are provided in the repositories [214] and [215].

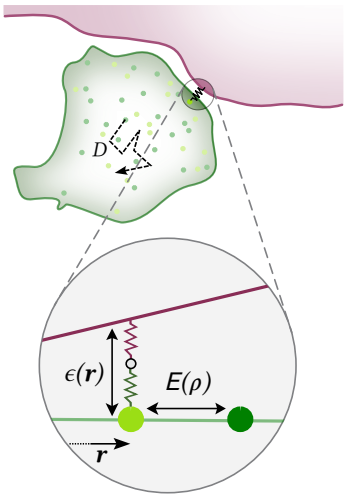


Figure 3.1: A membrane enclosed compartment (green shading) is populated with particles which may either be bound (light green circles) or unbound (dark green circles) from binding sites in the compartment's environment (pink shading), with binding rates determined by the binding energy field $\epsilon(r)$. The particles diffuse with a diffusion coefficient D and undergo particle-particle interaction accounted for by a density-dependent interaction potential $E(\rho)$. Adjusted from [213], licenced under CC BY 4.0.

3.1 Particle distributions guided by binding

Consider a set of diffusive particles situated on a two-dimensional membrane of area A, and interacting with their surroundings by attaching to and detaching from binding sites in their local environment (see Figure

3.1). Using a mean-field approximation, the distribution of these particles can be described by their densities ρ_i in the bound (i = b) and unbound (i = u) states, and the time-evolution of these densities is given by mass-conserving reaction-diffusion equations [137, 138, 172],

$$\partial_t \rho_i = \nabla \cdot \left[D_i \nabla \rho_i + \beta D_i \rho_i \nabla \tilde{\mu}_i \right] + \mathcal{R}_i \tag{3.1}$$

with the inverse thermal energy scale $\beta = (k_B T)^{-1}$ and, in general, density-dependent diffusion coefficients D_i. Dynamical equations such as Equation (3.1) were derived in Chapter 2 and can also be routinely derived from Langevin equations of the particle dynamics as detailed in [172]. The adjusted chemical potentials $\tilde{\mu}_i(\rho_u, \rho_b, \mathbf{r})$, which depend on the particle densities and the position $\mathbf{r}$, capture the energetic difference between the bound and unbound particle states, along with any particle-particle interactions (see also Chapter 2, Equation (2.42)), and the reaction terms $\mathcal{R}_i$ account for fluxes between the bound and unbound particle states according to $\mathcal{R}_u = -\mathcal{R}_b = f(\rho_u, \rho_b, \tilde{\mu}_u)(\rho_b - e^{-\beta\Delta\tilde{\mu}}\rho_u)$, where $\Delta\tilde{\mu} = \tilde{\mu}_b - \tilde{\mu}_u$ and $f(\rho_u, \rho_b, \tilde{\mu}_u)$ is a density dependent prefactor that has no effect on the equilibrium densities.

The equilibrium solution to the coupled dynamical equations given by Equation (3.1) can be found by imposing the condition of vanishing fluxes – including particles state fluxes such that $\mathcal{R}_i = 0$ and spatial fluxes such that $\boldsymbol{J} = D_i \nabla \rho_i + \beta D_i \rho_i \nabla \tilde{\mu}_i = 0$. Assuming no-flux boundaries, integrating the latter of these conditions and substituting into the former finds the formal solution for the equilibrium particle densities,

$$\rho_u = \rho_{b,0} e^{-\beta\tilde{\mu}_u}, \qquad \rho_b = \rho_{b,0} e^{-\beta(\Delta\tilde{\mu}+\tilde{\mu}_u)}, \tag{3.2}$$

where $\rho_{i,0}$ is an integration constant. Combined, these solutions yield the total particle density

$$\rho = \rho_u + \rho_b = \rho_{b,0} e^{-\beta\tilde{\mu}_u}(1 + e^{-\beta\Delta\tilde{\mu}}), \tag{3.3}$$

in which $\tilde{\mu}_u$ captures the influence of particle-particle interactions experienced by unbound particles, and $\Delta\tilde{\mu} = \tilde{\mu}_b - \tilde{\mu}_u$ accounts both for differences in energy between the bound and unbound particle states, as well as any differences in the particle-particle interactions between the particle states.

How do particle interactions influence the equilibrium distribution? Using Equation (2.46) from Chapter 2, the adjusted chemical potential is found to be

$$\tilde{\mu}_j = \frac{\delta F}{\delta \rho_j} - k_B T \ln \rho_j + \text{constant} \tag{3.4}$$

where $F = U - TS$ is the Helmholtz free energy, with Boltzmann entropy S, temperature T, and system energy U. Decomposing the system energy

into an energetic contribution $U_{\mathrm{I}}(\rho_{\mathrm{u}}, \rho_{\mathrm{b}})$ arising from particle-particle interactions, and an additional contribution $\int (\epsilon_{\mathrm{u}}\rho_{\mathrm{u}} + \epsilon_{\mathrm{b}}\rho_{\mathrm{b}})\, dA$ accounting for the energy ϵ_i contributed by each particle according to its state i, the resulting adjusted chemical potential is given, up to a constant, by

$$\tilde{\mu}_j = E_i(\rho_{\mathrm{u}}, \rho_{\mathrm{b}}, \mathbf{r}) + \epsilon_j(\mathbf{r}). \tag{3.5}$$

The interactions of the particles with their environment are therefore encapsulated by the binding energy fields $\epsilon_i(\mathbf{r})$ that describes the energy associated with particles in state i, and an interaction-potential field

$$E_i(\rho_{\mathrm{u}}, \rho_{\mathrm{b}}, \mathbf{r}) = \frac{\partial}{\partial \rho_j}\left[- k_B T \ln \omega + U_{\mathrm{I}}(\rho_{\mathrm{u}}, \rho_{\mathrm{b}})\right], \tag{3.6}$$

that provides a correction to the chemical potential of an ideal gas to account for particle-particle interactions. These corrections may arise either from particle arrangement constraints – described by the number of ways ω of arranging distinguishable particles (see Chapter 2) – and/or from the energetic contributions of the particle-particle interactions described by $U_{\mathrm{I}}(\rho_{\mathrm{u}}, \rho_{\mathrm{b}})$.

Assuming that the particle-particle interactions and arrangement constraints are independent of the binding status of the particles, the interaction potentials for both bound and unbound particles are equal and dependent only on the total density, $E_i(\rho_{\mathrm{u}}, \rho_{\mathrm{b}}, \mathbf{r}) = E(\rho)$[1]. Furthermore, to investigate how spatial heterogeneities in the environment influence the particle distribution, I consider a non-uniform particle-environment binding energy field $\epsilon(\mathbf{r}) = \epsilon_{\mathrm{b}}(\mathbf{r}) - \epsilon_{\mathrm{u}}$ and assume that ϵ_{u} is spatially homogeneous in order to focus on only one guiding field. With the constants $\rho_{\mathrm{b},0}$ and ϵ_{u} incorporated into the normalization factor $l^{-2} = \rho_{\mathrm{b},0} e^{-\beta\epsilon_{\mathrm{u}}}$, these assumptions result in the equilibrium total particle density

$$\rho = \frac{1}{l^2} e^{-\beta E(\rho)}(1 + e^{-\beta\epsilon}). \tag{3.7}$$

1: This assumption breaks down when the particle binding sites are inhomogeneous, or at a lower density than the maximum particle density. I explore the impact of relaxing this assumption in Section 3.2.4

The integration constant l is a length-scale set by the conservation of the total particle number N over the whole membrane area according to

$$N = \int_A \mathrm{d}A\, \rho. \tag{3.8}$$

This constraint determines the average separation distance between particles across the whole surface $\bar{d} = \sqrt{A/N}$, and makes Equation (3.7) a non-local mapping of the binding energy field to the density field.

Recovery of expected distributions If the energetic cost of binding is constant, Equation (3.7) recovers a homogeneous distribution of

particles as expected from unbiased diffusion. Conversely, if there is no particle repulsion, and as such no limit to the density of the particles, $E(\rho)$ is a constant, and we recover the density of particles obeying Boltzmann statistics.

3.1.1 Signal filters from guided particle distributions

As shown by Equation (3.7), the equilibrium distribution of the diffusive particles is 'guided' by the spatially heterogeneous binding energy field. As such the patterning process acts as a signal filter, for which Equation (3.7) details a non-linear, non-local, *response function* that maps spatial features of the input binding energy field ϵ into the output density field ρ.

Such a signal filter is particularly relevant for the case of membrane-enclosed biological compartments, where information must be processed and relayed from outside to within the compartment, or vice versa, to regulate biological functions such as cell migration or cell differentiation [212]. By reflecting symmetry-broken features of the compartment's environment, the binding energy field may act as an initial store of this information. The mapping of input energy signals to the output particle densities then permits the transmission of this input signal into the plane of the membrane, from which processes within the compartment may 'read' the stored information.

In general, either through noise or due to non-linear features of the response function, the response functions of the filter may be non-invertible, such that several input values would result in the same output density. In these cases, the output density signal is a compressed version of the input, where some of information in the original signal has been lost, or *filtered*, from the transmitted signal.

Signal filters for proximity sensing

Spatial heterogeneities in the binding energy field could arise, for example, from a non-homogeneous separation distance $h(\mathbf{r})$ of binding sites from the membrane surface. By converting this variation in separation to a spatially dependent binding energy field through an effective Hookean interaction,

$$\epsilon(\mathbf{r}) = \frac{k}{2} h(\mathbf{r})^2 + \epsilon_c, \tag{3.9}$$

the non-uniform interaction energy would capture spatial variations in binding site proximity. The response function given by Equation (3.7) would therefore constitute a signal filter for binding site proximity. Heterogeneities in the chemical or elastic properties of the environ-

ment might furthermore introduce a non-uniform chemical binding energy ϵ_c and/or effective spring constant k, whose heterogeneities would also be sensed by this mechanism.

3.2 Interaction potentials tune response functions

The exact form of the signal filter produced through the guided particle patterning process described in Equation (3.7) is determined by the inter-particle interactions comprising the interaction potential $E(\rho)$. Although this potential could take a limitless array of forms, in the interest of space I limit this section to detailing general filter features of the most relevant – i.e. repulsive and attractive – interaction types.

3.2.1 Repulsive interactions

A common category of inter-particle interactions are short-range repulsions. As shown in Equation (3.6), the impact of these interactions can be modelled by a potential that accounts for particle repulsion either via a density-dependent system energy or as a constraint on particle arrangements. In this section, I first derive expressions for $E(\rho)$ for steric interactions (i.e. 'Hard-core repulsion') in a lattice gas framework, and then similarly address continuous repulsive interactions (i.e 'Soft-core' repulsion), before investigating the response function of the resultant filter.

Hard-core repulsion

Steric interactions are a specific, analytically tractable case of short-range potentials. Such 'hard-core repulsions' describe particles with volume exclusion, where the particle-particle interaction potential is infinite at particle separations less than the particle diameter, and zero otherwise [146, Chapter 5].

Following a similar process to Section 2.2, I discretise the membrane surface into B boxes of size dx, indexed by j (see Figure 3.2). Within each box the steric interactions are modelled using a lattice gas approximation, where particles are confined to a lattice of Ω single-occupancy sites of size d_0, which imposes an effective particle size. For generality, I assume that binding sites are accessible on only a subset Q_j of the lattice sites in any box.

Assuming that all particle-particle interactions are hard-core, the system energy is given by $U = \sum_j \epsilon_{j,\mathrm{u}} K_j + \epsilon_{j,\mathrm{b}} M_j$, where M_j and K_j are the

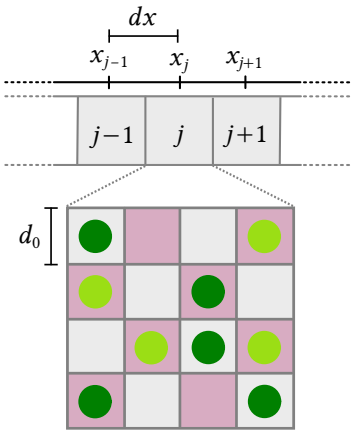


Figure 3.2: In a lattice model of hard-core repulsion, particles are distributed between square lattices of total size dx, which have single occupancy sites of size d_0 and are indexed by j. Q_j of the lattice sites (pink shading) permit particles to be in either a bound (light green circles) or unbound (dark green circles) state, whereas the remaining lattice sites permit only unbound particles.

numbers of bound and unbound particles respectively in each box. In contrast, the entropy

$$S = -\sum_j k_B \ln \frac{Q!(\Omega - M_j)!}{(Q - M_j)!(\Omega - M_j - K_j)!K_j!M_j!}, \tag{3.10}$$

is found from the number of microstates that make up the macrostate defined the values of M_j and K_j in all of the boxes. This is found by treating the particles as indistinguishable but the sites distinguishable and counting the number of distinct arrangements of the bound particles in the Q_j binding spaces, followed by the arrangements for the free particles in the remaining $\Omega - M_j$ spaces, for each box j .

From the Helmholts free energy, $F = U - TS$, using Stirling's approximation and taking the limit of small boxes, $dx \to 0$, the chemical potentials for the bound and unbound particles are found to be

$$\mu_\mathrm{u}(x) = \lim_{dx\to 0} \mu_{j,\mathrm{u}} = \epsilon_\mathrm{u}(x) - k_B T \ln\left(\frac{\rho_\mathrm{max} - \rho(x)}{\rho_\mathrm{u}(x)}\right) \tag{3.11}$$

and

$$\mu_\mathrm{b}(x) = \lim_{dx\to 0} \mu_{j,\mathrm{b}} = \epsilon_\mathrm{b}(x) - k_B T \ln\left(\frac{\rho_\mathrm{ext}(x) - \rho(x)}{\rho_\mathrm{b}(x)}\right). \tag{3.12}$$

where I have used $\rho_\mathrm{u}(x) = \lim_{dx\to 0} K_j$ and $\rho_\mathrm{b}(x) = \lim_{dx\to 0} M_j$, and where the maximum density and the external binding site density are defined as $\rho_\mathrm{max} = \lim_{dx\to 0} \Omega/dx^2 = 1/d_0^2$, and $\rho_\mathrm{ext}(x) = \lim_{dx\to 0} Q_j/dx^2$ respectively. By comparison of Equations (3.11) and (3.12) to Equation (3.5), I find the interaction potentials for hard-core repulsion

$$E_\mathrm{u}(\rho) = -k_B T \ln\left(\frac{\rho_\mathrm{max} - \rho}{\rho_\mathrm{max}}\right) \tag{3.13}$$

and

$$\Delta E(\rho) = E_\mathrm{b}(\rho) - E_\mathrm{u}(\rho) = -k_\mathrm{B} T \ln\left(\frac{\rho_\mathrm{ext} - \rho_\mathrm{b}}{\rho_\mathrm{max} - \rho_\mathrm{b}}\right). \tag{3.14}$$

The same result can also be recovered by deriving the equilibrium distribution directly through entropy maximisation (see Appendix A.2.2) and comparing to the equilibrium distribution in Equation (3.7). Under the further constraint that the binding site density is equal to the maximum particle density, $\rho_\mathrm{max} = \rho_\mathrm{ext}$, the change in interaction potential becomes $\Delta E(\rho) = 0$, and I recover the expected result that $E_i(\rho_\mathrm{u}, \rho_\mathrm{b}, r) = E(\rho)$.

Soft-core repulsion

An alternative option for modelling short-range repulsive interactions is to detail the energetic contribution explicitly. In this section, I outline this

approach, where there are no constraints on particle arrangements – such that the particles obey Boltzmann statistics – and instead *soft-core* particle interactions are introduced through continuous, repulsive, pairwise interaction potentials.

Assuming state-agnostic particle-particle interactions, the system energy contribution from all the pairwise interactions takes the form $U_\mathrm{I}(\rho_\mathrm{u}, \rho_\mathrm{b}) = U_\mathrm{I}(\rho)$. Since for particles obeying Boltzmann statistics the ω term in Equation (3.6) has no density dependence, we find that inter-particle interactions give rise to a density-dependent mean-field potential $E(\rho) = \partial_{\rho_i} U_\mathrm{I}(\rho)$[2]. For pairwise additive potentials, the system interaction energy is given by the sum over each inter-particle potential $u(d_{mk})$, between particles indexed by m and k and separated by a distance d_{mk}. Neglecting non-nearest neighbour interactions, and assuming a locally uniform distribution of particles, the separation distances between the particles can be approximated by the average local separation of particles – i.e. $d_{mk} \approx d = \sqrt{1/\rho}$. The interaction energy is then approximated as

$$U_\mathrm{I}(\rho) \approx \int \rho \frac{\xi}{2} u(d)\, dA, \tag{3.15}$$

where ξ denotes the number of nearest neighbours. From this energy, the interaction potential is found to be

$$E(\rho) = \frac{\delta U_\mathrm{I}(\rho)}{\delta \rho} \approx \frac{\xi}{2}\left(u(d) + \rho \frac{\mathrm{d}u(d)}{\mathrm{d}d}\frac{\mathrm{d}d}{\mathrm{d}\rho}\right) = \frac{\xi}{2}u\,(d) - \frac{\xi}{4}\frac{\mathrm{d}u\,(d)}{\mathrm{d}d}d. \tag{3.16}$$

This approach is, in principle, similar to the virial expansion, which is used to determine contributions to equations of state arising from particle interactions for low particle densities [146, Chapter 5.2], however it is, in contrast, suited to higher density systems where the assumed crystal-like particle distribution is more applicable.

2: This relation is also derived in the discrete limit in Appendix A.2, where I obtain $E(\rho) = \lim_{dx\to 0} \partial_{K_j} E_r(K_j + M_j)$ from the local many-body interaction $E_r(K_j + M_j)$ in the continuum limit of the coarse-graining box size $dx \to 0$.

Soft-core repulsion examples

Although Equation (3.16) holds for any short-range interaction, soft-core *repulsive* interactions are particularly interesting in the context of biological systems. As introduced in Chapter 1, examples of repulsive interactions include those arising from membrane-curvature, and shielded electrostatic interactions. Here, to demonstrate the validity of the approximations made above, I compare these two soft-core interaction types with the results of hard-core repulsion derived previously.

Although curvature mediated interactions are not pairwise additive, they can be approximated as such at low particle densities. Under

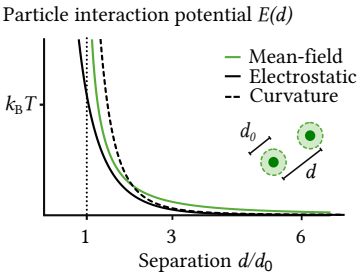


Figure 3.3: The hard-core interaction potential $E(\rho)$ from Equation (3.13) (green line) that accounts for steric interactions between particles of diameter d_0 takes a similar form to the potentials arising from soft-core repulsive curvature-mediated (solid black line) and shielded electrostatic interactions (dashed black line). Reproduced from [213], licenced under CC BY 4.0.

these conditions, the pairwise interaction potential is given by [147]

$$u_{\text{curvature}}(d_{mk}) = 8\pi\kappa\theta^2\left(\frac{d_0}{2d_{mk}}\right)^4, \tag{3.17}$$

where θ is the membrane contact angle, and κ denotes the membrane bending rigidity [216]. Using Equation (3.16), this results in the mean-field potential

$$E_{\text{curvature}}(d) = 24\pi\kappa\theta^2\frac{\xi}{2}\left(\frac{d_0}{2d}\right)^4. \tag{3.18}$$

Described by the Yukawa potential [151, 152], shielded electrostatic interactions similarly lead to an interaction energy

$$u_{\text{electrostatic}}(d_{mk}) = \frac{Q^2}{4\pi\epsilon_0}\frac{e^{-d_{mk}/\lambda_d}}{d_{mk}}, \tag{3.19}$$

with vacuum permittivity $\epsilon_0 = 55.2\text{e}^2\text{eV}^{-1}\mu\text{m}^{-1}$, which can be written in units of $K_\text{b}T$ as $\epsilon_0 = 2.14\times10^6\text{e}^2\text{k}_\text{b}\text{T}^{-1}\text{nm}^{-1}$, for $T = 300\,\text{K}$. The corresponding interaction potential then reads

$$E_{\text{electrostatic}}(d) = \frac{\xi}{2}\frac{Q^2}{4\pi\epsilon_0}\frac{e^{-d/\lambda_d}}{d}\left(\frac{3}{2}+\frac{d}{2\lambda_d}\right). \tag{3.20}$$

I compare these short-range soft-core repulsive interactions to hard-core repulsion (Equation (3.13)) in Figure 3.3, using $d_0 = 1\text{nm}$ as the typical order of magnitude for protein size [217, 218], and assuming the number of nearest neighbours to be $\xi = 6$. For the curvature-mediated interactions I further use a membrane bending rigidity of $\kappa = 20k_\text{B}T$ and membrane contact angle $\theta \approx \pi/24$, and for the electrostatic interactions I consider particles with four elementary charges, $Q \approx 4e$ and use a Debye length λ_d =0.8 nm [219, 220].

From this comparison I recover the expected result that such soft-core repulsive interactions can be well approximated by the hard-core lattice gas approximation. Given the suitability of the lattice framework for approximating soft-core short-range repulsion, I use Equation (3.13) to approximate all short-range repulsion in the remainder of this work.

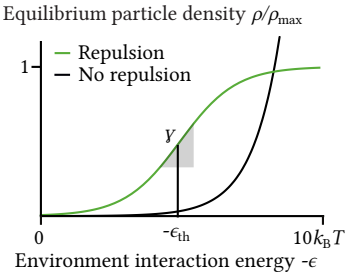


Figure 3.4: The response function resulting from short-range particle repulsion is a sigmoidal mapping (green line) between the binding energy field ϵ and the particle density ρ, which approaches a maximum density ρ_{max} for favourable binding energies in comparison to non-interacting particles (black curve). The gain γ and threshold of the filter ϵ_{th} are annotated by the shaded triangle and vertical line respectively. Reproduced from [213], licenced under CC BY 4.0.

Repulsion leads to sigmoidal filter

Having derived the form of the interaction potential for short range repulsion, it remains to determine the form of the resultant signal filter. Substituting the interaction potential from Equation (3.13), and using $\rho_{\text{max}} = 1/d_0^2$, the equilibrium particle distribution given by Equation (3.7)

becomes

$$\rho(\mathbf{r}) = \frac{1 + e^{-\beta\epsilon(\mathbf{r})}}{l^2 + d_0^2(1 + e^{-\beta\epsilon(\mathbf{r})})}. \tag{3.21}$$

This corresponds to a sigmoidal mapping between the binding energy field and the density field, as shown in Figure 3.4. At low density, the particles are distributed between the different positions according to ϵ. However, at high particle density, the volume-excluding nature of the particle-particle interactions limits the density, resulting in a plateaux for high input energy values. This is, in fact, akin to a Fermi-Dirac distribution where energy levels are spatially separated and the occupation numbers of these levels corresponds to the particle density distribution [146, Chapter 7].

The gain γ and the threshold ϵ_{th} that characterise this sigmoidal relation are given by

$$\gamma := \max\left(\left\|\frac{\partial\rho}{\partial\epsilon}\right\|\right) = \frac{\beta l^2}{4d_0^2(d_0^2 + l^2)}, \tag{3.22}$$

and

$$\epsilon_{\text{th}} := \operatorname{argmax}\left(\left\|\frac{\partial\rho}{\partial\epsilon}\right\|\right) = k_B T \ln\left(\frac{d_0^2}{d_0^2 + l^2}\right), \tag{3.23}$$

and depend on two length scales: the effective particle size d_0, and the average particle separation across the whole surface $\bar{d} = \sqrt{A/N}$, see the annotations in Figure 3.4. The second of these length scales is introduced through the integration constant l, which is set by the membrane area and the total number of embedded particles via the constraint in Equation (3.8). The explicit expression for l – for uniform ϵ – follows $l \propto \sqrt{\bar{d}^2 - d_0^2}$ in the hard-core interaction case.

The square-ratio of the two length-scales, $d_0^2/\bar{d}^2$, gives the fraction of the membrane surface covered by the particles – i.e. the membrane saturation. By tuning this saturation, the response function can be adjusted. As presented in Figure 3.5, using as an example a linear input binding energy field of $\epsilon(x) = k_B T(20x/L - 10)$ on a domain of length L, varying the membrane saturation shifts both the threshold and gain[3]. For small, or few, particles in the membrane the threshold energy is low, leading to small high particle density regions. In comparison large, plentiful, particles lead to an almost saturated membrane where the high and low density regions differ only slightly in particle density, leading to response functions with a low gain.

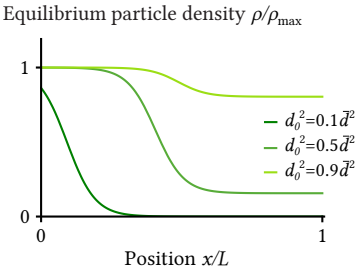


Figure 3.5: For an input binding energy field $\epsilon(x) = k_B T(20x/L - 10)$ on a domain of length L the distribution of short-range repulsive particles takes the form of a sigmoid that depends on the ratio of the effective particle size d_0 and the average particle separation $\bar{d}$.

In the context of signal processing, the particle distributions resulting from repulsive particle interactions can be understood to sigmoidally *filter* external heterogeneities, effectively providing a physical mechanism for binarising the binding-energy field. This is reminiscent of 'thresholding filters' in image analysis and computer vision algorithms, which calculate a binary compression of an initial input and are used to select features of

3: This choice of example input permits clear visualisation and considers a biologically achievable energy range, since typical chemical-interaction potentials between biological proteins are on the order of $-10k_BT$ [221].

the input image for further downstream processing [34]. As discussed in Chapter 1, such sigmoidal filters are useful and common both in biological systems and in synthetic signal-filtering applications [27, 33, 34].

3.2.2 Binding site distributions influence response functions

In addition to the length scales d_0 and $\bar{d}$, the sigmoidal response function given by Equation (3.21) also depends on the spatial organisation of the interaction sites, which in biological systems are commonly distributed non-uniformly. How would the thresholding capability of the filter described above be influenced by such changes in the binding site organisation?

Low-density uniform binding sites

The previous section detailed the sigmoidal signal filter that results from repulsive interactions when binding sites are uniformly distributed with the same density as the maximum density of the particles, such that each particle can exist in either a bound or unbound state (using the assumption $\rho_{\text{ext}} = \rho_{\max}$ in Section 3.2.1). One clear generalisation is to therefore relax this constraint, and consider uniformly distributed, but lower density, binding sites.

Introducing a uniform area fraction of binding sites $f_{\text{BS}} = \rho_{\text{ext}}/\rho_{\max}$, the substitution of Equations (3.14) and (3.13) into (3.2) yields the total general equilibrium particle density

$$\rho(x) = \frac{\rho_{\max}}{1 + e^{\alpha+\beta\epsilon_{\text{u}}}}\left(1 + f_{\text{BS}}\frac{e^{\alpha+\beta\epsilon_{\text{u}}}}{e^{\beta\epsilon(x)} + 1 + e^{\alpha+\beta\epsilon_{\text{u}}+\beta\epsilon(x)}}\right). \tag{3.24}$$

Comparing this result to Equation (3.21), again using a linear example input field for visualisation purposes, I find that reducing the binding site area fraction results in a smaller difference between the maximum and minimum densities, and a correspondingly smaller gain (Figure 3.6).

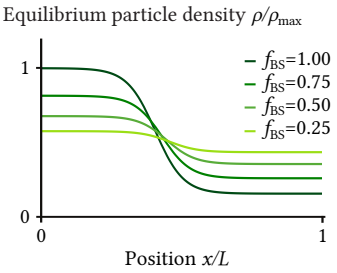


Figure 3.6: Decreasing the area fraction f_{BS} of the membrane covered by *uniformly* distributed binding sites leads to a reduction in the gain of the response function resulting from repulsive particle interactions, as visualised here for an example input energy field of $\epsilon(x) = (20x/L - 10)k_{\text{b}}T$ over a domain of length L, and length scales satisfying $d_0^2 = 0.5\bar{d}^2$. The $f_{\text{BS}} = 1$ condition corresponds to the medium green line in Figure 3.5. Reproduced from [213], licenced under CC BY 4.0.

Linear distributions of binding sites

Living systems rarely have uniform binding sites. Indeed, membrane-bound proteins often interact with slender polymers such as cytoskeletal filaments [222]. To identify, how signal transmission is influenced when particles bind to such filamentous structures, I consider binding sites

confined to ζ parallel lines of length L that cover a fraction $\phi = \zeta L d_0/A = \zeta d_0/L$ of the total membrane surface, where d_0 is the width of a single particle.

The line densities ρ_{1D} along these parallel strips of membrane are the one-dimensional equivalents to Equation (3.21),

$$\rho_{1D}(x) = \frac{1 + e^{-\beta\epsilon(x)}}{l + d_0(1 + e^{-\beta\epsilon(x)})}, \tag{3.25}$$

where the integration constant l is obtained, as before, by enforcing the particle conservation condition from Equation (3.8). Assuming the input energy is the same for all filaments, this condition takes the form

$$\frac{\phi}{L d_0}\left(\int_0^L dx \rho_{1D}(x)\right) + \frac{1-\phi}{l^2 + d_0^2} = \frac{1}{\bar{d}^2}, \tag{3.26}$$

in which x parametrizes the coordinates along the ζ lines, and the interaction energy along the binding lines is given by $\epsilon = \epsilon(x)$ and is infinite elsewhere. When $\phi = 1$, I therefore recover the same results as in the uniform binding site case discussed in Section 3.2.1.

Assuming the line densities will be later read out along their respective binding filaments, I can compare the signal filters for input signals encoded in filamentous binding site distributions with the two-dimensional case by normalising by the maximum densities $\rho_{\max,1D} = 1/d_0$ and $\rho_{\max} = 1/d_0^2$ respectively. In Figure 3.7, I make this comparison using, as before, a linear input binding energy field to aid visualisation. I find that lowering ϕ has a similar impact as decreasing $\bar{d}$, and as such results in a signal filter that is only achievable with more particles in the case of uniform binding sites. This occurs because the non-binding membrane region acts as a particle bath that buffers the particles used in the binding regions. Interestingly, these results imply that – in systems where there is a high cost to creating extra particles, as is the case for proteins in living systems – filaments should be favoured over surfaces for transmitting spatial information across interfaces due to the lower particle number requirement [223, 224].

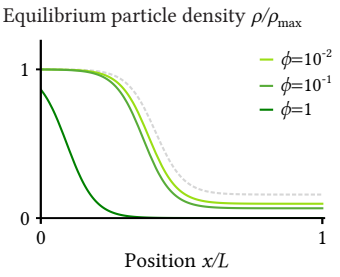


Figure 3.7: Filamentous binding site distributions enable similar response functions to be produced with a lower number of membrane-associated particles than for two-dimensional ($\phi = 1$) distributions. Shown, as before, using an example input energy field of $\epsilon(x) = (20x/L - 10)k_bT$ over a domain of length L, and with length scales satisfying $d_0^2 = 0.1\bar{d}^2$, reducing the fraction ϕ of the surface covered by lines of binding sites has the same effect as decreasing $\bar{d}$ or increasing d_0. The $\phi = 1$ condition corresponds to the dark green line in Figure 3.5, the $d_0^2 = 0.5\bar{d}^2$ line from the same figure is shown here for comparison (dashed grey line).

3.2.3 Attractive interactions

In the preceding sections, I have detailed how purely repulsive particle-particle interactions lead to sigmoidal signal filters, which binarise input binding energy fields. A natural extension to this is to ask how non-repulsive interactions influence particle-mediated signal filters.

Intuitively, attractive interactions between particles lead to their enhanced accumulation in high-density regions. Indeed, I demonstrate this en-

hancement by considering a system of particles undergoing short-range attraction coupled to shorter-range repulsion or volume exclusion. This shorter range repulsion prevents the unphysical divergence of particle density in energetically favourable regimes. Since introducing such interactions opens a wide field of possibilities, I will not attempt to address these in their full generality but will instead provide an illustrative example. Suppose that I have a set of Van der Waals particles [225, Chapter 17] that undergo particle-particle interactions as introduced in Chapter 1. Permitting the particles to also bind to an external structure with the binding energy field $\epsilon(\mathbf{r})$, the Van der Waals free energy density given by Equation (1.7) can be extended to account for free energy contributions from both the bound and unbound particle distributions, described by the densities ρ_b and ρ_u respectively. This results in a free energy density of the form [146, Chapter 5.5],

$$\beta f_\mathrm{VdW} = \rho_\mathrm{u}\left[\ln\left(\frac{C\rho_\mathrm{u}}{1-d_0^2\rho}\right)-1\right]+\rho_\mathrm{b}\left[\ln\left(\frac{C\rho_\mathrm{b}}{1-d_0^2\rho}\right)-1\right]-\beta a\rho^2+\beta\epsilon(\mathbf{r})\rho_\mathrm{b}, \tag{3.27}$$

as shown in Figure 3.8(a). As described in Chapter 1, in the Van der Waals description, d_0 defines the length scale of the repulsive interaction, and thus the maximum particle density $\rho_\mathrm{max} = 1/d_0^2$ as before, and a defines the magnitude of the attractive interaction. The constant C arises due to the discretisation of phase space, and does not effect the subsequent particle distributions considered here [146, Chapter 5.5]. The corresponding chemical potentials for the bound and unbound particles are given respectively by

$$\begin{aligned}\beta\mu_\mathrm{VdW,u} &= \left[\ln\left(\frac{C\rho_\mathrm{u}}{1-d_0^2\rho}\right)+\frac{d_0^2\rho}{1-d_0^2\rho}-2\beta a\rho\right]\\ \beta\mu_\mathrm{VdW,b} &= \left[\ln\left(\frac{C\rho_\mathrm{b}}{1-d_0^2\rho}\right)+\frac{d_0^2\rho}{1-d_0^2\rho}-2\beta a\rho\right]+\beta\epsilon(\mathbf{r}).\end{aligned} \tag{3.28}$$

As shown in Equation (3.3), the equilibrium ratio of bound to unbound particles is set by the difference in the adjusted chemical potentials $\Delta\tilde{\mu} = \epsilon(\mathbf{r})$ such that $\rho_\mathrm{b} = \rho_\mathrm{u}e^{-\beta\epsilon(\mathbf{r})}$.

Assuming that the binding and unbinding rates are much faster than particle diffusion, such that the ratio of bound and unbound particles is given by the equilibrium condition, the free energy can be given in terms of the total density ρ and $\epsilon(\mathbf{r})$ (see Figure 3.8(b)). Taking the partial derivative of this free energy with respect to the total particle density ρ is then governed by the effective chemical potential

$$\beta\mu_{\mathrm{VdW},\rho} = \ln\left(\frac{C\rho}{1-d_0^2\rho}\right)-\ln\left(1+e^{\beta\epsilon(\mathbf{r})}\right)+\frac{d_0^2\rho}{1-d_0^2\rho}-2\beta a\rho+\beta\epsilon(\mathbf{r}) = \text{constant}. \tag{3.29}$$

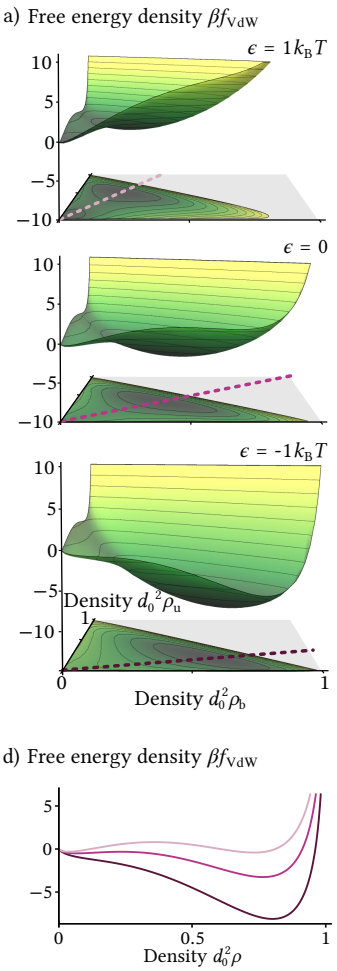


Figure 3.8: The free energy density of Van der Waals particles at any given point in the membrane domain depends on the binding energy at that location. a) Free energy densities for binding energies $\epsilon = -1k_\mathrm{B}T$, 0, $1k_\mathrm{B}T$ [top to bottom] – which depend on both the bound ρ_b and unbound ρ_u particle densities – and their projections onto the ρ_b-ρ_u plane. The pink dashed lines indicate the equilibrium ratio of bound to unbound particles for each energy condition. b) The free energy density for rapidly equilibrating binding dynamics depends on the total density $\rho = \rho_\mathrm{b}+\rho_\mathrm{u}$ and the binding energy. The line colours indicate match the dashed lines in the (a) corresponding to the same binding energy. Parameters: $a = 5d_0^2k_\mathrm{B}T$, $C = 50d_0^2$

Short-range attractive interactions are known to result, under certain parameter configurations, in the phase separation of particle systems into high- and low-density regions. The conditions for phase separation can be found by determining high- and low-density values that satisfy the coexistence requirements $\mu_{\mathrm{VdW},\rho}(\rho_1) = \mu_{\mathrm{VdW},\rho}(\rho_2)$ and $p(\rho_1) = p(\rho_2)$, where

$$p = \rho\mu_{\mathrm{VdW},\rho}(\rho) - f(\rho) = \frac{k_B T\rho}{1 - d_0^2\rho} - a\rho^2 = \frac{\rho(1 + a\beta d_0^2\rho^2)}{1 - d_0^2\rho} \tag{3.30}$$

is the pressure of the particle gas [146, Chapter 5.5]. These conditions trace out the 'binodal' curve (sold line, Figure 3.9(a)), which encloses the region in which two phases are permitted in the system. Similarly, the region of bi-stability, where spontaneous phase separation occurs, is enclosed within the 'spinodal' curve (dashed line, Figure 3.9(a)), defined by the extrema of the chemical potential [226]. Systems starting with a uniform particle density, ρ_{uniform}, that fall within the phase-separating region of parameter space will automatically separate into two regions with densities on the binodal curve. This uniform density corresponds to the number of particles in the membrane divided by the membrane area, and is therefore determined by the length-scale $\bar{d}$ giving the average particle separation in the membrane, i.e. $\rho_{\mathrm{uniform}} = N/A = 1/\bar{d}^2$. Therefore, as in the previous sections, the length-scale $\bar{d}$ plays an important role in determining the form of the particle-mediated signal filter.

Interestingly, the pressure given by Equation (3.30) is independent of the binding energy, and changes in the binding energy only changes the chemical potential by a magnitude shift. Therefore, both the binodal and spinodal curves are unaffected by the binding energy field. This means that the conditions for phase separation are not affected by the guiding of the phase separation by $\epsilon(\mathbf{r})$. Indeed, without particle binding (i.e taking the limit $\epsilon \to \infty$) results in identical binodal and spinodal curves to those presented in Figure Figure 3.10. Furthermore, the size of the resulting phase regions is also unchanged by the binding energy field, since it is determined by the number of particles in the system (as shown in Figure 3.9 (b)) – similar to how the integration constant l is determined by total particle number in Section 3.1. In contrast, the binding interactions do influence the *positions* of the phase regions. Numerically solving $\beta\mu_{\mathrm{VdW},\rho} = 0$ for the equilibrium particle density field given various input interaction energy fields $\epsilon(\mathbf{r})$, I find that the high- and low-density phases coincide with the low-and-high binding energy regions respectively, see Figure 3.10. Compared to the results of Section 3.2.1 (Figure 3.5), I find that the attractive particle interactions in this system result in a higher gain, which would enable a more effective binarisation of the input field.

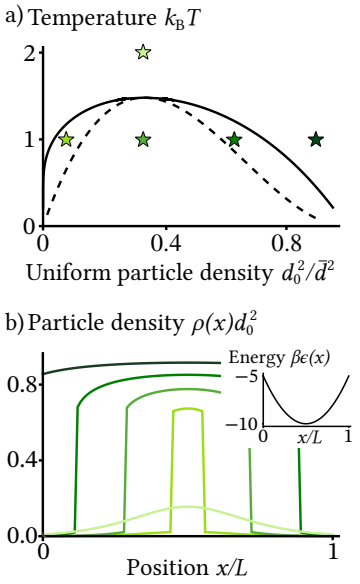


Figure 3.9: In the presence of binding interactions, Van der Waals particles phase separate into a high density and low density region, with the position of the high density region determined by the input binding energy field. a) Binodal (solid line) and spinodal (dashed line) curves for the Van der Waals gas. The stars indicate the locations for the example systems explored in (b). b) Example equilibrium particle distributions for a quadratic input of $\epsilon = (20(x/L - 0.5)^2 - 10)k_B T$ (shown in inset) and initial uniform particle densities of $1/\bar{d}^2 = 0.075/d_0^2, 0.33/d_0^2$, $0.63/d_0^2$, and $0.9/d_0^2$ at a thermal energy of $1k_B T$, and $1/\bar{d}^2 = 0.33/d_0^2$ at a thermal energy of $2k_B T$, indicated by colours ranging from black to light green respectively. Phase separation occurs below the binodal curve and the size of the high-density region is determined by the total particle density. Parameters: $a = 5d_0^2 k_B T$, $C = 50d_0^2$

Without binding interactions, attractive particles are known to spontaneously separate into these two phases, with the initial locations of the phases determined by the random fluctuations that drive the initial phase separation. Over time, these phases undergo Ostwald ripening until only one high- and one low-density region are left [227], whose location depends on both the random initial patterning and the boundary conditions of the system. In contrast, when binding interactions are present, the phases are positioned to minimise the system free energy, and as such the translational symmetry of the system is broken with the high-density phase pinned to the energetically favourable regions. This process arrests the Ostwald ripening of the phases when the effective surface tension from particle attraction is combatted by the effective attraction of particles to energetically favourable regions. This phase separating process, when coupled to heterogeneous features in the environment, is called controlled or *guided* phase separation, and has been discussed both in the context of small-scale pattern design and in regards to understanding phase-separation mediated biological processes including DNA strand repair and cellular division [160, 162, 163, 228]. My results suggest that such phase separating systems could act as a binarising surface filter in spatially-resolved signal processing cascades, and would perform better than filters arising from purely repulsive interactions. However, I find that guided phase separating systems arising from the inclusion of attractive interactions do not produce more complex particle-mediated response functions than purely-repulsive ones discussed above.

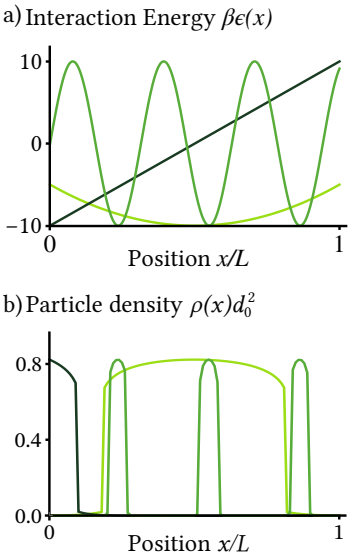


Figure 3.10: Guided phase separation thresholds various input binding energy fields a) Example input binding energies of liner $\beta\epsilon = 20x/L - 10$ (dark green), sinusoidal $\beta\epsilon = 10\sin(20x/L)$ (medium green), and quadratic $\beta\epsilon = 20(x/L - 0.5)^2 - 10$ (light green) used to explore guided phase separation of Van der Waals gas, on a domain of length L. b) Output density profiles resulting from the inputs in (a), with corresponding line colours, for the total domain particle density $1/\bar{d}^2 = 0.33/d_0^2$. Parameters: $a = 5d_0^2 k_B T$, $C = 50d_0^2$

3.2.4 Edge detection by particle distributions

I have demonstrated that particle interactions can lead to sigmoidal signal filters, but is it possible to design particle-density mappings that are non-monotonic in binding energy such that we could select other features of the input binding energy signal? Such a mechanism would be crucial in the design of useful particle-mediated signal processors.

Total density measurements result in monotonic response functions

To have non-monotonic filters, it is required that two values of the binding energy ϵ result in the same density value. Assuming that only the total particle density is accessible by downstream processes, and also that the inter-particle interactions depend only on the total density, I return to Equation (3.3) and, using Equation (3.5), find that

$$e^{\beta\Delta E(\rho)}(\rho e^{\beta E_u(\rho)} - 1) = e^{-\beta\epsilon}, \tag{3.31}$$

where the right-hand-side is a monotonic function of the binding energy.

As defined in Chapter 2, Equation (2.46), the adjusted chemical potential $\tilde{\mu}$ is the functional derivative of the system free energy minus $k_B T \log(\rho)$. As such, the contributions $E_i(\rho)$ must be single-valued for any value of the density ρ. This means that the left hand side of equation (3.31) is single valued for any particular value of ρ. Combined with the monotonic form of the right-hand-side, this implies that the mapping from ϵ to ρ will always be monotonic, and as such could at most only provide a thresholding-like filter.

Distinguishing ρ_u and ρ_b allows for edge detection

However, when relaxing the condition that the bound and unbound particles are indistinguishable to downstream processes, it is possible to form non-monotonic filters. The downstream differentiation of bound and unbound particles is in fact common in biological systems, where molecules frequently exhibit allosteric properties, whereby the process of binding leads to a conformational change in the molecule and alters other molecule properties [229]. This allosteric regulation can lead to bound and unbound particles exhibiting different interaction properties with other molecules that could constitute the readout mechanism for the signal filter.

From Equation (3.2), the equilibrium particle distributions for the bound and unbound particles in equilibrium are

$$\rho_u = \frac{\rho}{1 + e^{-\beta\epsilon}}, \qquad \rho_b = \frac{\rho}{1 + e^{\beta\epsilon}}. \tag{3.32}$$

Therefore, even though species-agnostic interaction potentials lead to a monotonic mapping of the input energy ϵ to total particle density ρ, the mappings to the unbound particle density ρ_u are not necessarily monotonic. For instance, the examples in Figure 3.10 result in – albeit small – peaks of unbound particles at the edges of the high-density region, see Figure 3.11. This constitutes a form of edge detection. The peaks of unbound particles have a larger density, although with less prominent peaks, when considering less favourable binding dynamics where particles spend more time in their unbound states (Figure 3.12). This edge detection filter is therefore most suited to situations in which the edges have interaction energies close to zero. Other biologically inspired edge detection methods – including an algorithm using multiple genetic circuits in ecoil [230], and light-sensitive neuronal networks [231–233] – have focused on mechanisms to identify image boundaries at the multi-cell scale. However, at the sub-cellular level this has been less well

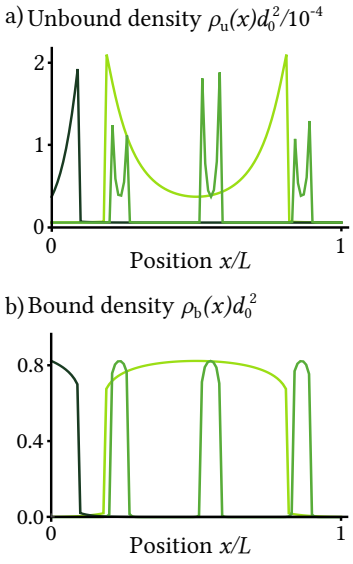


Figure 3.11: Distinguishing between bound and unbound particles permits edge detection. a) The unbound particle densities resulting from the same example conditions as in Figure 3.10 (colours indicating the same input energy profiles) display peaks at the edges of the high particle density regions. b) Due to small magnitudes of unbound particles, the bound particle densities closely resemble the total particle densities in Figure 3.10(b).

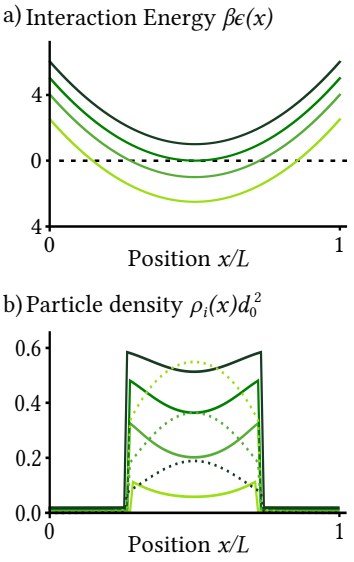


Figure 3.12: The density of edge-detecting unbound particles is larger for less favourable binding energies. a) Various examples of quadratic binding energy inputs of the form $\beta\epsilon = 20(x/L - 0.5)^2 - \beta\epsilon_0$ where $\beta\epsilon_0 = -10, -5, 0, 5, 10$, coloured from dark to light green respectively. b) Unbound (solid) and bound (dashed) particle densities resulting from these input energy profiles, with corresponding colours, for a total membrane density of $1/\bar{a}^2 = 0.33/d_0^2$. Parameters: $a = 5d_0^2 k_B T$, $C = 50d_0^2$

explored. Such a mechanism therefore provides a possible means for spatially-resolved signal computation at subcellular scales.

It would be interesting to ask if more complicated or more tunable filters arise when considering other particle-particle interactions, and it would be an exciting future project to investigate how edge-detecting, and other complex filter systems, could arise from species-dependent interactions that further couple ρ_u and ρ_b through $E_i(\rho_u, \rho_b)$.

3.3 Concluding remarks

In summary, I have detailed how simple interactions between membrane-associated particles and binding sites in their surroundings can form signal filters that permit the selective transmission of environmental heterogeneities into particle density distributions, and have explored how the non-linear amplification of the transmitted signals is controlled by the interactions between particles in the plane of the interface.

In particular, I find that short-range repulsive particle interactions produce sigmoidal mappings between non-uniform binding energy fields and the particle densities that form in response – providing a physical mechanism for the *binary thresholding* of the input energy field. These filters are shown to depend on the effective interaction range of the particles and the number of particles relative to interface size. Such filters may play key roles in the subcellular processing of spatial signals, akin to how similar thresholding filters are used in control circuits and neural networks to introduce non-linearities and compress outputs [30, 31].

By considering how changes in the distribution of binding sites influence the thresholding nature of these mappings, I demonstrate that a reduction of the dimension of the input from a two-dimensional distribution of binding sites to filamentous input structures improves the feasibility of these filters in a biological context. The thresholding nature of these filters is further shown to be improved by additional attractive particle interactions that lead to phase separating systems.

Extending beyond monotonic response functions, I demonstrate that the allosteric nature of typical biological molecules further permits edge detection by facilitating the distinction between bound and unbound particle states.

These results hint at the possibility of diverse particle-mediated signal filters, which, when coupled to further downstream signal processing layers may permit subcellular systems to undertake complex computations in order to respond to external stimuli. The investigation of such downstream cellular processes would be an interesting extension of this

work, perhaps leading to the development of pattern recognition-like functions at compartment interfaces. With the advent of methods for developing synthetic vesicle systems, and progress in creating increasingly small soft-robotic systems, these subcellular signalling pathways may aid in the design of synthetic pattern-based computation and in the bottom up study of natural signal processing mechanisms [234].

Quantifying selective information processing

4

As introduced in the previous chapter, the formation of inhomogeneous protein distributions – guided by the interactions of the particles with structures in their surroundings – gives rise to signal filters that selectively transmit signals encoded within particle-environment interaction energy fields into particle density distributions. These signal filters are described by the mean-field mappings, i.e. the filter response functions, between the input energy and output density fields. However, it is important to assess how effectively these particle-mediated signal filters transmit information when accounting for the noise in the innate particle dynamics. In this chapter I therefore quantify the information that output densities carry about their input interaction energies, given both noisy particle dynamics and the compression resulting from the filter's response function. In particular, I focus on the repulsion-mediated signal filter derived in Section 3, whose response function, Equation (3.21), I provide here for reference:

$$\rho(\mathbf{r}) = \frac{1 + e^{-\beta\epsilon(\mathbf{r})}}{l^2 + d_0^2(1 + e^{-\beta\epsilon(\mathbf{r})})},$$

where d_0 is the effective particle size due to repulsion and determines the maximum particle density $\rho_{\max} = 1/d_0^2$, and l is an integration constant set by the average particle separation in the membrane $\bar{d}$ through Equation (3.8). The fraction of membrane covered by particles is given by the ratio $d_0^2/\bar{d}^2$, which can therefore be interpreted as a measure for how saturated the membrane is.

Part of the results presented in this chapter are included in a paper accepted at PRL [213]. The computations for in this chapter were implemented in Mathematica [235] and Python, the notebooks and scripts for which are provided in the repositories [214] and [215].

4.1 Mutual Information

The thresholding filter introduced in Chapter 3 is a non-local response function, such that the output field depends on the input signal over the whole membrane domain. Therefore, to evaluate the level of signal compression by the particle-mediated sigmoidal filter, it is informative to ask how much of the information in the *whole* input signal is transmitted to the *whole* output field. As discussed in Chapter 1, the amount of information transmitted by a filter is given by the mutual information, Equation (1.1), which for the continuous fields ρ and ϵ is defined as:

Definition 4.1.1 (Continuous mutual information) *The continuous mutual information between two random fields $\boldsymbol{\epsilon}$ and $\boldsymbol{\rho}$ is given by*

$$I = \iint \mathcal{D}\boldsymbol{\epsilon}\mathcal{D}\boldsymbol{\rho}\, p(\boldsymbol{\rho},\boldsymbol{\epsilon}) \ln \frac{p(\boldsymbol{\rho},\boldsymbol{\epsilon})}{p(\boldsymbol{\rho})p(\boldsymbol{\epsilon})}, \tag{4.1}$$

where the integral is a functional integral over all possible input and output fields, and $p(\boldsymbol{\rho},\boldsymbol{\epsilon})$ is the joint probability density of observing an output field $\boldsymbol{\rho}$ and an input field $\boldsymbol{\epsilon}$ from which the marginal probability densities $p(\boldsymbol{\rho}) = \int \mathcal{D}\boldsymbol{\epsilon}\, p(\boldsymbol{\rho},\boldsymbol{\epsilon})$ and $p(\boldsymbol{\epsilon}) = \int \mathcal{D}\boldsymbol{\rho}\, p(\boldsymbol{\rho},\boldsymbol{\epsilon})$ may be calculated.

This 'whole-field' quantification differs from many common information quantification approaches in, for example, biological pattern-forming systems where studies have focused on the information transmitted by biochemical signals at singular locations in the field [16]. Such single-location quantifications are relevant in the context of cell fate signalling, where chemical gradients on the tissue scale inform individual cells of their position in the tissue through only a local measurement of the chemical signal. In contrast, the signal filters described in Chapter 3 output particle distributions that depend on the whole of the input interaction field, rather than at one location. The whole-field mutual information therefore provides a means to assess the information transmission in these non-local cases.

4.1.1 Membrane discretisation

Although the continuous mutual information in Equation (4.1) provides a means of quantifying information transmission, the functional integrals – in combination with the constraint on particle number in Equation (3.8) – present an analytical challenge. This is overcome by discretising both the spatial domain and the possible values of the input and output fields.

Discretising the two-dimensional membrane into boxes indexed by $j = 1, 2, \ldots, B$, I represent the particle density and interaction energy fields as random vectors which take the realizations $\{\rho_j\}$ and $\{\epsilon_j\}$, whose elements are sampled from finite sets of density values $\mathscr{P}$ and energy states $\mathscr{E}$, see Figure 4.1. The mutual information between the random vectors is then given by the discrete mutual information [82, Chapter 2] .

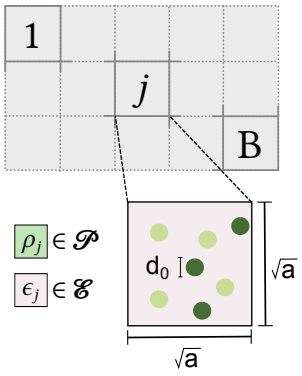


Figure 4.1: A membrane surface (grey shading) is discretised into B boxes of area a indexed by j. In each box, the interaction energy and particle density may take any value from a finite set of possible values, $\mathscr{E}$ and $\mathscr{P}$ for energy and particle densities respectively.

Definition 4.1.2 (Discrete mutual information) *The discrete mutual*

information between random vectors $\{\epsilon_k\}$ and $\{\rho_j\}$ is given by

$$I = \sum_{\{\epsilon_k\}} \sum_{\{\rho_j\}} P(\{\rho_j\}, \{\epsilon_k\}) \ln \frac{P(\{\rho_j\}, \{\epsilon_k\})}{P(\{\rho_j\}) P(\{\epsilon_k\})}, \tag{4.2}$$

where $\sum_{\{\epsilon_k\}} = \sum_{\epsilon_1} \sum_{\epsilon_2} \dots \sum_{\epsilon_B}$ and $\sum_{\{\rho_j\}} = \sum_{\rho_1} \sum_{\rho_2} \dots \sum_{\rho_B}$ are the equivalent of the functional integrals in Equation (4.1) and denote sums over all possible realizations of the input and output random vectors respectively. The marginal probabilities $P(\{\rho_j\}) = \sum_{\{\epsilon_k\}} P(\{\rho_j\}, \{\epsilon_k\})$ and $P(\{\epsilon_k\}) = \sum_{\{\rho_j\}} P(\{\rho_j\}, \{\epsilon_k\})$, are given by the sum of the joint probability over all possible input or output vectors respectively.

Equation (4.2) gives the information that is gained about the input field by observing the output density field, providing a clear measure for signal filter performance.

4.2 Probability distributions describe filter noise

Since the mutual information quantifies the ability of a system to distinguish between different inputs by measuring the output $\{\rho_j\}$, the event space of possible inputs must be the same as the support of the input distribution $P(\{\epsilon_k\})$ which spans all the possible input signals that the system may receive. This distribution would therefore take a system-specific form. For example – in the context of biological signal processing – a cell may need to distinguish between different polarity directions in which case both directions must be included within the support of $P(\{\epsilon_k\})$.

In the following, I confine the input values ϵ_k to a finite, discrete set of interaction energies $\mathfrak{E}$. To remain as system-agnostic as possible, I then consider a uniform distribution of the input vectors, such that $P(\{\epsilon_k\}) = \prod_k P(\epsilon_k)$ with the local values ϵ_k independently sampled from a uniform distribution $P(\epsilon_k) = 1/|\mathfrak{E}|$ over the set $\mathfrak{E}$ of possible energy values. This allows me to assess the capability of the filter to transmit information, without bias for a particular input form.

Thermal input noise It is important to note that the distribution of inputs is not accounting for thermal fluctuations of the input, but rather differences in scenarios a system may experience. Thermal fluctuations would be expected to lead to slight deviations about the inputs considered here, and may be mapped into an additional 'channel noise' of the filter. This extension is not addressed here.

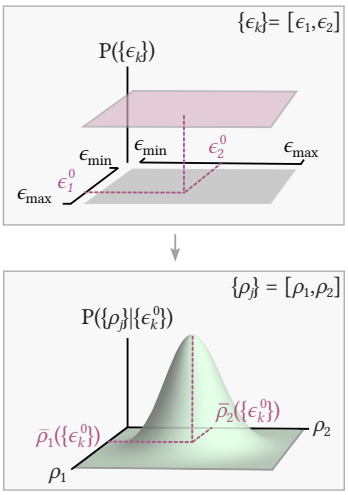


Figure 4.2: To calculate the mutual information, the realisations of the input random vector of interaction energies $\{\epsilon_k\}$ are distributed uniformly over the space of all possible inputs, given the set of energy values [pink, shown here for $B = 2$]. The corresponding conditional distribution of particle density vector realisations $P(\{\rho_j\}|\{\epsilon_k^0\})$ takes the form of a multi-dimensional Gaussian distribution about the density expected for the particular input realisation $\{\epsilon_k^0\}$of the energy vector [green]. Reproduced from [213], licenced under CC BY 4.0.

4.2.1 Particle-density noise

The mean-field signal-filter models in Chapter 3 provide the expected output density profile for a given input. In the discrete framework outlined above, this may be denoted $\{\bar{\rho}_j\} = \{\bar{\rho}(\epsilon_j, l)\}$, where $\bar{\rho}(\epsilon_j, l)$ is given by Equation (3.21) and the integration constant l introduces a dependence on the full vector $\{\epsilon_j\}$ through the particle conservation constraint, Equation (3.8). In the discrete limit this constraint takes the form

$$\sum_j \bar{\rho}_j(\epsilon_j, l) = \frac{B}{\bar{d}^2}, \tag{4.3}$$

where, as before, $\bar{d} = \sqrt{A/N}$ sets the number N of particles present in the whole membrane surface of area A. However, the observed particle distribution will vary from this expected value due to, for example, the thermally driven rearrangements of particles. In the context of signal transmission, these particle density fluctuations act as a 'channel noise' that limits the capacity of the filter to transmit information. Measured realisations of the density vector can then be considered as having been sampled form a distribution with a mean density profile $\{\rho_j\}$.

In this section, I first derive an approximate conditional probability distribution for the density vectors given an input vector $\{\epsilon_k\}$, and then assess the validity of my results by estimating particle distributions using Metropolis Hastings sampling.

Gaussian approximation of particle fluctuations

Denoting the number of particles in each box by $N_j = \rho_j a$, where a is the coarse-graining area, the system-dependent conditional probability $P(\{\rho_j\}|\{\epsilon_j\}) = P(\{N_j\}|\{\bar{N}_j\})$ is given by the probability of having box occupation numbers $\{N_j\}$ given expected occupation numbers $\{\bar{N}_k\}$, where $N_j = a\rho_j$ and $\bar{N}_j = a\bar{\rho}_j$. In this section, I derive this conditional probability, starting from the distribution of occupation number $P(\{N_j\})$. In the microcanonical ensemble, the distribution of occupation numbers $P(\{N_j\}) \propto e^{S/k_\mathrm{B}}$ is given by the total – system and reservoir – entropy $S = k_\mathrm{B} \ln W$ [146, Chapter 4]. As in Chapter 3, the entropy is given by the number of microstates W that make up the macrostate defined by $\{N_j\}$.

Assuming the 'readout' process that follows downstream of the particle patterning does not distinguish between bound and unbound particles, the difference between the particle species may be neglected and the number of microstates given by

$$W = \prod_j \frac{\Omega!}{(\Omega - N_j)! N_j!}, \tag{4.4}$$

where $\Omega = a/d_0^2$ is the number of spaces a particle can occupy in the box. Applying the Stirling approximation, this results in the conditional probability

$$P(\{N_j\}) \propto \prod_j e^{\Omega(\ln\Omega-1)-(\Omega-N_j)(\ln(\Omega-N_j)-1)-N_j(\ln N_j-1)}. \tag{4.5}$$

The conditional distribution may then be found by applying a saddle-point approximation, in the limit of large occupation numbers, about the expected occupation number. By imposing that the extrema of Equation (4.5) is given by $\{\bar{N}_j\}$, this expansion effectively accounts for the interaction energy contributions which were introduced explicitly in Chapter 3, and allows for the estimation of the fluctuations about this expected value[1]. This results in the conditional distribution obeys [236]

$$\ln P(\{N_j\}) = \sum_j \frac{1}{2}\left(\frac{\partial^2 \ln P}{\partial N_j^2}\bigg|_{N_j=\bar{N}_j}(N_j-\bar{N}_j)^2\right), \tag{4.6}$$

in which the first-order derivatives vanish, as well as the cross-term $\partial \ln P(\{N_j\})/(\partial N_j \partial N_{k\neq j})$.

1: Indeed, in later Chapter 5, I present a more involved derivation of density probability that does include both bound and unbound particles, and indeed recovers the results of this section in the correct limit.

Evaluating the partial derivatives of Equation (4.5) with respect to N_j and recalling that $\Omega = a/d_0^2$, $N_j = a\rho_j$, and $\bar{N}_j = a\bar{\rho}_j$, results in the conditional probability

$$P\big(\{\rho_j\}|\{\epsilon_k\}\big) \propto \prod_j P\big(\rho_j|\{\epsilon_k\}\big) \tag{4.7}$$

where the *local* conditional probability, $P\big(\rho_j|\{\epsilon_k\}\big)$, of a density ρ_j in box j is independent of the other boxes, and takes the form of a Gaussian distribution,

$$P\big(\rho_j|\{\epsilon_k\}\big) \propto \exp\left[-\frac{\big(\rho_j-\bar{\rho}_j\big)^2}{2\sigma_{\rho,j}^2}\right], \tag{4.8}$$

with mean $\bar{\rho}_j = \rho\big(\epsilon_j, l\big)$, given by Eq. (3.21), and standard deviation

$$\sigma_{\rho,j}(\bar{\rho}_j) = \sqrt{\frac{(1-d_0^2\bar{\rho}_j)\bar{\rho}_j}{a}}. \tag{4.9}$$

It is important to note that the introduction of spatial couplings between particles at different positions would further modulate the probability given by Equation (4.8) and introduce correlations that we have neglected here. In particular, during this derivation we have neglected the constraint on total particle number, which would introduce correlations between the density fluctuations at different positions. In the low noise limit, this would have little influence on the resultant distribution. However, for noisy systems where the fluctuations are large relative to the mean particle density, this – and higher order expansion terms in Equation (4.6) – would

be expected to contribute significantly to the density probabilities.

Continuum limit For completeness, I note that in the continuum limit $a \to 0$ and $B \to \infty$ Equation (4.7) yields a functional

$$p(\boldsymbol{\rho}, \boldsymbol{\epsilon}) \propto \exp\left[-\int_A dA \, \frac{(\boldsymbol{\rho} - \bar{\boldsymbol{\rho}})^2}{2(1 - d_0^2\bar{\boldsymbol{\rho}})\bar{\boldsymbol{\rho}}}\right]$$

where the integral is taken over the area A of the domain, $\boldsymbol{\epsilon}$ and $\boldsymbol{\rho}$ are the input and output fields, and the mean output field $\bar{\boldsymbol{\rho}}$ is defined by Equation (3.21).

Metropolis sampling

To assess the validity of Equation (4.8), I compare predictions of the conditional probability $P\big(\rho_j|\{\epsilon_k\}\big)$ to estimations made by performing Metropolis-Hastings sampling of particle distributions for fixed input energy profiles.

In brief, Metropolis-Hastings sampling permits the estimation of a probability distribution, $P(X)$, through the construction of a Markov chain with a steady state given by the 'target' probability distribution [237–239]. Applicable to cases where the target distribution is proportional to a known function of the system state, the Markov chain is constructed by proposing a new system state Y, with probability $g(Y|X)$ of being proposed, given the current state X. The proposed sample of the system is accepted if a randomly generated number, between 0 and 1, is smaller than the acceptance ratio, given by

$$\alpha = \frac{g(X|Y)P(Y)}{g(Y|X)P(X)}. \tag{4.10}$$

If the proposed change is not accepted, the current state is recounted. Since the target distribution is proportional to a known function, α may be calculated even if the full form of $P(X)$ is unknown. As such, each successive element in the chain – after an initial sampling period needed to converge to the steady state – resembles a sample from the target distribution. Such methods enable the efficient counting of the statistical weights between sampled states without having to generate many samples, and are therefore often used to estimate probability distributions when other sampling methods are deemed inefficient [240].

For my system, the sampling was performed on a 100-by-100 lattice, where each site could hold a maximum of one particle. The particles were assumed to interact with some external structure with possible interaction energies discretised into 241 evenly-spaced values between

$-12k_BT$ and $12k_BT$. The sampling process then consisted of iteratively choosing two lattice points: one occupied by a particle, and the other unoccupied; and proposing a new state with the particle swapped from the occupied to free lattice site and randomly allocated to either the bound or unbound state.

Identifying the initially occupied and unoccupied sites with the index $j \in 1, 2$ respectively, and corresponding lattice coordinates $r_j = (v_j, w_j)$, the original state X was described by the interaction energies $\epsilon_{1,X}$ and $\epsilon_{2,X}$, and the binding status of the particle at position 1 – which had an energy $\epsilon_{\text{particle},X} = \epsilon_{1,X}$ if bound and $\epsilon_{\text{particle},X} = 0$ if unbound. Similarly, the proposed state Y was described by the interaction energies $\epsilon_{1,Y}$ and $\epsilon_{2,Y}$, and the binding status of the particle, now in position 2, with energy $\epsilon_{\text{particle},Y} = \epsilon_{2,Y}$ if bound and $\epsilon_{\text{particle},Y} = 0$ if unbound.

To account for thermal fluctuations, the interaction energies for the *proposed* state, at both sites, were sampled anew from a discrete Gaussian energy distribution

$$p_\epsilon(\epsilon_{j,X}|v_j, w_j) = Ne^{-\frac{(\epsilon_{j,X}-\epsilon(v_j,w_j))^2}{2\sigma_\epsilon^2}}, \tag{4.11}$$

with a linear mean energy profile $\epsilon(v_j, w_j) = (2v_j/10 - 10)k_BT$ and standard deviation $\sigma_\epsilon = 1K_BT$, where N is a normalisation factor. Effectively, by including this, I introduce minor fluctuations in the interaction energy over time. This allowed me to assess how much the thermal fluctuations influence the conditional particle-density distribution. The probability of having proposed state Y given a current state X was then given by the product of the probabilities of the new interaction energies at these two locations

$$g(Y|X) = p_\epsilon(\epsilon_{1,X}|r_{1,X})p_\epsilon(\epsilon_{2,X}|r_{2,X}). \tag{4.12}$$

In comparison, the probability that this proposed state was sampled from the target distribution was proportional to the product of the probabilities of the interaction energies at these two locations, and the Boltzmann weight of the state according to the particle energy, taking the form

$$p(Y) \propto p_\epsilon(\epsilon_{1,Y}|r_{1,Y})p_\epsilon(\epsilon_{2,Y}|r_{2,Y})e^{-\epsilon_{\text{particle},Y}}. \tag{4.13}$$

The acceptance ratio therefore became the Boltzmann weight of the transition from X to Y, $\alpha = e^{-(\epsilon_{\text{particle},Y}-\epsilon_{\text{particle},X})}$.

Following an initial 'cool-down' sequence of samples to allow for convergence to a steady state, I took 10000 samples of the particle distribution. Since the example input was independent of the index w, I averaged the occupation boolean, where a value of 1 indicated an occupied site and 0 and unoccupied site, over this coordinate to find the density in each coarse-grained column of area $a = 100d_0^2$. This provided a discrete set

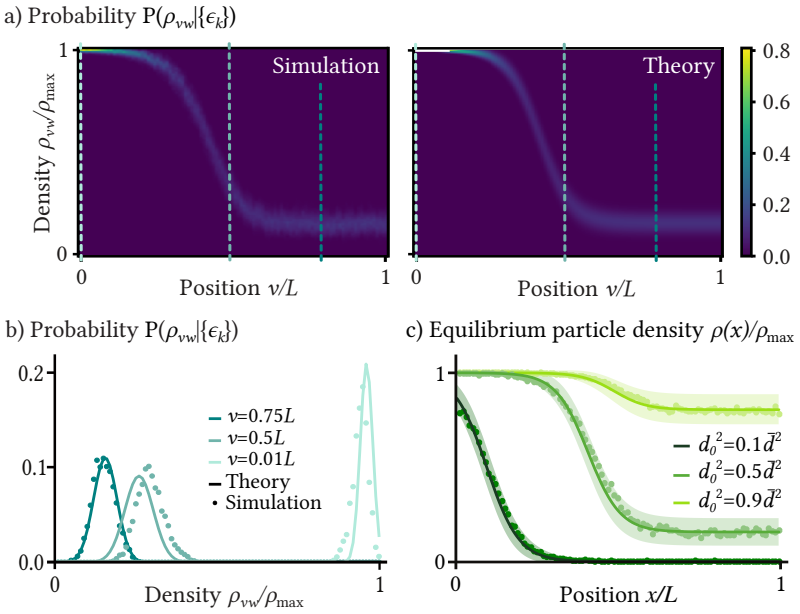


Figure 4.3: The conditional probability $P(\rho_j|\{\epsilon_k\})$ is well characterised by Equation (4.9), for an input energy $\epsilon_{vw} = (20v/L-10)k_BT$. a) The Gaussian approximation (right) of $P(\rho_{vw}|\{\epsilon_{vw}\})$ is compared to the distribution estimated using Metropolis-Hastings sampling on a L-by-L lattice (left) for $L = 100$, a membrane saturation of $(d_0/\bar{d})^2 = 0.5$. Coloured dashed lines show the coordinates plotted in (b), and white shading indicates values above 0.8. b) Estimated conditional distribution compared to the Equation (4.8). c) The estimated mean density profiles sampled for 1000, 5000, and 9000 particles recover the sigmoidal form expected from Equation (3.21) (and presented in Figure 3.5) for the corresponding membrane saturations $d_0^2/\bar{d}^2 = 0.1$, 0.5, and 0.9 for an interaction energy $\epsilon(x) = k_BT(20x/L - 10)$ on a domain of length L. Shading indicates two standard deviations given by Equation (4.9). Reproduced from [213], licenced under CC BY 4.0.

of possible densities $\mathcal{P} = \{0, 0.01, ..., 0.99, 1\}$. The conditional probability of a density given a particular v index (and therefore input energy ϵ_{vw}) was identified by finding the fraction of the samples that have this density for each of the v coordinates. The accuracy of this estimated distribution was improved by repeating the sampling process twelve times and averaging the results.

Using this set-up, I sampled the particle distributions for 1000, 5000, and 9000 particles, which in terms of $(d_0/\bar{d})^2$ corresponded to the values 0.1, 0.5, and 0.9 respectively. The resulting mean density at position $x = d_0v$, found by taking the weighted mean $\bar{\rho}_x = \sum_\rho \rho P(\rho|x)$ over all possible density values, accurately reflected the mean density expected from Equation (3.21), see Figure 4.3(c). Comparing the sampled local conditional probability for the $(d_0/\bar{d})^2 = 0.5$ case with Equation (4.8) (Fig. Figure 4.3(a) and (b)), the high-density region is found to have a narrower density distribution than the low-density region. Although discrepancies between the analytical and sampled distributions within the intermediate density region are apparent – most likely due the impact of particle conservation on the density noise profile, which has been neglected – , I find that Equation (4.9) adequately approximates the noise profile in the density in the low noise limit of input energy fluctuations.

4.2.2 Joint probability

Having derived and verified the form of the local conditional probability $P(\rho_j|\{\epsilon_k\})$, the joint probability

$$P(\{\rho_j\},\{\epsilon_k\}) = \frac{1}{n^B}\prod_j^B P(\rho_j|\{\epsilon_k\}). \tag{4.14}$$

of observing a particular set of density values $\{\rho_j\}$ together with an input energy vector $\{\epsilon_k\}$ may be found as the product of the conditional probability $P(\{\rho_j\}|\{\epsilon_k\}) = \prod_j P(\rho_j|\{\epsilon_k\})$ and the marginal probability of the input energy vector $P(\{\epsilon_k\})$.

4.3 Transmitted information depends on membrane parameters

Given the channel noise described above, how well can the repulsion-mediated sigmoidal filter described by Equation (3.21) transfer information into a membrane particle-density field? Using the joint probability, Equation (4.14), I quantify the achievable information transmission by evaluating the discrete mutual information over a range of length-scale ratios, $\bar{d}/d_0$. This is a computationally expensive process in which the probabilities for each possible output density set given each possible input energy set must be calculated. I therefore limit the evaluation to small systems discretised into $B = 5$ boxes of coarse-graining area $a = 10d_0^2$, resulting in the set of all possible density values $\mathcal{P} = \{0, d_0^2/a, 2d_0^2/a, ..., 1\}$. Similarly, to define the set of possible energy values, I assume a minimum interaction energy of $-10k_BT$ (a typical energy for the chemical binding of biological proteins) and an upper cutoff of $+10k_BT$. Discretising the possible energy levels to integer values in units of $2k_BT$ then gives the set of possible input energies $\mathcal{E} = \{-10k_BT, -8k_BT, ..., 10k_BT\}$.

Numerically evaluating the mutual information under this discretisation results in an optimal membrane particle saturation for particle-mediated information transmission [Fig. Figure 4.4(b)]. This optimal saturation corresponds to filters that achieve a balance between having threshold energies close to the middle of the input range (i.e. $\epsilon_{min} < \epsilon_{th} < \epsilon_{max}$) and gains large enough to overcome the channel noise. Specifically, the mutual information achievable by a sigmoidal filter can be understood as a measure of how well the filter can distinguish between different inputs [13]. Approximating the sigmoidal filters as perfectly binarising filters, this is best done by filters that separate the inputs into two equally-sized groups [13]. Therefore, sigmoidal filter systems with a low membrane saturation, i.e. with $\bar{d} >> d_0$, result in poor information

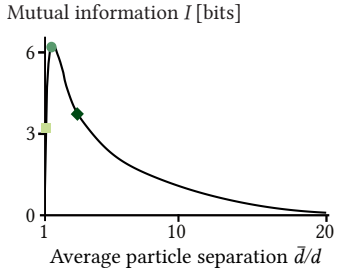


Figure 4.4: The mutual information I shared between the random vectors of input energy $\{\epsilon_k\}$ and output density $\{\rho_k\}$ exhibits a peak as a function of $\bar{d}/d_0$ where the membrane saturation is sufficient to maintain the threshold energy within the input range while also low enough to ensure a large difference in densities between the high and low density plateaux. This was computed for a discrete system of $B = 5$ boxes of area $a = 10d_0^2$. The diamond, circle, and square icons indicate the parameter values used in Figure 4.3(c), $d_0^2/\bar{d}^2 = 0.1$, 0.5, and 0.9 respectively. Adjusted from [213], licenced under CC BY 4.0.

transmission because they have very low threshold energies and as such cannot effectively distinguish between a large range of inputs (e.g. dark green diamond in Figure 4.4). In comparison, systems with a saturated membrane, i.e. with $\bar{d} \approx d_0$, suffer from a rapidly reducing gain upon increases in membrane saturation (e.g. light green square in Figure 4.4). This leads to the high- and low-density particle regions that are closer in density value, and as such more susceptible to the particle density noise.

Notably, the magnitude of the mutual information at the peak is higher than values reported in other biological contexts [13, 67, 88, 241]. The primary reason for this is that I consider the information stored in the full energy and density fields, rather than the more common approach of considering the value at only one location. As such, the signals contain more information. Indeed, mutual information is limited by the entropy of the input, which in this system is approximately 17 bits. A further reason for the high magnitude of mutual information is my choice of energy and density discretisation. Specifically, the magnitude of these results is set by the size of sets $\mathscr{E}$ and $\mathscr{P}$, in addition to the coarse-graining length-scale, because the magnitude of discrete mutual information depends on the number of possible inputs and outputs.

4.3.1 Binding site distributions influence information transmission

In the biological context, the membrane-adjacent structures that encode the initial input signal may take a variety of forms. Indeed, in the experimental context, quantifications of the structure and distribution of a variety of subcellular objects – including lipid bilayers, cytoskeletal filaments, the extracellular matrix and many others [242] – have been used as observational proxies for cellular and system states. As shown in Chapter 3 Section 3.2.2, such diversity in binding site arrangements leads to variations in the response functions of the corresponding signal filters. How do such differences in binding site density influence the information transmission by the resulting signal filters?

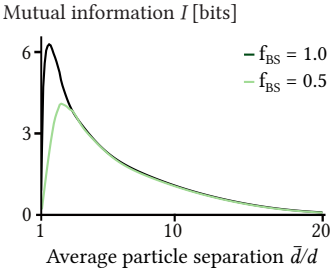


Figure 4.5: The mutual information between the input energy and output density vectors is reduced when the density of uniform binding sites is reduced from $\rho_{ext} = \rho_{max}$ ($f_{BS} = 1$, black, repeat from Figure 4.4) to $\rho_{ext} = 0.5\rho_{max}$ ($f_{BS} = 0.5$, green). The trends were computed, as in the previous figure, for a system discretised into $B = 5$ boxes of area $a = 10d_0^2$. Reproduced from [213], licenced under CC BY 4.0.

Low-density uniform binding sites

In Chapter 3 Section 3.2.2 I found that lowering the *uniform* density of binding sites leads to a reduction in the gain of the signal filter, and causes the high- and low-density plateaux of the filter to be closer in absolute density value. By evaluating Equation (4.2) using a mean density profile determined by Equation (3.26), I find that these changes leads to a reduction in the information transfer achieved by the filter

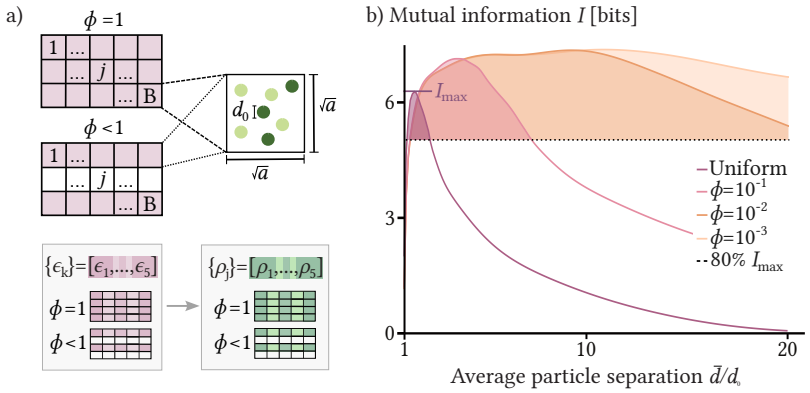


Figure 4.6: Lowering the area fraction of filaments, ϕ, leads to an increase in the optimal information transmission regime of the filter ($I > 0.8I_{max}$). a) The system was discretised into 5 boxes of area $a = 10d_0^2$ per line, either with (purple shading) or without (no shading) binding sites, with ϕ the area fraction of the binding region. b) Comparing the mutual information for $\phi = 10^{-3}$, 10^{-2}, 10^{-1}, and 1 – corresponding to the light orange, dark orange, pink, and dark purple lines respectively – shows that optimal information transfer for lower membrane saturations (i.e. larger $\bar{d}/d_0$) for filamentous input structures compared to uniform-binding-site surfaces. Reproduced from [213], licenced under CC BY 4.0.

(see Figure 4.5). One can intuitively understand that this arises because the distinction between the maximum and minimum density values is reduced relative to the channel noise. Indeed this is why the influence of the reduction in binding site density is more apparent for saturated membranes, where – as previously discussed – the gain is already low.

Linear binding site distributions

In comparison, how is information transmission influenced by input energy fields arising from particles binding to filamentous structures? Considering systems where the potential only varies along one axis (see Figure 4.6(a)) permits a comparison between the mutual information transmitted in the case of uniform binding sites and the case of filamentous binding site distributions. Discretising the axis over which the potential varies into 5 boxes to permit a comparison to the 2D case above and introducing – as in Section 3.2.2 – parallel lines of binding sites covering a fraction ϕ of the total membrane surface, I find that effective information transfer is achievable with fewer membrane particles in the case of filamentous inputs compared to the case of uniform binding sites, as seen in Figure 4.6(b) where reducing ϕ leads to a wider peak in the mutual information at higher $\bar{d}/d_0$.

Specifically, in the case of filamentous binding site distributions, the binding-free membrane region acts as a particle bath that buffers the particles used in the binding region. This particle buffering reduces the influence of particle number limits on the formation of the sigmoidal mapping, and leads to larger threshold values for the same number of particles in the membrane domain. This therefore increases the mutual information for low membrane saturation, i.e. $\bar{d} >> d_0$, by improving how well the filter can distinguish between all the possible inputs.

Importance of readout mechanism It is, however, important to note that this quantification of mutual information implies the mechanism that reads out the particle distribution is able to readout the particle *line*-densities, rather than the two-dimensional membrane density. If the latter is done, the line densities would be averaged with the low particle densities in the adjacent non-binding regions, leading to a reduction in the net gain of the sigmoid filter and subsequently a further loss of information. Such line-density readouts are indeed achievable in cellular contexts, where protein distributions along quasi-1D structures influence many subcellular processes such as the tension generation by motor proteins along filaments, or molecular events at cell-cell contact lines [243–247].

4.3.2 Optimal information transmission regimes

Membrane saturations close to the value for peak information transmission trace out regions in d_0-$\bar{d}$ parameter space where the filter can be considered to operate optimally. Using a limit of 80% of the uniform binding site case for illustration processes, I compare how these optimal regimes depend on the microscopic system parameters captured in the length-scales $\bar{d}$ and d_0, by plotting the zones corresponding to the optimal regimes for different filament area fractions ϕ in Figure 4.7. As expected from the result of the previous section, this shows that lower filament densities lead to larger optimal regions, in a more accessible region of the particle space.

Do real surface-associated protein complexes and other macromolecular structures in biological cells fall within these optimal regions? With the challenges associated with visualising nm-scale spatial arrangements of proteins along interfaces starting to be overcome by bioimaging technologies, the effective particle sizes and typical cellular concentrations of particles in several systems have been reported or can be inferred from other measurements [143, 144]. Aiming to compile an exhaustive list of proteins and subcellular structures for which this data is currently available, I identified values for $\bar{d}$ and d_0, which had been measured or or could be estimated from literature. Making no selection beyond restricting the search to *membrane-bound* particle-like structures, this was achieved for a range of protein complexes and macromolecular complexes, provided together with the corresponding sources in Appendix A.3, Table A.1.

Overlaying these values on the optimal information processing regimes in Figure 4.7, I find that several proteins that are known to interact with neighbouring cell membranes (connexons [248, 249]) or subcellular membraneous structures (ERMES [250, 251]) have sizes and surface

densities that position them within the optimal region for uniform binding sites. In contrast, other particles fall within the lower filament area fraction regions. In particular, several proteins that to associate with actin fibers or bundles (AChR [252–254], E-Cadherin [255], TCR [256], ICAM-1 [257, 258]) fall within the region where $\phi = 10^{-2}$. This area fraction is indeed approximately the fraction of actin-filaments, as estimated using measurements of cortical actin filament densities obtained from high-resolution cryo-ET tomograms in fibroblasts [259]. These results therefore support the idea that particle-mediated signal filters are a plausible mechanism for information transfer in biological, subcellular, membrane systems.

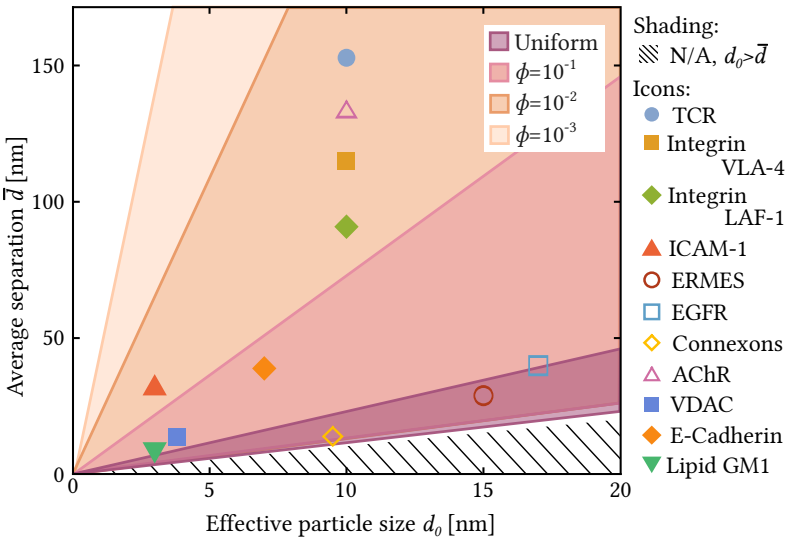


Figure 4.7: Macromolecular complexes in living cells fall within the optimal information transmission regions of d_0-$\bar{d}$ parameter space traced out by the regimes presented in Figure 4.6 (coloured accordingly). In particular, complexes known to associate to filamentous structures in the actin cytoskeleton (e.g., AChR, E-Cadherin,TCR) fall within the regime for an actin-like area fraction $\phi = 10^{-2}$ (dark orange), which is much larger than for the case of uniform binding sites (dark purple). Reproduced from [213], licenced under CC BY 4.0.

4.4 Selective information transmission

A key question when considering how information is passed through a filter is to ask how much of the permitted information is *relevant* to the downstream task. Indeed, in many cases signal filters are used, not only to transmit information, but to select particular information from the input.

To quantify the performance of selective signal filters, one must specify the relevant signal s that will be needed in downstream processes, and which is encoded within the input signal. A good filter acts as a 'bottleneck' to the flow of information, compressing the input to produce an output that contains the specific task-relevant signal, and little other information from the input [76, 83, 101, 110]. This permits the output to encapsulate the information that the input carried about s while discarding information that is irrelevant to the downstream task. In this section I quantify how effective the repulsion-mediated signal filter is at selectively transmitting

information, with a focus on downstream readout mechanisms for which a binarising filter would be suitable.

4.4.1 What is a typical readout?

What would be a suitable signal s with which to assess performance of the repulsion-mediated sigmoidal filters? In biological contexts, the relevant signals would of course be system specific, and may take the form of a spatially resolved field — written as a vector $\{s\}$ in the discrete limit – or may alternatively be a scalar variable describing a global property of the system. To keep the following computations and analysis tractable, I take – as an example – the relevant information to be a scalar that quantifies the fraction of the membrane area where interaction energy is greater than a threshold value ϵ_s, such that

$$s = \sum_i \Theta(\epsilon_i - \epsilon_s) \tag{4.15}$$

where $\Theta(\epsilon)$ is a Heaviside function. Such a signal would be relevant, for example, in the case of buckling-mediated readout mechanisms, as used in computational soft materials [51, 260], where a transition from a non-buckled to buckled state provides a binary output that depends on the size of the high particle density region.

As previously discussed, the repulsion-mediated sigmoidal filters transmit an approximately binarised version of their input signal, where the noisy particle response to the input interaction energy amplifies energetic differences close to the sigmoid's threshold, but neglects small energy variations far from the threshold value. These filters are therefore especially suited to selectively transmit the relevant signal defined in Equation (4.15), allowing me to quantify the signal performance for cases where the mechanism is likely to be useful.

It is important to note that, by defining the relevant signal to be a deterministic many-to-one function of the input $s = f(\{\epsilon\})$, I impose that the input signal $\{\epsilon\}$ perfectly captures the relevant signal. As such, the conditional probability $P(s|\{\epsilon\}) \in \{0, 1\}$ of a relevant signal s given a particular input $\{\epsilon\}$ is binary. Although common in, for example, machine learning classification tasks, this is contrary to other treatments of selective information transmission in biological contexts, where the input signal is considered a noisy, and thus imperfect, representation of the relevant signal [83, 261]. Such approaches are useful when considering a specific, well-quantified biological system. However, in this thesis, the deterministic mapping permits me to address how well *certain features of the input signal* are transmitted, without specifying the noise-profile of the steps preceding the signal filter.

4.4.2 Limits to information transmission

In analogy to the definition of mutual information in Equation (4.2), the amount of relevant information that is transmitted to the output field by the filter is given by the discrete mutual information

$$I_{\mathrm{R}}(\{\rho_j\}, s) = \sum_s \sum_{\{\rho_j\}} P(\{\rho_j\}, s) \ln \frac{P(\{\rho_j\}, s)}{P(\{\rho_j\})P(s)}. \tag{4.16}$$

Here, the sum $\sum_s$ is over all possible values of the relevant signal; the joint probability of a relevant signal s and an output ρ is given by $P(\{\rho_j\}, s) = \sum_{\{\epsilon_k\}} P(\{\rho_j\}, \{\epsilon_k\})P(s|\{\epsilon_k\})P(\{\epsilon_k\})$; and the marginal probabilities $P(\{\rho_j\}) = \sum_s P(\{\rho_j\}, s)$ and $P(s) = \sum_{\{\epsilon_k\}} P(s, \{\epsilon_k\})$ may be found from the joint probabilities $P(\{\rho_j\}, s)$ and $P(s, \{\epsilon_k\}) = P(s|\{\epsilon_k\})P(\{\epsilon_k\})$ respectively [101]. By construction, all the information that $\{\rho_j\}$ contains about s must come from $\{\epsilon_k\}$. This is a statement of the data processing inequality, which constrains the information $I(\{\rho_j\}, s)$ to be less than or equal to the information the output shares with the input,

$$I_{\mathrm{R}}(\{\rho_j\}; s) \leq I(\{\rho_j\}; \{\epsilon_k\}), \tag{4.17}$$

with equality when *only* relevant information is transmitted by the filter. Similarly, the data processing inequality also implies that the information the input signal carries about the relevant signal,

$$I_{\mathrm{I}}(\{\epsilon_k\}; s) = \sum_{\{\epsilon_k\}} \sum_s P(s, \{\epsilon_k\}) \ln \frac{P(s, \{\epsilon_k\})}{P(s)P(\{\epsilon_k\})}, \tag{4.18}$$

must be greater than or equal to the information the output has about the relevant signal

$$I_{\mathrm{R}}(\{\rho_j\}; s) \leq I_{\mathrm{I}}(\{\epsilon_k\}; s). \tag{4.19}$$

Equality in Equation (4.19) occurs when all the relevant information initially stored in the input signal is transmitted to the output signal. Comparing the performance of a filter in the *information plane*, where the mutual information $I(\{\rho_j\}; \{\epsilon_k\})$ is shown against $I_{\mathrm{R}}(\{\rho_j\}, s)$, to the limits of ideal selective filters given by Equations (4.17) and (4.19) therefore permits the assessment of the effectiveness of the filter in selecting the relevant signal s.

Information Bottleneck Method As discussed in Chapter 1, in many systems – and particularly in biological contexts – the input is an intermediary signal in a larger computational process, and therefore a noisy encoding of the relevant signal. In these cases, the data processing inequality does not adequately reflect the true constraints on information transmission. Instead, the ideal bound for the optimal

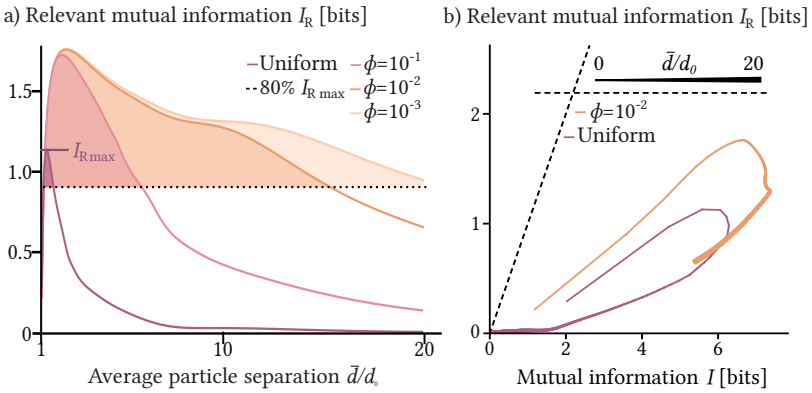


Figure 4.8: Relevant transmitted information shows an optimum regime for input binarisation. a) Relevant transmitted information, calculated for the same conditions as Figure 4.6 and for the desired output defined in Equation (4.15) with $\epsilon_s = 0$, exhibits an optimal information transmission regime ($I_R > 0.8 I_{Rmax}$) that is increased in size as the filamentous area fraction ϕ is decreased. b) The filter with filamentous distributions of binding sites, $\phi = 10^{-2}$ (dark orange), outperforms the uniform-binding-site filter (dark purple). Dashed lines indicate limits from the data processing inequality and line thickness indicates $\bar{d}/d_0$ value, corresponding to following the trends in (a).

encoding of the signal into the output lies within the inequality regime of constraints (4.17) and (4.19). As introduced by Tishby et al. [101] under the name 'Information Bottleneck Method', this true bound may be found by identifying the mapping of input to output that transmits the most relevant information possible, for the smallest transmission total information. Such mappings maximise the Lagrangian,

$$\mathcal{U} = I_R(\{\rho_j\}; s) - \lambda_{IB} I(\{\rho_j\}; \{\epsilon_k\}), \tag{4.20}$$

where λ_{IB} is a Lagrange multiplier that limits the overall transmitted information during the optimisation process, and may be calculated using an iterative algorithm introduced in the same paper.

The information bottleneck bound, however, as been shown by Kolchinsky et al. [261] to saturate to the DPI inequality bounds when the relevant signal is a deterministic function of the input. This is indeed the case considered in this chapter, and therefore the true limits of the filter, for the relevant signal defined by Equation (4.15), are given by Equations (4.17) and (4.19).

4.4.3 Where does this signal filter fall?

Using the discretised framework introduced in Section 4.3, I assess how effective the repulsion-mediated signal filters are at transmitting the relevant signal defined by Equation (4.15), for $\epsilon_s = 0$. Calculating $I_R(\{\rho_j\}, s)$ under the same conditions used in Figure 4.6, I find that the filters exhibit a peak in the transmitted relevant information at an optimal membrane saturation (see Figure 4.8) similar that observed for general information transmission in Section 4.3 (c.f. Figure 4.6). In the case of filamentous binding site distributions (see Section 4.3.1), reducing the area fraction of filaments, ϕ, leads to a larger optimal value of the length-scale ratio $\bar{d}/d_0$ than in the uniform binding site case.

Mirroring the approach in Section 4.3, I define the optimal relevant information transmission regime as being where the relevant information is within 80% of the maximum relevant information in the uniform binding site case (i.e. the black dashed line in Figure 4.8(a)). Although the optimal regime for relevant information transmission increases in size as ϕ is reduced, it remains smaller and confined to lower $\bar{d}/d_0$ values than the optimal general information transmission regime in Figure 4.6 – a difference which is exacerbated as ϕ is reduced. This means that, although reducing the line density of the filaments relaxes the need for large numbers of membrane bound particles for adequate information transmission, the best transmission of relevant information for $\epsilon_s = 0$ still occurs at reasonably high membrane saturations.

Focusing, for ease of visual comparison, on the uniform case and the filamentous case where $\phi = 10^{-2}$, I compare the performance of the filter to the ideal performance limits in the information plane (see Figure 4.8(b)). I find that neither filter achieves an optimal transmission of relevant information, however, the filter with filamentous binding site distributions performs better than the one with uniform distributions – both overall and for the same amount of total transmitted information (i.e. for the same value of $I(\{\rho_j\};\{\epsilon_k\})$). This further supports the result that filamentous structures give rise to better binarising signal filters than uniform structures. The imperfect nature of these filters has several origins. Most intuitively, these filters enact noisy compressions, and as such will be limited in their signal transmission by to the noisy nature of the particle dynamics. Furthermore, the response functions of the filters, as clearly seen in Figure 4.3(c), are not perfect binary mappings but are instead sigmoidal, with a finite gain. As such, they are not able resolve the boundary between inputs just above and just below the filter's threshold perfectly. Finally, there is the more nuanced feature of the filters that the threshold of the filter is set by the particle number and the input field, and is therefore not necessarily equal to ϵ_s. One must therefore tune the membrane saturation, i.e. tune $\bar{d}/d_0$, to best match the filter threshold with the desired threshold for as many of the inputs as possible. The filters therefore selectively transmit information close to their threshold, and may be tuned by varying the particle size or number of particles, similar to how how resistors may be tuned to shape non-linear mappings between voltages in electronic audio processing [262]. Even when optimally tuned, however, the threshold will not be the same for all inputs, and the filter would therefore not able to enact the desired threshold specified by Equation (4.15) perfectly across the full range of possible inputs.

This dependence of optimal information transmission on the membrane saturation can be further illustrated by considering different desired

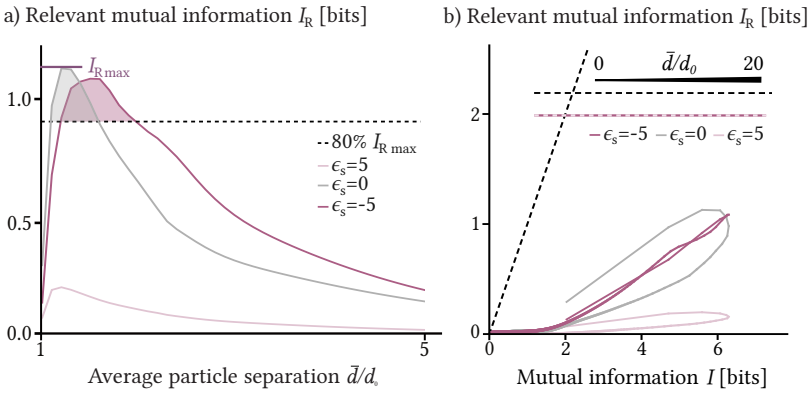


Figure 4.9: Tuning $\bar{d}/d_0$ selects transmitted information for uniform binding sites. a) Good relevant information transmission by the uniform-binding-site filter is achieved for both $\epsilon_s = 0$ (grey, from Figure 4.8) and $\epsilon_s = -5$ (dark purple) but not $\epsilon_s = 5$ (light purple). Shading indicates optimal relevant information transmission ($I_R > I_{Rmax}$, as defined in Figure 4.8(a)) b) Although the horizontal data processing limit given by Equation (4.19) for $\epsilon_s = \pm 5$ (light and dark pink dashed line) is reduced compared to $\epsilon_s = 0$ (horizontal black dashed line). The limit from Equation (4.17) is unchanged (diagonal dashed line).

signal thresholds. Taking as an example the thresholds $\epsilon_s = \{-5, 0, 5\}$ I repeat the quantifications of $I_R(\{\rho_j\}, s)$ and $I_R(\{\rho_j\}; \{\epsilon_k\})$ for both uniform binding sites and filamentous distributions of binding sites with $\phi = 10^{-2}$, presented in Figure 4.9 and Figure 4.10 respectively. It is important to note that when changing the desired threshold ϵ_s, one also changes the information that the inputs contain about the relevant signal, $I_I(\{\epsilon_k\}; s)$, and as such the limit given by Equation (4.19). $I_I(\{\epsilon_k\}; s)$ is maximal for a desired threshold equal to the median value of the possible interaction energies – in this case $\epsilon_s = 0$ – as indicated with a gray dashed line in Figure 4.9(a) and Figure 4.10(a). The cases of $\epsilon_s = -5$ and $\epsilon_s = 5$ both have lower, and equal, values of $I_I(\{\epsilon_k\}; s)$, as indicated by the coloured dashed lines in Figure 4.9(a) and Figure 4.10(a)

In the uniform case, Figure 4.9(a), I find that the filters can be tuned to adequately select the relevant information if the desired thresholds are close to, or less than, zero. For desired threshold much larger than zero, the filter does not effectively transmit the desired information, even though the general transmitted information is unchanged. This can be understood through comparisons to Figure 3.5, where the filter threshold is shown to be unable be set significantly higher than the median possible input energy. Instead, attempts to tune the length-scales $\bar{d}$ and d_0 for large, positive desired thresholds results in filter thresholds only slightly over the median value, and small filter gains.

These results are more striking when considering filamentous structures. Similar to the uniform case, the filters perform poorly for large positive desired thresholds. However, for both the $\epsilon_s = -5$ and $\epsilon_s = 0$ cases the filter perform well. The value of $\bar{d}/d_0$ for which a relevant signal threshold of $\epsilon_s = -5$ is selectively transmitted is significantly larger than for $\epsilon_s = 0$. This highlights how the tuning of $\bar{d}/d_0$ determines which features of the input signal are selected by the filters. Interestingly, for the $\epsilon_s = 0$ case, increasing the ratio $\bar{d}/d_0$ leads to a regime in which the transmitted relevant information decreases while the total transmitted information increases. This means that the filter loses effectiveness in

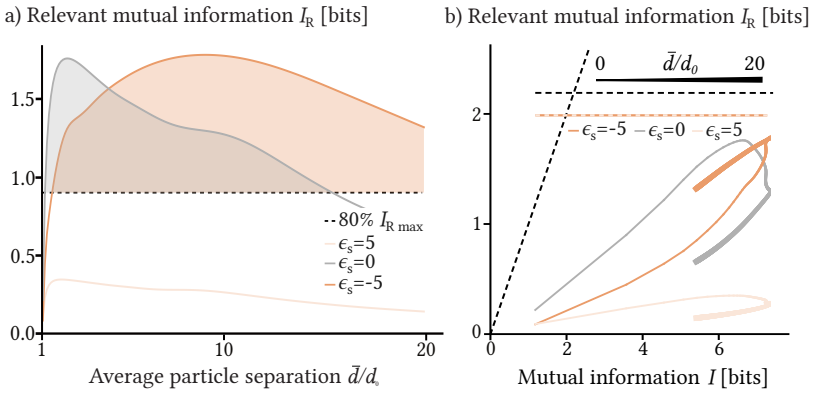


Figure 4.10: Tuning $\bar{d}/d_0$ selects transmitted information for inputs filamentous distributions of binding sites with $\phi = 10^{-2}$. a) Optimal relevant information transmission for $\epsilon_s = 0$ (grey, from Figure 4.8) and $\epsilon = -5$ (dark orange) is achieved by tuning $\bar{d}/d_0$. The filter performs poorly for $\epsilon_s = 5$ (light orange). Shading indicates the optimal transmission regime ($I_R > I_{Rmax}$, as defined in Figure 4.9(a)). b) The optimal regimes are close to the ideal data processing limit (dashed lines). The coloured dashed line indicates the reduced horizontal data processing limit for $\epsilon_s = \pm 5$.

distinguishing the desired signal while still appearing to increase the information it transmits about the input signal in general. In comparison, for a desired threshold of $\epsilon_s = -5$ the general transmitted information starts to decrease at the same value of $\bar{d}/d_0$ as the relevant information, implying that the optimal membrane saturation for general information transmission, and relevant information transmission, are the same in this case.

Overall, I find that the repulsion-mediated sigmoidal filters introduced in Chapter 3 selectively transmit binarised information about their input, with the threshold specified by the physical parameters of the membrane-particle system.

4.5 Concluding remarks

In this chapter I have assessed the effectiveness of particle-mediated sigmoidal filters in selectively transmitting information encoded within input interaction energy fields. Using a biophysically motivated noise profile – verified with Metropolis-Hastings sampling – I introduced a discrete mutual information framework to account for information transfer by spatially resolved fields, and identified an optimal information transmission regime determined by the average separation and effective size of the particles in the membrane. Extending the framework to treat inputs encoded in filamentous structures, I built on the results of Chapter 3 to show that a reduction in the dimension of the structure leads to more effective and less metabolically costly filters, and demonstrated that these optimal regions are biologically attainable through the comparison with example protein-membrane systems. By specifying which features of the input signal should be selected by the filter, I further showed that these sigmoidal filters selectively amplify features in the input signal, and may be tuned by varying effective particle size and average particle separation. Such findings may help in the design of synthetic filters, or to guide investigations into the role of that spatial particle distributions

play in biological subcellular signalling systems, by providing an outline of the capabilities and limitations to these signalling modalities.

An important caveat to the results presented here is that I have assumed throughout this chapter that the probability distribution of possible inputs is uniform. If the inputs are known to be sampled from some other distribution, the conclusions drawn may differ. In particular, the optimal values of the length-scale ratio required to select different input features would depend on the distribution of possible inputs between which the filter is intended to distinguish. Nevertheless, the overall finding that distributions of repulsive particles can act as sigmoidal filters of input interaction energy fields, and may be tuned by varying the average particle separation and effective particle size, would still hold.

Signal filters with feedback | 5

The sensing achieved by the particle-mediated signal filters addressed in this thesis requires direct interactions between particles and their surroundings. However, these interactions can affect the environmental signal that the filter is supposed to be sensing. In soft systems, this 'observer effect' introduces a feedback from the particle distributions onto the interaction energies, changing how the input signal is processed. Although in typical experimental scenarios such measurement effects are considered a nuisance, they can also provide great benefit under the correct circumstance [141]. Indeed, in control theory feedback loops are used to maintain and manipulate complex system processes [30, 263, 264]. Similarly, in living systems feedback is integral to innumerable functions, including giving rise to biochemical waves, switching behaviours, and ensuring cellular processes are robust to perturbations [142, 265, 266]. It is therefore important how such feedbacks affect information transmission, and how they may be utilised in the design of desired filters. Feedback loops pose a challenge for information theoretic approaches, since it can be difficult to define a direction to the flow of information. However, in many cases if the feedback is reasonably local in the information processing cascade, it can be treated as part of the signal filter and be incorporated when calculating the corresponding response function.

Previous studies including feedback in information processing cascades have focused on system robustness or spontaneous organisation, and have shown that feedback can facilitate information both through the formation of information-carrying patterns and by minimising the impact of noise [69, 121, 202, 267–272]. However, less work has focused on quantifying the information transmission in these systems [23, 80, 273]. Inspired by how particle distributions can affect changes on membrane and environment shapes, I consider how feedback from the particle distributions onto the interaction energy field influences the response function of the repulsion-mediated sigmoidal filter (Equation (3.21)). Assessing how the feedback influences the noise in the filter, I then compare the information transmission to the results for the no-feedback presented in the previous chapter.

The notebooks and scripts used to produce the results of this chapter are provided in repository [215].

5.1 Feedback on input energy

Feedback from the distribution of particles onto the input interaction energy could have several origins, for example perhaps the most intuitive feedback mechanism is where the binding of the particles leads to a local reduction in interaction energy, forming a positive feedback loop. The binding of a particle could cause a deformation in the environment, bringing the particle's binding site – and others in the local vicinity – closer to the membrane. Such a change in proximity would lower the energetic barrier for other particles to bind to the neighbouring sites, and would therefore constitute a positive feedback loop, whereby high-density regions promote lower interaction energies, which in turn favour even higher densities.

To keep my approach broadly applicable, I model this feedback through perturbations to the interaction energy, without directly specifying the underlying feedback mechanism. I assume that, free from particle binding, the environment surrounding the sensing membrane has a minimum free energy configuration that gives rise to an initial (i.e before feedback) interaction energy field $\epsilon_0(\mathbf{r})$. This constitutes the initial store of information in the environment, acting as an input signal that determines the post-feedback density and energy fields. Assuming a stiff environment, such that perturbations to the interaction energy are small, I use a saddle-point approximation to describe the energetic penalty arising from perturbations to the environment. This results in a quadratic contribution to the system free energy, $\int dA\lambda(\beta\epsilon_0 - \beta\epsilon)^2/2$, where $\epsilon(\mathbf{r})$ is the resulting – 'true' – interaction energy following the binding-induced changes to the environment. The environmental stiffness is parametrised by the elastic coefficient λ, which has units of energy density and describes the energetic cost per area for perturbations to the interaction energy field. Under this construction, the output equilibrium density and interaction energy fields are functions of the input interaction energy $\epsilon_0(\mathbf{r})$.

Considering, as before, a signal filter operating at equilibrium, the output fields minimise the system free energy,

$$F = \int dA \left[\lambda \frac{(\beta\epsilon_0 - \beta\epsilon)^2}{2} + \mu_\mathrm{b}\rho_\mathrm{b} + \mu_\mathrm{u}\rho_\mathrm{u} \right] + \text{const.} \tag{5.1}$$

where $\rho_\mathrm{b}(\mathbf{r})$ and $\rho_\mathrm{u}(\mathbf{r})$ describe the bound and unbound particle densities respectively. The free energy considered in Chapter 4 is recovered in the limit of stiff environments, i.e. large λ. The chemical potentials for the bound and unbound particles are given in Section 3.1, Equations (3.4) and (3.5), as $\mu_i = E_i + \epsilon_i + k_\mathrm{B}T \ln \rho_i$ for $i \in \{\mathrm{b}, \mathrm{u}\}$ where $\epsilon_\mathrm{b} = \epsilon + \epsilon_\mathrm{u}$ and ϵ_u is a constant. Assuming the particles experience only short-range

repulsive interactions, with effective particle size d_0 and uniform binding sites at the maximum particle density $\rho_{\text{ext}} = 1/d_0^2$, Equations (3.13) from Section 3.2.1 give $E_\text{b} = E_\text{u} = -k_B T \ln(1 - d_0^2\rho)$, where $\rho = \rho_\text{b} + \rho_\text{u}$ is the total particle density. The conditions for equilibrium then become

$$\mu_\text{b}(x) = \epsilon - k_B T \ln\left(\frac{1 - d_0^2\rho}{\rho_\text{b}}\right) + \text{const.} = 0 \tag{5.2}$$

and

$$\mu_\text{u}(x) = -k_B T \ln\left(\frac{1 - d_0^2\rho}{\rho_\text{u}}\right) + \text{const.} = 0, \tag{5.3}$$

from which the equilibrium total particle density is found to be the same as in Section 3.2.2, given by Equation (3.21) (repeated here for convenience),

$$\rho(\mathbf{r}) = \frac{1 + e^{-\beta\epsilon(\mathbf{r})}}{l^2 + d_0^2(1 + e^{-\beta\epsilon(\mathbf{r})})}.$$

As before, the integration constant l is set by the total number of particles N within the whole membrane area A, through Equation (3.8). This introduces a dependence on the average particle separation in the membrane, $\bar{d} = \sqrt{A/N}$. This equilibirum density field could, of course, have been derived from the particle dynamics as presented in Chapter 3. Conversely, the equilibrium true interaction energy field, found by minimising the free energy, Equation (5.1), takes the form

$$\beta\epsilon = \beta\epsilon_0 - \frac{\rho}{\beta\lambda(1 + e^{\beta\epsilon})}, \tag{5.4}$$

where I have used the relation $\rho_\text{b} = \rho/(1+e^{\beta\epsilon})$ found from Equations (5.2) and (5.3).

Equations (3.21) and (5.4) give the output of the signal filter determined by the input field ϵ_0 and system parameters d_0, $\bar{d}$, and λ. Comparing to the results of Chapter 3, and using as an example a one-dimensional linear input profile $\epsilon_0(x) = (20x/L - 10)k_\text{B}T$, I find that feedback from the particle distribution to the interaction energy leads to an increase in the gain of the signal filter (see Figure 5.1). This is expected for such a positive feedback loop, where more particles are able to bind where there are already bound particles – since their presence lowers the interaction energy in their neighbourhood. Due to particle conservation, this leads to fewer particles in the low density region, and a correspondingly lower reduction in the interaction energy.

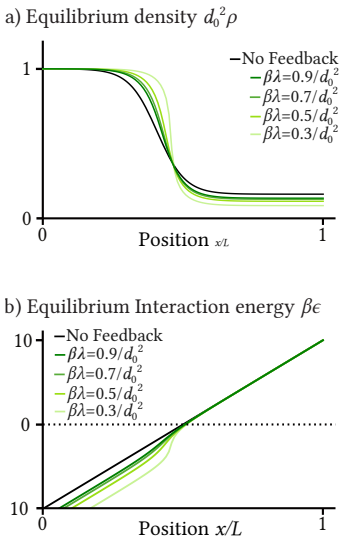


Figure 5.1: Resultant particle density distributions (a) and interaction energy fields (b) for an initial input interaction energy $\beta\epsilon_0 = 20x/L - 10$ show that feedback between particle density and particle-environment interaction energy leads to an increased filter gain relative to the no feedback case (black line). Softer environments (small λ, light green) lead to a greater increases in the gain, and larger perturbations in the interaction energy away from the initial input value.

5.2 Feedback increases filter noise

Introducing feedback into the particle-pattering mechanism introduces another degree of freedom that may fluctuate in the system, contributing

to a more noisy signal filter. Following an approach similar to Section 4.2.1, in this section I quantify the channel noise for the sigmoidal signal filter in the presence of feedback. In contrast to the previous chapter, here the fluctuations in the bound and unbound particles are treated directly, as well as the fluctuations in the true energy field ϵ.

Operating in the discrete limit with coarse graining area a, as defined in Section 4.1.1, the conditional probability,

$$P(\{\rho_{\mathrm{b}j}\},\{\rho_j\},\{\epsilon_j\}|\{\epsilon_{0j}\}) \propto e^{-\beta F}, \tag{5.5}$$

of the random vector realisations $\{\rho_{\mathrm{b}j}\}$, $\{\rho_j\}$, $\{\epsilon_j\}$ given an input $\{\epsilon_{0j}\}$ – for the bound particle density, unbound particle density, and interaction energy respectively – is proportional to the exponential of the discrete system free energy,

$$F = \sum_j a\left[\lambda\frac{(\beta\epsilon_{0j}-\beta\epsilon_j)^2}{2} + \mu_{\mathrm{b}j}\rho_{\mathrm{b}j} + \mu_{\mathrm{u}j}\rho_{\mathrm{u}j}\right] + \text{const.}. \tag{5.6}$$

By applying a saddle-point approximation about the equilibrium fields $\{\bar{\rho}_{\mathrm{b}j}\}$, $\{\bar{\rho}_j\}$ and $\{\bar{\epsilon}_j\}$, the probability distribution at each site j is found to be independent of the other sites, such that

$$P(\{\rho_{\mathrm{b}j}\},\{\rho_j\},\{\epsilon_j\}|\{\epsilon_{0j}\}) \approx \prod_j P(\rho_{\mathrm{b}j},\rho_j,\epsilon_j|\{\epsilon_{0j}\}) \tag{5.7}$$

where the *local* conditional probability is given by

$$P(\rho_{\mathrm{b}j},\rho_j,\epsilon_j|\{\epsilon_{0j}\}) \approx N_j \prod_{X,Y} e^{-\frac{\beta}{2}A_{XY}(X-\bar{X})(Y-\bar{Y})} \tag{5.8}$$

with normalization constants N_j. Here the partial derivatives for the fields $X, Y \in \{\rho_j, \rho_{\mathrm{b}j}, \epsilon_j\}$ – whose equilibrium values are denoted by $\bar{X}, \bar{Y} \in \{\bar{\rho}_j, \bar{\rho}_{\mathrm{b}j}, \bar{\epsilon}_j\}$ – are denoted by

$$A_{XY} = \left.\frac{\partial^2 F}{\partial X \partial Y}\right|_{\mathrm{eq.}}, \tag{5.9}$$

with $|_{\mathrm{eq.}}$ indicating that the function is evaluated at the equilibrium system state.

Focusing on the cases where downstream mechanisms only read out the total density profile, the local conditional probability $P(\rho_j|\{\epsilon_{0j}\})$ for the total density is found by integrating Equation (5.8) over the fluctuations in $\rho_{\mathrm{b}j}$ and ϵ_j. For small fluctuations relative to the equilibrium values, the tails of the distribution can be neglected and the integrals taken to span an infinite domain. This results in Gaussian integrals, leading to

the local probability

$$P(\rho_j|\{\epsilon_{0j}\}) = P_{0,j}e^{-\frac{1}{2}\frac{(\rho_j-\bar{\rho}_j)^2}{\sigma^2_{\rho j}}}, \tag{5.10}$$

where I have incorporated the non-exponential factors into a single normalization constant $P_{0,j}$. Here, the inverse of the variance – after evaluating the partial derivatives of the free energy (see Appendix A.4), and using the relation $\bar{\rho}_{\mathrm{b}j} = \bar{\rho}_j/(1 + e^{\beta\bar{\epsilon}_j})$ – is given by

$$\sigma^{-2}_{\rho j} = \frac{a}{\bar{\rho}_j(1 - d_0^2\bar{\rho}_j)} - \frac{a}{(\beta\lambda(1 + e^{\beta\bar{\epsilon}_j})^2 - \bar{\rho}_j e^{\beta\bar{\epsilon}_j})}. \tag{5.11}$$

The Gaussian integrals are valid only in the stiff environment, small fluctuations, limit where $\beta\lambda \geq \bar{\rho}_{\mathrm{b}j}(\bar{\rho}_j - \bar{\rho}_{\mathrm{b}j})/\bar{\rho}_j$ (see Appendix A.4). As such, I constrain the following analysis to parameter regimes where $\beta\lambda \geq \rho_{\max}/4$.

The standard deviation given by Equation (5.11) over all possible density values, is shown for various environment stiffnesses in Figure 5.2. I find that the inclusion of feedback leads to an increase in the particle density noise relative to the no feedback case. In the context of a signal filter, this implies a trade-off between the increased gain provided by the inclusion of feedback in the filter mechanism, and the increased noise that arises as a consequence.

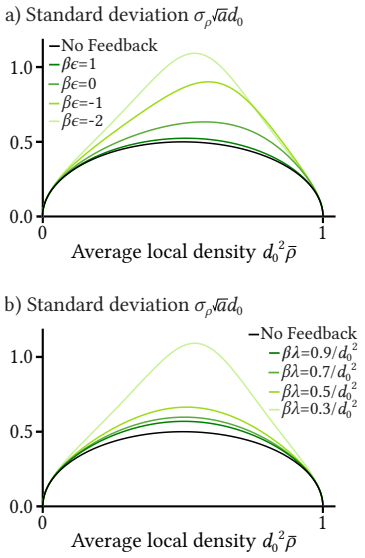


Figure 5.2: Feedback causes a density-dependent increase in the density distribution noise. Decreasing the interaction energy (a), i.e. making binding more favourable, or the environment stiffness (b) leads to an increase in the noise for intermediate density values.

The magnitude of the increase in standard deviation is influenced by the environmental stiffness, the local particle density, and the input interaction energy. Specifically, low – i.e. favourable – input interaction energies lead to larger density fluctuations than high interaction energies. This can be understood by noting that at higher interaction energies, fewer particles are in the bound state, and therefore the coupling to the now-variable true interaction energy ϵ is less significant. Similarly, a low environment stiffness leads to greater density fluctuations because a soft environment permits more variations in ϵ, which may then influence the local particle density. For very soft inputs the standard deviation diverges because the approximation of assuming the integral domain is infinite breaks down, but it is expected that in this regime the environment is too malleable for the initial input interaction energy to influence the output fields. In contrast, the feedback becomes negligible in the stiff limit $\lambda \to \infty$ and Equation (5.11) recovers the results from Section 4.2.1.

The standard deviation also tends to zero in the limits of maximum or vanishing density. Intuitively, this arises as a consequence of the saddle-point approximation, where it is assumed that the noise profile is symmetric about the mean density. Since there are hard limits to the density fluctuations in the extreme density limits, this constraint sets the fluctuations to zero in this model. In a true system, the density

distribution is likely to be non-symmetric at these limits, which may lead to a worse communication channel than presented here. This framework therefore allows me to calculate an upper bound for the performance of the signal filters with feedback.

5.3 Feedback improves selective information transmission

In the presence of feedback, how does the balance between increased filter noise and increased gain influence the capability of the sigmoidal filters to transmit information? In the previous chapter, I showed that such signal filters – without feedback – selectively transmit information from the input signal, and are tuned by varying $\bar{d}$ and d_0. In this section, the signal is encoded within the *initial* energy $\{\epsilon_{0j}\}$ and I use the discrete mutual information between this input and the output density $\{\rho_j\}$, and relevant information between the density and a desired signal s, – as defined in Chapter 4 Sections 4.3 and 4.4 respectively – to quantify the influence of feedback on this information transmission.

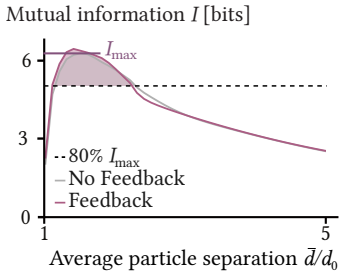


Figure 5.3: Feedback has little effect on the mutual information between the input energy field $\{\epsilon_{0j}\}$ and output density field $\{\rho)j\}$. Shown for a feedback system with uniform binding sites and environmental stiffness $\beta\lambda = 0.5/d_0^2$ (purple) and compared to the no feedback case from Figure 4.6 and Figure 4.8 (grey). Shading indicates the optimal information transmission regimes – where I is within 80% of the maximum for the no feedback case – which are almost unchanged by the inclusion of feedback.

Uniform binding sites

Starting with the case of uniform binding sites, and choosing an intermediate environmental stiffness of $\beta\lambda = 0.5/d_0^2$ for visualisation purposes, I find that introducing feedback produces negligible changes to the mutual information between $\{\rho_j\}$ and $\{\epsilon_{0j}\}$, Figure 5.3. Under these conditions, the benefit to information transmission that arises through the increased separation between the density magnitudes in the high and low density regions is mostly balanced by the increased noise in the system. Furthermore, there is an additional penalty to information transmission arising due to the increased gain, which not only increases the distinction between high and low densities, but also reduces the capacity of the filter to transmit non-binary information (such as to distinguish between input energy values that reside on the same side of the sigmoid's threshold).

In comparison, assuming that the desired readout of the filter is the fraction of the membrane area in which the input energy is greater than a threshold $\epsilon_s = 0$ (as introduced in Section 4.4.1), I find that the filter with feedback transmits a significantly larger proportion of *relevant* information than without feedback (see Figure 5.4(a)). This difference is particularly pronounced within the optimal information transmission regime (which, as defined in Section 4.3.2, contains all systems with length-scale parameters $\bar{d}$ and d_0 for which the information transmission is within 80% of the maximum achieved with uniform binding sites).

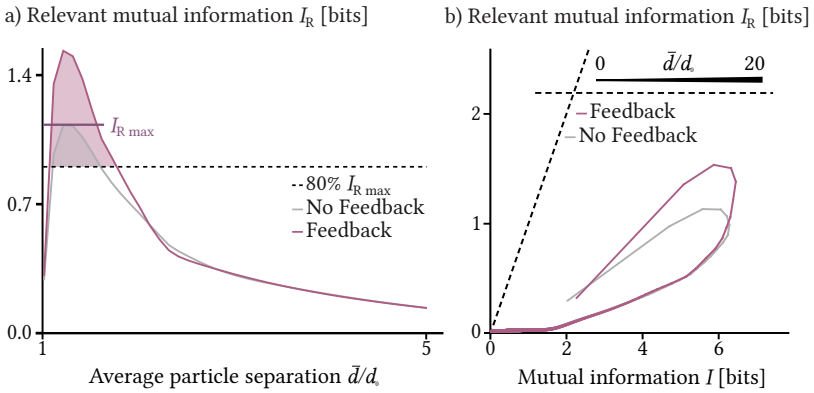


Figure 5.4: Feedback improves the filter's ability to binarise input interaction energies with uniform binding sites. a) Comparing a feedback system withenvironmental stiffness $\beta\lambda = 0.5/d_0^2$ (purple) to the no feedback case from Figure 4.6 and Figure 4.8, the transmission of relevant information for a desired binary signal with a threshold energy of $\epsilon_s = 0$ is increased. Shading indicates that the optimal regime (80% of the maximum relevant information, $I_{R\,max}$, in the no feedback case) is larger with feedback. b) This feedback filter remains far from the ideal specified by the data processing inequality (dashed lines).

A natural explanation of this observation is that feedback results in a sigmoidal filter with larger gain, which acts more accurately like a binarising filter. Since the desired output is indeed a function of the binarised input field, the feedback makes the filter more suitable to this particular readout. However, even with feedback, this filter does not achieve optimality for this desired readout, Figure 5.4(b).

Linear binding site distributions

Considering sigmoidal filters applied to filamentous structures, I find the mutual information between the output density and input interaction energy is also unchanged by feedback. Since the magnitudes of the densities in the high and low density regions are already well separated in the filamentous case, the further improvement coming from the feedback does not provide a sufficient improvement to information transmission to overcome the reduction in information caused by the more effective input binarisation and increased noise. The bumps visible in the mutual information in Figure 5.5 – particularly the $\beta\lambda = 0.5/d_0^2$ case – are likely an artifact arising from the discretisation chosen for the possible energy and density values.

On the other hand, the transmission of relevant information (defined equivalently to the uniform binding sites case above), is increased with feedback. In fact, for sufficiently soft environments, the filter perfectly transmits the desired information, assuming an adequately tuned value of $\bar{d}/d_0$, while discarding a large portion of the irrelevant signal (see Figure 5.6). Under these conditions, the near-perfect binarisation of the input signal achievable with feedback, accompanied with sufficiently distinct high- and low-density values, greatly outweighs the impact of increased noise in the filter.

The fact that this greater transmission of relevant information occurs alongside a lack of change in overall information transmission may at first appear unintuitive. However, such a trend can be understood by noting

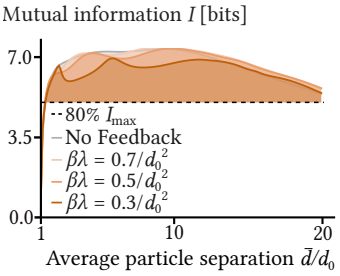


Figure 5.5: Similar to Figure 5.3, the mutual information is mostly unaffected by feedback. Apparent artifacts due to energy space discretisation appear as waves near the maximum I, and are particularly visible for soft materials. Shown for three environmental stiffnesses (light to dark orange indicating stiff to soft environments) and the no feedback case presented in Figure 4.6 and Figure 4.8 (grey). Shading indicates the optimal information transmission regime defined as the region with I within 80% of the maximum for uniform binding sites and no feedback (i.e. 80% of I_{max} from Figure 5.3).

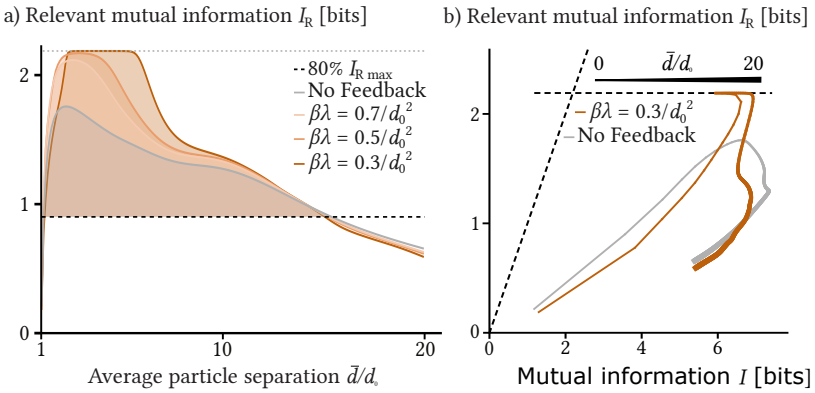


Figure 5.6: Feedback improves transmission of information relevant to signal binarisation for filamentous inputs. a) Decreasing environmental stiffnesses (light to dark orange indicating stiff to soft environments) leads to increases in relevant information transmission, relative to the no feedback case presented in Figure 4.6 and Figure 4.8 (grey). Shading indicates the optimal region defined by 80% of the uniform no-feedback case ($I_{R\,max}$ from Figure 5.4). b) For a stiffness of $\beta\lambda = 0.3/d_0^2$, the relevant transmitted information reaches the data processing limit (black dashed lines, and grey dashed line in (a)).

that the improvements to the capability of the filter to binarise the input is responsible for both of these consequences: improving the binarising nature of the filter makes it particularly suitable for selecting thesholded information, however by necessity this same adjustment makes the filter worse at distinguishing between intermediate energy values, resulting in a worse over all information transmission.

5.4 Feedback-facilitated garbage disposal

I have had the pleasure to co-supervise a master's student at the University of Heidelberg, Johannes Jung, whose project on particle sorting I detail briefly below. A manuscript presenting the results of this project is currently under preparation.

Feedback processes in signal filters need not necessarily loop back onto the input signal. Instead they may form a link between the output signal and internal features of the filter. This is the case for the garbage disposal mechanism addressed by Johannes Jung in his master's thesis [274].

Garbage disposal is the process of removing unnecessary or malfunctioning objects from a system. In living organisms, this homeostatic process is important for the maintenance of correct cellular function [275, 276]. Focusing on the maintenance of membrane composition, this project set out to investigate how vesiculation – the process of removing a section of membrane via a topological transition – could be used to sort dysfunctional from functional particles, and remove them from the surface of the original – or 'parent' – compartment. We defined functional particles as ones that can bind to an adjacent structure within the parent compartment, and dysfunctional particles as those that cannot. A key feature of the signal processing mechanism is that this bindable structure is absent from the vesicle, breaking the symmetry between the parent and vesicle surfaces. The binding affinity of the particles then acts as a sorting cue by which to distinguish the two particle species. From a signal

processing perspective, this binding affinity constitutes an input signal, which is processed by the sorting mechanism to output the distribution of particles between the parent and the vesicle compartments.

Considering a vesiculation process driven by an active contractile ring, Johannes found that particles were sorted primarily due to the difference in binding affinity. However, entropic effects from particle sorting introduced positive feedback between particle retention in the parent compartment and the compartment's resulting size. This meant that, compared to processes in the absence of entropic effects, a greater proportion of the possible initial membrane conditions would result in good particle sorting with large particle-rich parent compartments and small particle-poor vesicles. This therefore provides an effective mechanism for selectively disposing of dysfunctional particles while retaining as much of the the membrane surface in the parent compartment as possible.

A further feedback contribution from particle distribution to cell shape was then considered through the introduction of curvature-inducing particles. We considered particles that favour tightly curved membranes, described via a density-dependent spontaneous membrane curvature. Requiring that both functional and dysfunctional particles induced the same membrane curvature, this permitted the investigation into the role that feedback between particle distributions and membrane shape plays in particle sorting. Johannes found that, in addition to promoting vesiculation, this feedback led to larger discarded vesicles, compared to in the absence of curvature effects, and the retention of a are greater fraction of the functional particles in the parent compartment. As such, curvature-inducing particles were found to facilitate garbage disposal, at the cost of requiring larger portions of the original membrane to be lost to the discarded vesicle.

Through the identification of a sorting score, Johannes identified an optimal procedure for particle sorting. Interestingly, it was also found that the strength of the curvature-mediated feedback by the particles could be used to tune the sorting between desired outcomes. For low feedback strength, the sorting mechanism favours particle retention, with only a small number of dysfunctional particles being successfully discarded. Whereas, for large curvatures the removal of dysfunctional particles is prioritised, at the cost of losing some functional particles and a larger proportion of the membrane. The feedback therefore permits the tuning between the selective removal of dysfunctional particles, and their broad disposal.

Overall, this project identified a physical mechanism for particle garbage disposal in deformable membranes, and demonstrated how membrane shape dynamics can aid the removal of unwanted particles. Its treatment

of feedback from output signals to internal components of signal filters highlights how feedback can be used to tune signal filter response functions according to specific application priorities.

5.5 Concluding remarks

In summary, in this chapter I have outlined how feedback can be beneficial to signal processing systems, both as a means of improving signal selection and as a means of tuning filter performance.

Focusing on feedback between particle distributions and input energy signals for the sigmoidal signal filter, I found that feedback leads to increases in the filter gain, the separation between the density plateaux, and the filter's noise. Surprisingly the effects of these changes on general information transmission were found to cancel out, leading to a similar mutual information between input energy and output density fields as in the no feedback case discussed in the previous chapter. In contrast, I find that feedback improves the performance of binarising filters, with a greater transmission of information relevant to this task. For soft environments and filamentous binding site distributions, this improvement was shown to result in ideal relevant information transmission. In comparison to this input-directed feedback, I briefly outlined the work of Johannes Jung, which showed how feedback from output signals to internal components of a signal filter provides a mechanism for tuning signal filters according to different system priorities.

In both of the systems discussed here, it is important to note that although feedback can be beneficial, very large feedbacks (e.g. soft environments) may decouple the input and output signals entirely, leading to poor information transmission where the output signal is dominated by the feedback dynamics rather than the original input signal. Furthermore, I have confined this section to address feedback within directional information cascades. It is an ongoing area of research to address how interconnected feedback systems may be treated within an information processing framework when such directions are not clear.

Overall, these finding support the growing literature detailing the various benefits of feedback, and explicitly link these to an information theoretic framework in which such improvements may be assessed. Such findings may prove beneficial in informing how feedback may be used to facilitate information processing during the design of synthetic soft signal processing systems [264].

Surface segmentation in *Sphearoforma arctica*

6

Are there any examples of particle-mediated signal filters known to play a role in real biological systems? In Chapter 4, sigmoidal filters arising from repulsive particles were shown to exhibit an optimal parameter regime for information transmission, with large particles interacting with filamentous structures presenting the most metabolically accessible regime due to the low number of particles required. One biological particle species that falls within this range are nuclear pore complexes (NPCs)[277]. These large protein structures, each composed of more than 550 proteins [278], are embedded within the nuclear envelope (NE) – an intracellular membrane that forms the physical and regulatory interface between the cytoplasm and the nuclear interior [279, 280]. With a diameter of approximately 100 nm, these protein complexes are well suited to forming the sigmoidal signal filters described in the preceding chapters [278].

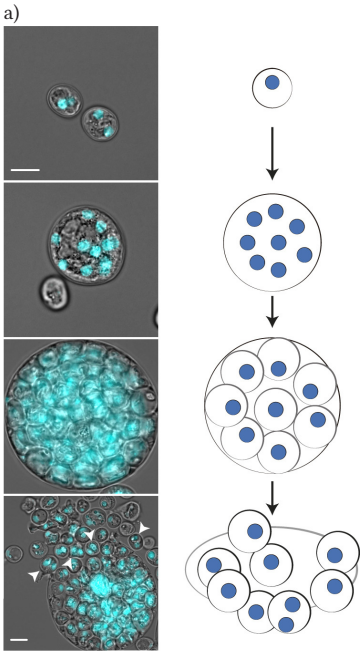
In this chapter I compare the predicted sigmoidal filters to distributions of these protein complexes in the single cellular organism *Sphearoforma arctica*, where NPCs have been observed to interact with biological filaments called microtubules (MTs), that comprise the extranuclear cytoskeleton.

The work presented in this chapter was completed in collaboration with Hiral Shah and Gautam Dey, who carried out expansion microscopy and electron tomography imaging of S. arctica nuclei. The results of this chapter are included in the paper accepted at PRL [213].

6.1 NPCs form aster distributions in *S. arctica*

NPCs are large protein assemblies that act as protein-scaffolds for pores in the NE, allowing the selective transport of molecules into and out of the nucleus [278]. With diameters that are orders of magnitude larger than typical membrane-bound proteins (100 nm compared to typical sizes of 1 nm [217, pg. 45]) – and with turnover timescales longer than the nuclear division cycle [281–283], such that particle conservation over the course of a cell cycle period may be assumed [145, 284] – NPCs are an interesting and observationally tractable particle species to consider for evidence of particle-distribution mediated information transfer. Indeed, in addition to controlling nucleocytoplasmic transport, NPCs are known to play key roles in regulating internuclear states such as gene expression [285–288]

a)



b)

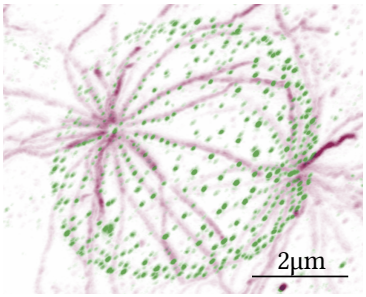


Figure 6.1: The *S. arctica* coenocytic (i.e. multi-nucleate) cell cycle. a) Representative images (left, blue labels nuclear content) and sketches (right) of cell life cycle stages of *S. arctica* starting from newborn single cell (row 1) which undergoes multiple nuclear divisions to result in a multi-nuclear 'coenocyte' (row 2) before cellularising (row 3), and eventually bursting to produce many daughter cells (row 4). Scale bar for first to third row: 10 μm, for fourth row: 20 μm b) Maximum intensity projection of an ultrastructure expansion microscopy image an *S. arctica* nucleus, with immunofluorescently labelled MTs (purple) and NPCs (green). (a) and (b) are reproduced from [295] and [213] respectively, licenced under CC BY 4.0.

through interactions with the extranuclear cytoskeleton. However, the role of the *spatial distribution* of NPCs in information transfer remains unclear [289–294].

The regulatory role of NPCs distributions has typically been overlooked, since in many species – including humans – NPCs are known to bind strongly to a fixed polymer network within the nucleus called the 'nuclear lamina' which prevents NPC diffusion. However, in species were lamina is not present, NPCs are able to diffuse in the membrane and can form interesting patterns [296, 297]. An example of such patterning has been observed in the nuclei of the species *Sphaeroforma arctica*. This single-cellular, multi-nucleate species from the protist clade Ichthyosporea is the subject of much interest in the study of the evolution of mitosis (nuclear division) and its driving factors [295], Figure 6.1(a). In this species, NPCs have been observed by Hiral Shah of the Dey group in EMBL Heidelberg interact with an external microtubular array to form high-density lines in an aster-like configuration, coinciding with microtubules emanating from microtubule organising centers (MTOCs) located on the outer nuclear surface [145], see Figure 6.1(b).

6.1.1 Input and output signals

The NPC patterns in *S. arctica* provide an interesting example case for the processing of spatial information in biology. The asters have been suggested to play a role in coordinating nuclear division and appear to segment the nuclear surface into three distinct regions: two caps deliniated by the high-density prongs of NPCs, and a middle region of low density and apparently uncoordinated NPCS [145]. Focusing on how the structure of the MT network influences the formation of these caps, I consider the separation distance between the NE and an adjacent MT as the input of the particle patterning process. Modelling the interaction as mediated by a Hookean elastic spring, this proximity may be mapped to an input interaction energy,

$$\epsilon(\mathbf{r}) = \frac{k}{2}(h(\mathbf{r}) - h_0)^2 + \epsilon_c, \tag{6.1}$$

in which h is the shortest distance – at the surface coordinate $\mathbf{r}$ – between the NE and the MT, ϵ_c is the energy required to bind an NPC to the MT, and k and h_0 are the spring constant and resting length of the effective spring. This effective interaction represents the combined elastic contributions arising from local deformations of the NE, stretching of the linker complex, and deflection of the MT. It is important to note that by assuming the input interaction energy takes the form given in Equation (6.1), I imply that the spatial variations in the interaction energy arise primarily due to differences in the proximity of filaments to the

NE, and neglect any contributions that may arise from heterogeneities in chemical binding or system deformability.

Under this construction, the separation h stores the original input signal, which may encode information on the status of the cytoplasm (and the other nuclei), and be subsequently transcribed into the NPC distribution to be later read off by processes within the nucleus. Indeed, the proximity of such biological filaments regulate nuclear processes in other systems, for example by influencing mitosis [298, 299] and nuclear import [300, 301], as well as causing the rearrangement of chromatin, leading to changes in gene expression [302, 303]. In particular, in *S. arctica* extranuclear MTs have been shown to influence nuclear positioning and morphology, including during mitosis [145, 304].

By measuring the separation distance h from immunostained expansion microscopy images of *S. arctica* nuclei, in this chapter I test the predictions of output NPC densities made by substituting Equation (6.1) into the filter response function given by Equation (3.21), and assess whether the arrangement of NPCs corresponds to the output of a filter for the proximity of the adjacent MTs through comparisons of this prediction to the probability of having an NPC located in the NE at that point measured along the input MTs.

6.2 Image analysis pipeline

The MT-NE distance, and the NPC coordinates, were estimated from expansion microscopy images with fluorescent labelling of the MTs, NPCs, and membranes. The imaging was carried out by Hiral Shah, following the experimental setup and data acquisition protocol described in [145]. This section focuses on the theoretical and computational analysis of these images. Expansion microscopy is an imaging method that combines optical and physical sample magnification, using a swellable hydrogel [305], to provide higher resolution static 3D image stacks than achievable using conventional confocal microscopy. This permitted the identification of individual NPCs and MTs in the imaged nuclei (Figure 6.2), from which the separation distance h and NPC distribution were extracted using a custom image analysis pipeline that I developed, briefly detail below and available in [306]. The original expansion microscopy data used in this chapter, and the corresponding post-processed image data is available at [307].

Nuclear Envelope segmentation

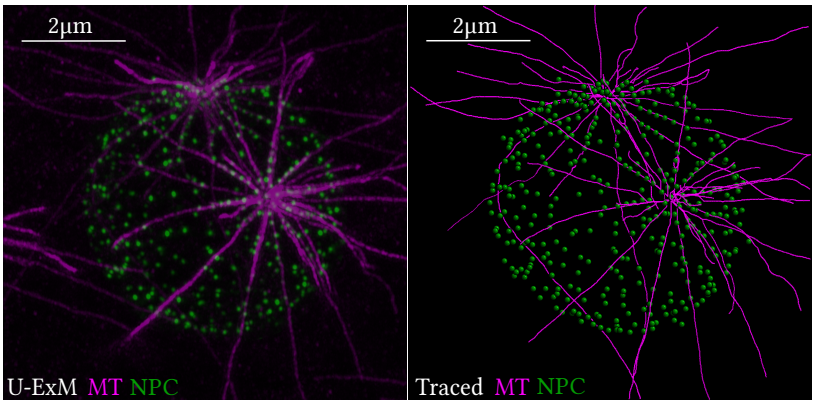


Figure 6.2: (Right) the custom image analysis pipeline accurately traces MTs (purple) and identifies NPC coordinates (green) from ultrastructure expansion microscopy images of *S. arctica* nuceli, shown here for an example image stack (left, 3D rendering). Reproduced from [213], licenced under CC BY 4.0.

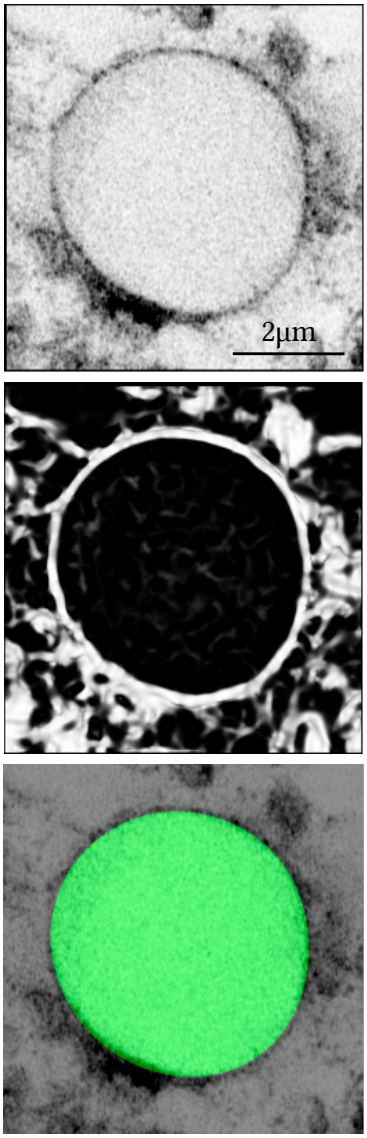


Figure 6.3: The nuclear volume is identified from the membrane channel signal (top) using a Hessian filter to identify surface-like structures (middle) from which the nuclear volume may be segmented (bottom), shown for an single mid-nucleus image slice.

The MT-NE separation distance was measured by first forming a binary image mask of the nuclear volume. This was performed using the fluorescent membrane marker, which labelled all phospholipid membranes in the expanded sample (Figure 6.3 top). Due to the variable intensity of this fluorescence labelling, I used a Hessian filter that amplified surface-like structures in the image to identify the nuclear surface, following the approach laid out by Frangi et al. in [28] (Figure 6.3 middle). This filter amplified image regions with one large and two small eigenvalues of the image Hessian, identifying surface-like regions where the image intensity remains approximately constant in two orthogonal directions, and changes rapidly in another. Similar image processing filters based on [28] exist both within standard imaging tools such as Fiji [308] and within image analysis packages such as SciKitImage [309]. However, these implementations focus on segmenting filamentous structures, and discard the other conclusions of the paper highlighting how similar algorithms may be used to amplify surface-like features in an image. I therefore produced a custom implementation of this surface filter. This was beneficial over simple thresholding methods in correcting for patchy and grainy labelling, and artifacts from inhomogeneous lighting conditions.

Following the identification of membrane surfaces, a sequence of thresholding and morphological operations on the resultant binary image permitted the segmentation of the inner nuclear volume, and from which the nuclear surface was identified (Figure 6.3 bottom). These follow-up procedures corrected for small Hessian-filter artifacts that protruded into and out of the NE. As such this protocol was well suited to the smooth surface of the *S. arctica* NE, but may struggle if applied to rough membrane topologies in other cellular environments.

MT Tracing

To overcome the challenges in data extraction presented by the patchy and uneven labelling in expansion microscopy data (Figure 6.4 top), the

positions of MTs in the three-dimensional image stacks were identified using the following custom pipeline.

Firstly, irrelevant microtubular signals from within the nuclear volume were removed from the analysis by masking the nuclear interior. These signals arise because, during nuclear division, a network of MT polymers called the mitotic spindle forms within the nucleus and, although in this study we considered only interphase (non-dividing) nuclei, some low-level MT signal occasionally appear within the nuclear volume [145]. Next, following standard noise-reduction treatments, filamentous structures were identified using skimage.filters's Sato and Frangi 'tubeness' filters. Although – similar to the surface filter described above – both filters utilise the image Hessian, they have different competencies [310], and I found that the sequential use of the Sato filter followed by Frangi filter was most affective in identifying filaments in the MT data (Figure 6.4 middle).

Having identified tube-like regions in the images, MT 'skeletons' – single pixel width traces of objects in a binary image – were formed by taking a threshold of the amplified signal and applying skimage.morphology's skeletonize_3d function. Due to patchy labelling and as a consequence of the tubeness filters, this skeletonizing process resulted in a collection of skeleton branches broken into slightly separated parts. A custom algorithm was then used to identify proximal, well-aligned branch ends, and join them by tracing the highest signal intensity path through the original image between the two end points using skimage.graph's route_through_array function. Having connected the disjointed branches, full MTs were traced by combining connected, aligned branches into single paths with preferential order determined by the alignment at the branch joint. This permitted the identification of individual MTs in the case where two branches crossed one another.

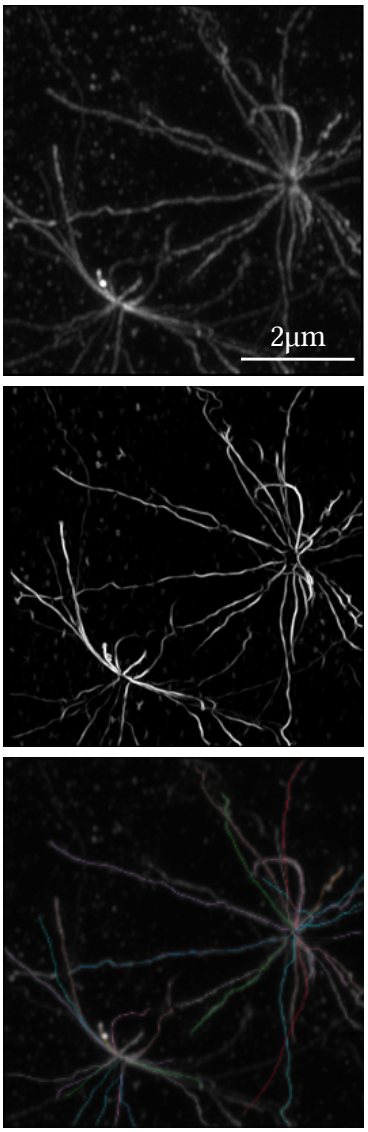


Figure 6.4: The MTs were traced from the MT image channel (top) using Sato and Frangi filters to identify tube-like structures (middle) from which the MT traces were identified (bottom), shown as a 3D rendering for an example image.

Finally, to reduce the influence of errors arising in the high MT density region around the MTOCs, which were identified manually, traces of MTs within a radius of 0.6 μm from either MTOC were discarded and replaced with the highest signal intensity path from the end of each MT trace to the MTOC coordinate. This identified the most likely path of the MT, given it originates from the MTOC, and resulted in a collection of MT coordinates that traced out the path of the MTs connected to each of the MTOCs in the 3D image stack (Figure 6.4 bottom). Having proved particularly successful in treating the patchy nature of MT labelling in expansion microscopy, this tracing protocol is now being applied to trace MTs in other systems.

Neglecting MTs from all nuclei with major segmentation or tracing errors, the separation distance h for each MT was calculated by identifying, for

each coordinate in the MT's trace, the closest point on the nuclear surface resulting in a MT 'shortest projection line' on the nuclear surface. The separation between these coordinate pairs gave the value of h. Since this chapter concerns MTs that influence NPC distributions, I discarded MTs that were either too short, or did not pass along the nuclear surface. This was achieved by applying a selection criteria on the total length of the filaments (>1.1 μm) and their proximity to the NE (<0.3 μm over the nearest 1.1 μm to the MTOC). The remaining filaments were used to estimate the average area-fraction of NE-proximal MT filaments on the NE, $\phi = (0.21 \pm 0.03)$ (used in Figure 6.9).

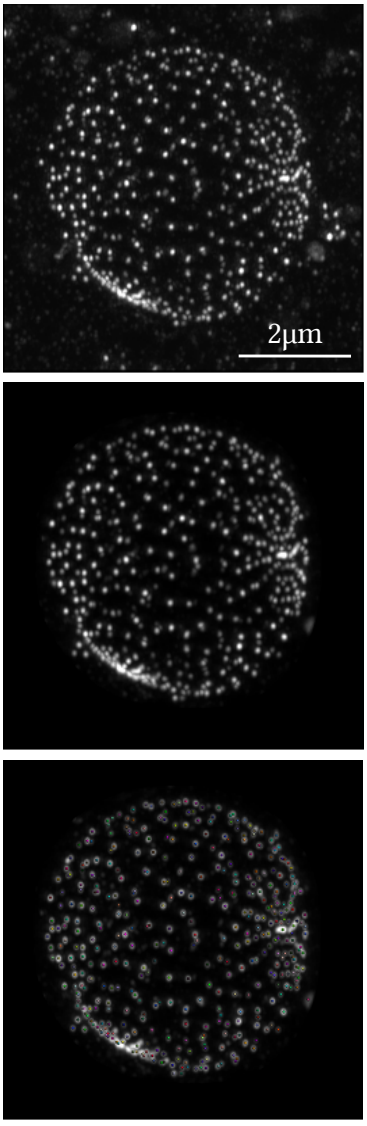


Figure 6.5: The NPC coordinates were identified from the NPC-labelled image channel (top) by first masking the non-nuclear signals (middle) and then identifying the NPCs using a Difference-of-Gaussian method (bottom), shown here as a 3D rendering of an example image.

NPC point extraction

The coordinates of approximately all the NPCs on the nuclear surfaces were identified by first masking the NPC signal (Figure 6.5 top) to keep only signals pertaining to NPCs at the nuclear surface (Figure 6.5 middle). This prevented erroneous contributions from artifacts in the cytoplasm and nuclear volume. Following noise reduction treatments, the coordinates of the NPCs were then identified using ski.features's blob_dog function, which implements a Difference-of-Gaussian method to identify peaks in signal intensity greater than 5% of the maximum intensity (Figure 6.5 bottom). Calculating the nuclear surface area from the segmented nuclear volume, the average spacing of NPCs over all of the nuclei was then found to be $\bar{d} = (440 \pm 50)$ nm. To measure the one-dimensional distribution of NPCs along the MT filaments, NPCs within a distance of 160 nm from each filament's shortest-distance projection line were treated as occupying the closest coordinate along this line. The NPC line-density along the length of each MT was then found by coarse-graining each filament to boxes of size 0.26 μm.

6.3 Sigmoidal NPC distributions

Using the extracted MT-NE separation distances, I compare predictions of NPC line-density to the measured one-dimensional NPC distributions. In this section, I describe how the MTs were grouped into clusters with similar separation distance profiles, across which average NPC densities could be found, and used to fit predicted NPC distributions to the measured values to find estimates of the effective particle size d_0 and effective spring constant k.

6.3.1 MT-NE separation increases along filament length

The accuracy of the NPC density measurements was improved by averaging over clusters of MTs with similar MT-NE separation profiles. The clusters were identified by first discarding all filaments of lengths less than 3.4 μm, so that the remaining filaments were of similar lengths. This resulted in 110 individual separation profiles. Using a region of interest (ROI) defined as the MT segments between 0.2 μm and 3.8 μm from the MTOC, the similarity between pairs ij of MTs was quantified using a dissimilarity score,

$$\text{Dissimilarity Score} = \frac{\sum_{k}^{\text{Min}(N_{pi},N_{pj})} |(h_{i,k} - h_{j,k})|}{\text{Min}(N_{pi}, N_{pj})} \tag{6.2}$$

where k indexes the pixels and N_{pi} denotes the total length of the ith profile h_i. Computing a dissimilarity matrix of these scores, the MTs were sorted into nine groups using an agglomerative clustering algorithm. Clusters containing fewer than five filaments, or with a combined dissimilarity score larger than 4000, were considered too dissimilar to be appropriate for further analysis and were therefore discarded, resulting in five suitable filament clusters.

This process revealed that the MTs had increasing separation distances from the NE from their anchor point at their respective MTOC to towards the nuclear equator (Figure 6.6, pale lines). Performing a rational-exponential re-parametrization of the measured height profiles in each cluster resulted in a differentiable representation of the typical separation profile h for each cluster (Figure 6.6, dark lines), which could then be used to make a prediction of the NPC line density along the filaments in the cluster using Equations (6.1) and (3.21). The corresponding average NPC line-density for the cluster was found by averaging over the different filaments in the cluster. Similarly, the average NPC separation $\bar{d}$ and average number of filaments per nuclear area ζ/A (see Chapter 3 Section 3.2.2) were also found independently for each cluster, as presented in Table 6.1, by averaging the values for each nucleus weighted by the number of MTs that each nucleus contributed to the cluster. The existence of five distinct clusters, and thus five distinct pairs of MT-NE separation profiles and NPC line-densities, permitted a perturbative study of the sigmiodal signal filter applicability to the *S. arctica* system by assessing the performance of the NPC line-density predictions made by Equation (3.21) for a variety of different inputs.

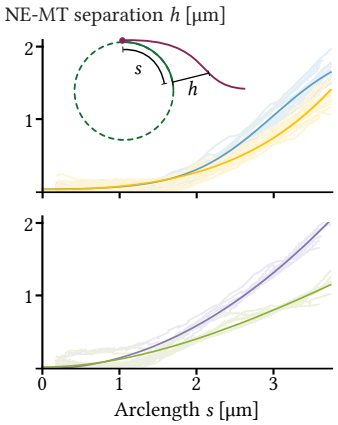


Figure 6.6: The individual tracks (light colours) making up the distinct MT clusters, and rational-exponential reparametrisation of the tracks in each cluster (dark colours), show that the MT-NE shortest separation distance increases as a function of the arc-length s away from the MTOC. Top: Clusters 1 (yellow) and 2 (blue), Bottom: Clusters 3 (green) and 4 (lilac), Cluster 5 shown in Figure 6.7). Inset reproduced from [213], licenced under CC BY 4.0.

6.3.2 Binarizing filter describes NPC line densities

Predictions of NPC line-densities were made by substituting the re-parametrized height profiles for each cluster into Equation (6.1) and using (3.21). I approximated the chemical interaction energy between an NPC and its binding site by a chemical interaction energy $\epsilon_{\mathrm{NPC}} = 25 k_{\mathrm{B}}T$ [311]. Hiral Shah provided independent measurements for the resting spring length $h_0 = (84 \pm 15)\,\mathrm{nm}$ from electron microscopy images by measuring the minimum NPC-MT separation distance within high-density regions of our ROIs from electron microscopy images from Shah et al.[145], using the image processing software Fiji [308], see Appendix A.5 Table A.2.

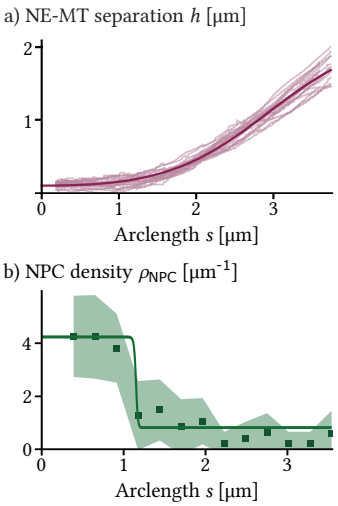


Figure 6.7: NPCs form sigmoidal line densities along MT filaments. a) MT-NE shortest separation distance for cluster 5, with individual MT traces in light purple, and the rational-exponential reparameterization of these tracks in dark purple. b) Fitting (solid green line) the NPC density predicted from the reparameterisation of the separation distance yielded parameter estimates for the effective spring constant k and the effective NPC size d_0 (see Table 6.1). Shaded area: 95% confidence interval. Fitted profiles for clusters 1 to 4 are available in Appendix A.5. Reproduced from [213], licenced under CC BY 4.0.

Estimates of the effective spring constant k and the effective particle size d_0 for each cluster were found by fitting the line-density predictions to the measured line-densities, using the average NPC separation $\bar{d}$ and average number of filaments per nuclear area ζ/A found for the corresponding cluster. Indexing each coarse-grained location by i, the fit was achieved independently for each cluster by minimizing the objective function

$$J(\rho, \hat{\rho}; k, d_0) = \sum_i (\hat{\rho}_i - \rho(\hat{s}_i; k, d_0))^2, \tag{6.3}$$

in which $\hat{\rho}$ denotes the measured density at the arc-length position $\hat{s}_i$, and ρ was computed by evaluating Equation (3.21) with re-parametrized separation distance $h(\hat{s}_i)$. Averaging the results of these fits, and taking the standard deviation as an estimate of the error, I thereby obtain estimates for the effective NPC size $d_{0,\mathrm{NPC}} = (260 \pm 50)\,\mathrm{nm}$, and the spring constant characterizing the elastic interaction between NPCs and microtubule filaments $k_{\mathrm{NPC}} = (0.04 \pm 0.03)\,\mathrm{pN\,nm^{-1}}$. An example of the fit results for cluster 5 is presented in Figure 6.7, and the remaining four are available in Appendix A.5. The fit results for each cluster are summarized, along with the corresponding $\bar{d}$ and ζ/A values, in Table 6.1.

Table 6.1: Estimates of *S. arctica* nuclear parameter estimates found from measurement or through minimisation of the objective function Equation (6.3) for the five distinct MT clusters. Reproduced from [213], licenced under CC BY 4.0.

Cluster	Average particle separation $\bar{d}$ [nm]	Average number of filaments per nuclear area ζ/A [µm^{-2}]	Fitted effective particle size d_0 [nm]	95% CI d_0 [nm]	Fitted effective spring constant k [pN nm^{-1}]	95% CI k [pN nm^{-1}]
1	443	0.27	310	{350, 260}	0.036,	{0.029, >0.109}
2	451	0.27	310	{340, 280}	0.024	{0.020, >0.109}
3	437	0.30	230	{260, 200}	0.090	{0.084, >0.109}
4	452	0.30	200	{240, 180}	0.030	{0.028, 0.032}
5	466	0.28	240	{250, 220}	0.026	{0.025, 0.059}

Effective particle size, d_0

While Equation (3.21) captures the sigmoidal relationship between the input separation and output density profiles of each cluster effectively, the resultant estimate of the minimal NPC separation distance is larger than the known NPC diameter (approximately 100 nm [278]). To assess the validity of the estimate for the effective particle size, I compared the result to independent measurements, made by Hiral Shah, of minimum NPC separations in the high-density regions of electron microscopy images from [145], see Figure 6.8 and Appendix A.5 Table A.2. These measurements provided an independent effective particle size estimate, $d_0 \approx (190 \pm 40)$ nm (mean ± standard deviation, Table A.2), close to the fitted estimates attained above. The larger effective particle sizes from both of these approaches suggest that the NPCs experience additional, non-steric repulsion effects, perhaps due to interactions mediated by NPC-induced curvature of the nuclear-envelope, or alternatively from additional proteins binding to the outer edge of the NPCs.

Effective spring constant, k

In comparison to the effective particle size, for which the fits were well-constrained, for some clusters the cost functions of the fit exhibited shallow regions that extended in the direction of the effective spring constant, implying that – over the explored range – the fit was of limited sensitivity to variations in this parameter. Despite this, the resulting estimates for the spring constants are consistent across the five distinct MT datasets (see Table 6.1).

Overall, I find that the sigmoidal signal filter model describes the observed NPC patterns well across all five of the MT clusters. Since each cluster has a distinct input separation profile, these results lend strong support to the interpretation that this particle pattern is guided by the spatially-varying separation between the nuclear envelope surface and the MTs that pass adjacent to it, and controlled by repulsive interactions between the NPCs.

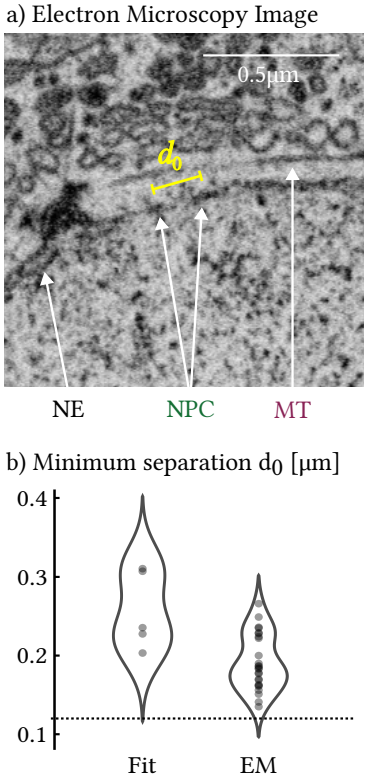


Figure 6.8: Fit values for the effective NPC size d_0 are supported by independent measurements of minimum NPC-NPC separation distances from electron microscopy (EM) images. a) EM images showing NPCs in the high-density region of the filament close to the MTOC. The separation distance was measured as centre-to-centre distance between adjacent NPCs along the filament. b) Comparison between the fit and EM-measured values for d_0. Dashed line indicates the diameter of the NPC ($120nm$). Reproduced from [213], licenced under CC BY 4.0.

6.4 NPCs fall within optimal information transmission regime

Having identified the MT-guided NPC patterns in *S. arctica* as a possible biological example of particle-mediated signal filtering, it is informative to ask whether sufficient information could be transmitted by these filters. I address this by comparing the measured values of the effective

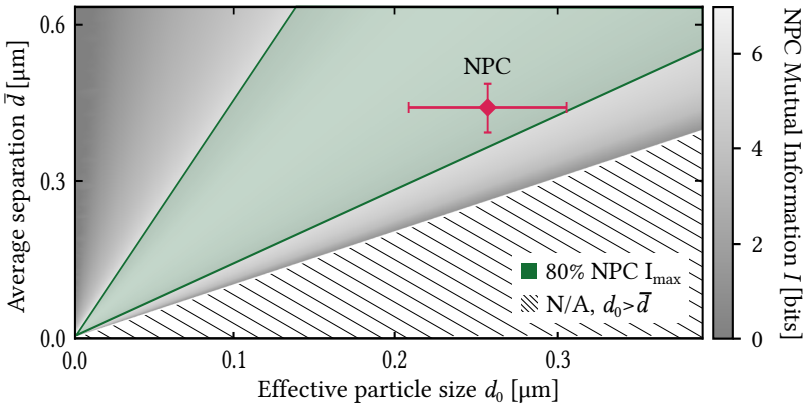


Figure 6.9: The fitted effective NPC size d_0, combined with the measured average NPC separation distance $\bar{d}$, indicate that the NPC distributions can binarise the MT proximity signal effectively, given the measured filament density $\phi = 0.21 \pm 0.03$. Reproduced from [213], licenced under CC BY 4.0.

particle size, d_0, and average NPC separation, $\bar{d}$, to the predicted region of optimal information transmission, calculated as described in Chapter 4 for filamentous inputs with an area fraction of $\phi = 0.21$ as measured for the *S. arctica* system (see Section 6.2). As shown in Figure 6.9, this reveals that the biophysical parameters of the *S. arctica* system allow the NPC pattern to efficiently transmit thresholded information about the MT proximity profiles into the NE, close to the optimal information transfer possible for this filter. Inspired by observations that MT-nucleus interactions can influence nuclear functions such as gene expression, mitotic remodelling, and chromatin organisation, this compressed filament proximity signal could be expected to coordinate subsequent processes within the nucleus [145, 298–304].

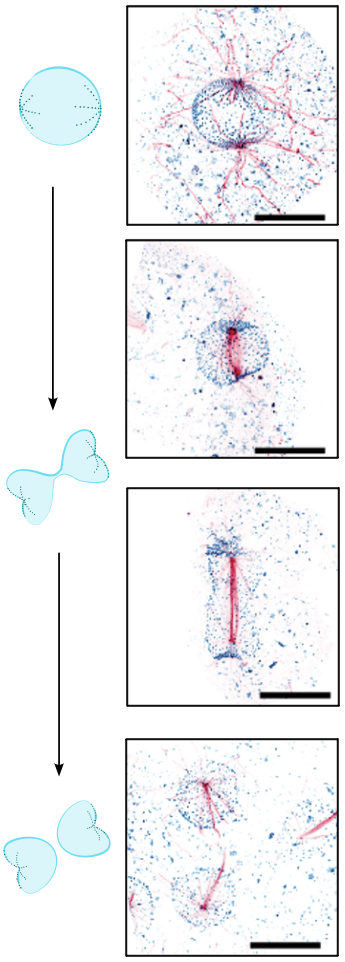


Figure 6.10: *S. arctica* nuclei undergo a 'closed mitosis' nuclear division cycle. Sketches (right) and maximum intensity projection expansion microscopy images (left, NPCs in blue, MTs in red) of this division process show that this cycle starts with the buckling of the MTOCs towards the nuclear centre. Images reproduced from [145], licenced under CC BY 4.0. Scale bars, 5 µm.

6.5 A possible readout mechanism

Throughout this thesis, it has been assumed that particle distributions facilitate information transmission by permitting processes within membrane-enclosed compartments to access information from the material surrounding the compartment. A key remaining question is therefore how such particle distributions could be read out by down stream tasks, the answer of which is likely to be system specific. While the following discussion is purely speculative, the NPC patterns discussed in this chapter may point towards a possible mechanism for reading out the information that they encode.

It has been observed that during the nuclear division cycle in *S. arctica*, the MTOCs dimple towards the centre of the nucleus, countering an outward force produced by the formation of the mitotic spindle (Figure 6.10) [145]. Occurring at approximately the same time in all of the nuclei of the cell, this dimpling may act as a trigger coordinating the subsequent nuclear division, perhaps by initialising the formation of the mitotic spindle. I suggest that such dimpling may be caused by a buckling

transition actuated by active stress imbalance between the MT and NPC layers of the pattern. In various species, interactions between NPCs and MTs are facilitated by the molecular motor dynein, or one of its homologues, which drive the active motion of cargo towards the minus end of MTs [312–315]. If the interaction presented in this chapter is mediated by a molecular motor like dynein, there would be an active force pulling the NPCs along the MT towards the MTOC. It is the subject of ongoing and future research to uncover the consequences of such active processes, however, it can be expected that this would lead to an increased gain of the filter and – more importantly in this case – produce a bi-layered system with one contractive layer of MT-associated dynein connected to an incompressible layer of high-density NPCs.

In brief, the NPCs would be stochastically pulled by the molecular motor, but their motion would eventually become arrested by their repulsion and result in maximum-density rods of NPCs. However, adjacent to the NPCs, the molecular motor may continue to walk along the microtubule, resulting in an effective active contraction. The resulting stress imbalance between the incompressible NPC rod on the nuclear surface and the MT-motor complex attached to it may lead to a buckling transition, akin to buckling in epithelial layers [316], causing a dimple to form when the stress imbalance is enough to overcome the surface tension and bending resistance of the membrane. Assuming this critical strain depends on the size of the high-density regions, this would permit the buckling transition to enact a binary readout of the size of the NPC asters. Such stress imbalance buckling transitions have indeed been observed in several biological systems, and have also been implemented in many synthetic systems, such as dome-patterned mechanical sensors and memory-storing hysteretic spring networks, to perform computations and binary readouts of continuous signals [19, 20, 51, 260, 316–324]. It would be an exciting future direction to investigate whether such a mechanism is at play in this biological system.

6.6 Concluding remarks

In summary, in this chapter I have presented an example biological system in which protein distributions may act as sigmoidal signal filters. Treating the structure of the microtubule cyctoskeleton as a fixed store of initial information in *S. arctica* cells, I demonstrated that distributions of NPCs embedded within the nuclear surface provide a thresholded readout of microtubule proximity to processes that occur within the nucleus, as such providing a means to coordinate intranuclear functions. Although I discussed briefly a possible mechanism for how such processes are read out, it remains an open question to determine whether this is truly a

mechanism of control in this system. As previously noted, the assumption of an effective elastic interaction and fixed input structure neglects any changes to the input that may be caused by the binding of the NPCs. It is, however, likely that such binding would influence the position of the microtubules, leading to a positive feedback loop between microtubule morphology and NPC distribution. Extending the approaches presented in this chapter to investigate the influence of feedback between the NPC distribution and the separation profile would be an interesting and pertinent extension of this work. As discussed in Section 5, such feedback is expected to lead to a sharper transition between the high and low density regions of the NPC distribution, although the resolution at which it is currently possible to estimate the NPC densities likely prevents observation of this difference. Overall, this chapter provides supporting evidence of pattern-based information transmission in real, living systems, and suggests a possible scenario in which to further test the implications of such signalling processes.

7 Summary and outlook

The development of responsive smart materials which mimic the life-like properties of biological systems, such adaptivity and responsivity, is expected to solve several of the challenges facing current information and computation technologies [1]. For the benefits of computation in such soft systems to be realised, several key constraints need to be overcome. For small-scale applications, the role of noise is expected to be particularly important [7]. Living cells provide a good inspiration for physical principles of computation in such noisy environments. In particular, the development of bottom-up synthetic cells and soft nanorobots has been facilitated by the understanding of biochemical and genetic signalling cascades [6]. However, despite evidence of their relevance in sensing and regulation in cells, spatially resolved particle distributions form a yet under-explored means of signal processing in noisy, soft, living materials. In this thesis, I have evaluated the performance of particle distributions in transmitting information across membrane compartment boundaries through the emergence of non-linear mappings which act as sigmoidal filters that can binarize heterogeneities in the environment.

Inter-particle interactions are known to play important roles in the emergence of protein patterns. Extending beyond chemical interactions, I therefore addressed how physical interactions between particles influence their spatial distribution. To place interaction-modulated particle distributions in a statistically supported non-equilibrium framework, I first presented an extension to current maximum caliber methods to address the impact that microscopic constraints on particle arrangements have on system dynamics. By considering state-transition statistics, I derived the local detailed balance and fluctuation theorem for continuous-time jump processes, as such providing further support for caliber approaches as the non-equilibrium equivalent to equilibrium ensemble methods. Through this approach, I accounted for general interacting particles, arising from both constraints on particle arrangements and from density-dependent interaction potentials, and showed how such interactions contribute to mean-field particle dynamics as density-dependent potentials. Currently, the microscopic caliber approach presented in this thesis is limited due to its heuristic argument for the network scaling dependence of discrete state escape rates, and a lack of treatment of particle-mediated energy fluxes. A future research direction could be to address these limitations, perhaps by explicitly accounting for ballistic movements in an approach that addresses both potential and kinetic energy.

Following the identification of contributions to mean field particle dynamics from physical particle interactions, I assessed the role of these interactions in determining equilibrium membrane-associated particle distributions, and controlling the transmission of spatially-resolved signals across biological interfaces. Considering input signals encoded within a particle-environment interaction energy field, I showed that inter-particle interactions lead to signal filters with non-linear response functions that map input signals to output particle density fields. Taking – as a key example – short range repulsive interactions, I identified a non-local sigmoidal filter mechanism, which enacts a binarisation of the input signal across a membrane and, using a van der Waals gas model, I further showed how such thresholding filters may be improved by longer-range attractive interactions. By considering these systems through the lens of signal processing, it may be possible to understand more about how patterns contribute to robust biological functions. It is exciting to think about whether such processes may also be utilized in synthetic cells.

Inspired by the diversity of structures in living cells, I further investigated how the shape of the input structure influences the response function of the particle-mediated thresholding filter. I found that this function depends on the distribution of interaction sites within the environment, with regions with no binding sites acting as a particle bath. Considering, in particular, linear output densities along filaments of binding sites, I found that a reduction in the area fraction of the membrane covered by the filaments leads to a reduction in the number of particles required to achieve the same form of response function. As such, encoding input signals in filamentous structures provides a key benefit over two-dimensional surfaces by reducing the need for large membrane particle saturations. Filaments are ubiquitous in subcellular biology, and are implicated in a wide range of signalling and regulatory processes [222]. My results suggest an explicit mechanism by which they may, when coupled to surface-mediated signal processing, provide a means of storing and communicating spatially resolved signals. Such 'dimension reduction' has been shown to be beneficial in other systems – including centriole formation and pattern emergence – and my results join this growing list of literature highlighting the importance of multi-dimensional interactions [137, 325]. Beyond these binarising filters, I further demonstrated that it is possible to develop more complex signal filters through allosteric mechanisms for distinguishing bound and unbound particles, specifically edge-detecting filters. This is exciting as it illustrates how complex calculations and pattern recognition may be achieved by guided particle distributions.

These results emphasise the role that hierarchical pattern forming systems

may play in signal transduction in living cells, suggesting a mechanism for signal processing that could be instigated to enact control in synthetic vesicles [6]. The integration of time-varying signals might allow additional signal processing modalities [87]. Addressing the forms of response functions in a dynamical setting would be an interesting and pertinent extension to this work, clarifying how actively driven systems transmit spatial signals.

The signal filter response functions describe the mean-field particle distributions resulting from input interaction energy fields. However, in a real system the density fluctuates about this value, impacting how well information can be transferred through the filter. The number of bits of transferred information is important for determining how reliably decisions can be made using the output signal. I therefore assessed the sensitivity of the binarising filters to particle noise by quantifying the information transmission that they achieve. Introducing a discrete mutual information that accounts for whole-field signals, I extend the existing mutual information formalisms to permit the quantification of information transmitted by non-local response functions. Although similar to the approach used in [16] to address information content in multiple genes, this is distinct in its application to non-local response functions. I identified optimal information transmission regimes, for both uniform and filamentous binding site distributions. This confirms my mean-field results that filaments provide an improved signal transmission compared to surfaces. By comparing with biological examples I demonstrated that these regimes are achievable in living systems. These results indicate that information transmission by particle distributions is plausible, and could be suitable for implementation in synthetic vesicle control. Unfortunately, computational capabilities limit my quantifications of information to small, discrete systems. More computational capacity would be needed to address this in a larger system where contributions from discretisation-induced artifacts would be less impactful.

By defining relevant features in the input signal, I also computed the transmission of relevant information, and in doing so assessed the level of signal compression resulting from both non-linear mappings and inherent channel noise. The resultant parameter regimes for relevant information transmission were found to approximately coincide with optimal regimes for general information transmission, with inputs encoded in filamentous structures performing better than those with uniform binding site distributions over a wider range of parameters. This process emphasised the tunability of the filter, with relevant signal features selected by tuning membrane saturation. Typically the selectivity of a filter would be assessed using the Fisher information [82, Chapter 11]. However,

this proves unwieldy and challenging for the case of spatially-resolved input signals, where the resultant matrix is multi-dimensional. The approach presented here provides an alternative option, albeit with the drawback of requiring the identification of a relevant signal. This desired readout, and the probability distribution of input signals used in the mutual information calculation, are system specific. Despite efforts made here to remain as general as possible, in applications to any particular system the analysis should be repeated with suitably chosen distributions and readouts.

Feedback is an important and unavoidable part of soft matter systems, and extensive work has shown that it can provide a means of system control, and aid in producing robust signalling cascades [140, 142, 265, 266]. By introducing feedback from the output density field back to the input energy fields, I found that particle-mediated signal filters also follow this trend, producing highly selective filters. This was shown to be particularly beneficial when combined with filamentous binding site distributions. However, interestingly, the total amount of information remained approximately constant, with the increased distinction between high and low density regions balanced by an increase in channel noise. These findings may help in the design process for particle-mediated signal filters by guiding the design of feedback-based control systems for sensing and signal transmission.

One of the best ways to know if a signal processing mechanism is practical is to see it in action. I therefore closed this thesis by presenting a biological system where this sensing mechanism may be at play. In collaboration with Hiral Shah and Gautam Dey, I used expansion microscopy data to investigate the origins of nuclear pore complex (NPC) asters in *S. arctica*, and showed that the NPC distributions followed the sigmoidal form expected for particle distributions acting as a binarising filter. Using parameter results found by fitting these distributions – verified by independent measurements – I showed that this system falls within the optimal information transmission regime. Although not a proof, these results lend weight to the interpretation of particle distributions as a signal processing medium. To address if this is indeed the case in *S. arctica*, future work is needed to identify the downstream consequences of these distributions, as well as to investigate what happens when information flow is interrupted or perturbed. In a step towards addressing this, I have suggested a possible readout mechanism for this system in the form of a buckling transition, by which the density-stored signal may contribute to coordinating the timing of the nuclear division.

While this thesis focused primarily on binarising filters, spatial information processing is likely at play in many subcellular contexts and it would be interesting to investigate how the guided formation of more

complex particle distributions may permit higher-level pattern recognition and information selection in these systems. Future work addressing how downstream processes read out particle densities and how larger feedback networks influence signal transmission are likely to shed light on such complex signal filters. A particularly interesting direction would be to extend the framework to consider input and output signals encoded in other spatially-varying system properties. Specifically, in this thesis the output signal was encoded within the particle density field, however, as indicated in the discussion of buckling, information can also be stored in other fields such as heterogeneous tensions. By extending the information-theoretic approaches here to the development of spatial order in other fields, a general principle for spatial signal processing may emerge. In the context of designing computational, small-scale metamaterials, the signal processing mechanism presented in this thesis may be enhanced through coupling to a memory-storage mechanism, which I have partially addressed by imposing a distinct fixed input and directionality of information cascade. Indeed, the geometry of subcellular structures, such as the cytoskeleton, have been suggested to store cellular memories [14]. However, true learning and complex communication may need a more explicit store of signals and a means of integrating signals from several sources. It is an interesting future research direction to consider how input information is stored and updated in such structures.

Overall, I present the formation of guided protein distributions along membranes as a signal filter mechanism and show that these distributions may be used to transmit information across membrane-enclosed compartment boundaries. This signal processing framework could aid the design of programmable synthetic vesicles or bio-compatible sensors by providing a new way for nanoscopic computational logic to be encoded in synthetic materials.

A

Appendices

A.1 Alternative derivation of the caliber for volume excluding particles

In compliment to the approach presented in the example of Section 2.1, here I provide an alternative method for constructing the caliber Ω for volume excluding particles, by iteratively addressing the particles entering a state from each of the possible origin states indexed by $i \in \{1, ..., L\}$. Starting with the n_l^{r} particles that remain in the same state, given that the particles cannot occupy the same lattice site, there are

$$\Omega_{l\to l} = \frac{g_l!}{(g_l - n_l^{\mathrm{r}})! n_l^{\mathrm{r}}!}$$

ways of arranging the remaining particles into the lattice sites. Similarly, given n_l^{r} sites are now occupied, there are

$$\Omega_{1\to l} = \frac{(g_l - n_l^{\mathrm{r}})!}{(g_l - n_l^{\mathrm{r}} - n_{1\to l})! n_{1\to l}!}$$

possible paths by which particles starting in state $i = 1$ can move into the new state l. Continuing this for all the initial states in the neighbourhood $\mathcal{N}(l)$ of state l, I find that the total number of paths to move all the particles into l is given by

$$\begin{aligned}\Omega_l = &\frac{g_l!}{(g_l - n_l^{\mathrm{r}})! n_l^{\mathrm{r}}!} \cdot \frac{(g_l - n_l^{\mathrm{r}})!}{(g_l - n_l^{\mathrm{r}} - n_{1\to l})! n_{1\to l}!} \cdot \\ &\ldots \cdot \frac{(g_l - n_l^{\mathrm{r}}) - \sum_{i\neq L} n_{i\to l})!}{(g_l - n_l^{\mathrm{r}}) - \sum_{i\in\mathcal{N}(l)} n_{i\to l})! n_{L\to l}!}.\end{aligned}$$

By symmetry, that it doesn't matter which order I chose to address the initial states i. Finally, this is repeated for all the final states l, giving the total path count as

$$\Omega = \prod_l \frac{g_l!}{(g_l - n_l^{\mathrm{r}} - \sum_{i\in\mathcal{N}(l)} n_{i\to l})! n_l^{\mathrm{r}}! \prod_{i\neq l} n_{i\to l}!},$$

as given in Equation (2.12).

A.2 Equilibrium interaction potentials from entropy maximization

This appendix section details methods and approaches included within the published article [213]

In the main text, I present a derivation of the contributions to the chemical potential arising from short-range inter-particle repulsion, first using a hard-core lattice gas model and second from a soft-core potential. In this appendix section I present an alternative approach for finding the chemical potential at equilibrium.

A.2.1 Soft-core interactions

Starting with the soft-core interactions, I re-derive the equilibrium solution through entropy maximization, which provides insight into the origins of the interaction potentials $E_i(\rho_\mathrm{u}, \rho_\mathrm{b})$. Consider a system of *indistinguishable* particles with continuous, pairwise-additive interactions, which are distributed amongst B multi-occupancy boxes of size dx (i.e. area dx^2) indexed by $j > 0$. For simplicity, I consider a one-dimensional system where these boxes are arranged in a line. The $j = 0$ box is considered as a reservoir, acting as a heat bath exchanging energy, but not particles, with the system. The particles are able to enter a bound state in any of the boxes. Identifying the number of unbound K_j and bound M_j particles in each box, I assign box degeneracies g_K and g_M for the unbound and bound particles – with energies ϵ_u (constant) and $\epsilon_{b,j}$ (box-dependent) – respectively. In addition to the single-particle energy contributions ϵ_b and ϵ_u arising from the particle state, I also introduce an energy term $E_r(K_j, M_j)$ to account for local interactions between the particles within a box.

Working in the microcanonical framework, I count the number of microstates W which share the macrostate defined by $\{K_j, M_j\}$. This is given by $W = \prod_{j\geq 0} W_j$, where W_j is given by the Maxwell-Boltzmann statistics [326, Chapter 13]

$$W_{j\geq 0} = \frac{g_K^{K_j}}{K_j!}\frac{g_M^{M_j}}{M_j!}. \tag{A.1}$$

Imposing constraints on the total energy U and total number of particles N, and the is number N_0 of particles in bath, the equilibrium state is found by maximising the Boltzmann entropy $S_\mathrm{B}/k_\mathrm{B} = \ln W$. Using the Lagrange multipliers α, α_0, β to apply the constraints on N, N_0, and U, respectively, the objective function

$$f(\{K_j\},\{M_j\}) = \ln W + \alpha\left(N - N_0 - \sum_{j>0}(K_j + M_j)\right) + \alpha_0(N_0 - K_0 - M_0) + \beta\left(U - \sum_{j\geq 0}\left(K_j\epsilon_\mathrm{u} + M_j\epsilon_{b,j} + E_r(K_j + M_j)\right)\right), \tag{A.2}$$

is extremised by requiring that its derivatives are equal to zero,

$$\frac{\partial f}{\partial K_j} = \ln g_K - \ln K_j - \delta_{0j}\alpha_0 - (1-\delta_{0j})\alpha - \beta\epsilon_\mathrm{u} - \beta\frac{\partial E_r(K_j + M_j)}{\partial K_j} = 0, \tag{A.3}$$

$$\frac{\partial f}{\partial M_j} = \ln g_M - \ln M_j - \delta_{0j}\alpha_0 - (1-\delta_{0j})\alpha - \beta\epsilon_{b,j} - \beta\frac{\partial E_r(K_j + M_j)}{\partial M_j} = 0. \tag{A.4}$$

Here, I apply the Stirling approximation, and use the Kronecker δ_{jk}. From Equations (A.3) and (A.4), the equilibrium box occupation numbers are found to be give by the Boltzmann distribution of particles over energy levels with the chemical potentials of the bath $\mu_0 = -\alpha_0/\beta$ and of the system $\mu_s = -\alpha/\beta$, i.e.

$$K_0 = g_K e^{\beta\left(\mu_0 - \epsilon_\mathrm{u} - \frac{\partial E_r(K_0 + M_0)}{\partial K_0}\right)}, \qquad M_0 = g_M e^{\beta\left(\mu_0 - \epsilon_{b,0} - \frac{\partial E_r(K_0 + M_0)}{\partial M_0}\right)} \tag{A.5}$$

$$K_{j>0} = g_K e^{\beta\left(\mu_s - \epsilon_\mathrm{u} - \frac{\partial E_r(K_j + M_j)}{\partial K_j}\right)}, \qquad M_{j>0} = g_M e^{\beta\left(\mu_s - \epsilon_{b,j} - \frac{\partial E_r(K_j + M_j)}{\partial M_j}\right)}. \tag{A.6}$$

The reservoir part of the solution, Equation (A.5)—which I keep for completeness, plays no role in the following.

I next divide Equation (A.6) by the box area dx^2, and take the continuum limit $dx \to 0$ as the number of

boxes increases to infinity, yielding the particle densities at $x = \lim_{dx\to 0} j\,dx$:

$$\lim_{dx\to 0} \frac{K_j}{dx^2} = \rho_{\mathrm{u}}(x) = g_{\mathrm{u}} e^{\beta\mu_s - \beta\epsilon_{\mathrm{u}} - \beta E(\rho)}, \tag{A.7}$$

$$\lim_{dx\to 0} \frac{M_j}{dx^2} = \rho_{\mathrm{b}}(x) = g_{\mathrm{b}} e^{\beta\mu_s - \beta\epsilon_{\mathrm{b}}(x) - \beta E(\rho)}, \tag{A.8}$$

with $\rho = \rho_{\mathrm{u}} + \rho_{\mathrm{b}}$ and the implied limits $\lim_{dx\to 0} \epsilon_{b,j} = \epsilon_{\mathrm{b}}(x)$,

$$\lim_{dx\to 0} \frac{g_K}{dx^2} = g_{\mathrm{u}}, \quad \lim_{dx\to 0} \frac{\partial}{\partial(K_j/dx^2)} \frac{1}{dx^2} E_r \left[dx^2 \left(\frac{K_i}{dx^2} + \frac{M_j}{dx^2} \right) \right] = \frac{\partial}{\partial \rho_{\mathrm{u}}} E_\rho(\rho_{\mathrm{u}} + \rho_{\mathrm{b}}) = \frac{\partial \rho}{\partial \rho_{\mathrm{u}}} \frac{\partial E_\rho(\rho)}{\partial \rho} = E(\rho), \tag{A.9}$$

$$\lim_{dx\to 0} \frac{g_M}{dx^2} = g_{\mathrm{b}}, \quad \lim_{dx\to 0} \frac{\partial}{\partial(M_j/dx^2)} \frac{1}{dx^2} E_r \left[dx^2 \left(\frac{K_j}{dx^2} + \frac{M_j}{dx^2} \right) \right] = \frac{\partial}{\partial \rho_{\mathrm{b}}} E_\rho(\rho_{\mathrm{u}} + \rho_{\mathrm{b}}) = \frac{\partial \rho}{\partial \rho_{\mathrm{b}}} \frac{\partial E_\rho(\rho)}{\partial \rho} = E(\rho). \tag{A.10}$$

Through comparison with Equations (3.2), I identify $\rho_{b,0} = g_u e^{\beta\mu_s}$, $\tilde{\mu}_u = \epsilon_u + E(\rho)$, and $\tilde{\mu}_b \epsilon_b + E(\rho) + k_B T \ln(g_u/g_b)$. Assuming, as equivalently assumed in the main text, that the degeneracy of the boxes is the same for bound and unbound particles, I recover the expected relation $E_i(\rho_u, \rho_b, x) = E(\rho)$, where $E(\rho)$ is defined through the limit in Equation (A.9).

A.2.2 Hard-core repulsion

Now considering the hard-core interactions, I derive the corrections to the chemical potential at equilibrium arising from particle exclusion through maximum entropy calculations using a lattice gas framework. Treating, as before, a system of particles distributed along a one-dimensional set of B boxes of area dx^2 and indexed by j, I now introduce a lattice within each box with a total of Ω sites, of which a subset Q permit the binding of particles. These sites are considered to be distinguishable. I then compute the entropy of the particles by counting the the number of microstates W that make up the macrostate defined by $\{K_j, M_j\}$, starting by counting the arrangements of the bound particles. Since these can only be in one of the Q binding-permitting sites, and no site can have more than one particle, the resulting number of combinations of the M_j indistinguishable particles is given by

$$W_{M,j} = \frac{Q!}{(Q - M_j)! M_j!}. \tag{A.11}$$

Next, the possible arrangements for the free particles are counted. These particles can be in any of the remaining $\Omega - M_j$ unoccupied sites, and therefore the number of arrangements for the free particles is given by

$$W_{K,j} = \frac{(\Omega - M_j)!}{(\Omega - M_j - K_j)! K_j!}. \tag{A.12}$$

The number of microstates for one box is then the product of Equations (A.11) and (A.12), giving

$$W_j = W_{K,j} W_{M,j} = \frac{Q!(\Omega - M_j)!}{(Q - M_j)!(\Omega - M_j - K_j)! K_j! M_j!}. \tag{A.13}$$

The equilibrium distribution of particles is found by extremizing $\ln W = \sum_{j\geq 0} \ln W_j$ subject to the same

energy and particle number constraints as before, yielding

$$K_j = (\Omega - M_j - K_j)e^{-\delta_{0j}\alpha_0} - (1-\delta_{0j})\alpha - \beta\epsilon_{\mathrm{u}}, \tag{A.14}$$

$$M_j = (\Omega - M_j - K_j)\frac{Q - M_j}{\Omega - M_j}e^{-\delta_{0j}\alpha_0} - (1-\delta_{0j})\alpha - \beta\epsilon_{\mathrm{b},j}. \tag{A.15}$$

Taking the continuum limit $dx \to 0$, following the process applied for Equations (A.7) and (A.8), I obtain the densities

$$\rho_{\mathrm{u}}(x) = \big(\rho_{\max} - \rho_{\mathrm{b}}(x) - \rho_{\mathrm{u}}(x)\big)e^{-\alpha-\beta\epsilon_{\mathrm{u}}} \tag{A.16}$$

$$\rho_{\mathrm{b}}(x) = \big(\rho_{\max} - \rho_{\mathrm{b}}(x) - \rho_{\mathrm{u}}(x)\big)\frac{\rho_{\mathrm{ext}} - \rho_{\mathrm{b}}(x)}{\rho_{\max} - \rho_{\mathrm{b}}(x)}e^{-\alpha-\beta\epsilon_{\mathrm{b}}(x)} \tag{A.17}$$

where $\rho_{\max} = 1/d_0^2 = \lim_{dx\to 0}\Omega/dx$ and $\rho_{\mathrm{ext}} = \lim_{dx\to 0} Q/dx$ are the maximum-packing and binding-site densities respectively. Using equations (3.2) and (3.5), the I recover the results for the interaction potential fields in Equations (3.13) and (3.14),

$$E_{\mathrm{u}}(\rho) = -k_{\mathrm{B}}T\ln\left(1 - \frac{\rho}{\rho_{\max}}\right) \tag{A.18}$$

and

$$\Delta E(\rho) = E_{\mathrm{u}}(\rho) - E_{\mathrm{b}}(\rho) = -k_{\mathrm{B}}T\ln\left(\frac{\rho_{\mathrm{ext}} - \rho_{\mathrm{b}}}{\rho_{\max} - \rho_{\mathrm{b}}}\right), \tag{A.19}$$

for $\rho_{b,0} = e^{-\alpha}\rho_{\max}$.

A.3 Biological system parameters for comparison to optimal regimes

Table A.1: Parameters reported in or estimated from literature for particle-like proteins and macromolecular structures at cellular and subcellular membrane surfaces, compared to optimal regimes for information transfer in Figure 4.7. Reproduced from [213], licenced under CC BY 4.0.

Protein structure	Average density $1/\bar{d}^2$ [μm^{-2}]	Average separation $\bar{d}$ [nm]	Interaction length d_0 [nm]
Integrin LFA-1	120 [327]	91	10 [328]
Integrin VLA-4	75 [327]	115	10 [328]
E-Cadherin	630 [329]	40	7 [329]
TCR	42.4 [330, 331, BNID 103567]**	154	10 [332] *
ICAM-1	900 [333]	33	3 [334] *
Connexons	5000 [335]	14	9.5 [336]
Piezo-1	0.52 [337]	1387	24 [338]
Caveolae	0.04 [339]	5000	75 [340]
VDAC	5000 [341]	14	3.8 [342]
ERMES	1200 [250]	29	15 [250]
Lipid GM1	17000 [343]**	8	3 [344]
AChR	55 [345, 346]**	135	10 [347, 348]
Fcγ-1	11 [349, 350] **	302	2.7 [351, BNID 117058]*
EGFR	636 [352]	40	17 [353]
Focal adhesions	0.14 [354, 355]**	2673	700 [354]
NPCs	-	440 ***	260 ***

* $d_0 = 2r_{\min}$ with $r_{\min}$ calculated from molecular mass value using [356]. ** Calculated as number of protein complexes per area of a spherical cell of a given diameter. *** Measured in our study.

A.4 Fluctuation calculations for particles densities with feedback

In Chapter 4, I derived the fluctuations in particle density expected in the presence of feedback. In this appendix section, I give the partial derivatives required to evaluate this probability, and show how the limit in λ arises from the integral process.

The non-zero partial derivatives of the free energy in Equation (5.6) are given by

$$\begin{aligned}
A = A_{\rho_{b,j}\rho_{b,j}} &= \frac{\partial^2 F}{\partial \rho_b^2}|_{\text{eq.}} = \frac{a\bar{\rho}_j}{\beta\bar{\rho}_{bj}(\bar{\rho}_j - \bar{\rho}_{bj})}, \\
B = A_{\rho_j\rho_j} &= \frac{\partial^2 F}{\partial \rho^2}|_{\text{eq.}} = \frac{a(\bar{\rho}_{\max} - \bar{\rho}_{bj})}{\beta(\bar{\rho}_j - \bar{\rho}_{bj})(\bar{\rho}_{\max} - \bar{\rho}_j)}, \\
C = A_{\epsilon_j\epsilon_j} &= \frac{\partial^2 F}{\partial \epsilon^2}|_{\text{eq.}} = a\beta^2\lambda, \\
D = A_{\rho_{bj}\epsilon_j} &= -\frac{\partial^2 \ln P}{\partial\epsilon\partial\rho_b}|_{\text{eq.}} = a, \\
E = A_{\rho_j\rho_{bj}} &= \frac{\partial^2 F}{\partial\rho\partial\rho_b}|_{\text{eq.}} = -\frac{a}{\beta(\bar{\rho}_j - \bar{\rho}_{bj})}.
\end{aligned} \tag{A.20}$$

Recalling the local conditional probability in Equation (5.8),

$$P(\rho_{b,j}, \rho_j, \epsilon_j|\{\epsilon_{0j}\}) \approx N_j \prod_{X,Y} e^{-\frac{\beta}{2}A_{XY}(X-\bar{X})(Y-\bar{Y})}, \tag{A.21}$$

the probability distribution for particle densities ρ_j can be found by integrating over the other variables $\rho_{b,j}$ and ϵ_j. Making the assumption that the fluctuations are small compared to the mean values $\bar{\rho}_{b,j}$ and ϵ_j, the integral limits are approximated as spanning an infinite domain, resulting in Gaussian integrals. Taking at first the integral over $\rho_{b,j}$ results in the conditional probability for ρ_j and ϵ_j,

$$P(\rho_j, \rho_{b,j}|\{\epsilon_{0j}\}) = N\sqrt{\frac{2\pi}{\beta A_{\rho_{bj}\rho_{bj}}}} e^{-\frac{\beta}{2}\left[A_{\rho_j\rho_j} - \frac{A^2_{\bar{\rho}_j\rho_{bj}}}{A_{\rho_{bj}\rho_{bj}}}\right](\rho_j-\bar{\rho}_j)^2} e^{-\frac{\beta}{2}\left[A_{\epsilon_j\epsilon_j} - \frac{A^2_{\bar{\rho}_{bj}\epsilon_j}}{A_{\rho_{bj}\rho_{bj}}}\right](\epsilon_j-\bar{\epsilon}_j)^2 + \frac{\beta A_{\rho_{bj}\epsilon_j}A_{\rho_j\rho_{bj}}}{A_{\rho_{bj}\rho_{bj}}}(\rho_j-\bar{\rho}_j)(\epsilon_j-\bar{\epsilon}_j)}, \tag{A.22}$$

with the constraint that $A_{\rho_{b,j}\rho_{b,j}} > 0$. Comparing to the results for the partial differentials in Equation (A.20) this is shown to hold.

Repeating this approach and integrating over ϵ_j gives the local conditional probability distribution for the densities,

$$P(\rho_j|\{\epsilon_{0j}\}) = N\frac{2\pi}{\sqrt{A_{\rho_{b,j}\rho_{b,j}}A_{\epsilon_j\epsilon_j} - A^2_{\bar{\rho}_{bj}\epsilon_j}}} e^{-\frac{\beta}{2}\left[A_{\rho_j\rho_j} - \frac{A^2_{\bar{\rho}_j\rho_{bj}}}{A_{\rho_{b,j}\rho_{b,j}}} - \frac{A^2_{\bar{\rho}_{bj}\epsilon_j}A^2_{\bar{\rho}_j\rho_{bj}}}{A_{\rho_{b,j}\rho_{b,j}}\left(A_{\epsilon_j\epsilon_j}A_{\rho_{b,j}\rho_{b,j}} - A^2_{\bar{\rho}_{bj}\epsilon_j}\right)}\right](\rho_j-\bar{\rho}_j)^2}. \tag{A.23}$$

Collecting the non-exponential terms in a single normalization constant $P_{0,j} = N\frac{2\pi}{\sqrt{A_{\rho_{b,j}\rho_{b,j}}A_{\epsilon_j\epsilon_j} - A^2_{\bar{\rho}_{bj}\epsilon_j}}}$, and substituting the partial differentials from Equation (A.20) results in the local conditional probability,

$$P(\rho_j|\{\epsilon_{0j}\}) = P_{0,j}e^{-\frac{1}{2}\frac{(\rho_j-\bar{\rho}_j)^2}{\sigma^2_{\rho_j}}}, \tag{A.24}$$

where, using $\bar{\rho}_b = \bar{\rho}_j/(1+e^{\beta\bar{\epsilon}_j})$,

$$\sigma_{\rho j}^{-2} = \frac{a}{\bar{\rho}_j(1-d_0^2\bar{\rho}_j)} - \frac{a}{(\beta\lambda(1+e^{\beta\bar{\epsilon}_j})^2 - \bar{\rho}_j e^{\beta\bar{\epsilon}_j})}, \tag{A.25}$$

as given in Equations (5.10) and (5.11).

The second Gaussian integral, over ϵ_j, is valid for $A_{\epsilon_j\epsilon_j} \geq A^2_{\rho_{bj}\epsilon_j}/A_{\rho_{b,j}\rho_{b,j}}$. Substituting into this constraint the results of the partial differentials in Equation (A.20), I find that this integral is only valid if $\beta\lambda \geq \bar{\rho}_{bj}(\bar{\rho}_j - \bar{\rho}_{bj})/\bar{\rho}_j$, and as such I constrain the subsequent analysis to the stiff limit in which $\beta\lambda \geq \rho_{max}/4 \geq \bar{\rho}_{bj}(\bar{\rho}_j - \bar{\rho}_{bj})/\bar{\rho}_j$.

A.5 Parameters and fits for sigmoidal NPC distributions

Table A.2: Parameter estimates for the effective spring rest length h_0 and the effective particle size d_0 for the NPC-MT system attained by Hiral Shah from electron tomograms of *S. arctica* nuclei available in [145]. Reproduced from [213], licenced under CC BY 4.0.

	Values measured from electron tomography images [nm]																							
A	71	90	89	110	75	80	87	81	71	85	85	88	65	70	94	83	85	88	91	83	110	91	108	57	66
B	266	236	222	183	156	161	163	141	224	249	235	186	228	178	187	135	151	169	177	229	200	162	190	170	184

A: Effective spring rest length h_0
B: Effective particle size d_0

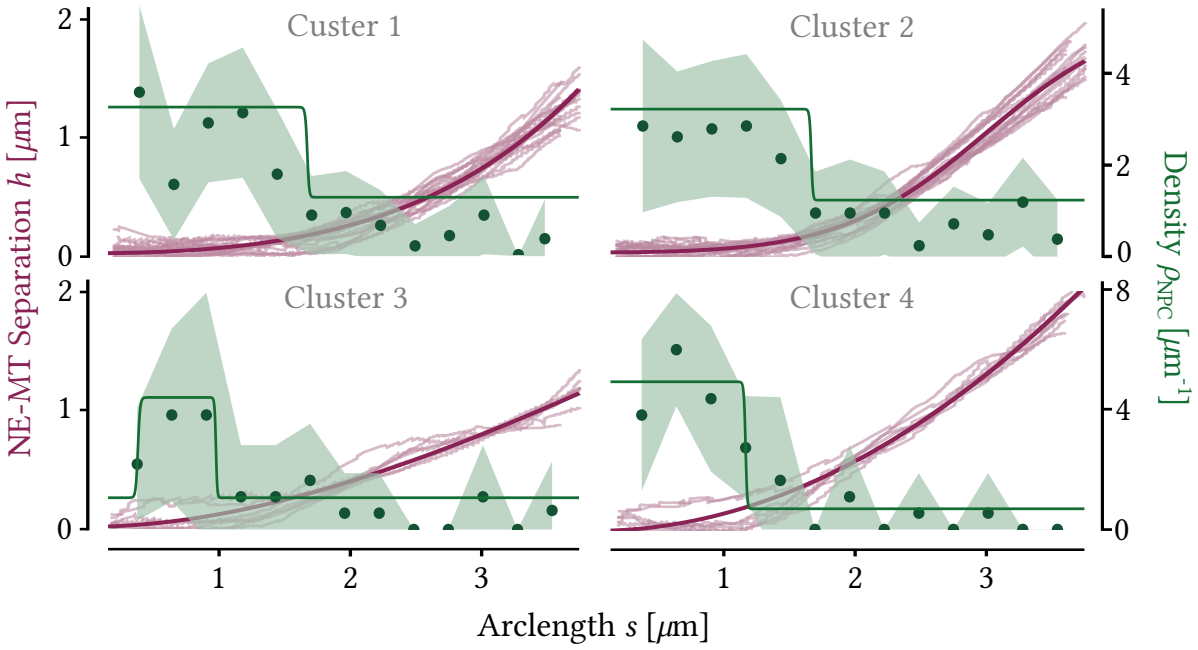


Figure A.1: NPC distributions form sigmoidal line densities along MT filaments. Independent fits of the response function (3.21) (solid green line) to measured NPC line densities (green dots with 95% confidence interval indicated by shaded area) shown here for clusters 1 to 4, with the fifth shown in the main text. The fits were made using the differentiable reparametrization of the separation distance h between the NE and the microtubules (dark purple) of the tracks (light purple) in each cluster. Each fit resulted in an estimate of the separation d_0 and the effective spring constant k. Reproduced from [213], licenced under CC BY 4.0.

List of publications

1	Jan Rombouts*, **Jenna Elliott***, Anna Erzberger **Forceful patterning: theoretical principles of mechanochemical pattern formation** *EMBO reports (2023) 24: e57739* doi.org/10.15252/embr.202357739	Chapter 1
2	**Jenna Elliott**, Hiral Shah, Roman Belousov, Gautam Dey, Anna Erzberger **Repulsive particle interactions enable selective information processing at cellular interfaces** *Accepted at Physical Review Letters* doi.org/10.1103/ywr6-mzz5	Chapter 3 Chapter 4 Chapter 6
3	Roman Belousov, **Jenna Elliott**, Florian Berger, Lamberto Rondoni, Anna Erzberger **Microcanonical ensemble out of equilibrium** *Under review at Physical Review E* doi.org/10.48550/arXiv.2504.20693	Chapter 2
4	Johannes Jung, **Jenna Elliott**†, Anna Erzberger† **Vesiculation-driven disposal on deformable membranes** *In preparation*	Chapter 5
5	**Jenna Elliott**, Anna Erzberger Working title: **Selective compression of spatial signals by protein patterns** *In preparation*	Chapter 3 Chapter 4 Chapter 5
6	**Jenna Elliott**, Roman Belousov, Anna Erzberger Working title: **Continuous and discrete diffusion models from microcanonical caliber methods** *In preparation*	Chapter 2
7	**Jenna Elliott**, Anna Erzberger Working title: **A surface and filament segmentation tool for three-dimensional image stacks** *In preparation*	Chapter 6

* These authors contributed equally to this work

† Co-corresponding authors

Bibliography

1. Alù, A. *et al.* Roadmap on embodying mechano-intelligence and computing in functional materials and structures. *Smart Materials and Structures* **34,** 063501 (2025).
2. Laschi, C. *et al.* Soft robotics: what's next in bioinspired design and applications of soft robots? *Bioinspiration & Biomimetics* (2025).
3. Volpe, G. *et al.* Roadmap for Animate Matter. *Journal of Physics: Condensed Matter* **37,** 333501 (2025).
4. Kaspar, C., Ravoo, B. J., van der Wiel, W. G., Wegner, S. V. & Pernice, W. H. P. The rise of intelligent matter. *Nature* **594,** 345–355 (2021).
5. Wei, M. & Kraft, D. J. *Autonomous life-like behavior emerging in active and flexible microstructures* arXiv:2506.15198. (2025).
6. Chen, S. *et al.* A roadmap for next-generation nanomotors. *Nature Nanotechnology* **20,** 990–1000 (2025).
7. Yasuda, H. *et al.* Mechanical computing. *Nature* **598,** 39–48 (2021).
8. Giaveri, S. *et al.* Building a Synthetic Cell Together. *Nature Communications* **16,** 7488 (2025).
9. Pedraza, J. M. & van Oudenaarden, A. Noise Propagation in Gene Networks. *Science* **307,** 1965–1969 (2005).
10. Kriebisch, C. M. *et al.* A roadmap toward the synthesis of life. *Chem* **11,** 102399 (2025).
11. Bradley, R. W. & Wang, B. Designer cell signal processing circuits for biotechnology. *New Biotechnology* **32.** European Congress of Biotechnology - ECB 16, 635–643 (2015).
12. Wolpert, L. Positional information and pattern formation in development. *Developmental Genetics* **15,** 485–490 (1994).
13. Tkačik, G. & Gregor, T. The many bits of positional information. *Development* **148,** dev176065 (2021).
14. Kalukula, Y. *et al.* The actin cortex acts as a mechanical memory of morphology in confined migrating cells. *Nature Physics* **21,** 1451–1461 (2025).
15. Luxton, G. G. & Gundersen, G. G. Orientation and function of the nuclear–centrosomal axis during cell migration. *Current Opinion in Cell Biology* **23,** 579–588 (2011).
16. Bauer, M., Petkova, M. D., Gregor, T., Wieschaus, E. F. & Bialek, W. Trading bits in the readout from a genetic network. *Proceedings of the National Academy of Sciences* **118,** e2109011118 (2021).
17. Rombouts, J., Elliott, J. & Erzberger, A. Forceful patterning: theoretical principles of mechanochemical pattern formation. *EMBO reports* **24,** e57739 (2023).
18. Jaeger, H., Noheda, B. & van der Wiel, W. G. Toward a Formal Theory for Computing Machines Made out of Whatever Physics Offers. *Nature Communications* **14,** 4911 (2023).
19. Kwakernaak, L. J. & van Hecke, M. Counting and Sequential Information Processing in Mechanical Metamaterials. *Physical Review Letters* **130,** 268204 (2023).
20. Liu, J. *et al.* Controlled pathways and sequential information processing in serially coupled mechanical hysterons. *Proceedings of the National Academy of Sciences* **121,** e2308414121 (2024).
21. Mazzolai, B. *et al.* Roadmap on Soft Robotics: Multifunctionality, Adaptability and Growth without Borders. *Multifunctional Materials* **5,** 032001 (2022).

22. Catozzi, S., Di-Bella, J. P., Ventura, A. C. & Sepulchre, J.-A. Signaling cascades transmit information downstream and upstream but unlikely simultaneously. *BMC Systems Biology* **10,** 84 (2016).

23. Moor, A.-L. & Zechner, C. Dynamic information transfer in stochastic biochemical networks. *Physical Review Research* **5,** 013032 (2023).

24. Lewandowsky, J., Bauch, G. & Stark, M. Information Bottleneck Signal Processing and Learning to Maximize Relevant Information for Communication Receivers. *Entropy* **24,** 972 (2022).

25. De Cheveigné, A. & Nelken, I. Filters: When, Why, and How (Not) to Use Them. *Neuron* **102,** 280–293 (2019).

26. Bechhoefer, J. *Control theory for physicists* (Cambridge University Press, Cambridge, England, 2021).

27. Bonev, B., Escolano, F. & Cazorla, M. A. *A Novel Information Theory Method for Filter Feature Selection* in *MICAI 2007: Advances in Artificial Intelligence* (eds Gelbukh, A. & Kuri Morales, Á. F.) (Springer Berlin Heidelberg, Berlin, Heidelberg, 2007), 431–440.

28. Frangi, A. F., Niessen, W. J., Vincken, K. L. & Viergever, M. A. *Multiscale vessel enhancement filtering* in *Medical Image Computing and Computer-Assisted Intervention — MICCAI'98* (eds Wells, W. M., Colchester, A. & Delp, S.) (Springer Berlin Heidelberg, Berlin, Heidelberg, 1998), 130–137.

29. Taye, M. M. Theoretical Understanding of Convolutional Neural Network: Concepts, Architectures, Applications, Future Directions. *Computation* **11,** 52 (2023).

30. Astrom, K. J. & Murray, R. M. *Feedback systems* (Princeton University Press, Princeton, NJ, 2008).

31. Dubey, S. R., Singh, S. K. & Chaudhuri, B. B. Activation functions in deep learning: A comprehensive survey and benchmark. *Neurocomputing* **503,** 92–108 (2022).

32. Estrada, J., Wong, F., DePace, A. & Gunawardena, J. Information Integration and Energy Expenditure in Gene Regulation. *Cell* **166,** 234–244 (2016).

33. Rubens, J. R., Selvaggio, G. & Lu, T. K. Synthetic mixed-signal computation in living cells. *Nature Communications* **7,** 11658 (2016).

34. Seelaboyina, R. & Vishwakarma, R. *Different Thresholding Techniques in Image Processing : A Review* in *ICDSMLA 2021* (eds Kumar, A., Senatore, S. & Gunjan, V. K.) (Springer Nature Singapore, Singapore, 2023), 23–29.

35. Ryzhkov, N. V., Nikolaev, K. G., Ivanov, A. S. & Skorb, E. V. Infochemistry and the Future of Chemical Information Processing. *Annual Review of Chemical and Biomolecular Engineering* **12,** 63–95 (2021).

36. Kim, S., Laschi, C. & Trimmer, B. Soft robotics: a bioinspired evolution in robotics. *Trends in Biotechnology* **31,** 287–294 (2013).

37. Ding, Y., Yalagala, B. P., Li, H., Mughal, M. W. & Heidari, H. Physical reservoir computing for biomedical applications. *Neuromorphic Computing and Engineering* **5,** 032002 (2025).

38. Kawasaki, T. & Miyazaki, K. Unified Understanding of Nonlinear Rheology near the Jamming Transition Point. *Physical Review Letters* **132,** 268201 (2024).

39. Gupta, S., Soni, J., Kumar, A. & Mandal, T. Origin of the nonlinear structural and mechanical properties in oppositely curved lipid mixtures. *The Journal of Chemical Physics* **159,** 165102 (2023).

40. Zimmerberg, J. & Kozlov, M. M. How proteins produce cellular membrane curvature. *Nature Reviews Molecular Cell Biology* **7,** 9–19 (2006).

41. Agudo-Canalejo, J. & Golestanian, R. Pattern formation by curvature-inducing proteins on spherical membranes. *New Journal of Physics* **19,** 125013 (2017).

42. Wang, Y., Li, L., Hofmann, D., Andrade, J. E. & Daraio, C. Structured fabrics with tunable mechanical properties. *Nature* **596,** 238–243 (2021).

43. Gupta, A., Chawla, K., Maheswaran, B., Syrlybayev, D. & Thevamaran, R. Embracing nonlinearity and geometry: a dimensional analysis guided design of shock absorbing materials. *Nature Communications* **16,** 7148 (2025).

44. Wang, X. & Cichos, F. Harnessing synthetic active particles for physical reservoir computing. *Nature Communications* **15,** 774 (2024).

45. Song, A., Murty Kottapalli, S. N., Goyal, R., Schölkopf, B. & Fischer, P. Low-power scalable multilayer optoelectronic neural networks enabled with incoherent light. *Nature Communications* **15,** 10692 (2024).

46. Stern, M., Guzman, M., Martins, F., Liu, A. J. & Balasubramanian, V. Physical Networks Become What They Learn. *Physical Review Letters* **134,** 147402 (2025).

47. Nakajima, K., Hauser, H., Li, T. & Pfeifer, R. Information processing via physical soft body. *Scientific Reports* **5,** 10487 (2015).

48. Bertoldi, K., Vitelli, V., Christensen, J. & van Hecke, M. Flexible mechanical metamaterials. *Nature Reviews Materials* **2,** 17066 (2017).

49. Jiao, Z. *et al.* Reprogrammable Metamaterial Processors for Soft Machines. *Advanced Science* **11,** 2305501 (2023).

50. Coulais, C., Teomy, E., de Reus, K., Shokef, Y. & van Hecke, M. Combinatorial design of textured mechanical metamaterials. *Nature* **535,** 529–532 (2016).

51. Faber, J. A., Udani, J. P., Riley, K. S., Studart, A. R. & Arrieta, A. F. Dome-Patterned Metamaterial Sheets. *Advanced Science* **7,** 2001955 (2020).

52. Guo, X., Guzmán, M., Carpentier, D., Bartolo, D. & Coulais, C. Non-orientable order and non-commutative response in frustrated metamaterials. *Nature* **618,** 506–512 (2023).

53. Melio, J., van Hecke, M., Henkes, S. E. & Kraft, D. J. *Colloidal Pivots Enable Brownian Metamaterials and Machines* arXiv:2503.17196. (2025).

54. Fischer, P. & Nelson, B. J. Tiny robots make big advances. *Science Robotics* **6,** eabh3168 (2021).

55. Bialek, W. Physical limits to sensation and perception. *Annual Review of Biophysics and Biophysical Chemistry* **16,** 455–478 (1987).

56. Muiños-Landin, S., Fischer, A., Holubec, V. & Cichos, F. Reinforcement learning with artificial microswimmers. *Science Robotics* **6,** eabd9285 (2021).

57. Kam, L., Shen, K. & Dustin, M. Micro- and Nanoscale Engineering of Cell Signaling. *Annual Review of Biomedical Engineering* **15,** 305–326 (2013).

58. Lu, K. *et al.* Advanced Bioinspired Organic Sensors for Future-Oriented Intelligent Applications. *Advanced Sensor Research* **2,** 2200066 (2023).

59. Tkačik, G. & Bialek, W. Information Processing in Living Systems. *Annual Review of Condensed Matter Physics* **7,** 89–117 (2016).

60. Bartlett, S. *et al.* Physics of Life: Exploring Information as a Distinctive Feature of Living Systems. *PRX Life* **3,** 037003 (2025).

61. Ganesan, A. & Zhang, J. How cells process information: Quantification of spatiotemporal signaling dynamics. *Protein Science* **21,** 918–928 (2012).

62. Tkačik, G. & Walczak, A. M. Information transmission in genetic regulatory networks: a review. *Journal of Physics: Condensed Matter* **23,** 153102 (2011).

63. Tohme, T., Vergassola, M., Mora, T. & Walczak, A. M. *Fast decisions with biophysically constrained gene promoter architectures* arXiv:2507.03720. (2025).

64. Schick, L., Kramar, M. & Alim, K. Dynamic Cost Allocation Allows Network-Forming Forager to Switch Between Search Strategies. *PRX Life* **2,** 033005 (2024).

65. Wurtz, T. Nested information processing in the living world. *Annals of the New York Academy of Sciences* **1500,** 5–16 (2021).

66. Brückner, D. B. & Tkačik, G. Information content and optimization of self-organized developmental systems. *Proceedings of the National Academy of Sciences* **121,** e2322326121 (2024).

67. Dubuis, J. O., Tkačik, G., Wieschaus, E. F., Gregor, T. & Bialek, W. Positional information, in bits. *Proceedings of the National Academy of Sciences* **110,** 16301–16308 (2013).

68. Seguin, C., Sporns, O. & Zalesky, A. Brain Network Communication: Concepts, Models and Applications. *Nature Reviews Neuroscience* **24,** 557–574 (2023).

69. Dullweber, T. & Erzberger, A. Mechanochemical feedback loops in contact-dependent fate patterning. *Current Opinion in Systems Biology* **32–33,** 100445 (2023).

70. Palmquist, K. H. *et al.* Reciprocal cell-ECM dynamics generate supracellular fluidity underlying spontaneous follicle patterning. *Cell* **185,** 1960–1973.e11 (2022).

71. Tsiairis, C. D. & Aulehla, A. Self-Organization of Embryonic Genetic Oscillators into Spatiotemporal Wave Patterns. *Cell* **164,** 656–667 (2016).

72. Falk, M. J., Strupp, A. T., Scellier, B. & Murugan, A. Temporal Contrastive Learning through implicit non-equilibrium memory. *Nature Communications* **16,** 2163 (2025).

73. Hakim, M. A. & Alam, M. I. Biologically inspired neural network layer with homeostatic regulation and adaptive repair mechanisms. *Scientific Reports* **15,** 33903 (2025).

74. Levchenko, A. & Nemenman, I. Cellular noise and information transmission. *Current Opinion in Biotechnology* **28,** 156–164 (2014).

75. Rhee, A., Cheong, R. & Levchenko, A. The application of information theory to biochemical signaling systems. *Physical Biology* **9,** 045011 (2012).

76. Bauer, M. How does an organism extract relevant information from transcription factor concentrations? *Biochemical Society Transactions* **50,** 1365–1376 (2022).

77. Binder, P., Schnellbächer, N. D., Höfer, T., Becker, N. B. & Schwarz, U. S. Optimal ligand discrimination by asymmetric dimerization and turnover of interferon receptors. *Proceedings of the National Academy of Sciences* **118,** e2103939118 (2021).

78. Kerr, R., Jabbari, S. & Johnston, I. G. Intracellular Energy Variability Modulates Cellular Decision-Making Capacity. *Scientific Reports* **9,** 20196 (2019).

79. Brennan, M. D., Cheong, R. & Levchenko, A. How Information Theory Handles Cell Signaling and Uncertainty. *Science* **338,** 334–335 (2012).

80. Cheong, R., Rhee, A., Wang, C. J., Nemenman, I. & Levchenko, A. Information Transduction Capacity of Noisy Biochemical Signaling Networks. *Science* **334,** 354–358 (2011).

81. Shannon, C. E. A Mathematical Theory of Communication. *Bell System Technical Journal* **27,** 379–423 (1948).

82. T.M. Cover & Thomas, J. A. *Elements of Information Theory* 99th ed. (John Wiley & Sons, Nashville, 1991).

83. Bauer, M. & Bialek, W. Information Bottleneck in Molecular Sensing. *PRX Life* **1,** 023005 (2023).

84. Uda, S. Application of information theory in systems biology. *Biophysical Reviews* **12,** 377–384 (2020).

85. Kramar, M., Hahn, L., Walczak, A. M., Mora, T. & Coppey, M. *Single cells can resolve graded stimuli* arXiv:2410.22571. (2024).

86. Razo-Mejia, M. *et al.* First-principles prediction of the information processing capacity of a simple genetic circuit. *Physical Review E* **102,** 022404 (2020).

87. Hahn, L., Walczak, A. M. & Mora, T. Dynamical Information Synergy in Biochemical Signaling Networks. *Physical Review Letters* **131,** 128401 (2023).

88. McGough, L. *et al.* Finding the Last Bits of Positional Information. *PRX Life* **2,** 013016 (2024).

89. Flowers, G. P. & Crews, C. M. Remembering where we are: Positional information in salamander limb regeneration. *Developmental Dynamics* **249,** 465–482 (2020).

90. Kline, A. G., Koch-Janusz, M., Walczak, A. M., Mora, T. & Palmer, S. E. *Coarse-graining reveals collective predictive information in a sensory population* bioRxiv, doi: 10.1101/2025.10.18.683195. (2025).

91. Altan-Bonnet, G., Mora, T. & Walczak, A. M. Quantitative immunology for physicists. *Physics Reports* **849,** 1–83 (2020).

92. Mahuas, G., Buffet, T., Marre, O., Ferrari, U. & Mora, T. Strong, but not Weak, Noise Correlations are Beneficial for Population Coding. *PRX Life* **3,** 033012 (3 2025).

93. Barton, N. H. & Tkačik, G. *Evolution and information content of optimal gene regulatory architectures* bioRxiv, doi: 10.1101/2025.06.10.657849. (2025).

94. Borbély, R. & Tkačik, G. *Regulatory architectures optimized for rapid evolution of gene expression* bioRxiv, doi: 10.1101/2025.06.10.658850. 2025.

95. Chlubna, T. & Zemčík, P. Comparative survey of image compression methods across different pixel formats and bit depths. *Signal, Image and Video Processing* **19,** 981 (2025).

96. Joshi, S. P., Gargate, P. A., Apotikar, Y. P., Jaiswal, R. C. & Munot, M. V. *A Comprehensive Investigation and Implementation of Lossless Image Compression Techniques for Social Media Network* in *ICT Systems and Sustainability* (eds Tuba, M., Akashe, S. & Joshi, A.) (Springer Nature Switzerland, Cham, 2026), 103–111.

97. Hosseini, M. *A Survey of Data Compression Algorithms and their Applications* arXiv:2506.10000. (2025).

98. Ketshabetswe, K. L., Zungeru, A. M., Mtengi, B., Lebekwe, C. K. & Prabaharan, S. R. S. Data Compression Algorithms for Wireless Sensor Networks: A Review and Comparison. *IEEE Access* **9,** 136872–136891 (2021).

99. Ibraheem, M. K. I., Dvorkovich, A. V. & Al-khafaji, I. M. A. A Comprehensive Literature Review on Image and Video Compression: Trends, Algorithms, and Techniques. *Ingénierie des systèmes d information* **29,** 863–876 (2024).

100. Azeloglu, E. U. & Iyengar, R. Signaling Networks: Information Flow, Computation, and Decision Making. *Cold Spring Harbor Perspectives in Biology* **7,** a005934 (2015).

101. Tishby, N., Pereira, F. C. & Bialek, W. *The information bottleneck method* arXiv:physics/0004057. (2000).

102. Strouse, D. & Schwab, D. J. The Deterministic Information Bottleneck. *Neural Computation* **29,** 1611–1630 (2017).

103. Slonim, N. & Tishby, N. *Agglomerative Information Bottleneck* in *Advances in Neural Information Processing Systems* (eds Solla, S., Leen, T. & Müller, K.) **12** (MIT Press, 1999).

104. Zaidi, A., Estella-Aguerri, I. & Shamai (Shitz), S. On the Information Bottleneck Problems: Models, Connections, Applications and Information Theoretic Views. *Entropy* **22,** 151 (2020).

105. Witteveen, O., Rosen, S. J., Lach, R. S., Wilson, M. Z. & Bauer, M. *Optimizing information transmission in the canonical Wnt pathway* arXiv:2506.22633. (2025).

106. Tishby, N. & Zaslavsky, N. *Deep learning and the information bottleneck principle* in *2015 IEEE Information Theory Workshop (ITW)* (2015), 1–5.

107. Weingarten, N., Yakhini, Z., Butman, M. & Gilad-Bachrach, R. *Tighter Bounds on the Information Bottleneck with Application to Deep Learning* arXiv:2402.07639. (2024).

108. Zhang, Q., Lu, M., Yu, S., Xin, J. & Chen, B. An information bottleneck approach for feature selection. *Pattern Recognition* **164,** 111564 (2025).

109. Kawaguchi, K., Deng, Z., Ji, X. & Huang, J. *How does information bottleneck help deep learning?* in *Proceedings of the 40th International Conference on Machine Learning* (JMLR.org, Honolulu, Hawaii, USA, 2023).

110. Kleinman, M. *et al. The information bottleneck as a principle underlying multi-area cortical representations during decision-making* bioRxiv, doi: 10.1101/2023.07.12.548742. (2023).

111. Alim, K., Andrew, N., Pringle, A. & Brenner, M. P. Mechanism of signal propagation in Physarum polycephalum. *Proceedings of the National Academy of Sciences* **114,** 5136–5141 (2017).

112. Vallverdú, J. *et al.* Slime mould: The fundamental mechanisms of biological cognition. *Biosystems* **165,** 57–70 (2018).

113. Le Verge-Serandour, M. & Alim, K. Physarum polycephalum: Smart Network Adaptation. *Annual Review of Condensed Matter Physics* **15,** 263–289 (2024).

114. Xie, J., Tato, C. M. & Davis, M. M. How the immune system talks to itself: the varied role of synapses. *Immunological Reviews* **251,** 65–79 (2012).

115. Seminario, M.-C. & Bunnell, S. C. Signal initiation in T-cell receptor microclusters. *Immunological Reviews* **221,** 90–106 (2008).

116. Hartman, N. C., Nye, J. A. & Groves, J. T. Cluster size regulates protein sorting in the immunological synapse. *Proceedings of the National Academy of Sciences* **106,** 12729–12734 (2009).

117. Huppa, J. B. & Davis, M. M. T-cell-antigen recognition and the immunological synapse. *Nature Reviews Immunology* **3,** 973–983 (2003).

118. Haupt, A. & Minc, N. How cells sense their own shape – mechanisms to probe cell geometry and their implications in cellular organization and function. *Journal of Cell Science* **131,** jcs214015 (2018).

119. Rangamani, P. *et al.* Decoding Information in Cell Shape. *Cell* **154,** 1356–1369 (2013).

120. Wigbers, M. C. *et al.* A hierarchy of protein patterns robustly decodes cell shape information. *Nature Physics* **17,** 578–584 (2021).

121. Bischof, J. *et al.* A cdk1 gradient guides surface contraction waves in oocytes. *Nature Communications* **8,** 849 (2017).

122. Hamaguchi, M. S. & Hiramoto, Y. Protoplasmic movement during polar-body formation in starfish oocytes. *Experimental Cell Research* **112,** 55–62 (1978).

123. Klughammer, N. *et al.* Cytoplasmic flows in starfish oocytes are fully determined by cortical contractions. *PLOS Computational Biology* **14** (ed Asthagiri, A. R.) e1006588 (2018).

124. Burkart, T., Wigbers, M. C., Würthner, L. & Frey, E. Control of protein-based pattern formation via guiding cues. *Nature Reviews Physics* **4,** 511–527 (2022).

125. Zieske, K. & Schwille, P. Reconstitution of self-organizing protein gradients as spatial cues in cell-free systems. *eLife* **3,** e03949 (2014).

126. Wettmann, L. & Kruse, K. The Min-protein oscillations in Escherichia coli: an example of self-organized cellular protein waves. *Philosophical Transactions of the Royal Society B: Biological Sciences* **373,** 20170111 (2018).

127. De Boer, P. A., Crossley, R. E. & Rothfield, L. I. A division inhibitor and a topological specificity factor coded for by the minicell locus determine proper placement of the division septum in E. coli. *Cell* **56,** 641–649 (1989).

128. Lutkenhaus, J. Assembly Dynamics of the Bacterial MinCDE System and Spatial Regulation of the Z Ring. *Annual Review of Biochemistry* **76,** 539–562 (2007).

129. Hu, Z. & Lutkenhaus, J. Topological regulation of cell division in Escherichia coli involves rapid pole to pole oscillation of the division inhibitor MinC under the control of MinD and MinE. *Molecular Microbiology* **34,** 82–90 (1999).

130. Raskin, D. M. & de Boer, P. A. J. Rapid pole-to-pole oscillation of a protein required for directing division to the middle of Escherichia coli. *Proceedings of the National Academy of Sciences* **96,** 4971–4976 (1999).

131. Loose, M., Fischer-Friedrich, E., Ries, J., Kruse, K. & Schwille, P. Spatial Regulators for Bacterial Cell Division Self-Organize into Surface Waves in Vitro. *Science* **320,** 789–792 (2008).

132. Ivanov, V. & Mizuuchi, K. Multiple modes of interconverting dynamic pattern formation by bacterial cell division proteins. *Proceedings of the National Academy of Sciences* **107,** 8071–8078 (2010).

133. Denk, J. *et al.* MinE conformational switching confers robustness on self-organized Min protein patterns. *Proceedings of the National Academy of Sciences* **115,** 4553–4558 (2018).

134. Collier, J. Hierarchical Dynamical Information Systems With a Focus on Biology. *Entropy* **5,** 100–124 (2003).

135. Falasco, G., Rao, R. & Esposito, M. Information Thermodynamics of Turing Patterns. *Physical Review Letters* **121,** 108301 (2018).

136. Würthner, L. *et al.* Bridging scales in a multiscale pattern-forming system. *Proceedings of the National Academy of Sciences* **119,** e2206888119 (2022).

137. Brauns, F., Halatek, J. & Frey, E. Phase-Space Geometry of Mass-Conserving Reaction-Diffusion Dynamics. *Physical Review X* **10,** 041036 (2020).

138. Halatek, J. & Frey, E. Rethinking pattern formation in reaction–diffusion systems. *Nature Physics* **14,** 507–514 (2018).

139. Staddon, M. F. How the zebra got its stripes: Curvature-dependent diffusion orients Turing patterns on three-dimensional surfaces. *Physical Review E* **110,** 034402 (2024).

140. Banerjee, S., Gardel, M. L. & Schwarz, U. S. The Actin Cytoskeleton as an Active Adaptive Material. *Annual Review of Condensed Matter Physics* **11,** 421–439 (2020).

141. El-Samad, H. Biological feedback control—Respect the loops. *Cell Systems* **12,** 477–487 (2021).

142. Hannezo, E. & Heisenberg, C.-P. Mechanochemical Feedback Loops in Development and Disease. *Cell* **178,** 12–25 (2019).

143. Norman, R. X. *et al.* One step 4× and 12× 3D-ExM enables robust super-resolution microscopy of nanoscale cellular structures. *Journal of Cell Biology* **224,** e202407116 (2025).

144. Radmacher, N. *et al.* Molecular Level Super-Resolution Fluorescence Imaging. *Annual Review of Biophysics* **54,** 163–184 (2025).

145. Shah, H. *et al.* Life-cycle-coupled evolution of mitosis in close relatives of animals. *Nature* **630,** 116–122 (2024).

146. Kardar, M. *Statistical physics of particles* (Cambridge University Press, Cambridge, England, 2007).

147. Yolcu, C., Haussman, R. C. & Deserno, M. The Effective Field Theory approach towards membrane-mediated interactions between particles. *Advances in Colloid and Interface Science* **208,** 89–109 (2014).

148. Sitarska, E. *et al.* Sensing their plasma membrane curvature allows migrating cells to circumvent obstacles. *Nature Communications* **14,** 5644 (2023).

149. Sciortino, A. *et al.* Active membrane deformations of a minimal synthetic cell. *Nature Physics* **21,** 799–807 (2025).

150. Beber, A. *et al.* Membrane reshaping by micrometric curvature sensitive septin filaments. *Nature Communications* **10,** 420 (2019).

151. Liboff, R. L. Transport Coefficients Determined Using the Shielded Coulomb Potential. *The Physics of Fluids* **2,** 40–46 (1959).

152. Rowlinson, J. The Yukawa potential. *Physica A: Statistical Mechanics and its Applications* **156,** 15–34 (1989).

153. Xu, Z., Wang, W., Cao, Y. & Xue, B. Liquid-liquid phase separation: Fundamental physical principles, biological implications, and applications in supramolecular materials engineering. *Supramolecular Materials* **2,** 100049 (2023).

154. Klosin, A. *et al.* Phase separation provides a mechanism to reduce noise in cells. *Science* **367,** 464–468 (2020).

155. Riback, J. A. & Brangwynne, C. P. Can phase separation buffer cellular noise? *Science* **367,** 364–365 (2020).

156. Alston, H., Rouches, M., Murugan, A., Walczak, A. M. & Mora, T. *Optimal sensing through phase separation* arXiv:2507.19021. (2025).

157. Laghmach, R., Di Pierro, M. & Potoyan, D. A. The interplay of chromatin phase separation and lamina interactions in nuclear organization. *Biophysical Journal* **120,** 5005–5017 (2021).

158. Li, C., Li, Z., Wu, Z. & Lu, H. Phase separation in gene transcription control. *Acta Biochimica et Biophysica Sinica* **55,** 1052–1063 (2023).

159. Peng, L., Li, E.-M. & Xu, L.-Y. From start to end: Phase separation and transcriptional regulation. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms* **1863,** 194641 (2020).

160. Wang, Y.-L. *et al.* Liquid-liquid phase separation in DNA double-strand breaks repair. *Cell Death & Disease* **14,** 746 (2023).

161. Zhang, R., Liu, Y. & Gao, J. *Chapter Four - Phase separation in controlling meiotic chromosome dynamics* in *Meiosis in Development and Disease* (ed Cole, F.) **151** (Academic Press, 2023), 69–90.

162. Kurita, R. Control of pattern formation during phase separation initiated by a propagated trigger. *Scientific Reports* **7,** 6912 (2017).

163. Antal, T., Bena, I., Droz, M., Martens, K. & Rácz, Z. Guiding fields for phase separation: Controlling Liesegang patterns. *Physical Review E* **76,** 046203 (2007).

164. Yang, Y., Ru, Y., Zhao, T. & Liu, M. Bioinspired multiphase composite gel materials: From controlled micro-phase separation to multiple functionalities. *Chem* **9,** 3113–3137 (2023).

165. Jin, Y.-y. *et al.* Deformation-induced phase separation of active vesicles. *Physical Chemistry Chemical Physics* **26,** 24699–24708 (2024).

166. Peliti, L. & Pigolotti, S. *Stochastic thermodynamics* en (Princeton University Press, Princeton, NJ, 2021).

167. Seifert, U. Stochastic thermodynamics, fluctuation theorems and molecular machines. *Reports on Progress in Physics* **75,** 126001 (2012).

168. Crooks, G. E. Entropy production fluctuation theorem and the nonequilibrium work relation for free energy differences. *Phys. Rev. E* **60,** 2721–2726 (3 1999).

169. Chavanis, P. H. Nonlinear mean field Fokker-Planck equations. Application to the chemotaxis of biological populations. *The European Physical Journal B* **62,** 179–208 (2008).

170. Elliott, M. & Ginossar, E. Applications of the Fokker-Planck equation in circuit quantum electrodynamics. *Physical Review A* **94,** 043840 (2016).

171. Risken, H. & Frank, T. D. *The Fokker-Planck equation* 2nd ed. en (Springer, Berlin, Germany, 1989).

172. Gardiner, C. W. *Stochastic methods: A Handbook for the Natural and Social Sciences* (Springer, Berlin, 2010).

173. Belousov, R., Hassanali, A. & Roldán, É. Statistical physics of inhomogeneous transport: Unification of diffusion laws and inference from first-passage statistics. *Phys. Rev. E* **106,** 014103 (1 2022).

174. Jaynes, E. T. The Minimum Entropy Production Principle. *Annual Review of Physical Chemistry* **31,** 579–601 (1980).

175. Pressé, S., Ghosh, K., Lee, J. & Dill, K. A. Principles of maximum entropy and maximum caliber in statistical physics. *Reviews of Modern Physics* **85,** 1115–1141 (2013).

176. Jaynes, E. T. Information Theory and Statistical Mechanics. *Physical Review* **106,** 620–630 (1957).

177. Ghosh, K., Dixit, P. D., Agozzino, L. & Dill, K. A. The Maximum Caliber Variational Principle for Nonequilibria. *Annual Review of Physical Chemistry* **71,** 213–238 (2020).

178. Dixit, P. D. *et al.* Perspective: Maximum caliber is a general variational principle for dynamical systems. *The Journal of Chemical Physics* **148,** 010901 (2018).

179. Seifert, U. *Stochastic thermodynamics* (Cambridge University Press, Cambridge, England, 2025).

180. Ge, H., Pressé, S., Ghosh, K. & Dill, K. A. Markov processes follow from the principle of maximum caliber. *The Journal of Chemical Physics* **136,** 064108 (2012).

181. Lecomte, V., Appert-Rolland, C. & van Wijland, F. Thermodynamic formalism and large deviation functions in continuous time Markov dynamics. *Comptes Rendus. Physique* **8,** 609–619 (2007).

182. Filyukov, A. A. & Karpov, V. Y. Method of the most probable path of evolution in the theory of stationary irreversible processes. *Journal of Engineering Physics* **13,** 416–419 (1967).

183. Filyukov, A. A. & Karpov, V. Y. Description of steady transport processes by the method of the most probable path of evolution. *Journal of Engineering Physics* **13,** 326–329 (1967).

184. Rogers, D. M. & Rempe, S. B. Irreversible Thermodynamics. *Journal of Physics: Conference Series* **402,** 012014 (2012).

185. Hazoglou, M. J., Walther, V., Dixit, P. D. & Dill, K. A. Communication: Maximum caliber is a general variational principle for nonequilibrium statistical mechanics. *The Journal of Chemical Physics* **143,** 051104 (2015).

186. Davis, S. & González, D. Hamiltonian formalism and path entropy maximization. *Journal of Physics A: Mathematical and Theoretical* **48,** 425003 (2015).

187. Monthus, C. Non-equilibrium steady states: maximization of the Shannon entropy associated with the distribution of dynamical trajectories in the presence of constraints. *Journal of Statistical Mechanics: Theory and Experiment* **2011,** P03008 (2011).

188. Bolhuis, P. G., Brotzakis, Z. F. & Vendruscolo, M. A maximum caliber approach for continuum path ensembles. *The European Physical Journal B* **94,** 188 (2021).

189. Movilla Miangolarra, O., Eldesoukey, A., Movilla Miangolarra, A. & Georgiou, T. T. Maximum entropy inference of reaction–diffusion models. *The Journal of Chemical Physics* **162,** 194104 (2025).

190. Abadi, N. & Ruzzenenti, F. Maximum entropy in dynamic complex networks. *Physical Review E* **110,** 054308 (2024).

191. Stock, G., Ghosh, K. & Dill, K. A. Maximum Caliber: A variational approach applied to two-state dynamics. *The Journal of Chemical Physics* **128,** 194102 (2008).

192. Ghosh, K., Dill, K. A., Inamdar, M. M., Seitaridou, E. & Phillips, R. Teaching the principles of statistical dynamics. *American Journal of Physics* **74,** 123–133 (2006).

193. Dixit, P. D. Stationary properties of maximum-entropy random walks. *Physical Review E* **92,** 042149 (2015).

194. Landau, L. D. & Lifshitz, E. M. *Statistical Physics* 3rd ed. (Butterworth-Heinemann, Oxford, 1996).

195. Saunders, S. The concept 'indistinguishable'. *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics* **71,** 37–59 (2020).

196. Saunders, S. The Gibbs Paradox. *Entropy* **20,** 552 (2018).

197. Aldous, D. & Fill, J. A. *Reversible Markov Chains and Random Walks on Graphs* Unfinished monograph, recompiled 2014, available at `http://www.https://www.stat.berkeley.edu/users/aldous/RWG/book.html`. 2002.

198. Maes, C. Local detailed balance. *SciPost Physics Lecture Notes* (2021).

199. Spinney, R. & Ford, I. *Fluctuation Relations: A Pedagogical Overview* in *Nonequilibrium Statistical Physics of Small Systems* (Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, Germany, 2013), 3–56.

200. Vanag, V. K. & Epstein, I. R. Cross-diffusion and pattern formation in reaction–diffusion systems. *Phys. Chem. Chem. Phys.* **11,** 897–912 (2009).

201. Farnsworth, K. D., Nelson, J. & Gershenson, C. Living is Information Processing: From Molecules to Global Systems. *Acta Biotheoretica* **61,** 203–222 (2013).

202. Collinet, C. & Lecuit, T. Programmed and self-organized flow of information during morphogenesis. *Nature Reviews Molecular Cell Biology* **22,** 245–265 (2021).

203. Evans, C. G., O'Brien, J., Winfree, E. & Murugan, A. Pattern recognition in the nucleation kinetics of non-equilibrium self-assembly. *Nature* **625,** 500–507 (2024).

204. Banerjee, D. S. *et al.* Learning via mechanosensitivity and activity in cytoskeletal networks. *PRX Life* (2025).

205. Brennan, M. D., Cheong, R. & Levchenko, A. How Information Theory Handles Cell Signaling and Uncertainty. *Science* **338,** 334–335 (2012).

206. Nandan, A., Das, A., Lott, R. & Koseska, A. Cells use molecular working memory to navigate in changing chemoattractant fields. *eLife* **11,** e76825 (2022).

207. Gomes, E. R., Jani, S. & Gundersen, G. G. Nuclear Movement Regulated by Cdc42, MRCK, Myosin, and Actin Flow Establishes MTOC Polarization in Migrating Cells. *Cell* **121,** 451–463 (2005).

208. Vercruysse, E. *et al.* Geometry-driven migration efficiency of autonomous epithelial cell clusters. *Nature Physics* **20,** 1492–1500 (2024).

209. d'Alessandro, J. *et al.* Cell migration guided by long-lived spatial memory. *Nature Communications* **12,** 4118 (2021).

210. Ron, J. E. *et al.* Emergent seesaw oscillations during cellular directional decision-making. *Nature Physics* **20,** 501–511 (2024).

211. Cho, S., Irianto, J. & Discher, D. E. Mechanosensing by the nucleus: From pathways to scaling relationships. *Journal of Cell Biology* **216,** 305–315 (2017).

212. Gatenby, R. A. The Role of Cell Membrane Information Reception, Processing, and Communication in the Structure and Function of Multicellular Tissue. *International Journal of Molecular Sciences* **20,** 3609 (2019).

213. Elliott, J., Shah, H., Belousov, R., Dey, G. & Erzberger, A. Repulsive particle interactions enable selective information processing at cellular interfaces. *Phys. Rev. Lett.* (2025).

214. Elliott, J. *Elliott2025 Mathematica Files* `https://git.embl.de/elliot/Elliott2025-Mathematica-Files`. 2025.

215. Elliott, J. *Doctoral Thesis Additional Files* `https://git.embl.org/elliot/doctoral-thesis-additional-files`. 2025.

216. Karal, M. A. S., Billah, M. M., Ahmed, M. & Ahamed, M. K. A review on the measurement of the bending rigidity of lipid membranes. *Soft Matter* **19,** 8285–8304 (2023).

217. Milo, R. & Phillips, R. *Cell Biology by the Numbers* (Garland Science, New York, 2015).

218. Zhdanov, V. P. Conditions of appreciable influence of microRNA on a large number of target mRNAs. *Molecular BioSystems* **5,** 638 (2009).

219. Hu, M., Briguglio, J. J. & Deserno, M. Determining the Gaussian Curvature Modulus of Lipid Membranes in Simulations. *Biophysical Journal* **102,** 1403–1410 (2012).

220. Wennerström, H., Vallina Estrada, E., Danielsson, J. & Oliveberg, M. Colloidal stability of the living cell. *Proceedings of the National Academy of Sciences* **117,** 10113–10121 (2020).

221. Mobley, D. L. & Gilson, M. K. Predicting Binding Free Energies: Frontiers and Benchmarks. *Annual Review of Biophysics* **46,** 531–558 (2017).

222. Hohmann, T. & Dehghani, F. The Cytoskeleton—A Complex Interacting Meshwork. *Cells* **8,** 362 (2019).

223. Kafri, M., Metzl-Raz, E., Jona, G. & Barkai, N. The Cost of Protein Production. *Cell Reports* **14,** 22–31 (2016).

224. Kaleta, C., Schäuble, S., Rinas, U. & Schuster, S. Metabolic costs of amino acid and protein production in Escherichia coli. *Biotechnology Journal* **8,** 1105–1114 (2013).

225. Huang, K. *Statistical Mechanics* 2nd ed. (John Wiley & Sons, Nashville, TN, 1987).

226. Jülicher, F. & Weber, C. A. Droplet Physics and Intracellular Phase Separation. *Annual Review of Condensed Matter Physics* **15,** 237–261 (2024).

227. Voorhees, P. W. The theory of Ostwald ripening. *Journal of Statistical Physics* **38,** 231–252 (1985).

228. Ong, J. Y. & Torres, J. Z. Phase Separation in Cell Division. *Molecular Cell* **80,** 9–20 (2020).

229. Liu, J. & Nussinov, R. Allostery: An Overview of Its History, Concepts, Methods, and Applications. *PLOS Computational Biology* **12,** e1004966 (2016).

230. Tabor, J. J. *et al.* A Synthetic Genetic Edge Detection Program. *Cell* **137,** 1272–1281 (2009).

231. Kerr, D., Coleman, S., McGinnity, M., Wu, Q. & Clogenson, M. *Biologically inspired edge detection* in *2011 11th International Conference on Intelligent Systems Design and Applications* (2011), 802–807.

232. Luo, Z., Lin, C., Li, F. & Pan, Y. BLEDNet: Bio-inspired lightweight neural network for edge detection. *Engineering Applications of Artificial Intelligence* **124,** 106530 (2023).

233. Yang, D., Peng, B. & Wu, X. A bio-inspired edge and segment detection method by modeling multiple visual regions. *The Visual Computer* **41,** 3173–3188 (2024).

234. Adamala, K. P. *et al.* Present and future of synthetic cell development. *Nature Reviews Molecular Cell Biology* **25,** 162–167 (2024).

235. Inc., W. R. *Mathematica, Version 14.0* Champaign, IL, 2024.

236. Touchette, H. The large deviation approach to statistical mechanics. *Physics Reports* **478,** 1–69 (2009).

237. Hastings, W. K. Monte Carlo sampling methods using Markov chains and their applications. *Biometrika* **57,** 97–109 (1970).

238. Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H. & Teller, E. Equation of State Calculations by Fast Computing Machines. *The Journal of Chemical Physics* **21,** 1087–1092 (1953).

239. Chib, S. & Greenberg, E. Understanding the Metropolis-Hastings Algorithm. *The American Statistician* **49,** 327–335 (1995).

240. Dunson, D. B. & Johndrow, J. E. The Hastings algorithm at fifty. *Biometrika* **107,** 1–23 (2019).

241. Tkačik, G. & Wolde, P. R. t. Information Processing in Biochemical Networks. *Annual Review of Biophysics* **54,** 249–274 (2025).

242. Alberts, B. *et al. The Cytoskeleton and Cell Behavior* in *Molecular Biology of the Cell* 4th (Garland, New York, 2002).

243. Cho, Y. *et al.* Tricellulin secures the epithelial barrier at tricellular junctions by interacting with actomyosin. *Journal of Cell Biology* **221,** e202009037 (2022).

244. Resnik-Docampo, M. *et al.* Tricellular junctions regulate intestinal stem cell behaviour to maintain homeostasis. *Nature Cell Biology* **19,** 52–59 (2017).

245. Howard, J. Mechanical Signaling in Networks of Motor and Cytoskeletal Proteins. *Annual Review of Biophysics* **38,** 217–234 (2009).

246. Sun, X. & Alushin, G. M. Cellular force-sensing through actin filaments. *The FEBS Journal* **290,** 2576–2589 (2023).

247. Bosveld, F. *et al.* Epithelial tricellular junctions act as interphase cell shape sensors to orient mitosis. *Nature* **530,** 495–498 (2016).

248. Kirichenko, E. Y., Skatchkov, S. N. & Ermakov, A. M. Structure and Functions of Gap Junctions and Their Constituent Connexins in the Mammalian CNS. *Biochemistry (Moscow), Supplement Series A: Membrane and Cell Biology* **15,** 107–119 (2021).

249. Alberts, B. *et al. Cell Junctions* in *Molecular Biology of the Cell* 4th (Garland, New York, 2002).

250. Wozny, M. R. *et al.* In situ architecture of the ER–mitochondria encounter structure. *Nature* **618,** 188–192 (2023).

251. Cheema, J. Y., He, J., Wei, W. & Fu, C. The Endoplasmic Reticulum-Mitochondria Encounter Structure and its Regulatory Proteins. *Contact* **4,** 25152564211064491 (2021).

252. Dai, Z., Luo, X., Xie, H. & Peng, H. B. The Actin-Driven Movement and Formation of Acetylcholine Receptor Clusters. *The Journal of Cell Biology* **150,** 1321–1334 (2000).

253. Chen, Y. *et al.* Coronin 6 Regulates Acetylcholine Receptor Clustering through Modulating Receptor Anchorage to Actin Cytoskeleton. *The Journal of Neuroscience* **34,** 2413–2421 (2014).

254. Xing, G. *et al.* A mechanism in agrin signaling revealed by a prevalent Rapsyn mutation in congenital myasthenic syndrome. *eLife* **8,** e49180 (2019).

255. Troyanovsky, R. B., Indra, I. & Troyanovsky, S. M. Actin-dependent α-catenin oligomerization contributes to adherens junction assembly. *Nature Communications* **16,** 1801 (2025).

256. Dustin, M. L. & Cooper, J. A. The immunological synapse and the actin cytoskeleton: molecular hardware for T cell signaling. *Nature Immunology* **1,** 23–29 (2000).

257. Celli, L., Ryckewaert, J.-J., Delachanal, E. & Duperray, A. Evidence of a Functional Role for Interaction between ICAM-1 and Nonmuscle α-Actinins in Leukocyte Diapedesis. *The Journal of Immunology* **177,** 4113–4121 (2006).

258. Carpén, O., Pallai, P., Staunton, D. E. & Springer, T. A. Association of intercellular adhesion molecule-1 (ICAM-1) with actin-containing cytoskeleton and alpha-actinin. *The Journal of cell biology* **118,** 1223–1234 (1992).

259. Lembo, S. *et al. The distance between the plasma membrane and the actomyosin cortex acts as a nanogate to control cell surface mechanics* bioRxiv, doi: 10.1101/2023.01.31.526409. (2023).

260. Udani, J. P. & Arrieta, A. F. Taming geometric frustration by leveraging structural elasticity. *Materials &; Design* **221,** 110809 (2022).

261. Kolchinsky, A., Tracey, B. D. & Kuyk, S. V. *Caveats for information bottleneck in deterministic scenarios* in *International Conference on Learning Representations* (2019).

262. Mancini, R. *Chapter 6 - Development of the Non Ideal Op Amp Equations* in *Op Amps for Everyone (Second Edition)* (ed Mancini, R.) Second Edition, 67–75 (Newnes, Burlington, 2003).

263. Riccardi, L., Naso, D., Turchiano, B. & Janocha, H. Design of Linear Feedback Controllers for Dynamic Systems With Hysteresis. *IEEE Transactions on Control Systems Technology* **22,** 1268–1280 (2014).

264. Jones, R. D. *et al.* Robust and tunable signal processing in mammalian cells via engineered covalent modification cycles. *Nature Communications* **13,** 1720 (2022).

265. Ingolia, N. T. & Murray, A. W. Positive-Feedback Loops as a Flexible Biological Module. *Current Biology* **17,** 668–677 (2007).

266. Autorino, C. *et al. A closed feedback between tissue phase transitions and morphogen gradients drives patterning dynamics* bioRxiv, doi: 10.1101/2025.06.06.658228. (2025).

267. Moussaid, M., Garnier, S., Theraulaz, G. & Helbing, D. Collective Information Processing and Pattern Formation in Swarms, Flocks, and Crowds. *Topics in Cognitive Science* **1,** 469–497 (2009).

268. Lestas, I., Vinnicombe, G. & Paulsson, J. Fundamental limits on the suppression of molecular fluctuations. *Nature* **467,** 174–178 (2010).

269. Ben Isaac, E., Manor, U., Kachar, B., Yochelis, A. & Gov, N. S. Linking actin networks and cell membrane via a reaction-diffusion-elastic description of nonlinear filopodia initiation. *Physical Review E* **88,** 022718 (2013).

270. Wu, Z., Su, M., Tong, C., Wu, M. & Liu, J. Membrane shape-mediated wave propagation of cortical protein dynamics. *Nature Communications* **9,** 136 (2018).

271. Miao, Y. *et al.* Wave patterns organize cellular protrusions and control cortical dynamics. *Molecular Systems Biology* **15,** e8585 (2019).

272. Pezzotta, A. & Briscoe, J. Optimal control of gene regulatory networks for morphogen-driven tissue patterning. *Cell Systems* **14,** 940–952.e11 (2023).

273. Tang, Y. *et al.* Quantifying information accumulation encoded in the dynamics of biochemical signaling. *Nature Communications* **12,** 1272 (2021).

274. Jung, J. *The disposal of membrane bound proteins lacking interactions with adjacent anchoring structures via membrane fission* Master's thesis (Heidelberg University, Heidelberg, Germany, 2025).

275. Torday, J. Homeostasis as the Mechanism of Evolution. *Biology* **4,** 573–590 (2015).

276. Watson, H. Biological membranes. *Essays in Biochemistry* **59,** 43–69 (2015).

277. Dultz, E., Wojtynek, M., Medalia, O. & Onischenko, E. The Nuclear Pore Complex: Birth, Life, and Death of a Cellular Behemoth. *Cells* **11,** 1456 (2022).

278. Fernandez-Martinez, J. & Rout, M. P. One Ring to Rule them All? Structural and Functional Diversity in the Nuclear Pore Complex. *Trends in Biochemical Sciences* **46,** 595–607 (2021).

279. Bahmanyar, S. & Schlieker, C. Lipid and protein dynamics that shape nuclear envelope identity. *Molecular Biology of the Cell* **31** (ed Kozminski, K.) 1315–1323 (2020).

280. Balaji, A. K., Saha, S., Deshpande, S., Poola, D. & Sengupta, K. Nuclear envelope, chromatin organizers, histones, and DNA: The many achilles heels exploited across cancers. *Frontiers in Cell and Developmental Biology* **10** (2022).

281. Daigle, N. *et al.* Nuclear pore complexes form immobile networks and have a very low turnover in live mammalian cells. *The Journal of Cell Biology* **154,** 71–84 (2001).

282. Rabut, G., Doye, V. & Ellenberg, J. Mapping the dynamic organization of the nuclear pore complex inside single living cells. *Nature Cell Biology* **6,** 1114–1121 (2004).

283. Toyama, B. H. & Hetzer, M. W. Protein homeostasis: live long, won't prosper. *Nature Reviews Molecular Cell Biology* **14,** 55–61 (2013).

284. Varberg, J. M., Unruh, J. R., Bestul, A. J., Khan, A. A. & Jaspersen, S. L. Quantitative analysis of nuclear pore complex organization in Schizosaccharomyces pombe. *Life Science Alliance* **5,** e202201423 (2022).

285. Soheilypour, M., Peyro, M., Jahed, Z. & Mofrad, M. R. K. On the Nuclear Pore Complex and Its Roles in Nucleo-Cytoskeletal Coupling and Mechanobiology. *Cellular and Molecular Bioengineering* **9,** 217–226 (2016).

286. Biedzinski, S. *et al.* Microtubules control nuclear shape and gene expression during early stages of hematopoietic differentiation. *The EMBO Journal* **39,** e103957 (2020).

287. Geng, J. *et al.* Microtubule Assists Actomyosin to Regulate Cell Nuclear Mechanics and Chromatin Accessibility. *Research* **6,** 0054 (2023).

288. Maizels, Y. & Gerlitz, G. Shaping of interphase chromosomes by the microtubule network. *The FEBS Journal* **282,** 3500–3524 (2015).

289. Nair, A., Khanna, J., Kler, J., Ragesh, R. & Sengupta, K. Nuclear envelope and chromatin choreography direct cellular differentiation. *Nucleus* **16,** 2449520 (2025).

290. Goldberg, M. W. Nuclear pore complex tethers to the cytoskeleton. *Seminars in Cell &; Developmental Biology* **68,** 52–58 (2017).

291. Huang, P. *et al.* The nuclear pore Y-complex functions as a platform for transcriptional regulation of FLOWERING LOCUS C in Arabidopsis. *The Plant Cell* **36,** 346–366 (2024).

292. Guo, Y. & Zheng, Y. Lamins position the nuclear pores and centrosomes by modulating dynein. *Molecular Biology of the Cell* **26** (ed Magin, T. M.) 3379–3389 (2015).

293. Hoffman, L. M. *et al.* Mechanical stress triggers nuclear remodeling and the formation of transmembrane actin nuclear lines with associated nuclear pore complexes. *Molecular Biology of the Cell* **31** (ed Weaver, V. M.) 1774–1787 (2020).

294. Donnaloja, F., Jacchetti, E., Soncini, M. & Raimondi, M. T. Mechanosensing at the Nuclear Envelope by Nuclear Pore Complex Stretch Activation and Its Effect in Physiology and Pathology. *Frontiers in Physiology* **10,** 896 (2019).

295. Ondracka, A., Dudin, O. & Ruiz-Trillo, I. Decoupling of Nuclear Division Cycles and Cell Size during the Coenocytic Growth of the Ichthyosporean Sphaeroforma arctica. *Current Biology* **28,** 1964–1969.e2 (2018).

296. Steinberg, G. *et al.* Motor-driven motility of fungal nuclear pores organizes chromosomes and fosters nucleocytoplasmic transport. *Journal of Cell Biology* **198,** 343–355 (2012).

297. Zsok, J. *et al.* Nuclear basket proteins regulate the distribution and mobility of nuclear pore complexes in budding yeast. *Molecular Biology of the Cell* **35** (ed Corbett, A.) (2024).

298. Velasquez-Carvajal, D. *et al.* Microtubule reorganization during mitotic cell division in the dinoflagellate Ostreospis cf. ovata. *Journal of Cell Science* **137,** jcs261733 (2024).

299. Maiato, H., Sampaio, P. & Sunkel, C. E. *Microtubule-Associated Proteins and Their Essential Roles During Mitosis* in, 53–153 (Academic Press, 2004).

300. Roth, D. M., Moseley, G. W., Glover, D., Pouton, C. W. & Jans, D. A. A Microtubule-Facilitated Nuclear Import Pathway for Cancer Regulatory Proteins. *Traffic* **8,** 673–686 (2007).

301. Roth, D. M., Moseley, G. W., Pouton, C. W. & Jans, D. A. Mechanism of Microtubule-facilitated "Fast Track" Nuclear Import. *Journal of Biological Chemistry* **286,** 14335–14351 (2011).

302. Kelley, M. E. *et al.* Spastin regulates anaphase chromosome separation distance and microtubule-containing nuclear tunnels. *Molecular Biology of the Cell* **35** (ed Thery, M.) ar48 (2024).

303. Gerlitz, G., Reiner, O. & Bustin, M. Microtubule dynamics alter the interphase nucleus. *Cellular and Molecular Life Sciences* **70,** 1255–1268 (2013).

304. Dudin, O. *et al.* A unicellular relative of animals generates a layer of polarized cells by actomyosin-dependent cellularization. *eLife* **8,** e49801 (2019).

305. Wassie, A. T., Zhao, Y. & Boyden, E. S. Expansion microscopy: principles and uses in biological research. *Nature Methods* **16,** 33–41 (2018).

306. Elliott, J. *Elliott2025 Image Analysis S-arctica* `https://git.embl.de/elliot/Elliott2025-Image-Analysis-S-arctica`. 2025.

307. Elliott, J., Shah, H., Belousov, R., Dey, G. & Erzberger, A. *Image dataset used in 'Repulsive particle interactions enable selective information processing at cellular interfaces'* `https://www.ebi.ac.uk/biostudies/bioimages/studies/S-BIAD2081`. 2025.

308. Schindelin, J. *et al.* Fiji: an open-source platform for biological-image analysis. *Nature Methods* **9,** 676–682 (2012).

309. Van der Walt, S. *et al.* scikit-image: image processing in Python. *PeerJ* **2,** e453 (2014).

310. Bukenya, F., Bai, L. & Kiweewa, A. *A Review of Blood Vessel Segmentation Techniques* in *2018 1st International Conference on Computer Applications &; Information Security (ICCAIS)* (IEEE, 2018), 1–10.

311. Li, L., Alper, J. & Alexov, E. Cytoplasmic dynein binding, run length, and velocity are guided by long-range electrostatic interactions. *Scientific Reports* **6,** 31523 (2016).

312. Reck-Peterson, S. L., Redwine, W. B., Vale, R. D. & Carter, A. P. The cytoplasmic dynein transport machinery and its many cargoes. *Nature Reviews Molecular Cell Biology* **19,** 382–398 (2018).

313. Hu, D. J.-K. *et al.* Dynein Recruitment to Nuclear Pores Activates Apical Nuclear Migration and Mitotic Entry in Brain Progenitor Cells. *Cell* **154,** 1300–1313 (2013).

314. Bolhy, S. *et al.* A Nup133-dependent NPC-anchored network tethers centrosomes to the nuclear envelope in prophase. *Journal of Cell Biology* **192,** 855–871 (2011).

315. Splinter, D. *et al.* Bicaudal D2, Dynein, and Kinesin-1 Associate with Nuclear Pore Complexes and Regulate Centrosome and Nuclear Positioning during Mitotic Entry. *PLoS Biology* **8** (ed Pellman, D.) e1000350 (2010).

316. Fierling, J. *et al.* Embryo-scale epithelial buckling forms a propagating furrow that initiates gastrulation. *Nature Communications* **13,** 3348 (2022).

317. Pezzulla, M., Stoop, N., Steranka, M. P., Bade, A. J. & Holmes, D. P. Curvature-Induced Instabilities of Shells. *Physical Review Letters* **120,** 048002 (2018).

318. Ideses, Y. *et al.* Spontaneous buckling of contractile poroelastic actomyosin sheets. *Nature Communications* **9,** 2461 (2018).

319. Guha Ray, C. & Haas, P. A. Unbuckling Mechanics of Epithelial Monolayers under Compression. *Physical Review Letters* **134,** 118402 (2025).

320. Banerjee, D. S., Freedman, S. L., Murrell, M. P. & Banerjee, S. Growth-induced collective bending and kinetic trapping of cytoskeletal filaments. *Cytoskeleton* **81,** 409–419 (2024).

321. Shohat, D. & van Hecke, M. Geometric Control and Memory in Networks of Hysteretic Elements. *Physical Review Letters* **134,** 188201 (2025).

322. Wang, H., Liu, Y.-Q., Hang, J.-T., Xu, G.-K. & Feng, X.-Q. Dynamic high-order buckling and spontaneous recovery of active epithelial tissues. *Journal of the Mechanics and Physics of Solids* **183,** 105496 (2024).

323. Yang, D. *et al.* Buckling Pneumatic Linear Actuators Inspired by Muscle. *Advanced Materials Technologies* **1,** 1600055 (2016).

324. Wang, Q. *et al.* Transient Amplification of Broken Symmetry in Elastic Snap-Through. *Physical Review Letters* **132,** 267201 (2024).

325. Banterle, N. *et al.* Kinetic and structural roles for the surface in guiding SAS-6 self-assembly to direct centriole architecture. *Nature Communications* **12,** 6180 (2021).

326. Carter, A. H. *Classical and statistical thermodynamics* (Pearson Education, Philadelphia, 2000).

327. Robert, P. *et al.* Functional Mapping of Adhesiveness on Live Cells Reveals How Guidance Phenotypes Can Emerge From Complex Spatiotemporal Integrin Regulation. *Frontiers in Bioengineering and Biotechnology* **9,** 625366 (2021).

328. Ekerdt, B. L., Segalman, R. A. & Schaffer, D. V. Spatial organization of cell-adhesive ligands for advanced cell culture. *Biotechnology Journal* **8,** 1411–1423 (2013).

329. Truong Quang, B.-A., Mani, M., Markova, O., Lecuit, T. & Lenne, P.-F. Principles of E-Cadherin Supramolecular Organization In Vivo. *Current Biology* **23,** 2197–2207 (2013).

330. Janeway Jr., C. A., Travers, P., Walport, M. & et al. *Antigen recognition by T cells* in *Immunobiology: The Immune System in Health and Disease* 5th (Garland Science, New York, 2001).

331. Jiang, X. *et al.* MRI of tumor T cell infiltration in response to checkpoint inhibitor therapy. *Journal for ImmunoTherapy of Cancer* **8,** e000328 (2020).

332. Swamy, M., Dopfer, E. P., Molnar, E., Alarcón, B. & Schamel, W. W. A. The 450kDa TCR Complex has a Stoichiometry of $\alpha\beta\gamma\varepsilon\delta\varepsilon\zeta\zeta$. *Scandinavian Journal of Immunology* **67,** 418–420 (2008).

333. Ma, V. P.-Y. *et al.* The magnitude of LFA-1/ICAM-1 forces fine-tune TCR-triggered T cell activation. *Science Advances* **8,** eabg4485 (2022).

334. Bui, T. M., Wiesolek, H. L. & Sumagin, R. ICAM-1: A master regulator of cellular responses in inflammation, injury resolution, and tumorigenesis. *Journal of Leukocyte Biology* **108,** 787–799 (2020).

335. Kurz-Isler, G., Voigt, T. & Wolburg, H. Modulation of connexon densities in gap junctions of horizontal cell perikarya and axon terminals in fish retina: effects of light/dark cycles, interruption of the optic nerve and application of dopamine. *Cell & Tissue Research* **268,** 267–275 (1992).

336. Goodenough, D. A. & Revel, J. P. A fine structural analysis of intercellular junctions in the mouse liver. *The Journal of Cell Biology* **45,** 272–290 (1970).

337. Kuntze, A. *et al.* Protonation of Piezo1 Impairs Cell-Matrix Interactions of Pancreatic Stellate Cells. *Frontiers in Physiology* **11,** 89 (2020).

338. Mulhall, E. M. *et al.* Direct observation of the conformational states of PIEZO1. *Nature* **620,** 1117–1125 (2023).

339. Rizzo, V., Morton, C., DePaola, N., Schnitzer, J. E. & Davies, P. F. Recruitment of endothelial caveolae into mechanotransduction pathways by flow conditioning in vitro. *American Journal of Physiology-Heart and Circulatory Physiology* **285,** H1720–H1729 (2003).

340. Thomsen, P., Roepstorff, K., Stahlhut, M. & van Deurs, B. Caveolae Are Highly Immobile Plasma Membrane Microdomains, Which Are not Involved in Constitutive Endocytic Trafficking. *Molecular Biology of the Cell* **13** (ed Riezman, H.) 238–250 (2002).

341. Mannella, C. A. VDAC—A Primal Perspective. *International Journal of Molecular Sciences* **22,** 1685 (2021).

342. Hiller, S., Abramson, J., Mannella, C., Wagner, G. & Zeth, K. The 3D structures of VDAC represent a native conformation. *Trends in Biochemical Sciences* **35,** 514–521 (2010).

343. Lyu, Y. *et al.* Effect of GM1 concentration change on plasma membrane: molecular dynamics simulation and analysis. *Physical Chemistry Chemical Physics* **26,** 12552–12563 (2024).

344. Mojumdar, E. H., Grey, C. & Sparr, E. Self-Assembly in Ganglioside–Phospholipid Systems: The Co-Existence of Vesicles, Micelles, and Discs. *International Journal of Molecular Sciences* **21,** 56 (2020).

345. Evans, S., Goldman, D., Heinemann, S. & Patrick, J. Muscle acetylcholine receptor biosynthesis. Regulation by transcript availability. *Journal of Biological Chemistry* **262,** 4911–4916 (1987).

346. McMahon, D. K. *et al.* C2C12 cells: biophysical, biochemical, and immunocytochemical properties. *American Journal of Physiology-Cell Physiology* **266,** C1795–C1802 (1994).

347. Lo, M. M. S., Barnard, E. A. & Dolly, J. O. Size of acetylcholine receptors in the membrane. An improved version of the radiation inactivation method. *Biochemistry* **21,** 2210–2217 (1982).

348. Geng, L., Zhang, H. L. & Peng, H. B. The formation of acetylcholine receptor clusters visualized with quantum dots. *BMC Neuroscience* **10,** 80 (2009).

349. Kerntke, C., Nimmerjahn, F. & Biburger, M. There Is (Scientific) Strength in Numbers: A Comprehensive Quantitation of Fc Gamma Receptor Numbers on Human and Murine Peripheral Blood Leukocytes. *Frontiers in Immunology* **11,** 118 (2020).

350. Wacleche, V., Tremblay, C., Routy, J.-P. & Ancuta, P. The Biology of Monocytes and Dendritic Cells: Contribution to HIV Pathogenesis. *Viruses* **10,** 65 (2018).

351. Saji, F. Dynamics of immunoglobulins at the feto-maternal interface. *Reviews of Reproduction* **4.** BNID 117058, 81–89 (1999).

352. Zhang, F. *et al.* Quantification of Epidermal Growth Factor Receptor Expression Level and Binding Kinetics on Cell Surfaces by Surface Plasmon Resonance Imaging. *Analytical Chemistry* **87,** 9960–9965 (2015).

353. Abulrob, A. *et al.* Nanoscale Imaging of Epidermal Growth Factor Receptor Clustering. *Journal of Biological Chemistry* **285,** 3145–3156 (2010).

354. Horzum, U., Ozdil, B. & Pesen-Okvur, D. Step-by-step quantitative analysis of focal adhesions. *MethodsX* **1,** 56–59 (2014).

355. TruongVo, T. N. *et al.* Microfluidic channel for characterizing normal and breast cancer cells. *Journal of Micromechanics and Microengineering* **27,** 035017 (2017).

356. Erickson, H. P. Size and Shape of Protein Molecules at the Nanometer Level Determined by Sedimentation, Gel Filtration, and Electron Microscopy. *Biological Procedures Online* **11,** 32–51 (2009).

Acknowledgements

I want to express my deepest gratitude to my supervisor and mentor, Dr. Anna Erzberger, whose support and guidance have been unwavering. Your scientific curiosity and the exciting research environment that you have created have been inspirational throughout my doctoral journey, and your patience and kindness have been invaluable throughout this great adventure.

A heartfelt thank you to Prof. Ulrich Schwarz for his support both in my joining the University and during my PhD, and also to Prof. Frauke Gräter. I am grateful to you both for the advice and time you have given as members of my thesis advisory committee. I would also like to express my sincere thanks to Prof. Peer Fischer for his time and effort as a member of my thesis examination committee. Your perspective and evaluation of my work is greatly appreciated.

This research benefited enormously from my collaboration with Roman Belousov. Thank you for your dedication, clarity, and enthusiasm. It has been a pleasure to learn from you and I truly value our in-depth discussions.

I am also extremely grateful to my amazing collaborators, Hiral Shah and Gautam Dey, without whom this endeavour would not have been possible. Working with you has been amazing, and I really appreciate all the effort you have put in to providing such beautiful data, and your invaluable expertise and guidance.

A special thanks to Johannes Jung. Supervising you during your master's thesis was a pleasure and I have thoroughly enjoyed our stimulating chalk-board discussions. I wish you all the best in your continued scientific journey.

Many thanks to my colleagues and friends, including Ian Estabrook, Patrick Jentsch, Pamela Guruciaga, Laeschkir Würthner, Tim Dullweber, and all other members of the Erzberger lab both past and present for the stimulating discussions and thoughtful feedback. I have really valued working together and our shared meetings, retreats, and dinners have made this PhD journey an extremely enjoyable experience.

Lastly, I would like to thank my family for their support throughout my studies and for instilling in me an enthusiasm for science. To my fiancé, Max, thank you for being there for me and for providing constant support and love (...and tea!). Of everything that has happened over these years, sharing them with you is what I will cherish most.