Aggregation-Fragmentation Models for Transport in a Biological System

A Thesis

Submitted to the
Tata Institute of Fundamental Research, Mumbai
for the degree of Doctor of Philosophy
in Physics

by

Himani Sachdeva

School of Natural Sciences
Tata Institute of Fundamental Research
Mumbai

May, 2014
Final Version Submitted in August, 2014

DECLARATION

This thesis is a presentation of my original research work. Wherever contributions of others are involved, every effort is made to indicate this clearly, with due reference to the literature, and acknowledgement of collaborative research and discussions.

The work was done under the guidance of Professor Mustansir Barma, at the Tata Institute of Fundamental Research, Mumbai, and in collaboration with Professor Madan Rao from the National Centre for Biological Sciences (T.I.F.R.), Bangalore.

(Himani Sachdeva)

In my capacity as supervisor of the candidate's thesis, I certify that the above statements are true to the best of my knowledge.

(Mustansir Barma)

Date:

Acknowledgements

The six years I spent in TIFR were touched profoundly by many people and experiences. My most heartfelt gratitude is towards Mustansir whose gentle encouragement, unceasing care and support made it possible for me to get through difficult times and complete my thesis. His way of engaging with research, which was positive yet critical, persevering and relaxed at the same time, was very enabling for me, making it possible for me to find and sustain my own rhythms of working. I have been deeply touched by his sensitivity and commitment towards my overall welfare at each step, and his readiness to connect with and nurture the best in people and situations, and I carry a very strong sense of this as I leave. It was a privilege to learn from and work with him, and a great pleasure to converse with him on a range of things. I would like to thank him for making my time in TIFR so rewarding and meaningful. I also want to thank his wife Rashida for her warmth and hospitality on a number of occasions.

Much of this work was carried out in collaboration with Madan. I would like to thank him for introducing me to biology-oriented problems, and for his guidance and time. I would also like to thank Deepak for many helpful discussions and insights related to my work, and for the courses he taught, which were very useful. I had many interesting and enjoyable interactions with my officemates, especially Sambuddha, Sasi, Umesh, Rahul and Nikhil. I thank Tarun and Geet for helping me out with computer-related issues. People in the theory office and the technical support staff, especially Kapil were of great help, whenever any practical need arose.

My time in TIFR was enriched by many close bonds. In particular, I thank my friends Sangeetha, Vandna, Aditya, Naren and Rakesh for their companionship, warmth and care during the first half of my stay while they were at TIFR, and also for maintaining a sense of connectedness after leaving. I especially want to thank Aditya for his affection and unfailing support over the last six years. In the last couple of years, I formed strong bonds with Garima and Shalini, both of whom were a source of comfort during challenging times.

There were many people who continued to be significantly involved in my life during this period, in spite of not being in Bombay. I thank Minki, Aloke and Sujoy for their friendship and presence and the time they spent with me during their visits to Bombay. I would also like to thank Sanghamitra Ghosh who was available whenever I needed her advice. Much of my engagement with the challenges during this time would not have been possible without Honey Vahali. Her compassion, counsel and care have been deeply enabling for me, and inspire my utmost gratitude.

Finally, I would like to thank my family- my parents, Bua, brother and grandmother, for their unconditional love and for continuing to support my need to find my own paths, even if these take me a little away from them. I would especially like to acknowledge my Dada, my Nani and Nana, whose love, care and belief live on as quiet influences in my life.

This list would be incomplete without a thank you to the beautiful city of Bombay and its seafaces, rains and cafes, and the sense of space it provides in spite of its bustling crowds. I am grateful that I had the opportunity to live here and find myself in some of its rhythms.

Contents

Synopsis

1 Overview

Biological systems often consist of a large number of interacting components which self-organise over many spatio-temporal scales to form complex structures. A key feature is that these are highly organised, low-entropy structures which are maintained far from equilibrium by a constant flow of energy, mass etc. Non-equilibrium statistical physics thus emerges as a natural framework for addressing questions related to self-organisation and transport in biological systems. The interaction between biology and statistical physics can be thought of as a dual one– first, biological questions inspire new models which can give insight into non-equilibrium phenomena in general. Second, statistical physics approaches to modeling biological phenomena sometimes shed light on important aspects such as the role of fluctuations in biology and also highlight 'universal' behaviour across diverse length scales.

In recent years, there has been a lot of interest in understanding mechanisms of transport and self-assembly at the sub-cellular level. Cells possess multiple transport pathways, which are crucial for exchange of molecules between a cell and its environment, as well as between different organelles (compartments) within the cell. For instance, the secretory pathway involves movement of molecules from the endoplasmic reticulum (ER) to the cell surface via an organelle system called the Golgi apparatus. The Golgi apparatus itself consists of spatially and chemically distinct sub-compartments known as *cisternae*. Cargo molecules such as proteins and lipids arrive from the ER, enter through one face of the Golgi, undergo several chemical reactions (processing) as they move through the Golgi; the modified products then leave through the other face for target locations, mainly on the cell membrane [1].

How are cargo molecules transported *within* the Golgi? This question has attracted a great deal of interest as well as controversy. Two alternative hypotheses have been suggested [2]:
Vesicular Transport: In this view, the cisternae are stable structures with a fixed spatial location within the Golgi and a fixed set of resident enzymes. Cargo molecules shuttle from one stationary cisterna to the next in small sacs called vesicles, and get chemically modified by the enzymes present at each cisterna.
Cisternal Maturation: This theory envisages cisternae as highly dynamic structures that form by fusion of incoming vesicles and move as a whole, carrying cargo molecules with them. Specific enzymes get attached to a cisterna in different stages of its progression, and modify the molecules in it. The final cisterna eventually breaks up, releasing processed end products.

These two mechanisms differ essentially in whether cargo molecules are transported by the movement of large aggregates (cisternae) or small fragments (vesicles) that break off from these large aggregates. Independent studies on a variety of cells provide mixed evidence for various possibilities [3, 4]. It has also been suggested that the two mechanisms are not mutually exclusive, and transport may be occurring via a combination of both [5].

Not only the cargo molecules, but also the molecules that constitute the Golgi organelle, must be trafficked by the processes described above. This invites the following question: how are spatially and chemically distinct structures (cisternae) of the Golgi formed and dynamically maintained, given the rules of molecular trafficking, which broadly, may be described as: (i) localised influx and outflux of 'particles' (vesicles) (ii) transformation of particles from one species to the other, i.e. chemical processing of cargo molecules by enzymes (iii) transport of particles either by chipping, i.e., fragmentation and movement of single particles (corresponding to vesicle movement) or through movement of whole aggregates (corresponding to cisternal maturation).

In keeping with the reciprocal relationship between biological questions and statistical physics models, the aim of this thesis is also twofold. The first is to construct the simplest version of a model incorporating the processes involved in Golgi dynamics and analyse it in detail to understand the broad effects that emerge from their interplay. The second aim is to identify those regions in the parameter space of the model that yield Golgi-like structures and examine whether these structures bear distinctive signatures of the transport mechanism, which would be relevant for experiments on the Golgi.

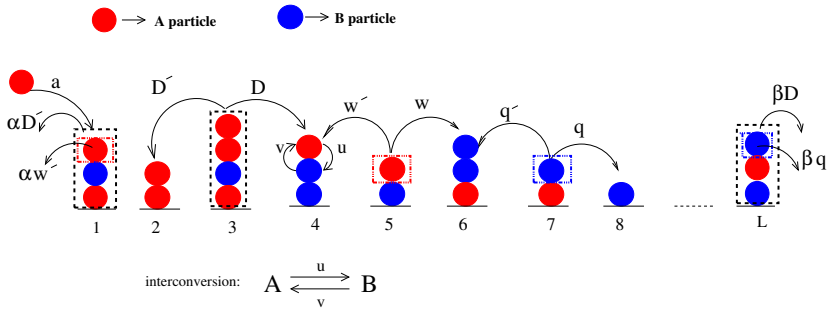


Figure 1: Illustration of allowed moves in the model: injection of single particles at site 1 at rate a; interconversion of single particles $A \to B$ at rate u and $B \to A$ at rate v; chipping of an A (B) particle and rightward hopping with rate w (q) or leftward hopping with rate w' (q'); movement of full stack to the right (left) with rate D (D'); outflux of full stacks or single particles from left and/or right boundary.

The model consists of multiple species (denoted by 'A', 'B', 'C', ...) of particles on a 1D lattice with L sites [1]. For simplicity, we focus on the two-species version of this model; the extension to more number of species is straightforward. The system evolves as follows– at each instant, a site i is chosen at random and one of the following moves (illustrated in fig. 1) occurs:

[1]Studying a 1D model can be justified on the grounds that there is a net direction of transport of molecules through the Golgi; the spatial variation of enzymes across the Golgi also occurs along this direction.

i. *Influx*: A single 'A' particle of unit mass is injected at the first site ($i = 1$) at rate a.

ii. *Interconversion*: With rate u, one of the A particles residing on site i converts to type B. Conversely, one of the B particles can also convert to type A with rate v.

iii. *Chipping of single particles*: With rate w (or w'), one of the A particles breaks off from the mass present at site i the site and hops to the site $i + 1$ (or $i - 1$). Similarly, with rate q (or q'), one of the B particles breaks off from site i and hops to $i + 1$ (or $i - 1$).

iv. *Movement and aggregation of full stacks*: With rate D (or D'), the full stack at site i, i.e. all A and B particles together, hops to site $i + 1$ (or $i - 1$) and adds to the mass there.

v. *Outflux from boundaries*: With rate βD (or $\alpha D'$), the full stack at site L (or site 1) exits the system; with rate βw (or $\alpha w'$), a single A particle breaks off and exits from site L (or site 1); with rate βq (or $\alpha q'$), a single B particle exits from site L (or site 1).

Since the parameter space is large, it is useful to first consider various reduced versions of the model which incorporate only a subset of all the allowed processes. This approach allows us to isolate the distinct effects that arise from the interplay of different processes. (i) A natural starting point is the single-species version of the model, obtained by setting all the interconversion rates to zero. This simply gives the 1D aggregation-fragmentation model with influx of single particles and outflux of aggregates at the boundaries. The key question that we address here is– does this system undergo a phase transition, as in the case of the closed periodic model [6], and what is the nature of the phases? (ii) Next, we consider the multi-species version of this model; the interesting limit here turns out to be the one with *no* movement of full stacks. The interplay of injection, chipping and interconversion in this system gives rise to a different sort of phase, which we analyse in detail. (iii) Finally we identify two limits of the full multi-species model, one which is aggregation-dominated and the other chipping-dominated, which exhibit steady state structures which are qualitatively similar to Golgi structures. Further, we bring out the differences between the structures generated in these two limits by analysing both static and dynamical properties.

The main results of the thesis can be summarised as follows:

i. *Aggregation-fragmentation model with open boundaries*: On increasing influx or decreasing the fragmentation rate, the system with unbiased motion of particles undergoes a phase transition from a normal (fragmentation-dominated) phase to a 'condensate' (aggregation-dominated) phase in which a finite fraction of the total mass accumulates in a macroscopic aggregate or condensate. The mass in the condensate has a broad distribution and shows giant number fluctuations, unlike the condensate in the closed periodic version of this model [6]. The presence of such a fluctuating condensate leads to several unusual properties: the total number of particles M in the system shows giant fluctuations and has a strongly non-Gaussian probability distribution. Further, the time-series $M(t)$ is not self-similar and shows intermittency, which we characterize in terms of appropriately defined structure functions and their ratios, as in turbulence phenomena [7]. The model with biased motion of aggregates also undergoes a phase transition to an aggregation-dominated phase with temporal intermittency of M. These

results are established using numerical simulations for non-zero fragmentation rates and analytical calculations in the case of zero fragmentation. In particular, for the zero fragmentation case, we explicitly calculate dynamical structure functions and demonstrate their anomalous scaling.

ii. *Multi-species model with interconversion between species*: A different sort of phase transition arises in the pure chipping limit ($D = D' = 0$) of the multi-species model. Depending on the parameters, this system may exist either in a steady phase, in which the average mass at each site attains a time-independent value, or in a 'growing' phase, in which the total mass grows indefinitely in time. The growing phases have interesting spatial structure, with unbounded growth of one or more species of particles occurring in spatial regions of tunable width and location, while steady state is attained in the rest of the system. We have analysed this model by solving equations for particle currents at each site.

iii. *Relevance to Golgi structure and transport*: We focus on two limits, namely, the only chipping limit and the only stack hopping limit of the full multi-species model and identify the conditions under which the model generates multiple, spatially and chemically distinct aggregates (which correspond roughly to cisternae in the Golgi). The only chipping model yields static aggregates whose position, size and composition remains stable in time. In the limit with biased stack movement as the predominant mode of transport, we obtain several moving aggregates whose size, composition and even number can vary significantly over time. The structures in these two limits have distinctive static and dynamical properties; our work suggests some possible probes for these.

2 Aggregation-Fragmentation Model with Open Boundaries

The single species version of the general model can be obtained by setting $u = 0$. This model allows for [see fig. 1] (i) influx of single particles at site 1 at rate a (ii) movement of whole stacks with rates D (or D') to the right (or left) and aggregation with the neighbouring mass (iii) chipping (fragmentation) of unit mass from a site and movement to the right or left with equal rates w (iv) outflux of mass from site L (or site 1) either via exit of whole stacks with rate βD (or $\alpha D'$) or exit of single particles after chipping with rate βw (or αw).

This model is the open-boundary version of a well-studied aggregation-chipping model [6] defined on a closed periodic lattice with fixed total mass. The closed model is known to undergo a phase transition as mass density is increased: when the total mass in the system exceeds a critical value, 'condensation' sets in, with the excess mass forming a single macroscopic condensate. The phenomenon is akin to Bose-Einstein condensation, but in real space. Condensation transitions also occur in other mass transport models [8] such as the zero range process and its variants [9]. All the above systems are, however, characterised by a *fixed* total mass. Does the condensation transition survive in an open system where the total mass is not conserved, but can, in principle, undergo large fluctuations due to the exit of aggregates of any size? This is the main question that we address in this part of the thesis.

We study this question for both unbiased ($D = D'$) and fully biased ($D' = 0$) movement of aggregates. Further, for the unbiased case, we consider two special but representative

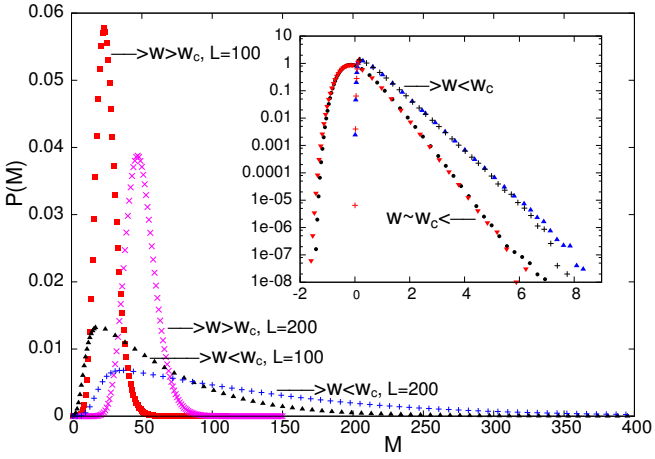


Figure 2: $P(M)$ vs. M for two different L in the normal phase (Gaussian tails) and condensate phase (long tails). Inset: Scaling collapse of tails on plotting $LP(M)$ vs. M/L in the condensate phase and $L^{2/3}P(M)$ vs. $(M - \langle M \rangle)/L^{2/3}$ near the critical point.

boundary conditions: (i) with outflux only from site L (corresponding to $\alpha = 0, \beta = 1$) (ii) with outflux from both site 1 and L ($\alpha = 1, \beta = 1$). Thus, we study the following three cases in all:
(A) *Unbiased stack hopping; outflux from both boundaries*: $D = D'$ and $\alpha = \beta = 1$.
(B) *Unbiased stack hopping; outflux only from right boundary*: $D = D'$ and $\alpha = 0, \beta = 1$.
(C) *Biased stack hopping; outflux from right boundary*: $D' = 0$ and $\alpha = \beta = 1$
In each of these three cases, injection occurs at rate a and chipping with symmetric rates w to the left and right. We study various properties of the models (A)-(C) using numerical simulations for non-zero chipping rates w and analytically in the $w = 0$ case. These properties are discussed in more detail below.

A. Unbiased stack hopping; outflux from both boundaries: As chipping rate w is decreased, the system undergoes a phase transition from a normal to a condensate phase. The two phases have very different static and dynamical properties which are discussed below.

Statics: In the *condensate (small w) phase*, the largest mass m_* in the system has an average value proportional to L, implying that the system contains a macroscopic condensate. This condensate is quite different from the one found in the closed periodic version of this model [6], in that it shows giant fluctuations, with $\Delta m_* \propto \langle m_* \rangle \propto L$. Further, the distribution of m_* is broad-tailed, and has the non-Gaussian form $P(m_*) \sim (1/M_0) \exp(-m_*/M_0)$ for large m_*, where M_0 is proportional to L. We refer to this exponential tail as the 'condensate tail' and find that it appears in the distribution of the mass contained in any region of the system, from the full system down to a single site. Specifically, the steady state distribution $P(M)$ of total mass M in the system behaves as $P(M) \propto 1/M_0 \exp(-M/M_0)$ at large M [fig. 2].

A natural consequence of the non-Gaussian mass distributions is that the root mean square (rms) fluctuations of various quantities show anomalous behaviour. The total mass M in the system shows giant rms fluctuations ΔM that scale as L rather than $\sqrt{L}$, such that $\Delta M/\langle M \rangle$ remains finite even as $L \to \infty$. Further, even for the mass m_j at a single site j, we find $\Delta m_j \propto \sqrt{L} g(j/L)$, implying that $\Delta m_j/\langle m_j \rangle$ actually increases as $L \to \infty$. This sort of strong dependence on L is very different from usual finite size effects which become increasingly irrelevant at large system sizes.

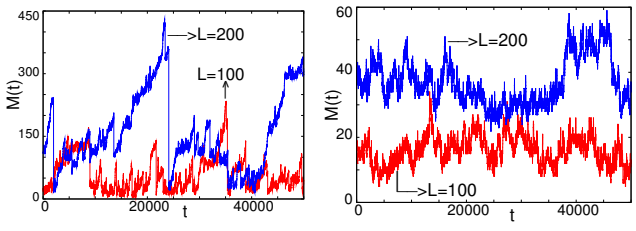
Several other systems such as vibrated needles [10] and passive particles on fluctuating interfaces [11] also show giant number fluctuations, but in these systems, the fluctuations in a region of linear size Δl are proportional to Δl [12], rather than the system size L, as found in our model. The emergence of L-dependent fluctuations in our model, or in condensation models in general, is due to the fact that there is no constraint on the number of particles per site, in contrast to the hard core constraints in systems studied in [10, 11].

In the *normal (large w) phase*, there is no macroscopic condensate, i.e., $\langle m_* \rangle / L \to 0$ as $L \to \infty$. The total mass has normal rms fluctuations $\Delta M \propto \sqrt{L}$, with the distribution for the rescaled mass variable $(M - \langle M \rangle)/\Delta M$, approaching a Gaussian form at large L. Further, unlike in the condensate phase, the moments of the mass at any site j depend only on the relative position j/L and not directly on system size L.

The transition from the normal to the condensate phase takes place at a *critical chipping rate* w_c, which increases with injection rate a (with D held constant). The quantity $\langle m_* \rangle / L$ can be thought of as an order parameter associated with this transition; w_c is thus the fragmentation rate at which this quantity goes to zero in the limit $L \to \infty$. At the critical point also, large fluctuations of the total mass are found: $\Delta M \propto L^{\theta_c}$ where $\theta_c \simeq 2/3$. The mass distribution has a tail of the form: $P(M) \sim \frac{1}{M_2} \exp\left(\frac{-(M - \langle M \rangle)}{M_2}\right)$ where $\langle M \rangle \sim L$ and $M_2 \sim L^{\theta_c}$ [inset, fig. 2].

Since this system is spatially inhomogeneous, the distribution of mass at a site or in a region depends on the location of the site/region. In the condensate phase, in particular, for intermediate values of the chipping rate w, we find evidence for qualitatively different behaviour at different spatial positions in the system. Preliminary numerics suggest that at intermediate w, the system contains a macroscopic condensate, which is restricted to *left lying* regions of the system, and cannot be found towards the right. This is reflected in the behaviour of $m_*^{(L)}[i, j]$, defined as the largest mass in the region $[i + 1, j]$ of a system of size L. For left-lying regions, $\langle m_*^{(L)}[i, j] \rangle \propto L$ for fixed $i/L, j/L$, whereas for regions near the right boundary, $\langle m_*^{(L)}[i, j] \rangle$ is sub-linear in L. Further, numerics suggest that this co-existence of condensate-like and normal behaviours in different regions of the system does not persist down to very small w. However, for conclusive evidence, it may be necessary to simulate very large systems, which we have not been able to do.

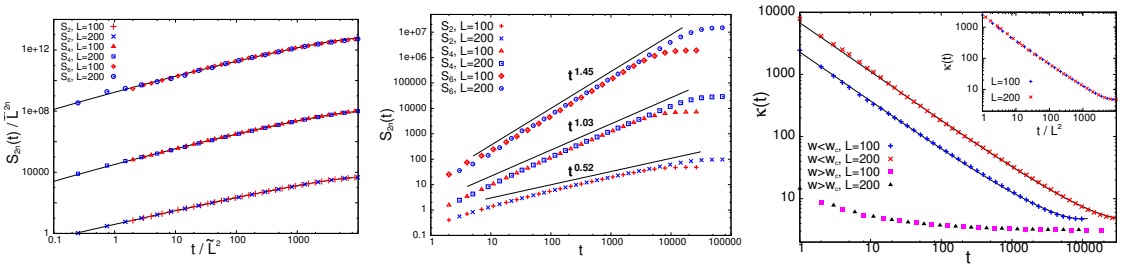
Dynamics: Contrasting signatures of the phases also appear in dynamical properties: $M(t)$ is self-similar in time in the normal phase but exhibits breakdown of self-similarity in the condensate phase. This difference is apparent in typical time series $M(t)$ of the total mass in the two phases [see fig. 3]. In the condensate phase, $M(t)$ shows large crashes over



(a) (b)

Figure 3: A typical realisation of $M(t)$ vs. t for two different system sizes L in the (a) condensate phase (b) normal phase

very short time scales, due to the exit of a condensate of $\mathcal{O}(L)$ mass; no such large crashes occur in the time series in the normal phase. The breakdown of self-similarity can be quantified by defining dynamical structure functions $S_n(t) = \langle [M(t) - M(0)]^n \rangle$, where $\langle ... \rangle$ denotes average over histories and $t = 0$ is an arbitrary time instant after the system attains stationarity. Self-similar signals typically show $S_n(t) \propto t^{u(n)}$ as $t/\tau \to 0$, where $u(n) \propto n$, and τ is a time scale having to do with the largest structures in the system. A deviation from the linear dependence of the exponents $u(n)$ on n reflects the breakdown of self-similarity and may occur, for example, if the signal $M(t)$ alternates between periods of quiescence (small or no activity) and bursts (sudden large changes) [7]. Such an alternation is characteristic of intermittency. A well-studied measure of intermittency is the flatness, defined as $\kappa(t) = S_4(t)/S_2^2(t)$, which is just the kurtosis (upto an additive constant) of the time-dependent probability distribution of $M(t) - M(0)$. For intermittent signals, $\kappa(t)$ diverges as $t/\tau \to 0$ [7].



(a) (b) (c)

Figure 4: (a) Structure functions $S_{2n}(t)$ in the condensate phase for different L: Good data collapse on plotting $S_{2n}(t)/\tilde{L}^{2n}$ vs $t/\tilde{L}^2$ (where the rescaling $\tilde{L} = L/100$ is just for displaying all three structure functions in the same plot clearly). Solid lines indicate fits to $yg_{2n}[log(y)]$ as described in text.
(b) In the normal phase, $S_{2n}(t)$ show no dependence on L and behave as $S_{2n}(t) \sim t^{u(n)}$ for small t, where $u(n) \sim n/2$.
(c) $\kappa(t)$ vs. t for two different L in the two phases. Solid lines for $w < w_c$ are fits to the form described in the text. Inset: Data collapse of $\kappa(t)$ vs. t in the condensate phase on rescaling time as t/L^2.

We measure the dynamical structure functions $S_n(t)$ using numerical simulations and find that they behave very differently in the two phases. In the *condensate phase*, the structure functions scale as: $S_n(t) \sim L^n f_n(t/L^2)$, where $f_n(y) \sim (-1)^n y g_n[\log(y)]$ for small y (with g_n chosen to be a polynomial) [2], and approaches an n-dependent constant value at large y [see fig. 4a]. Thus, the structure functions show an extreme form of anomalous scaling where the exponent $u(n)$ is actually independent of n, with a weak n-dependence of the scaling form f_n entering only through the multiplicative $\log t$ terms. This extreme anomalous scaling is referred to as *strong* intermittency and also occurs in Burgers turbulence [13], where spatial structure functions of the velocity field exhibit similar behaviour. From the small t form of $S_n(t)$, it follows that the flatness $\kappa(t)$ diverges at small times in a strongly L dependent way [fig. 4c]. In fact, it is a function of t/L^2 and shows a power law divergence as $t/L^2 \to 0$ [inset, fig. 4c].

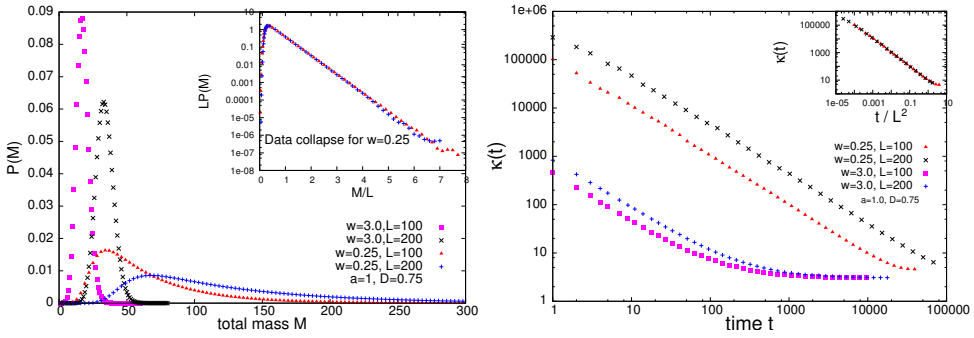
In the *normal phase*, by contrast, the structure functions are independent of L at small t and scale as $S_{2n}(t) \sim t^{u(n)}$, where the dependence of $u(n)$ on n is close to linear [see fig. 4b],

[2] The log terms are suggested by the analytical calculation for the $w = 0$ limit discussed later, where they can be obtained explicitly.

indicating that $M(t)$ is self-similar. The flatness $\kappa(t)$ approaches a finite, L independent value as $t \to 0$ [fig. 4c]. At the *critical point*, $M(t)$ again shows intermittency, with the flatness diverging as $t \to 0$ in an L dependent manner. However, there seems to be no simple scaling which collapses the curves for different L.

B. Unbiased stack hopping; outflux only from right boundary: In order to investigate the role of boundary conditions, we also study the case where influx and outflux occur at opposite ends, unlike in model A, where both were allowed at the same site. For the total mass M in this system to be an extensive quantity, injection rates of the sort $a = \tilde{a}/L$ [where $\tilde{a}$ is $\mathcal{O}(1)$] are required. Such $1/L$ rates also ensure that the *average* value of the particle current through the system is the same (and proportional to $1/L$) in both models A and B.

We study this system using numerical simulations and find that it also shows a condensation transition similar to that found for model A. At small chipping rates w, the system contains a macroscopic condensate, as evident from the behaviour of the largest mass: $\langle m_* \rangle \propto L$, while for large w, there is no condensate formation. As in the case of model A, at small w, the total mass M is characterised by a non-Gaussian distribution [fig. 5a] and shows giant $\mathcal{O}(L)$ rms fluctuations, while for large w, the distribution of M has a Gaussian form with normal $\mathcal{O}(\sqrt{L})$ fluctuations. Numerical measurements of the region-wise largest mass $m_*^{(L)}[i,j]$ suggest that model B also exhibits condensate-like and normal behaviours in different spatial regions for intermediate values of chipping rate w.



(a) (b)

Figure 5: Different static and dynamical properties of M in the two phases for model B.
(a) $P(M)$ has a long 'condensate' tail for small w and a Gaussian tail for large w. Inset: Data collapse of the condensate tails for small w on plotting $LP(M)$ vs. M/L.
(b) Flatness $\kappa(t)$ as a function of t: Divergence of $\kappa(t)$ on $\mathcal{O}(L^2)$ time scales for small w and on $\mathcal{O}(L)$ time scales for large w. Inset: Data collapse for small w on plotting $\kappa(t)$ vs. t/L^2.

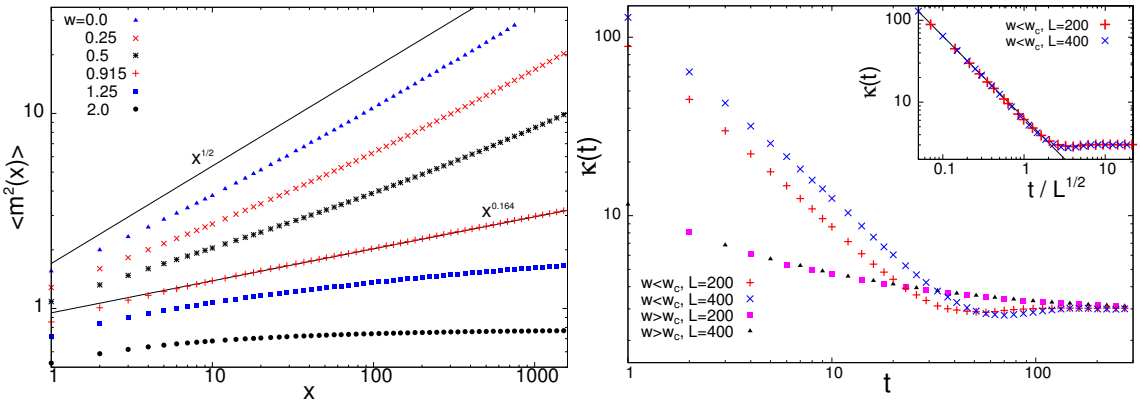
Distinctive signatures of the two phases also appear in the dynamical properties of M. At small w, the time series $M(t)$ shows strong intermittency on $\mathcal{O}(L^2)$ time scales: the dynamical structure functions behave as $S_n(t) \propto L^{n-2}t$, without the multiplicative $log(t)$ terms found in $S_n(t)$ for model A [3]. For large w, the structure functions show self-similar scaling $S_n(t) \propto t^{n/2}$ for $\tau_1 \ll t \ll \tau_2$ where $\tau_1 \propto L$ and $\tau_2 \propto L^2$. For $t \ll \tau_1$, the structure functions exhibit

[3]The different functional forms of structure functions in the condensate phase of model A and model B can be calculated analytically in the $w = 0$ limit

intermittency-associated anomalous scaling; this, however, is simply a consequence of $\mathcal{O}(1/L)$ injection rates, and does not signify condensate formation. As before, these differences are also reflected in the behaviour of the flatness $\kappa(t)$: for small w, the flatness shows a power law divergence on $\mathcal{O}(L^2)$ time scales, while for large w, the divergence occurs on much shorter $\mathcal{O}(L)$ time scales [fig. 5b].

In conclusion, the basic features of models A and B appear to be quite similar, indicating that the fact of a phase transition in this model and the nature of the condensate phase are robust to variations in boundary conditions.

C. Biased stack hopping; outflux from right boundary: To investigate the role of bias in this model, we study the case with fully asymmetric stack movement ($D' = 0$), but with symmetric chipping (and movement) of single particles [4]. This model also shows a phase transition from a fragmentation-dominated phase to an aggregation-dominated phase on decreasing w or increasing a. The aggregation-dominated phase in this case is characterised by formation and exit of aggregates with mass that is $\mathcal{O}(\sqrt{L})$, rather than $\mathcal{O}(L)$, as was found in the unbiased case. In the fragmentation-dominated phase, on the other hand, there are no L^α-sized aggregates (with $\alpha > 0$). This difference is reflected in the plots of $\langle m^2(x)\rangle$ vs. x in the two phases [fig. 6a]. For large w, the moment $\langle m^2(x)\rangle$ approaches a constant value at large x, indicating that there are no L-dependent aggregates in the system. For small w, the plots of $\langle m^2(x)\rangle$ vs. x bend upwards, consistent with an approach to $\sqrt{x}$ at large x (as in the limiting case $w = 0$, where this can be established analytically). The transition takes place at w_c (corresponding to the curve with no bending on the log-log plot), at which $\langle m^2(x)\rangle \sim x^{\alpha_c}$ with $\alpha_c \simeq 0.16$.



(a) (b)

Figure 6: (a) $\langle m^2(x)\rangle$ vs. x for $a = 1$, $D = 1.5$ and different w. Note the upward (downward) bending of curves on log-log plot in aggregation (chipping) dominated phase. There is no bending at w_c.
(b) $\kappa(t)$ vs. t for two different L in the aggregation-dominated and normal phases. Inset: Data collapse of $\kappa(t)$ vs. t for different L on scaling time as $t/\sqrt{L}$ in the aggregation-dominated phase. The divergence at small t roughly follows $\sim (t/\sqrt{L})^{-1}$ (solid line).

[4] There is no phase transition if movement of single particles has a bias. This was established exactly for the closed periodic model in [14], but the basic argument holds for the open case as well.

We also identify signatures of the two phases in the properties of the total mass M. Numerical simulations show that in both the phases, the total mass shows normal rms fluctuations $\Delta M \propto \sqrt{L}$ with the probability distribution of the rescaled mass, $(M - \langle M \rangle)/\Delta M$, approaching a Gaussian form at large L. However, in the aggregation-dominated phase, this is a consequence of sudden $\mathcal{O}(\sqrt{L})$ changes occurring sporadically at $\mathcal{O}(\sqrt{L})$ time separations, whereas in the fragmentation-dominated phase, it emerges from the addition of uncorrelated or weakly correlated $\mathcal{O}(1)$ variations in M at every time step. This distinction is captured in the dynamical properties of $M(t)$: in the aggregate phase, the total mass M shows intermittency on $\mathcal{O}(\sqrt{L})$ time scales, corresponding to a typical time interval of $\mathcal{O}(\sqrt{L})$ duration between large exit events. Flatness and hyperflatness are functions of $t/\sqrt{L}$ and diverge as power laws as $t/\sqrt{L} \to 0$ [see fig. 6b]. By contrast, for large w, the time series $M(t)$ is not intermittent, and there is no L-dependent increase in flatness as $t \to 0$ [fig. 6b]. Thus, in this model, it is the dynamical rather than static properties of M that show imprints of aggregate formation.

Analytic results in the $w = 0$ limit: For the pure aggregation limit $w = 0$, we are able to calculate various static and dynamical properties for the three cases (A)-(C), and demonstrate analytically the distinctive features that characterise the condensate phase, namely giant fluctuations and temporal intermittency. The results obtained from these calculations are in excellent agreement with numerics.

The analytical approach used for these calculations is based on a 'closure' property: the probability distributions of mass $M_{i,j}$ in any single continuous region $[i+1, j]$ of the system satisfy recursion relations that involve other *single-region* mass distributions only. This closed set of equations can be solved for the distribution or equivalently, various moments of $M_{i,j}$. The distribution of the total mass M, and of mass m_k at any site k can then be recovered as special cases of this. The main approximation involved in solving these equations in the unbiased cases (A) and (B) is the use of a spatial continuum limit which transforms recursion relations involving discrete lattice sites into partial differential equations in a continuous spatial variable. The PDE, thus obtained, is a Laplace equation on a right isosceles triangle, and can be solved by mapping onto the corresponding equation on a square. For the biased case (C), the continuum approximation is not self-consistent and we work directly with recursion relations.

The closure property also forms the basis of the calculation of dynamical quantities such as structure functions. These are related to various autocorrelation function of $M(t)$, e.g., $S_2(t) = 2(\Delta M)^2 - 2[\langle M(t)M(0)\rangle - \langle M(t)\rangle\langle M(0)\rangle]$. To calculate this auto-correlation function, we define the generalised spatio-temporal correlation function $\langle M_{i,j}(t)M(0)\rangle - \langle M_{i,j}(t)\rangle\langle M(0)\rangle$ and find that this satisfies a set of time-dependent equations which are also closed in i and j. By solving these, it is possible to obtain the auto-correlation function and hence $S_2(t)$. Higher order structure functions $S_n(t)$ for $n > 2$ can also be calculated similarly. The results explicitly demonstrate the anomalous, intermittency-associated scaling of structure functions with time, as well as the divergence of flatness in the small t limit.

This model provides a simple case where anomalous scaling exponents of structure functions can be calculated analytically. Moreover, unlike other turbulent systems such as the Burgers

fluid, the structure functions considered here are *temporal* structure functions and are related to dynamical correlation functions, which are usually harder to calculate. An important limitation of this approach is that it cannot be used for $w > 0$, as the closure property of the equations for $M_{i,j}$ is lost for non-zero fragmentation rates. Also, this method works only for one-dimensional systems where lattice sites have a natural ordering.

In conclusion, the principal result of this part of the thesis is to establish the existence of a condensate phase in aggregation-chipping models where the total mass is not conserved due to boundary fluxes. This phase is characterized by anomalous steady state fluctuations of the total mass, and by intermittency in the dynamics, which can be probed using structure functions, as in the case of fluid turbulence. Turbulence-like features of aggregation have been studied earlier in [15], where multi-scaling of n-point mass-mass correlation functions was demonstrated for an aggregation model. Our work focuses on a different connection with turbulence, which is related to *temporal* intermittency of mass and multi-scaling of dynamical structure functions. An interesting direction for further work is to investigate whether other systems with giant number fluctuations such as self-propelled particles also show temporal intermittency; dynamical structure functions, as used in this study, could be useful probes in these systems as well.

3 Multi-species Model with interconversion between species

The single species aggregation-fragmentation model discussed so far, can be generalised to include multiple species of particles, with interconversion between particles of different species at some specified rates [fig. 1]. To understand the new effects that arise due to interconversion, it is useful to study multi-species extensions of the two limits, namely the pure chipping limit ($D = D' = 0$) and the pure aggregation limit ($w = w' = q = q' = 0$). In the pure aggregation (zero chipping) limit, the only effect of having multiple, interconverting species is that the composition of aggregates becomes position dependent; the number, typical mass and spatial distribution of aggregates remain unaltered from the single-species case discussed in sec. 2.

In the pure chipping model (where the rate of stack movement is zero), a more interesting effect can arise, which is discussed in detail below. Let m_i^A (m_i^B) be the number of A (B) particles on a site i in a system of L sites. This system is evolved in time as follows: at any instant a site i is chosen randomly. If i is 1, then a single particle of type 'A' is injected into the system at this site. For a general i, a single particle of type A or B either chips off from site i and moves to a neighbouring location with some species-dependent rates, or the 'A' particle converts to a 'B' particle (or vice versa) at site i itself. Single particles of type A or B can also leave the system at the boundaries, again with species-dependent rates. No stack movement is allowed ($D = D' = 0$). Note that the rates of various chipping/interconversion processes are *independent of the number of particles at a site.*

What is the state of this system at long times? We address this question analytically and using numerical simulations and find that depending on various parameters, at long times, the system can either exist in a stationary state characterised by a constant average number of particles, or in a growing phase in which the total number of particles increases indefinitely in time. The growing phase can have interesting spatial structure, with regions in steady state

(finite mean mass) co-existing with regions with pile-ups (indefinitely growing mean mass), as depicted schematically in fig. 7.

To discuss spatial distributions of A (B) particles in the two phases, it is useful to define the occupation probability s_i^A (s_i^B) as the probability of finding at least one A (B) particle at a site i. Since chipping/interconversion rates are independent of the number of particles at a site, all particle currents can be expressed in terms of s_i^A and s_i^B. For example, the *net* 'A' particle current from site i to $i+1$ is simply $ws_i^A - w's_{i+1}^A$ while the net $A \to B$ interconversion current at site i is given by $us_i^A - vs_i^B$.

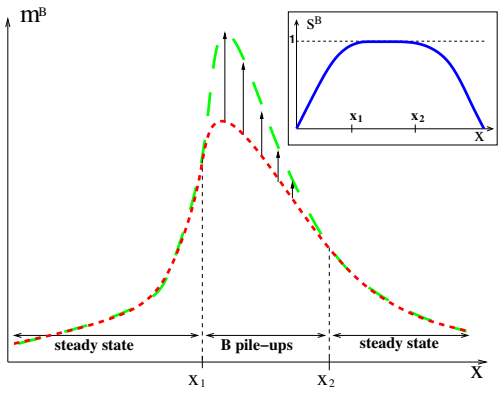
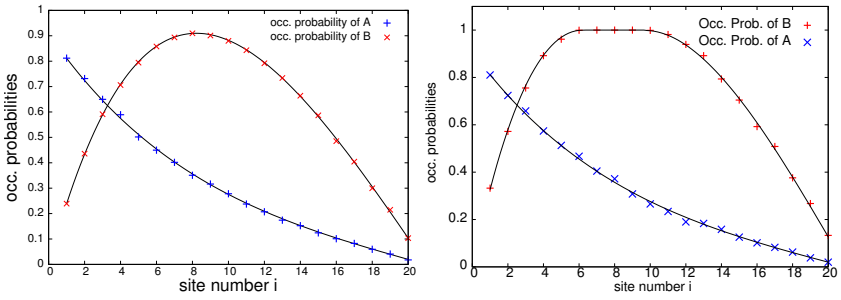


Figure 7: Schematic depiction of a growing phase with pile-ups of B when $q = q'$: $d\langle m^B(x)\rangle/dt > 0$ at all times and $s^B(x) = 1$ in the pile-up region $x_1 < x < x_2$; constant $\langle m^B(x)\rangle$ and $s^B(x) < 1$ everywhere else.

In the *steady phase*, s_i^A and s_i^B reach time independent values *less than 1*, at each site i. Also, the particle currents due to chipping, interconversion etc. balance out at each site, resulting in finite, time-independent values of $\langle m_i^A\rangle$ and $\langle m_i^B\rangle$ for all i. In the *growing phase*, on the other hand, there exist sites where the 'out' current always falls short of the 'in' current, causing the average A (B) mass at the site to grow indefinitely in time. We refer to this unbounded growth as an A (B) *pile-up* at the site. At the site with an A (B) pile-up, the mean occupancy $s^A(s^B)$ approaches 1 at long times. While an A pile-up can be found only at the first site, B pile-ups can also occur in the bulk of the system [5].



(a) (b)

Figure 8: (a)-(b): Occupation probabilities s_i^A and s_i^B as a function of site number i for a system in (a) the steady phase (b) the growing phase with B pile-ups. Data points obtained from simulations in good agreement with analytical predictions (solid lines).

The phase boundary between the steady and the growing phases can be found analytically, by analysing the equations for the time evolution of $\langle m_i^A\rangle$ and $\langle m_i^B\rangle$. These equations only

[5] For B pile-ups to occur a macroscopic (i.e. proportional to L) distance away from the left boundary in a system of size L, interconversion rates should be much smaller than chipping rates: specifically, we require $u/w + v/q$ to be $\mathcal{O}(1/L)$ (or $\mathcal{O}(1/L^2)$) for asymmetric (or symmetric) chipping rates.

involve s_i^A and s_i^B, which can thus be found by assuming that the system is in steady state. If $s_i^A < 1$ and $s_i^B < 1$ for all i, then this solution represents the true steady phase profiles of the occupation probabilities (see fig. 8a), while if $s_i^A > 1$ and/or $s_i^B > 1$ for any i, then this solution is unphysical, implying that the assumption of steady state must be invalid. A more careful analysis yields steady state occupation probabilities in the growing phase as well (see fig. 8b). The parameters for which the solutions s_i^A and/or s_i^B, just touch the value 1 for some i, lie on the critical surfaces separating steady and growing phases.

The occurrence of pile-ups in this model is a simple consequence of the chipping and interconversion rates being independent of m_i^A and m_i^B. For such rates, the particle currents are constrained by an upper bound and hence, may fail to balance in certain regions of the system for some parameters. This effect is fairly general, and also accounts for the formation of linearly growing pile-ups at the boundary sites in the single species version of the model [16]. The occurrence of B pile-ups *in the bulk of the system*, in our model, however, requires interconversion between different species. Interconversion causes the net A (B) current in any bond of the lattice to show non-trivial spatial variation, leading to a situation where this imbalance of currents can occur in some region in the bulk as well.

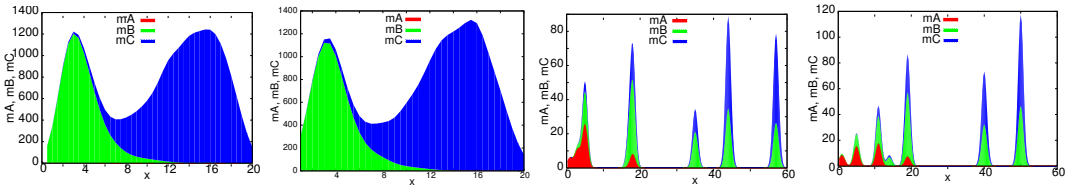
Versions of this model with more than two species of particles show very similar behaviour. For example, the $A \to B \to C \to D$.. model also exhibits growing phases with pile-ups of B and/or C and/or D occurring in overlapping or disjoint spatial regions. The model can also be generalised to allow for chipping and interconversion rates that depend on the number of particles m_i at a site. It follows from arguments related to the boundedness of currents, that this generalised model would also show pile-ups in specific regions of the system, as long as the chipping and interconversion rates are bounded for large m_i. Pile-ups are eliminated once stack movement is introduced, as the particle current out of site i now becomes proportional to $\langle m_i \rangle$ and is no longer bounded above.

4 Relevance to Golgi structure and transport

Experimental evidence in support of the two hypotheses, namely, vesicular transport and cisternal maturation, is mixed. While cisternal maturation has been directly visualised in yeast cells [3], there are other experiments measuring the exit statistics of molecules through the Golgi [4], which appear to refute the maturation hypothesis. A possible resolution is that the dominant mode of Golgi transport could be different in different types of cells. Devising measurements that can distinguish between various modes of transport thus becomes an important experimental challenge. Another interesting question is whether the two transport mechanisms give rise to Golgi cisternae that are structurally different, and whether this can provide an explanation for variations in Golgi structure observed across different types of cells [2].

Motivated by these questions, we focus on the two limiting cases of the general multi-species model– the only chipping limit and the only stack hopping limit, and identify for each limit, the conditions under which the system exhibits multiple, spatially separated aggregates (cisternae) with position-dependent chemical composition. Further, we ask (a) whether the steady state structures in the two limits are qualitatively different (b) what other numerical

measurements could distinguish between the different transport processes in the two limits of the model.



(a) (b) (c) (d)

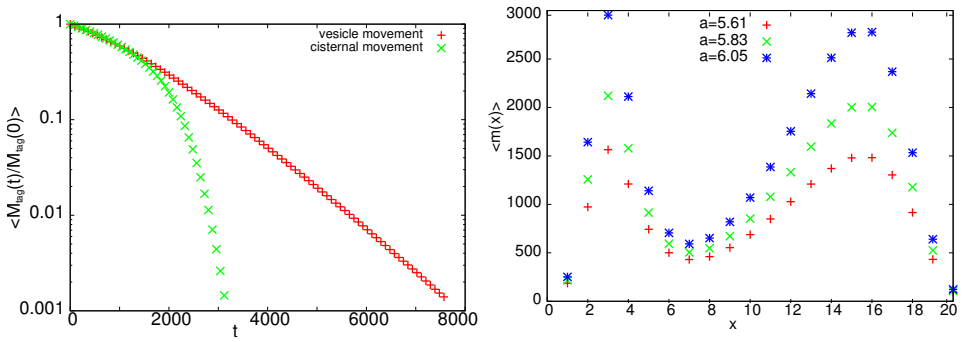
Figure 9: (a)-(b) Two snapshots of a system with static aggregates (obtained in the pure chipping limit): structures show little variation over time. (c)-(d) Two snapshots of a system with dynamic aggregates (obtained in the zero chipping limit): structures vary significantly with time.

We first consider the pure chipping model with $A \to B \to C$ conversion and take the chipping and interconversion rates at a site with m^Z number of 'Z' particles to be proportional to some function $f\left(m^Z\right)$ of m^Z. As discussed earlier, this system shows pile-ups for $f\left(m^Z\right)$ that are bounded for large m^Z. A large, sharply peaked but *finite* aggregate (instead of a pile-up) can be generated by tuning the system to be in the stationary phase but close to the phase boundary between the stationary and pile-up phases [6]. The mass profile obtained under these conditions has multiple, spatially resolvable peaks, with each peak corresponding to a different species. For constant rates [$f\left(m^Z\right)$ independent of m^Z], however, the aggregates obtained in this way are highly unstable, i.e. the mass m of any aggregate shows large fluctuations ($\Delta m/\langle m\rangle \sim 1$). Numerics show that for stability of aggregates, $f\left(m^Z\right)$ must increase with m^Z. The competing requirements of stability and sharpness of aggregates are satisfied in our model with rates of the Michelis-Menten sort, for which $f\left(m^Z\right)$ is an increasing function for small m^Z but saturates to a constant value for large m^Z. With such rates, the chipping model generates multiple aggregates that are somewhat sharp and stable, as shown in fig. 9a, 9b.

We next consider the model with *biased* stack movement, *no* chipping, and $A \to B \to C$ conversion. This model generically produces (without much fine-tuning) multiple, spatially separated aggregates with space-dependent chemical composition. The typical number, size, and spatial distribution of aggregates is the same as in the single-species version; only the chemical composition, i.e. the relative amounts of A,B,C are altered due to interconversion. Thus, there are typically $\mathcal{O}(\sqrt{L})$ number of aggregates in a system of length L; aggregates become sparser but larger on going from left to right, with the typical mass scaling as $\sqrt{x}$ with distance x from the point of injection. In addition, the composition of aggregates varies with position, changing from A-rich to B-rich to C-rich from left to right (see representative example in fig. 9c, 9d). These qualitative features are not altered in the presence of chipping, as long as the chipping rate is much smaller than the rate of stack movement.

[6] This way of obtaining sharp but finite aggregates requires fine tuning of rates. A more robust way is to introduce stack movement as a small perturbation to a system with pile-ups. This eliminates pile-ups, leading instead to the emergence of sharply peaked but finite aggregates in the same region where the pile-up occurred in the corresponding $D = 0$ model. Aggregates obtained in this way are however not very stable in time.

Figure 9 shows two snapshots (at two different time instants) each of structures obtained in the (a) chipping limit [figs. 9a,9b] (b) stack hopping limit [figs. 9c,9d]. The mass profiles in the two limits are quite different. The only chipping model generates static aggregates [figs. 9a,9b] whose spatial position, size and composition does not vary much between the two snapshots. Another distinctive feature is that the mass profile is continuous, with a significant amount of material in the region between the aggregates. By contrast, in the only stack-hopping limit, aggregates are highly dynamic [figs. 9c,9d]– their positions, composition and even number can show significant variation over time. These aggregates are typically 'mixed' and have particles of both types B and C, with the relative amounts of the species changing gradually over space. Also, the region between successive aggregates contains little or no mass.



(a) (b)

Figure 10: (a) Number of tagged particles in the system as a function of time for (i) static cisternae (exponential decay) and (ii) dynamic cisternae (linear decay). Unit of time for the two curves has been chosen differently in order to display both clearly in the same plot.
(b) $\langle m_i \rangle$ vs. i for 3 different values of influx rate a in the case of static cisternae: mass profile changes shape (becomes more peaked) as a is increased.

The structures in these two limits have other differences which can be probed using certain measurements. Motivated by experiments on the Golgi [4], we perform numerical simulations in which all the particles present in the system at a particular time instant are tagged, and the decay of the concentration of tagged particles in the system is monitored as a function of time. Numerics show that for the model with stack movement, the concentration decays linearly ($\sim 1 - t/\tau_A$), whereas for the only chipping model, the decay is nearly exponential ($\sim \exp[-t/\tau_B]$) [fig 10a]. Experiments on certain cell types have also reported an exponential decay [4]. Another possible probe could be to monitor the increase δm in the mass m of an aggregate, in response to an increase in the influx. We measure this for the two limits using simulations and find that the model with stack movement shows a linear response ($\delta m/m$ independent of m), while for the only-chipping model, the response is non-linear, ($\delta m/m$ higher for larger m). Thus, in the latter case (with static cisternae), the mass profile becomes more and more peaked as influx increases [fig. 10b]. This suggests that treating cells to alter the influx of molecules from the ER could be a potentially useful way of probing the Golgi.

To summarise, structures that bear some qualitative similarities to Golgi cisternae can be obtained in two different limits of the model. The static and dynamical properties of the

structures in these two limits are quite different; our study suggests some possible probes of these. A significant limitation of this study is that it does not explicitly take into account the role of enzymes in the processing of cargo molecules in the Golgi, and the complex dynamics of the enzymes themselves. Studying enzyme plus cargo models and examining the conditions under which they give rise to reduced models with effective rates is thus an important next step in extending this work towards more realistic modeling of the Golgi.

5 Conclusion

In this thesis, we have studied a simple model motivated by the processes involved in transport and structure formation in the Golgi apparatus in cells. The first main result of this work is to demonstrate the occurrence of a condensation transition in an open system with aggregation and fragmentation of masses. A distinctive feature of the condensate phase in this system is the intermittent dynamics of the total mass, which can be quantified using structure functions, and also studied analytically in a simplified limit. An interesting future direction could be to study whether temporal intermittency occurs more generally in particle systems with large fluctuations and clustering. The second main result of the work relates to the emergence of a non-stationary, growing phase in the multi-species model. This phase is characterised by the formation of indefinitely growing aggregates in spatial regions of tunable width and location. Finally, we identify two limits of the full model in which structures qualitatively similar to Golgi cisternae can be found. We study some properties of these structures, and discuss two different measurements which could be sensitive to the nature of the transport mechanism.

References

[1] B. Alberts et al, *Molecular Biology of the Cell*, 4/e (New York: Garland, 2003)

[2] B.S. Glick, A. Nakano, Annu Rev Cell Dev Biol **25**:113–132 (2009).

[3] E. Losev, C.A. Reinke, J. Jellen, D.E. Strongin, B.J. Bevis, B.S. Glick. Nature **441**, 1002 (2006); K. Matsuura, M. Takeuchi, A. Ichihara, K. Mikuriya and A. Nakano, Nature **441**, 1007 (2006).

[4] G.H. Patterson, K. Hirschberg, R.S. Polishchuk, D. Gerlich, R.D. Phair and J. Lippincott-Schwartz, Cell **133**, 1055 (2008).

[5] H. R. Pelham, J. E. Rothman, Cell **102**, 713 (2000).

[6] S. N. Majumdar, S. Krishnamurthy, and M. Barma, Phys. Rev. Lett. **81**, 3691 (1998); J. Stat. Phys. **99**, 1 (2000).

[7] U. Frisch, *Turbulence: The Legacy of A. N. Kolmogorov*, (Cambridge Univ. Press, Cambridge, 1995).

[8] S.N. Majumdar, Les Houches (2008) lecture notes, arXiv:0904:4097

[9] M. R. Evans and T. Hanney, J. Phys. A **38**, R195 (2005)

[10] V. Narayan, S. Ramaswamy, and N. Menon, Science 317, 105 (2007).

[11] D. Das and M. Barma, Phys. Rev. Lett. **85**, 1602 (2000); D. Das, M. Barma, and S.N. Majumdar, Phys. Rev. E **64**, 046126 (2001)

[12] S. Dey, D. Das, and R. Rajesh, Phys. Rev. Lett **108**, 238001 (2012).

[13] J. Bec, K. Khanin, Phys. Rep. **447**, 1 (2007).

[14] R. Rajesh and S. Krishnamurthy, Phys. Rev. E **66**, 046132 (2002).

[15] C. Connaughton, R. Rajesh, and O. Zaboronski, Phys. Rev. Lett. **94**, 194503 (2005); Physica D **222**, 97 (2006).

[16] E. Levine, D. Mukamel and G. M. Schütz, J. Stat. Phys. **120** 759 (2005).

Publications

- *Multi-species Model with Inter-conversion, Chipping and Injection*,
  H. Sachdeva, M. Barma and M. Rao, Phys. Rev. E **84**, 031106 (2011).
- *Condensation and Intermittency in an Open-Boundary Aggregation-Fragmentation Model*,
  H. Sachdeva, M. Barma, and M. Rao, Phys. Rev. Lett. **110**, 150601 (2013).
- *Analytical Study of Giant Fluctuations and Temporal Intermittency in an Aggregation Model*,
  H. Sachdeva and M. Barma, J. Stat. Phys. **154**, 950 (2014).

Chapter 1

Introduction

1.1 Biological Context: Golgi Structure and Transport

Biological systems often consist of a large number of interacting components which self-organise over many spatio-temporal scales to form complex structures. A key feature is that these are highly organised, low-entropy structures which are maintained far from equilibrium by a constant flow of energy, mass etc [1]. Non-equilibrium statistical physics thus emerges as a natural framework for addressing questions related to self-organisation and transport in biological systems. In fact, statistical physics approaches have proven fruitful in studying emergent behaviour at various levels of biological organisation, ranging from the sub-cellular to the population-wide [2, 3, 4].

The interaction between biology and statistical physics can be thought of as a dual one– first, biological questions often inspire models that typify and provide insight into interesting non-equilibrium behaviour. This has been the case with the Exclusion Process which was first introduced [1] as a model for ribosomes moving along RNAs [5], and has now emerged as a paradigmatic model for studying several general properties of non-equilibrium systems [7]. Second, statistical physics approaches to modeling biological phenomena sometimes shed light on new aspects such as the role of fluctuations in biology [8], the architecture and robustness of biological networks [9, 10], the occurrence of 'universal' features across diverse length scales [11]. An example of the last arises in the context of self-propelled entities, where the modeling of swimming organisms, flocks of birds and cytoskeleton filaments in cells, within the common framework of active hydrodynamics, suggests that all these systems are governed by similar equations. This may explain features such as giant density fluctuations, which are common to all these systems [12].

In recent years, there has been a lot of interest in decoding the mechanisms of transport and self-assembly at the sub-cellular level. Cells possess multiple transport pathways, which are crucial for exchange of molecules between a cell and its environment, as well as between different organelles (compartments) within the cell. For instance, the secretory pathway involves movement of molecules from the endoplasmic reticulum (ER) to the cell surface via an organelle system called the Golgi apparatus. The Golgi apparatus itself consists of spatially

[1] The Exclusion Process was also introduced independently in the mathematics literature in the context of interacting Markov Processes by Spitzer in [6]

and chemically distinct sub-compartments known as cisternae. The cisternae close to the ER constitute the cis Golgi, the cisternae on the other end close to cell membrane form the trans Golgi, while the intermediate cisternae are referred to as the medial Golgi. Different types of cargo i.e. protein and lipid molecules arrive from the ER, enter the Golgi through the cis face, traverse the Golgi in the cis to trans direction, and get sequentially modified by different enzymes in different regions of the Golgi; the modified products then leave through the trans face for target locations, mainly on the cell membrane [13].

How are cargo molecules transported *within* the Golgi? This question has attracted a great deal of interest as well as controversy [14]. Two alternative hypotheses have been suggested [13]:
Vesicular Transport: In this view, the cisternae are stable structures with a fixed spatial location within the Golgi and a fixed set of resident enzymes. Cargo molecules shuttle from one stationary cisterna to the next in small sacs called vesicles, and get chemically modified by the enzymes present at each cisterna.
Cisternal Maturation: This model envisages cisternae as highly dynamic structures that form by fusion of incoming vesicles and move as a whole, carrying cargo molecules with them. Specific enzymes get attached to a cisterna in different stages of its progression, and modify the molecules in it. The final cisterna eventually breaks up, releasing processed end products.

These two mechanisms differ essentially in whether cargo molecules are transported by the movement of large aggregates (cisternae) or small fragments (vesicles) that break off from these large aggregates. Independent studies on a variety of cells provide mixed evidence for various possibilities [15, 16]. It has also been suggested that the two mechanisms are not mutually exclusive, and transport may be occurring via a combination of both [17].

Not only the cargo molecules, but also the molecules that constitute the Golgi organelle, must be trafficked by the processes described above. This invites the following question: how are spatially and chemically distinct structures (cisternae) of the Golgi formed and dynamically maintained, given the rules of molecular trafficking, which broadly, may be described as: (i) localised influx and outflux of 'particles' (vesicles) (ii) transformation of particles from one species to the other, i.e. chemical processing of cargo molecules by enzymes (iii) transport of particles either by chipping, i.e., fragmentation and movement of single particles (corresponding to vesicle movement) or through movement of whole aggregates (corresponding to cisternal maturation).

In keeping with the reciprocal relationship between biological questions and statistical physics models, the aim of this thesis is also twofold. The first is to construct the simplest version of a model incorporating the processes involved in Golgi dynamics and analyse it in detail to understand the broad effects that emerge from their interplay, and their implications for statistical physics. The second aim is to identify those regions in the parameter space of the model that yield Golgi-like structures and examine whether these structures bear distinctive signatures of the transport mechanism, which would be relevant for experiments on the Golgi.

1.2 The Model

Motivated by the phenomenology of Golgi traffic, as described above, we define the following stochastic model. The model consists of multiple species (denoted by 'A', 'B', 'C', ...) of particles on a 1D lattice [2] with L sites, where each site can have an unrestricted number of particles. For simplicity, we focus on the two-species version of this model; the extension to a higher number of species is straightforward. The system evolves as follows: at each instant, a site i is chosen at random and one of the following moves (illustrated in fig. 1.1) occurs:

i. *Influx*: A single 'A' particle of unit mass is injected at the first site ($i = 1$) at rate a.

ii. *Interconversion*: With rate u, one of the A particles residing on site i converts to type B. Conversely, one of the B particles can also convert to type A with rate v.

iii. *Chipping of single particles*: With rate w (or w'), one of the A particles breaks off from the mass present at site i the site and hops to the site $i+1$ (or $i-1$). Similarly, with rate q (or q'), one of the B particles breaks off from site i and hops to $i+1$ (or $i-1$).

iv. *Movement and aggregation of full stacks*: With rate D (or D'), the full stack at site i, i.e. all A and B particles together, hops to site $i+1$ (or $i-1$) and adds to the mass there.

v. *Outflux from boundaries*: With rate βD (or $\alpha D'$), the full stack at site L (or site 1) exits the system; with rate βw (or $\alpha w'$), a single A particle breaks off and exits from site L (or site 1); with rate βq (or $\alpha q'$), a single B particle exits from site L (or site 1).

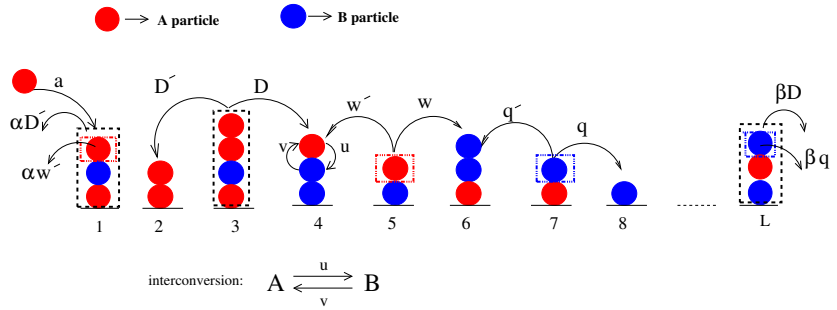


Figure 1.1: Illustration of allowed moves in the model: injection of single particles of type A at site 1 at rate a; interconversion of single particles $A \to B$ at rate u and $B \to A$ at rate v; chipping of an A (B) particle and rightward hopping with rate w (q) or leftward hopping with rate w' (q'); movement of full stack to the right (left) with rate D (D'); outflux of full stacks or single particles from left and/or right boundary.

The correspondence between the model and the Golgi is as follows: a single particle in the model can be viewed as the equivalent of a vesicle, while a stack i.e. any multi-particle aggregate corresponds to a cisterna. Chipping and movement of single particles thus corresponds to vesicle movement, and stack movement to cisternal progression. As in the Golgi, a net flux

[2] Studying a 1D model can be justified on the grounds that there is a net direction of transport of molecules through the Golgi; the spatial variation of enzymes across the Golgi also occurs along this direction.

is maintained through the system by having a source and sink of particles at different ends of the 1D system. The various species of particles in this model correspond to molecules in various stages of processing in the Golgi: 'A' particles thus correspond to unprocessed molecules entering the Golgi at the cis face, B particles to molecules in the next stage of processing, C particles to the subsequent stage and so on. Interconversion between particles of different species then becomes the equivalent of chemical processing in the Golgi.

In constructing this model, we make a number of simplifying assumptions. In reality, chemical processing in the Golgi is mediated by enzymes, which themselves may have complex spatial distributions and dynamics. Our model does not explicitly take this into account, but instead treats the enzyme-mediated process as an effective rate process. Second, all rates in the model are taken to be constant, e.g., the rate of aggregate movement and the single particle chipping rate are assumed to be independent of the mass of the parent aggregate. This may not be realistic in the context of the Golgi. Finally, a 1D model of this sort cannot address questions related to the morphology of Golgi cisternae. In subsequent chapters, we relax some of these assumptions, e.g. by allowing for chipping rates that depend on the number of particles at a site (Chapter 4). However, before considering these additional complications, it is useful to understand the behaviour of the simpler model defined above, as it also exhibits several non-trivial effects due to the interplay of different processes.

1.3 Processes involved in the model: a review of some known results

The model defined above incorporates a number of processes, namely diffusion and aggregation of masses, fragmentation of single particles and interconversion between various types of particles. The system is also open, in that there is influx and outflux of mass at the boundaries, leading to a net particle current through the system. Many of these elementary processes have been studied separately in some detail in various contexts in statistical physics, and are associated with different sorts of phases and phase transitions. A question that we address in this thesis is whether all these processes acting together can lead to new and interesting effects. This question thus provides an additional motivation for studying the model, apart from its possible relevance to Golgi traffic.

Before discussing the new effects that arise in our model due to the interplay of all these processes, it is useful to review the features of some simple models where only a few of these processes are at play and involved in shaping the state of the system. The models that are most pertinent to the present study are discussed below.

1.3.1 Boundary fluxes

Many systems in nature are *open*, in that they exchange mass or energy with their surroundings, typically at the boundaries of the system. Boundary fluxes can give rise to a non-equilibrium, *current-carrying* stationary state of the system, where the net current (of energy or mass) through the system is determined by the rate of influx and outflux at the boundaries. The presence of boundary fluxes typically destroys translational invariance, resulting in system-wide spatial gradients of various quantities. A simple model of an open

system with boundary fluxes is the Open-Boundary Asymmetric Exclusion Process which consists of particles moving uni-directionally while subject to hard core constraints on a 1D lattice, with influx and outflux of particles at the two ends of the lattice [7]. This system exhibits boundary-induced phase transitions, where the phase *in the bulk* of the system undergoes a transition from a high-density to a low-density state in response to a change in the rate of boundary influx and/or outflux [18].

1.3.2 Diffusion and Aggregation

Aggregation processes are ubiquitous in nature and play an important role in diverse phenomena ranging from structure formation in astrophysics [19], formation of clouds and aerosols [20], to polymer dynamics [21] and biological assembly. Valuable insight into the physics of aggregation phenomena has been gained by the study of simple models of the reaction-diffusion sort, which consist of particles that diffuse on some underlying lattice and react, e.g., aggregate on contact to form larger masses which also diffuse, aggregate, and so on, leading to a formation of larger and larger aggregates. While some of these models treat aggregates as spatially extended objects with the aim of studying their morphology [22], in other studies, aggregates are assumed to be point objects, and the main focus is on studying the kinetics of the aggregation process and distribution of aggregate masses. Within the latter class of models also, some models treat aggregation as an $A + A \to A$ reaction [23, 24], while in others, the mass of the reacting aggregates is also taken into account [25, 26, 27]. The following discussion is limited to aggregation models which consider aggregates as point objects with mass, as these are most relevant to the present work.

Dimensionality plays an important role in aggregation phenomena. Most aggregation models have an upper critical dimension $d_c = 2$. For dimensions greater than 2, the behaviour of aggregation-diffusion models is captured by the mean field Smoluchowski equation which describes the time evolution of the mass distribution of aggregates, ignoring spatial distributions altogether (see [28]). For $d \leq 2$, however, it becomes important to explicitly account for the movement of aggregates in space, as the behaviour of the system is diffusion-limited with spatial (anti-) correlations between particles being strong enough to cause deviation from mean field predictions. This point is clearly illustrated in the behaviour of models of aggregation without any external input of particles. In the absence of influx, no steady state is reached, and a kind of coarsening occurs, wherein the mass gets concentrated in increasingly larger but fewer aggregates. In fact the concentration of aggregates in the system decays as a power law $t^{-\phi}$ in time, where ϕ is predicted correctly by a mean field theory for $d > 2$ but is smaller than the mean field exponent for $d \leq 2$. This emergence of a critical dimension $d_c = 2$ in the simplest reaction diffusion system can be traced to the change in the recurrence properties of the underlying random walk process at $d_c = 2$.

More complex behaviour can arise in models where aggregation occurs in conjunction with a steady input of mass. This was studied in [25] by Takayasu et al, who considered diffusion and aggregation with spatially uniform injection of unit masses. They found that at large times, the interplay of injection and aggregation gives rise to a quasi-stationary state in which the mean mass at any site grows indefinitely, while the mass distribution of aggregates approaches a stationary power-law decaying form $\sim m^{-\theta}$ with a time-dependent cutoff. In 1D, the power law exponent could be calculated exactly ($\theta = 4/3$) and was found to be different

from the mean field value ($\theta = 3/2$)– another manifestation of diffusion-limited behaviour in lower dimensions. The Takayasu model provides a simple example of a system which shows self-organised criticality– the probability distribution of aggregate masses always decays as a power law, without any fine tuning of parameters. It is also one of the few models where time dependent quantities including spatio temporal correlation functions of mass can be calculated exactly [29, 30].

The above system with aggregation and injection can be viewed as a turbulent system in a generalised sense, as suggested in [31]. Turbulent systems, e.g., a turbulent fluid, are typically non-linear systems which have energy (or mass) injected at a particular 'driving' scale and dissipated at a different scale. The driving and dissipation scales are usually well-separated, and the system is characterised by a cascade of energy/mass in the inertial range between the two scales. In the aggregation-injection system, mass is the cascading quantity, with the injection of unit masses being the driving. If in addition, aggregates with mass above a certain cutoff value are removed from the system, then the dissipation scale is set by this cutoff which can even be taken to be infinity. Motivated by this view of aggregation as a turbulence-like process, Connaughton et al studied n-point mass-mass correlation functions in this model and demonstrated that these show multi-fractal scaling. They used field-theoretic techniques to calculate the n-point functions and found that for $d \leq 2$, these have an anomalous dependence on mass that deviates from the prediction of a Kolmogorov-like, self-similarity hypothesis [31].

All the above models possess translational invariance, which makes it relatively simple to study them analytically. In systems with localised injection of mass, translational invariance is broken, as manifested in position-dependent mass distributions. In general, these are difficult to calculate analytically, but in 1D, some exact results have been obtained [23, 24, 26]. For example, Cheng et al considered diffusing and aggregating masses in a semi-infinite system with a source of particles at the origin and explicitly calculated the steady state distribution of aggregate size as a function of distance from the source [26]. A noteworthy feature of systems with localised injection is that bias in movement of aggregates can lead to behaviour which is qualitatively different from the unbiased case [26]. This was studied in detail by Jain et al in [32] and Reuveni et al in [33], where various mass distributions of the system were analysed and shown to be different from those in the unbiased case.

In summary, the simplest diffusion-aggregation models also show non-trivial features, especially in lower dimensions when the behaviour is diffusion-limited. In models of aggregation with input, in particular, a quasi-stationary state is reached which shows self organised criticality in the sense of power law distribution of aggregate sizes [25], and turbulence-like behaviour in the sense of multi-scaling of higher correlation functions [31]. Further, localised injection leads to breakdown of translational invariance and non-trivial position-dependent properties [26, 32]. Many of these models can be analysed using exact techniques in 1D [25, 27].

1.3.3 Fragmentation

In the previous section, we considered a mode of mass exchange between spatial locations wherein the entire mass on a site diffuses to a neighboring location, joining with the mass already present there. However, in other contexts, it is possible that only a fragment of the

mass breaks off and moves. This fragment could either have a fixed mass or it could be a random fraction of the total mass [34]. In the following discussion, we focus on the former scenario, specifically on the Zero Range Process (ZRP) (see [35] for a review). The ZRP is a lattice model where each site of the lattice is occupied by an integer-valued mass. At each time step, a fragment of unit mass is allowed to break off from a site with a rate that is a function of the mass present at that site. The unit mass fragment then hops on to the adjoining site and adds to the mass present there.

This simple mass transfer model also has a number of interesting features [35]. First, the steady state of the model is exactly known and has a factorised form, i.e., it is the product of single-site probabilities (in the grand canonical ensemble). Second, for a class of fragmentation rates, the system undergoes a condensation transition: for total mass greater than a certain critical value, the excess mass accumulates in a single macroscopic aggregate. This aggregate containing a finite fraction of the total system mass is referred to as the condensate. Third, the ZRP is related to several other models, e.g, it can be mapped exactly to the asymmetric exclusion process; it also provides a coarse-grained description of the time-evolution of domains in many driven systems [36]. Further, because of the simple factorised form of the steady state of this system, it is an ideal starting point for studying more complicated effects such as the role of disorder [37, 38, 39], generalisations to multiple species of particles [40, 41], and increasing the range of interactions [42].

An interesting variant of the ZRP is the case where in addition to fragmentation and movement, there is also influx and outflux of single particles from an external reservoir, so that the total mass is not conserved. In particular, the model most relevant to the present work is the one where the exchange with the reservoir takes place at the boundaries of a 1D system [43]. This system, in addition to showing non-conservation of total mass, also lacks translational invariance. In [43], it was shown that for low rates of influx at the boundaries, this system attains a steady state which also has a factorised form. For strong boundary influx, by contrast, the system fails to achieve steady state and the total number of particles grows indefinitely with time: linearly growing aggregates may form on one or both boundary sites, while bulk sites attain steady state or show slower growth, depending on the injection, hopping rates, and asymmetry of hopping [43].

1.3.4 Aggregation-fragmentation dynamics

Non-trivial effects can arise in a system which allows for both aggregation and fragmentation. The interplay between these two processes was studied by Majumdar et al for a closed periodic system with fixed total mass [44]. In this model, masses diffuse and aggregate on contact to form larger masses which also diffuse and aggregate and so on. This tendency to form large aggregates is countered by allowing single particles to fragment or 'chip off' from an aggregate at a constant rate and move to a neighbouring location. The steady state behaviour of the system in the two limiting cases is somewhat trivial. In the pure aggregation limit, the system goes to a steady state in which the entire mass in the system comes together in a single aggregate. Conversely, in the pure chipping limit (zero aggregation), there is no macroscopic aggregate, the mass distribution at each site has an exponentially decaying form

and is independent of the other sites[3].

What is the behaviour of the system at finite rates of aggregation and chipping? The mean field analysis of [44] shows that the system undergoes a phase transition from an aggregation-dominated phase to a fragmentation-dominated phase as the fragmentation rate is increased. In the aggregation-dominated phase, a finite fraction of the total mass is gathered in a single macroscopic aggregate or 'condensate', while the remaining mass is contained in finite aggregates with a power law distribution. The mass in the condensate decreases as the fragmentation rate is increased and goes to zero at a critical fragmentation rate which depends on the mass density in the system. At this critical point, the single site mass distribution is purely a power law with no condensate peak. For fragmentation rates higher than this critical rate, the system exists in a disordered phase; there is no macroscopic aggregate and the mass distribution at any site decays exponentially. The phase transition in this system can be thought of as a real space version of Bose-Einstein condensation. The analysis in [45] shows that the critical value of density, as predicted by mean field theory, is exact and independent of dimension.

A number of variants of this model have been studied and are found to exhibit some interesting features. For instance, if the diffusion of aggregates occurs with a mass-dependent diffusion rate proportional to $m^{-\alpha}$ with α positive, then there is no condensation transition at a finite mass density [46]. Condensate formation is also curbed if the chipping process is biased in one direction, but not if bias is introduced in movement of aggregates alone [47].

1.3.5 Multi-species Models

A salient feature of our model is that there are multiple types or species of particles, with each species having different properties, e.g., different chipping rates etc. Multi-species models have been studied in several contexts in statistical physics; due to the non-equilibrium character of these models, the introduction of even one or a few particles of a different species into the system can alter the macroscopic state of the system [40].

An important class of multi-species models is one where the number of particles of each species is conserved. These include a model of sedimentation which reduces to a system of two coupled species of spins [48], the ABC model [49], and the AHR model [50]. In these models, the coupling between the species can lead to effective long-range interactions and novel effects such as phase separation and long range order in 1D. Another example of this class of models is the two-species ZRP, studied in [40, 41], which involves particles of species A (or B) hopping on the lattice with rates which depend on the number of both A and B particles at the departure site. Depending on the nature of the coupling between the two species, as encoded in the form of the hopping rates, condensates of one or both species can form at high densities. At a heuristic level, several of these two-species models can be conceptualised as particles of one species undergoing time evolution over a dynamic landscape determined by the distribution of the other species [40, 48].

[3]Note that the chipping move is a special case of the ZRP where the hopping rate is independent of the number of particles at the departure site. Such rates do not admit condensation in the ZRP.

Other kinds of effects arise in models with multi-species interactions which do not conserve the number of particles of each species. Multi-species models that allow for inter-conversion have been studied earlier in the context of transport in single [51, 52, 53] and multiple [54, 55, 56] open channel(s). For example, in [55], Reichenbach et al study the two lane open Asymmetric Exclusion Process (ASEP) with switching of particles between lanes. The lanes can be thought of as the equivalent of species in our model, and lane switching as analogous to interconversion. Lane switching in this model gives rise to phase co-existence: a high density phase occurs in a limited region of one channel, even as the rest of the system remains in a low density phase [55, 56].

1.3.6 Themes and Questions Relevant to the Present Work

The model studied in this thesis incorporates all the processes discussed above in sections 1.3.1-1.3.5, namely, boundary fluxes, aggregation-fragmentation dynamics and multi-species interactions. Studies dealing with these processes suggest several questions that may be relevant to the present work.

First, it can be seen that by setting interconversion rates to zero in our model, we obtain a single-species model with aggregation and fragmentation of masses in an open system with boundary fluxes. The competition between aggregation and fragmentation has been studied earlier in closed systems with mass conservation and is known to drive a condensation phase transition in these systems [44, 45]. This motivates the following question– Does the condensation transition survive in a system when the total mass is not conserved, but can undergo large fluctuations due to the exit of clusters of all sizes from the boundaries? And if yes, then how is condensate formation in the open-boundary system qualitatively different from that in the closed case? Other issues that turn out to be relevant in this context include the role of boundary conditions, the effect of directionality or bias in movement of aggregates and the emergence of position-dependent behaviour due to the lack of translational invariance.

More insight into aggregation-fragmentation dynamics in an open system can be gained by looking at the limiting cases of the model, i.e. the pure aggregation limit and the pure fragmentation limit in the presence of boundary fluxes. In particular, it is useful to examine where the open-boundary aggregation model can be situated within the broader literature on aggregation models (see sec. 1.3.2). Analytical techniques used to analyse 1D aggregation models [25] also turn out to be applicable here.

The full model, as defined in sec. 1.2, is a multi-species model where the number of particles of each species is not conserved, even in the bulk of the system, due to interconversion between species. As before, it is useful to consider the two limits– the pure aggregation and the pure fragmentation limit, and generalise them to include multiple interconverting species of particles. The aim is to understand whether the multi-species models exhibit new features, which do not arise with just a single type of particles. This was the case for the multi-species variant of the open boundary exclusion process, which shows strong spatial effects and phase co-existence, not present in the single species model [55]. A related effect turns out to be quite relevant to our model.

Thus, the broad approach we use to study this many-parameter model, is to first analyse various reduced versions of the model which incorporate only a subset of all the allowed processes. This approach allows us to isolate the distinct effects that arise from the interplay of different processes.

1.4 A Brief Statement of the Results of the Thesis

A natural starting point for our study is the single-species version of the full model, obtained by setting interconversion rates to zero. This simply gives the 1D aggregation-fragmentation model with influx of single particles and outflux of aggregates at the boundaries. We study this system using numerical simulations; a detailed account of these results is presented in Chapter 2. In Chapter 3, we consider the pure aggregation limit of the single-species model– this limiting case turns out to be analytically tractable, and is also representative of one phase of the model. In Chapter 4, we consider the multi-species version of the model and again look at limiting cases. The interesting limit here turns out to be the one with *no* movement of stacks as a whole. The interplay of injection, chipping and interconversion in the system gives rise to a different sort of phase, which is the main focus of this chapter. In Chapter 5, the relevance of our model to structure formation and transport in the Golgi is explored. We identify two limits of the full multi-species model, one which is aggregation-dominated and the other chipping-dominated, which exhibit steady state structures which are qualitatively similar to Golgi structures in different types of cells. Further, we bring out the differences between the structures generated in these two limits by analysing both static and dynamical properties. Finally, in Chapter 6, we discuss the conclusions of the thesis and outline interesting future directions.

A brief statement of the main results of the thesis is presented below:

i. *Aggregation-fragmentation model with open boundaries*: On increasing influx or decreasing the fragmentation rate, the system with unbiased motion of particles undergoes a phase transition from a normal (fragmentation-dominated) phase to a 'condensate' (aggregation-dominated) phase in which a finite fraction of the total mass accumulates in a macroscopic aggregate or condensate. The mass in the condensate has a broad distribution and shows giant number fluctuations, unlike the condensate in the closed periodic version of this model [44]. The presence of such a fluctuating condensate leads to several unusual properties: the total number of particles M in the system shows giant fluctuations and has a strongly non-Gaussian probability distribution. Further, the time-series $M(t)$ is not self-similar and shows intermittency, which we characterize in terms of appropriately defined structure functions and their ratios, as in studies of turbulence [57]. The model with biased motion of aggregates also undergoes a phase transition to an aggregation-dominated phase with temporal intermittency of M. These results are established using numerical simulations for non-zero fragmentation rates and analytical calculations in the case of zero fragmentation. In particular, for the zero fragmentation case, we explicitly calculate analytically dynamical structure functions and demonstrate their anomalous scaling.

ii. *Multi-species model with interconversion between species*: A different sort of phase transition arises in the only-chipping limit ($D = D' = 0$) of the multi-species model. De-

pending on the parameters, this system may exist either in a steady phase, in which the average mass at each site attains a time-independent value, or in a 'growing' phase, in which the total mass grows indefinitely in time. The growing phases have interesting spatial structure, with unbounded growth of one or more species of particles occurring in spatial regions of tunable width and location, while steady state is attained in the rest of the system. We have analysed this model by solving equations for particle currents at each site.

iii. *Relevance to Golgi structure and transport*: We focus on two limits, namely, the only-chipping limit and the only-stack-hopping limit of the full multi-species model, which correspond respectively to the vesicular transport and cisternal progression scenarios discussed in Section 1.1. For both these limits, we identify the conditions under which the model generates multiple, spatially and chemically distinct aggregates (which correspond roughly to cisternae in the Golgi). The only-chipping model yields static aggregates whose position, size and composition remains stable in time. In the only-stack-movement limit (with uni-directional movement of stacks), we obtain several moving aggregates whose size, composition and even number can vary significantly over time. The structures in these two limits have distinctive static and dynamical properties; our work suggests some possible probes for these.

The ensuing chapters give a detailed account of these results.

Chapter 2

Single Species Aggregation-Fragmentation Model with Open Boundaries

2.1 Model and Main Results

In this chapter, we discuss the *single-species* version of the full model defined in Chapter 1. The single-species limit is important to study, as it enables us to focus on the interplay of aggregation-fragmentation processes and boundary fluxes and the new features that emerge from this interplay, without the additional complication of multi-species effects. This work is also described in [58].

To recapitulate, this model allows for [see fig. 2.1]

i. influx of single particles of unit mass at site 1 at rate a

ii. movement of the entire stack (i.e. all the particles) present at a site to the right (or left) neighbour with rate D (or D'), and aggregation with the mass present there

iii. chipping (fragmentation) of a single particle from a stack with rate $2w$ and movement of the 'chipped' particle to the right or left neighbour with equal probability

iv. outflux of mass from site L (or site 1) either via exit of whole stacks with rate βD (or $\alpha D'$) or exit of single particles after chipping with rate βw (or αw).

Before studying the model in detail, it is useful to consider how each of the above processes acts individually on a collection of particles. Suppose we have a closed periodic system with a fixed number of particles, that undergo diffusion and aggregation on contact. As time progresses, the process of aggregation creates larger masses which also diffuse, aggregate, and so on, leading to a steady transfer of mass from smaller to larger aggregates. This system has a trivial steady state at long times, with the entire mass in the system accumulating in a single aggregate.

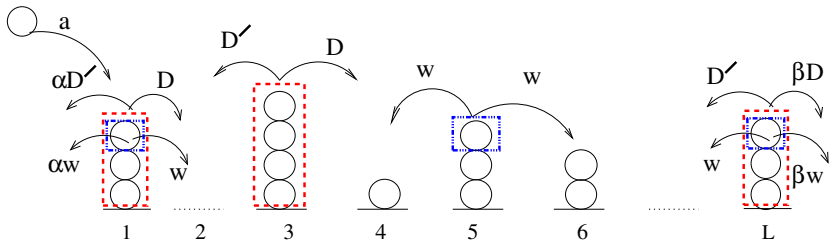


Figure 2.1: Illustration of allowed moves in the model: injection of single particles at site 1 at rate a; movement of entire stack to the right (left) with rate D (D'); chipping of single particles and movement to left or right with equal rates w; outflux of full stacks or single particles from left and/or right boundary.

If in addition to aggregation, the masses also undergo fragmentation or 'chipping' of single particles, then the long time behaviour of the system is somewhat different. Chipping acts counter to aggregation, causing large masses to shred, leading to mass transfer in the reverse direction, i.e. from large aggregates to small aggregates. In [44], Majumdar et al studied the steady state of such a closed periodic system with aggregation *and* single particle fragmentation (chipping) using a mean field analysis. They found that the system undergoes a phase transition as the chipping rate (or the mass density) is varied– at low chipping rates, the system exists in a 'condensate' phase, in which a finite fraction of the total mass is contained in a single aggregate or condensate; at high chipping rates, no such macroscopic aggregate forms. This phase transition also manifests itself in the properties of the mass distribution $P(m)$ of aggregates in the system. In the condensate phase, the mass distribution decays as a power law and has an additional sharply peaked part, corresponding to the macroscopic condensate. As the chipping rate is increased, the fraction of total mass contained in this condensate decreases, and goes to zero at a critical chipping rate. At the critical point, $P(m)$ exhibits a pure power-law decay, with no condensate peak. In the disordered or 'normal' phase, i.e., at high chipping rates, $P(m)$ decays exponentially with m.

The condensation transition in the closed system can be thought of as a real space analogue of Bose-Einstein condensation; when the total mass (or equivalently the mass density) in the system exceeds the maximum value that can be accommodated in the bulk, the excess mass 'condenses' into a single macroscopic aggregate. Condensation transitions also occur in other mass transport models [59] such as the zero range process and its variants [35]. All these models, however, are characterised by a *fixed total mass*.

The model studied in this chapter is the open-boundary version of the closed periodic aggregation-chipping model [44]. An important consequence of open boundaries is that the total number of particles or *total mass in the system is not conserved*; it varies with time and is characterised by a probability distribution. This leads to the following question– does the condensation transition survive in an open system where the total mass is not conserved, but can, in principle, undergo large fluctuations due to the exit of aggregates of any size?

As in the closed periodic case, another focus of this chapter is to study the mass distribution of aggregates in the system; this now also depends on the position of the aggregate, as the open-boundary system lacks translational invariance. As before, diffusion and aggregation

shift the mass distribution of aggregates towards higher and higher values of mass, while the fragmentation of single particles has the reverse effect. In addition, at high rates of influx, we expect the average density of the system to be high, and the typical mass of aggregates to shift to larger values. However, since aggregates exit the system on reaching the boundaries, this puts a limit on the mass that can be accreted by an aggregate, since the typical residence time of an aggregate in the system is finite. What is the mass distribution of aggregates that emerges in the presence of these effects and does it change qualitatively as different parameters of the model are varied? This is another key question addressed in the chapter.

We study these questions for both unbiased ($D = D'$) and fully biased ($D' = 0$) movement of aggregates. Further, for the unbiased case, we consider two special but representative boundary conditions: one in which outflux occurs from both site 1 and L at the same rate (corresponding to $\alpha = 1, \beta = 1$) and the other where outflux occurs only from site L (corresponding to $\alpha = 0, \beta = 1$). This gives rise to the following three cases in all:

A. *Unbiased stack hopping; outflux from both boundaries*: $D = D'$ and $\alpha = \beta = 1$.

B. *Unbiased stack hopping; outflux only from right boundary*: $D = D'$ and $\alpha = 0, \beta = 1$.

C. *Biased stack hopping; outflux from right boundary*: $D' = 0$ and $\alpha = \beta = 1$.

In each of these three cases, injection occurs at rate a and chipping with symmetric rates w to the left and right.

We study various properties of the models (A)-(C) using numerical simulations for non-zero chipping rates w and analytically in the $w = 0$ case (analytical results are discussed in Chapter 3). Our main finding is that the open boundary system with unbiased stack movement *does* undergo a condensation transition upon increasing the influx or decreasing the chipping rate. However, the nature of the condensate is very different from that in the closed model– the condensate mass shows giant fluctuations and is characterised by a long-tailed, non-Gaussian probability distribution, in contrast with the narrow, peaked distribution found in the closed system [44, 45]. The strongly fluctuating condensate gives rise to novel signatures of the total mass M such as giant number fluctuations and strong intermittency in time. Intermittency, which is a feature of many turbulent systems, is quantified in our study using dynamical structure functions of the total mass, in analogy with structure functions of the velocity field in fluid turbulence [57].

We also find that the fact of a phase transition and the main features of the condensate phase do not depend on boundary conditions (as encapsulated in the values of α and β) as long as stack movement is unbiased, although some minor differences emerge on varying α and β. The introduction of spatial bias in the model has a more significant effect. For biased movement of aggregates, the model still shows a phase transition from a normal phase to aggregation-dominated phase [1]. However, the aggregation-dominated phase is different from that in the unbiased models, in that the total mass exhibits temporal intermittency but no giant fluctuations.

[1]Directional bias in single particle movement destroys the phase transition (further discussion in section 5). Thus, in all the three cases (A)-(C), we assume that a particle, upon chipping, moves to the left or right with the same rate w.

2.2 Heuristic Argument for Phase Transition in the Open System

Since the system is open only at the boundaries, it is reasonable to postulate that for a sufficiently large system, the local behaviour at a site which is *far from the boundaries*, can be approximated by equations governing the closed periodic system [45]. In this section, we investigate whether this approach can be used to make some qualitative prediction about the steady state of the open-boundary system. With this aim, some of the main results for the closed system are reviewed below.

The closed system is characterised by a fixed total mass or equivalently a fixed mass density ρ. Thus, the parameters of this model are the density ρ and the chipping rate w (with the rate D set equal to 1). Let $s = s(\rho, w)$ denote the probability that a randomly chosen site in the system is non-empty. The exact analysis in [45] for the case of *unbiased stack movement*, shows that for small mass density ρ, the occupation probability s increases as ρ is increased, reaching its maximum value $s_c(w)$ at a critical value $\rho_c(w)$ of the density. If ρ is increased beyond this critical value, the occupation probability does not increase any further, but remains pegged at $s = s_c(w)$. In other words, the excess mass $(\rho - \rho_c)L$ does not affect the average occupancy of sites, as all of it accumulates in a single macroscopic aggregate or condensate. Thus, for $\rho > \rho_c(w)$, the system goes into a condensate phase in which (i) the average occupancy is fixed at $s = s_c(w)$ (ii) a critical value $\rho_c L$ of the total mass is accommodated in the bulk of the lattice (iii) the excess mass given by $(\rho - \rho_c)L$ collects in a single, moving aggregate. The critical values s_c and ρ_c were derived exactly in [45] and are given by:

$$s_c(w) = (\sqrt{1+w} - 1)/(\sqrt{1+w} + 1) \tag{2.1a}$$

$$\rho_c(w) = \sqrt{1+w} - 1 \tag{2.1b}$$

Let us see whether we can adapt these results to the open-boundary model with unbiased stack movement. For convenience, we henceforth set $D = 1$. The open-boundary model is different from the closed version in several key ways. First, the density is not a fixed parameter of the model, but a random variable with a distribution that is governed by the rates a and w. Second, all quantities such as mean density and occupancy vary with position and must be treated as functions ρ_j, s_j of site number j. In fact, a relation between ρ_j and s_j can be derived simply from the balance of particle currents at each site. This relation is given by:

$$\rho_j + w s_j = a\,(1 - j/L) \tag{2.2}$$

In spite of these differences, we make the assumption that in the limit of large L, sufficiently far from the boundaries, the behaviour of the open-boundary model is similar to that of the closed periodic model. Specifically, we assume that any part of the open system is locally in the condensate phase if the conditions for condensate formation, as derived for the closed model, are satisfied. Thus, if a site j belongs to a region which is locally in the condensate phase, then $s_j = s_c(w)$ and $\rho_j > \rho_c(w)$ where s_c and ρ_c are given by eq. (2.1). Setting

$s_j = s_c(w)$ and $\rho_j > \rho_c(w)$ in eq. (2.2) thus gives the following condition for a (re-scaled) location $x = j/L$ within the system to be locally in the condensate phase:

$$x < 1 - \left[\frac{1 + w - \sqrt{1+w}}{a}\right] \tag{2.3}$$

The above inequality implies that for a given set of parameters (w, a), there is a critical position in space $x_c(w, a)$ such that regions to the left of this location ($x < x_c$) support condensation while regions to the right of this location ($x > x_c$) are in a 'normal' phase with no condensate. Clearly, $x_c(w, a)$ is given by the expression on the right hand side of eq. (2.3).

From this argument, we can estimate the critical chipping rate at which the system undergoes a phase transition from a phase where there is a macroscopic condensate *somewhere* in the system to a phase where there is no condensate anywhere. The critical rate w_c can be obtained by setting $x_c(w_c, a) = 0$. It is given by:

$$w_c = \left(2a - 1 + \sqrt{4a+1}\right)/2 \tag{2.4}$$

To summarise, the above argument makes two main predictions. The first prediction is that the open-system with diffusive (unbiased) stack movement and chipping exhibits a condensate phase at small w and a normal phase at large w. The argument also gives an estimate for the critical chipping rate w_c at which the phase transition occurs. Second, it suggests that the condensate phase is not spatially uniform– the condensate is confined to a finite fraction of the system while the remaining fraction shows normal (disordered) behaviour.

However, this argument is based on the assumption that far from the boundaries, the open-boundary system can be described by the results derived for the closed system. This is an ad-hoc assumption, which needs to be tested by studying the model using other methods. In subsequent sections, we report the results of extensive numerical simulations and investigate whether the predictions of this simple argument are borne out by these.

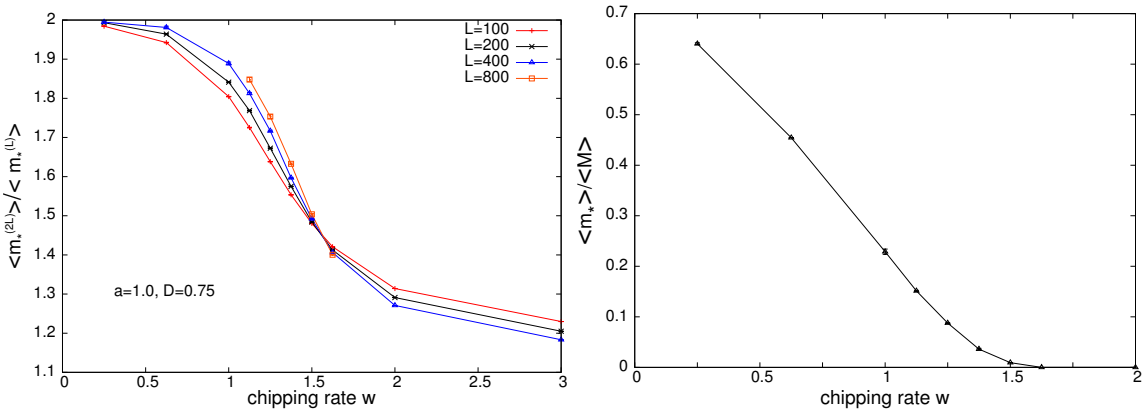
2.3 Model A: Unbiased stack hopping; outflux from both boundaries

In this section, we consider the case where both stack movement and single particle chipping occur with equal rates to the left and right. Further, while injection of single particles takes place only at the left boundary, outflux in the form of exit of single particles as well as entire stacks, occurs at *both* boundaries. This model is characterised by three parameters: the injection rate a, the chipping rate w, and the system size L.

We study the system by simulating the dynamics of the model for different values of the above parameters using the following Monte Carlo algorithm: at each time step, a site i is chosen randomly. If the site is occupied or if $i = 1$, we draw a random number from a uniform distribution between 0 and 1. For a site in the bulk, if this number lies between 0 and $2w/(a + 2w + 2D)$, then we allow a single particle to chip off from site i and move to site $i + 1$ or site $i - 1$ with equal probability. Similarly, if the random number lies between

$2w/(a+2w+2D)$ and $(2w+2D)/(a+2w+2D)$, then the entire mass at site i is made to move to the left or right neighbouring site (with equal probability). At the left (right) boundaries, single particles/stacks that hop leftwards (rightwards) leave the system irreversibly. selected site is site 1, and the random number is greater than $(2w+2D)/(a+2w+2D)$, then a single particle is added at site 1. In order to calculate statistics of various quantities, we first allow the system to attain steady state, and then sample the state of the system at regular time intervals. Average values are computed by performing both time averages (by sampling at $\sim 10^9$ time instants) and ensemble averages (over 10-20 independent copies of the system).

In order to probe condensate formation, we monitor the mass $m_*^{(L)}$ of the *largest aggregate* in the system. Numerics show that at small values of the chipping rate w, the largest aggregate, on an average, contains a finite fraction of the total mass $M^{(L)}$, i.e., $\langle m_*^{(L)} \rangle \propto \langle M^{(L)} \rangle \propto L$ for fixed a, w. In other words, there is a *macroscopic* aggregate or 'condensate' present in the system. At large w, by contrast, $\langle m_*^{(L)} \rangle$ is sublinear in L, implying that there is no condensate. The fraction $\langle m_*^{(L)} \rangle / \langle M^{(L)} \rangle$ or equivalently $\langle m_*^{(L)} \rangle / L$, can thus be taken as an order parameter, and the critical point identified with the chipping rate w_c at which $\langle m_*^{(L)} \rangle / \langle M^{(L)} \rangle$ goes to zero in the limit $L \to \infty$ [2].



(a) (b)

Figure 2.2: Determination of w_c for model A (for $a = 1.0$, $D = 0.75$) using (a) Ratio method: the ratio $\langle m_*^{(2L)} \rangle / \langle m_*^{(L)} \rangle$ increases with L for small w and decreases with L for large w. The point of intersection of various curves (at which point the ratios do not change with L) corresponds to w_c (b) Intercept method: The fraction $\langle m_* \rangle / \langle M \rangle$, as obtained by extrapolation of finite L data, goes to zero at the critical chipping rate w_c. Estimated value of w_c using both methods is between 1.5 and 1.625, consistent with the theoretical estimate (see eq. 2.4) of $w_c = 1.5687$.

We determine the critical chipping rate from our numerical data using two different methods, which will be referred to as the ratio method and the intercept method from now on. The *ratio method* involves plotting the ratio $\langle m_*^{(2L)} \rangle / \langle m_*^{(L)} \rangle$ vs. w for different values of L, as shown in fig. 2.2a. For small w, the ratio *increases* with increasing L and appears to be approaching the value 2, consistent with $\langle m_*^{(L)} \rangle \propto L$. This behaviour corresponds to the condensate phase. For large w, the ratio $\langle m_*^{(2L)} \rangle / \langle m_*^{(L)} \rangle$ *decreases* with increasing L and

[2]From now on, we omit the superscript (L), for the total mass $M^{(L)}$ in order to keep the notation simple.

approaches a value less than 2 (this value could be 1). This behaviour is characteristic of a system in the normal phase, where there is no condensate formation. The point of intersection of the curves corresponding to different L, yields the critical chipping rate w_c, as this is the chipping rate for which there is neither any increase nor decrease in the value of the ratio $\langle m_*^{(2L)}\rangle/\langle m_*^{(L)}\rangle$ as a function of L [3].

The other method for critical point determination, namely the *intercept method* is based on the assumption that in the condensate phase, the fraction $\langle m_*^{(L)}\rangle/\langle M\rangle$ approaches its asymptotic value 'c' as $L \to \infty$, with the approach being of the form $\langle m_*^{(L)}\rangle/\langle M\rangle \sim c + gL^{-\phi}$. This asymptotic value can be obtained for each w by plotting $\langle m_*^{(L)}\rangle/\langle M\rangle$ vs. $L^{-\phi}$ (where ϕ is that exponent for which the above plot becomes a straight line) and measuring the y-intercept of the straight-line plot. In practice, due to the small number of data points, we directly fit the data to get the parameters c, g and ϕ for each w. Figure 2.2b shows $\langle m_*^{(L)}\rangle/\langle M\rangle$, as found using this method, as a function of w. The critical chipping rate w_c is thus the rate for which the fitting parameter c becomes zero. The estimates of w_c obtained using the ratio and the intercept methods agree with each other and also with the theoretically estimated value in eq. (2.4) to within 10%.

The qualitatively different behaviour of $\langle m_*^{(L)}\rangle$ in the large w and small w limits establishes the existence of the two phases– the normal and the condensate phase. The two phases have very different static and dynamical properties which we discuss below.

2.3.1 Statics

2.3.1.1 Condensate phase

Largest mass $m_*^{(L)}$: In the condensate (small w) phase, the largest mass $m_*^{(L)}$ in the system has an average value proportional to L, implying that the system contains a macroscopic condensate. This condensate is, however, qualitatively different from the one found in the closed periodic version of this model [44], in that the mass in the condensate can fluctuate widely and shows root mean square (rms) fluctuations $\Delta m_*^{(L)} \equiv \sqrt{\langle [m_*^{(L)}]^2\rangle - \langle m_*^{(L)}\rangle^2}$ that scale as the mean condensate mass itself, i.e., $\Delta m_*^{(L)} \propto \langle m_*^{(L)}\rangle \propto L$. The distribution of the condensate mass also has a broad, strongly non-Gaussian character and for large $m_*^{(L)}$, is given by $P(m_*^{(L)}) \sim (1/M_0)\exp\left(-m_*^{(L)}/M_0\right)$ where M_0 is proportional to L (fig. 2.3a). We refer to this exponential tail, with characteristic mass M_0, as the 'condensate tail'.

The presence of the condensate leaves a strong imprint on all steady state properties of the system, with the non-Gaussian condensate tail appearing in the probability distributions of many quantities, including the total mass in the system.

[3] Note that this method of determining the critical point is equivalent to plotting $\langle m_*^{(L)}\rangle$ as a function of L on a log-log plot for different w, and studying the trend of these curves. If the $\langle m_*^{(L)}\rangle$ vs. L curve for a particular w bends upwards on the log-log plot, then this signifies that the system is in the condensate phase for that value of w. Conversely, curves that show downward bending correspond to the normal phase. The curve that shows no bending on the log-log plot is the critical curve, and the chipping rate it corresponds to, the critical chipping rate w_c.

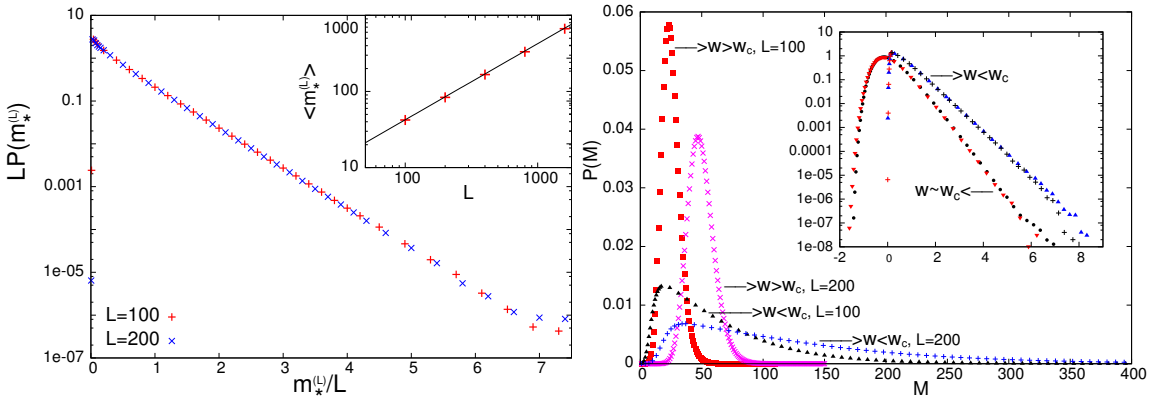


Figure 2.3: (a) Probability distribution $P(m_*^{(L)})$ of the largest mass in the system in the condensate phase for two different L. Data collapse of the exponentially decaying tails on plotting $LP(m_*^{(L)})$ vs. $m_*^{(L)}/L$. Inset: The average value $\langle m_*^{(L)} \rangle$ is proportional to L.
(b) $P(M)$ vs. M for two different L in the normal phase (Gaussian tails) and condensate phase (long tails). Inset: Scaling collapse of tails on plotting $LP(M)$ vs. M/L in the condensate phase and $L^{2/3}P(M)$ vs. $(M - \langle M \rangle)/L^{2/3}$ near the critical point.
Parameters used: $a = 1$, $D = 0.75$, $w = 0.25$ (condensate phase); $a = 1$, $D = 0.75$, $w = 2.0$ (normal phase); $a = 1$, $D = 0.75$, $w = 1.5$ (near critical point).

Total mass M: The steady state distribution $P(M)$ of total mass M in the system also has a condensate tail, and behaves as $P(M) \propto 1/M_0 \exp(-M/M_0)$ at large M [fig. 2.3b]. As a result, the rms fluctuations ΔM of the total mass exhibit anomalous behaviour, scaling as L rather than $\sqrt{L}$, so that $\Delta M/\langle M \rangle$ remains finite even as $L \to \infty$. The occurrence of such *giant* fluctuations implies that the system is not self-averaging even in the thermodynamic limit [4].

Position-dependent properties: Equation (2.2) suggests that the spatial co-ordinate that governs the behaviour at a site is not the site number i, but the relative position $x = i/L$. This expectation is also borne out by numerics, which show that at large L, local properties, e.g., moments of the mass at a particular location within the system depend on x and not i, except very close to the left boundary. Thus, in the analysis of position-dependent properties, we will monitor various quantities at a *fixed* relative position x, while varying system size L.

Recall that the heuristic argument in sec. 2.1 predicts that in the condensate phase ($w < w_c$), there is a critical (relative) position $x_c(w)$ corresponding to each chipping rate w, such that for $x < x_c$, the system is locally in the condensate phase, whereas for $x > x_c$, the system is locally in the normal or disordered phase. In order to investigate whether the system exhibits this kind of 'phase coexistence' for $w < w_c$, we monitor the *region-wise largest mass* $m_*^{(L)}[x_i, x_j]$, i.e., the mass of the largest aggregate in the region bound by the relative co-ordinates x_i, x_j in a system of size L. If the region $[x_i,x_j]$ is locally in the condensate phase, then the macroscopic condensate can be found in this region for a finite fraction of the time, so

[4] This would imply that the typical values of M observed at any instant would not in general be close to the ensemble-averaged (or time-averaged) value $\langle M \rangle$, even in the $L \to \infty$ limit.

that $\langle m_*^{(L)}[x_i, x_j]\rangle \propto L$. If the region in question is in the normal phase, then the macroscopic aggregate visits it with vanishing probability, so that $\langle m_*^{(L)}[x_i, x_j]\rangle$ is sub-linear in L.

Region-wise largest mass $m_*^{(L)}[x_i, x_j]$*:* In practice, it is convenient to divide the system into 10 regions, and measure the average value $\langle m_*^{(L)}[x_i, x_j]\rangle$ of the region-wise largest mass in each region, for different L, keeping x_i and x_j fixed. As before, we use two different methods, namely, the ratio method and the intercept method to extract the L-dependence of $\langle m_*^{(L)}[x_i, x_j]\rangle$.

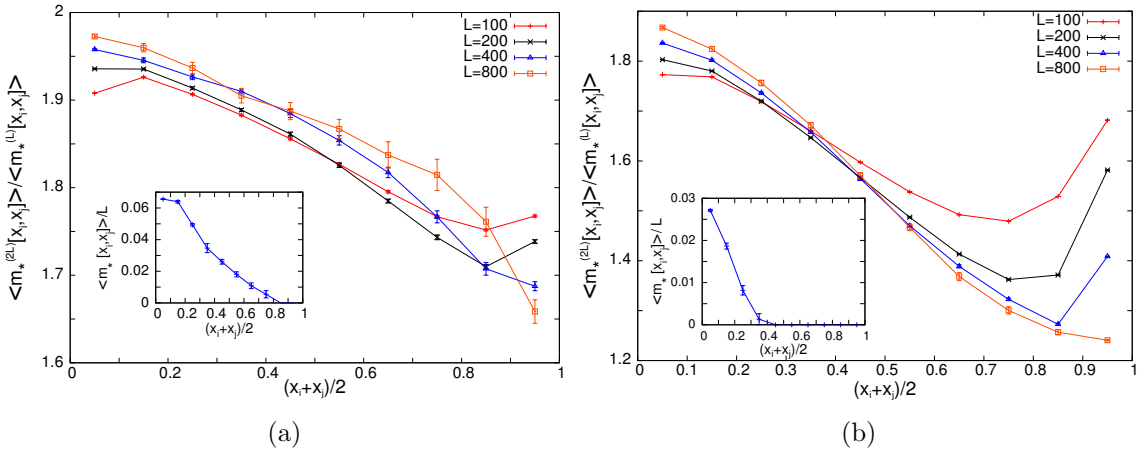


Figure 2.4: Data for average region-wise largest mass $\langle m_*[x_i, x_j]\rangle$ for (a) small w ($a = 1$, $D = 0.75$, $w = 0.625$) (b) intermediate w ($a = 1$, $D = 0.75$, $w = 1.125$).
(a) Small w: The ratios $\langle m_*^{(2L)}[x_i, x_j]\rangle / \langle m_*^{(L)}[x_i, x_j]\rangle$ increase with L for all regions, except for $x > 0.9$. Inset: The asymptotic $L \to \infty$ value of $\langle m_*^{(L)}[x_i, x_j]\rangle / L$ obtained by extrapolation (see text) vs. $(x_i + x_j)/2$. The fraction goes to zero for $x > 0.8$. The evidence for phase co-existence is inconclusive.
(b) intermediate w: The ratios $\langle m_*^{(2L)}[x_i, x_j]\rangle / \langle m_*^{(L)}[x_i, x_j]\rangle$ increase/decrease with L for left/right lying regions. Inset: $\langle m_*^{(L)}[x_i, x_j]\rangle / L$ goes to zero for $x > 0.3$. There is clear evidence for phase co-existence.

Both these methods indicate that at *intermediate* values of w (with $w < w_c$), there is phase coexistence, i.e., there is a position $x_c(w)$, such that $\langle m_*^{(L)}[x_i, x_j]\rangle$ is macroscopic for any $x_j < x_c$, while $\langle m_*^{(L)}[x_i, x_j]\rangle / L \to 0$ as $L \to \infty$ for $x_i > x_c$. This is depicted in figure 2.4b, which shows the ratios $\langle m_*^{(2L)}[x_i, x_j]\rangle / \langle m_*^{(L)}[x_i, x_j]\rangle$ as a function of the region mid-point $(x_i + x_j)/2$ for different L. For left-lying regions, the ratios increase monotonically with L, consistent with an approach to the value 2. For right-lying regions, the ratios decrease with L, and appear to be converging towards a value less than 2. This provides evidence for the existence of condensate-like and normal behaviours in left-lying and right-lying regions of the system respectively. A similar conclusion emerges if we use the intercept method to estimate the fractions $\langle m_*^{(L)}[x_i, x_j]\rangle / L$ for different regions (see inset, fig. 2.4b).

At *small values of* w, the behaviour is less clear as there is some discrepancy between the results obtained using the two methods. The ratio method indicates that condensate-like

behaviour prevails in almost all regions (note the upward trend of the ratio plots in all regions except the last in fig. 2.4a). On the other hand, the fraction $\langle m_*^{(L)}[x_i, x_j]\rangle/L$, as calculated using the intercept method, goes to zero beyond a critical location x_c, which however, is also close to 1 (see inset, fig. 2.4a). Thus, for small w, it becomes difficult to ascertain whether or not there is a *finite* fraction of the system that shows normal behaviour, associated with the absence of the condensate.

In general, due to the presence of strong finite size effects in the system, we expect the intercept method, which is based on the extrapolation of finite L data to the $L \to \infty$ limit, to be more reliable for extraction of the 'true' behaviour of the system in the thermodynamic limit. Another point in favour of the intercept method is that the $x_c(w)$ values it yields are close to those predicted by the analytical argument. Thus we conjecture that the coexistence of condensate-like and normal regions may be occurring at *any non-zero* chipping rate $w < w_c$, but this would manifest itself only in the large L limit, especially for small w. This, however, is a conjecture, and to determine conclusively whether or not there is phase coexistence at small w, it is necessary to simulate very large systems (which we have not been able to do because of the high computation times involved) or find an appropriate analytical description of the system for non-zero w.

Other mass distributions: The condensate tail appears in the probability distribution of mass in any region that is locally in the condensate phase. This is true even if the region consists of just a single site. Thus, the mass distribution $p_x^{(L)}(m)$ at a site with relative position x behaves as $p_x^{(L)}(m) \sim \frac{1}{L} f(x) \left(\frac{1}{M_0} \exp\left(-\frac{m}{M_0}\right)\right)$ for large m (see fig. 2.5). Here, the $1/L$ factor arises simply because the condensate is shared among an extensive number of sites, whereas the function $f(x)$ encodes the different probabilities with which the condensate visits different sites in the system.

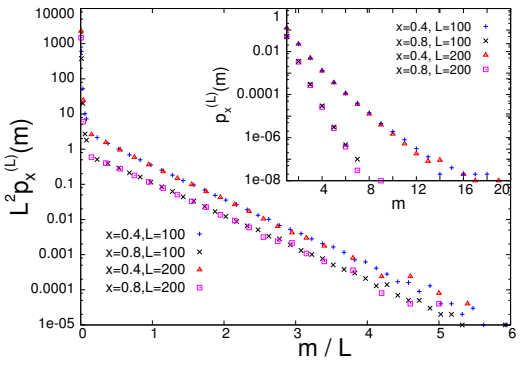


Figure 2.5: Single site mass distributions: In the condensate phase ($a = 1$, $D = 0.75$, $w = 0.25$), $p_x^{(L)}(m)$ shows condensate tails with the same characteristic M_0 for different x. There is good scaling collapse of $L^2 p_x^{(L)}(m)$ vs. m/L for a given x. Inset: In the normal phase ($a = 1$, $D = 0.75$, $w = 3.0$), $p_x^{(L)}(m)$ for a given x is independent of L to leading order and shows no condensate tail.

Another quantity of interest is the distribution $P_{exit}^{left}(m)$ of masses m exiting from the left boundary. The distribution for left-exiting masses also has a condensate tail of the form: $P_{exit}^{left}(m) \sim \frac{1}{L^2} \left(\frac{1}{M_0} \exp\left(-\frac{m}{M_0}\right)\right)$, for large m, for all w less than w_c.

The appearance of this kind of a condensate tail in all these mass distribution reflects the fact that a large fluctuation in any region of the system arises due to the visit of a macroscopic aggregate whose distribution is governed by the mass scale $M_0 \propto L$. The exponential form $\sim (1/M_0) \exp[-m/M_0]$ of these distributions just mirrors the probability distribution of mass

of this macroscopic aggregate. How often a particular region or site is visited by the aggregate is encoded in the non-universal pre-factor of the distributions.

We also monitor the distribution $P_{exit}^{right}(m)$ of right-exiting masses and the single-site mass distribution $p_x^{(L)}(m)$ in regions which are close to the right boundary (and are hence expected to be locally in the normal phase, based on the heuristic argument in sec. 2.2). For small w, the distribution $P_{exit}^{right}(m)$ and the single-site distributions in right lying regions do appear to have a condensate tail, at least for the largest system size ($L = 1600$) used in our simulations. For intermediate values of w (closer to w_c), there is a remnant of such a tail at small L, but it vanishes at large L. As before, it is difficult to establish conclusively whether this is a genuine distinction between small w and large w regimes of the condensate phase, or merely a finite size effect. If there is indeed a finite condensate-free region for any non-zero w (as we have conjectured earlier), then the condensate tail observed in distributions such as $P_{exit}^{right}(m)$ at small w, is likely to be a finite size remnant, which will vanish in the $L \to \infty$ limit.

Giant number fluctuations: A natural consequence of the non-Gaussian mass distributions is that the rms fluctuations of various quantities are anomalously large. This has already been discussed in the context of the total mass M, which shows $\Delta M \propto \langle M \rangle \propto L$. The region-wise mass (for a region in the condensate phase) also shows giant rms fluctuations $\Delta M_{x_i,x_j}^{(L)}$ that are proportional to L rather than $\sqrt{L}$, for fixed x_i, x_j. Further, even for the single site mass m_x, we find $\Delta m_x \propto \sqrt{L}$, implying that $\Delta m_x / \langle m_x \rangle$ actually increases for fixed x as $L \to \infty$.

Several other systems such as vibrated needles [60] and passive particles on fluctuating interfaces [61] also exhibit giant number fluctuations. However, there is an important difference between these systems and the model we study here. In the systems studied in [60], [61], the fluctuations in a region of linear size Δl are proportional to Δl [62], rather than the system size L. By contrast, in our model, the fluctuations at just a single site also show a leading order dependence on L. The emergence of L-dependent fluctuations in our model, or in condensation models in general, is due to the fact that there is no constraint on the number of particles per site, in contrast to the hard core constraints in systems studied in [60, 61].

2.3.1.2 Normal phase

In the *normal (large w) phase*, there is no macroscopic condensate, i.e., $\langle m_*^{(L)} \rangle / L \to 0$ as $L \to \infty$. As a result, the various probability distributions have properties that are very different from those in the condensate phase. These differences are enumerated below.

i. The total mass has normal rms fluctuations $\Delta M \propto \sqrt{\langle M \rangle} \propto \sqrt{L}$, with the distribution for the rescaled mass variable $(M - \langle M \rangle)/\Delta M$, approaching a Gaussian at large L ($P(M)$ shown in fig. 2.3b).

ii. The mass in any macroscopic region of the system is also normally distributed, with $\Delta M^{(L)}[x_i, x_j]$ scaling as $\sqrt{L}$ for fixed x_i and x_j.

iii. The single-site mass distribution $p_x^{(L)}(m)$ depends only on m and the relative position x and not directly on system size L (see inset, fig. 2.5). Thus, in contrast to the

condensate phase, all moments of mass at a fixed relative location x are independent of L.

iv. $P_{exit}(m)$: To leading order in L, the distribution $P_{exit}^{left}(m)$ of left-exiting masses is independent of L, while $P_{exit}^{right}(m) \sim (b/L)^m$ where b is a constant.

In fact, the behaviour of various quantities is qualitatively similar to that in the limiting case $D = 0$, where the above features can be derived analytically from the steady state measure found in [43].

2.3.1.3 Critical Point

The transition from the normal to the condensate phase takes place at a critical chipping rate w_c, which increases with injection rate a if D is held constant (see eq. (2.4) for a theoretical estimate). At this point, the mean largest mass in the system is found to scale as $\langle m_*^{(L)} \rangle \propto \sqrt{L}$. Numerics show that the total mass also exhibits anomalously large fluctuations, with $\Delta M \propto L^{\theta_c}$ where $\theta_c \simeq 2/3$. The mass distribution has a tail of the form: $P(M) \sim \frac{1}{M_2} \exp\left(\frac{-(M-\langle M \rangle)}{M_2}\right)$ where $\langle M \rangle \sim L$ and $M_2 \sim L^{\theta_c}$ [inset, fig. 2.3b].

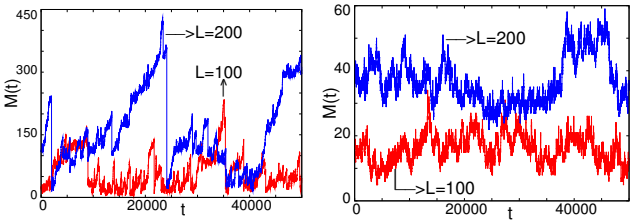
2.3.2 Dynamics

Contrasting signatures of the phases appear in various dynamical properties, and are even evident in a typical time series $M(t)$ of the total mass in the two phases [see fig. 2.6]. In the normal phase, $M(t)$ is self-similar in time, i.e., a small stretch of the time-series (on magnifying and rescaling) has the same statistical properties as a larger stretch. This is no longer true for $M(t)$ in the condensate phase (see fig. 2.6b), due to the occurrence of extreme changes over very short time scales, corresponding to the exit of a condensate with $\mathcal{O}(L)$ mass.

The breakdown of self-similarity can be quantified using dynamical structure functions $S_n(t)$, which may be defined in analogy with structure functions of the velocity field in fluid flow [57]. The dynamical structure functions are defined as follows:

$$S_n(t) = \langle [M(t) - M(0)]^n \rangle \qquad (2.5)$$

Here, $\langle ... \rangle$ denotes average over histories and $t = 0$ is an arbitrary time instant after the system attains stationarity.



(a) (b)

Figure 2.6: A typical realisation of $M(t)$ vs. t for two different system sizes L in the (a) condensate phase (b) normal phase. Parameters same as in fig. 2.5.

For self-similar signals, the structure functions typically behave as $S_n(t) \propto t^{u(n)}$ as $t/\tau \to 0$, where $u(n) \propto n$, and τ is a time scale having to do with the largest structures in the system [5]. A deviation from the linear dependence of the exponents $u(n)$ on n reflects the breakdown

[5]For instance, in our model, τ is proportional to L^2 and is the average time taken by a particle to traverse the system; it is also (upto a multiplicative factor) the typical time interval between exit events involving the condensate, and consequently the auto-correlation time associated with the total mass.

of self-similarity and may occur, for example, if the signal $M(t)$ alternates between periods of quiescence (small or no activity) and bursts (sudden large changes) [57]. Such an alternation is known as intermittency, and is a characteristic feature of turbulent flow, where the spatial structure functions of the velocity field exhibit anomalous dependence on the space variable. Other useful measures of intermittency are the flatness $\kappa(t)$ and hyperflatness $h(t)$, defined respectively as [57]:

$$\kappa(t) = S_4(t)/S_2^2(t) \tag{2.6a}$$

$$h(t) = S_6(t)/S_2^3(t) \tag{2.6b}$$

Note that $\kappa(t)$ is just the kurtosis (upto an additive constant) of the time-dependent probability distribution of $\Delta M(t)$. For intermittent signals, $\kappa(t)$ and $h(t)$ diverge as $t/\tau \to 0$. The behaviour of structure functions in the two phases is discussed below.

2.3.2.1 Condensate Phase

We use numerical simulations to measure the auto-correlation function $C(t) = \langle M(t)M(0)\rangle - \langle M(t)\rangle\langle M(0)\rangle$ and find that it decays as $C(t) \sim L^2\mathcal{C}(t/\tau)$, where the autocorrelation time scale τ is proportional to L^2. We also measure the dynamical structure functions and find that they scale as: $S_n(t) \sim L^n\mathcal{F}_n(t/\tau)$, where $\mathcal{F}_n$ is consistent with the form $\mathcal{F}_n(y) \sim (-1)^n y\mathcal{G}_n[\log(y)]$ for small y, (with $\mathcal{G}_n$ chosen to be a polynomial) and approaches an n-dependent constant value at large y [6]. This demonstrates the intermittent behaviour of $M(t)$ by establishing that for $n \geq 2$, all structure functions $S_n(t)$ scale as t at small t, independently of n, with a weak n-dependence arising only because of the multiplicative $\log(t)$ corrections. This extreme form of anomalous scaling, where the scaling exponent $u(n)$ is actually independent of n, is referred to as strong intermittency. As is evident from the small t expression for $S_n(t)$, the anomalous time scale, which results in breakdown of self-similarity is simply the auto-correlation time scale $\tau \propto L^2$. It follows that $\kappa(t)$ and $h(t)$ diverge at small times in a strongly L dependent way [fig. 2.7c and fig. 2.7d]. In fact, they are functions of t/τ and diverge as $t/\tau \to 0$ [inset, fig. 2.7c and fig. 2.7d].

2.3.2.2 Normal Phase

The time scale τ associated with the decay of the auto-correlation function of $M(t)$ is proportional to L^2 in this phase as well. However, the structure functions are now independent of L for $t \ll \tau$ and scale as $S_{2n} \propto L^n\mathcal{G}(t/\tau)$ for $t \gg \tau$. At small t, the structure functions behave as $S_{2n}(t) \sim t^{u(2n)}$ where $u(n) \sim n/4$ (see fig. 2.7b). The linear dependence of $u(n)$ on n indicates that the time series $M(t)$ [fig. 2.6b] is self-similar. The flatness $\kappa(t)$ and hyperflatness $h(t)$ approach a finite, L-independent value as $t \to 0$ [fig. 2.7c and fig. 2.7d].

2.3.2.3 Critical point w_c:

$M(t)$ continues to show intermittency at the critical point with flatness and hyperflatness diverging as $t \to 0$ in an L-dependent manner. However, there seems to be no simple scaling which collapses the curves for different L.

[6] In the limiting case $w = 0$, these can be calculated analytically, as discussed in Chapter 3. In fact, the $\log t$ terms are suggested by this analytical calculation

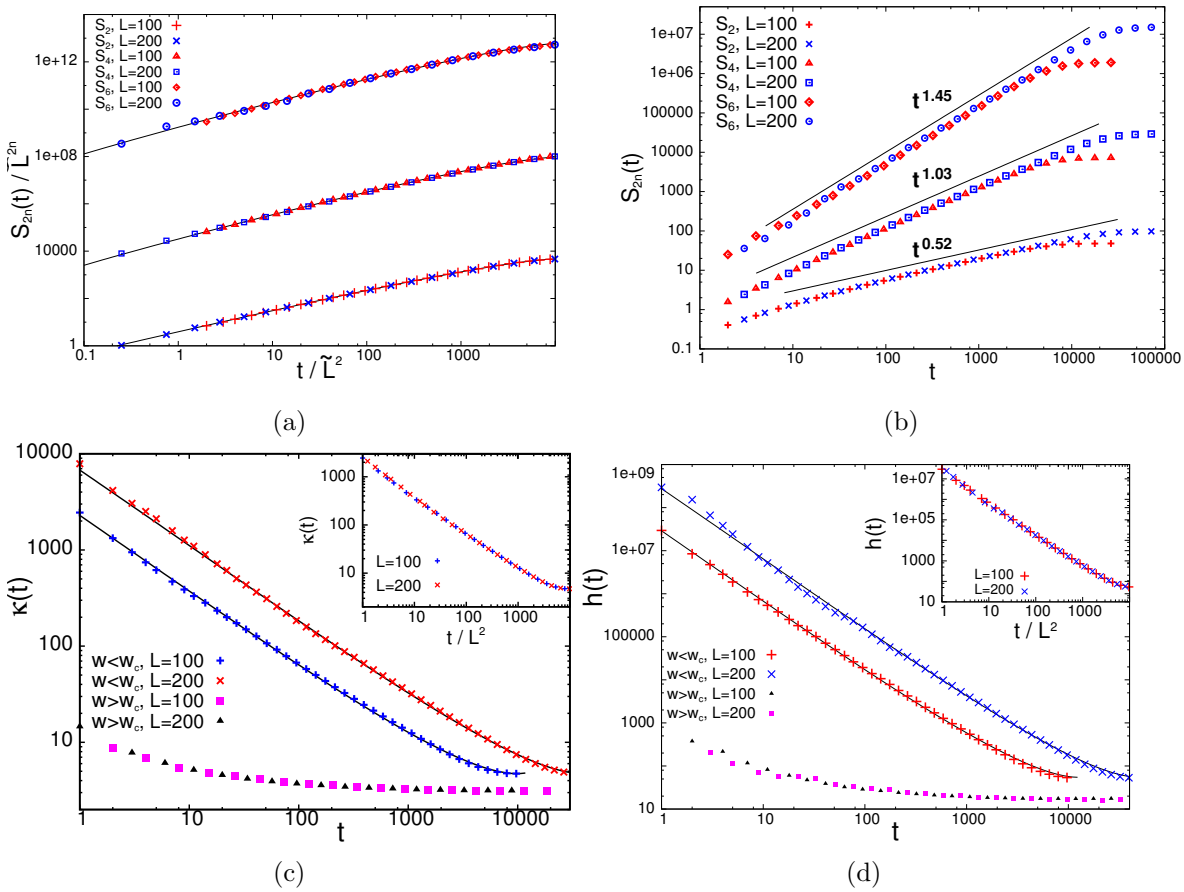


Figure 2.7: (a) Structure functions $S_{2n}(t)$ in the condensate phase for different L: Good data collapse on plotting $S_{2n}(t)/\tilde{L}^{2n}$ vs $t/\tilde{L}^2$ (where the rescaling $\tilde{L} = L/100$ is simply to display all three structure functions in the same plot clearly). Solid lines indicate fits to $y\mathcal{G}_{2n}[log(y)]$ as described in text.
(b) In the normal phase, $S_{2n}(t)$ show no dependence on L and behave as $S_{2n}(t) \sim t^{u(2n)}$ for small t, where $u(n) \sim n/4$.
(c) $\kappa(t)$ vs. t for two different L in the two phases. Solid lines for $w < w_c$ are fits to the form described in the text. Inset: Data collapse of $\kappa(t)$ vs. t in the condensate phase on rescaling time as t/L^2.
(d) $h(t)$ vs. t for two different L in the two phases. Inset: Data collapse of $h(t)$ vs. t in the condensate phase on rescaling time as t/L^2.
Model parameters for condensate and normal phase same as in figs. 2.5 and 2.6

2.4 Model B: Unbiased stack hopping; outflux only from right boundary

In order to investigate the role of boundary conditions, we also study a variant of the unbiased model in which influx of single particles occurs at the left boundary, while outflux of aggregates occurs only at the opposite (right) boundary. For the total mass M in this model to be an extensive quantity, injection rates of the sort $a = \tilde{a}/L$ [where $\tilde{a}$ is $\mathcal{O}(1)$] are required. With such $\mathcal{O}(1/L)$ injection rates, the *average* value of the particle current through the system becomes proportional to $1/L$, and is now the same as the average particle current in model A which was studied in the previous section. However, while in model B, the $\mathcal{O}(1/L)$ current is a consequence of Poissonian injection at rate $\tilde{a}/L$, in model A, it is proportional

to $1/L$ due to the almost complete cancellation of in and out currents at the left boundary over long time intervals spanning many injection and exit events. This suggests that higher statistics of the current may be different in the two cases, leading to some differences between the two models.

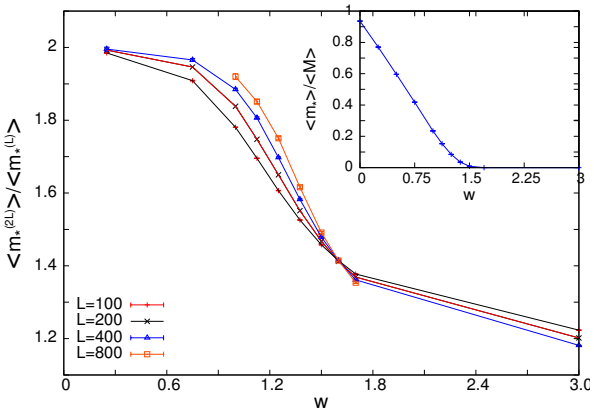


Figure 2.8: Determination of w_c for model B (for $a = 1$, $D = 0.75$) using the ratio method and the intercept method (inset). The point of intersection of the different $\langle m_*^{(2L)}\rangle/\langle m_*^{(L)}\rangle$ vs. w curves gives an estimate for w_c. Inset: The fraction $\langle m_*\rangle/\langle M\rangle$ goes to zero at w_c. Estimated value of w_c using both methods is between 1.5 and 1.625, consistent with the theoretical estimate (see eq. 2.4) of $w_c = 1.5687$.

The heuristic argument presented in sec. 2.2 is valid for this case as well, as the balance of particle currents at each site again yields eq. (2.2) (for $\mathcal{O}(1/L)$ injection rates). Thus, we expect this system to exhibit a phase transition between a condensate and a normal phase. As before, we probe this transition by measuring the largest mass $m_*^{(L)}$ in the system for different w, and analysing its dependence on system size L using both the ratio method and the intercept method. The two methods yield estimates of the critical chipping rate w_c (see fig. 2.8), which are close to each other, as well as the corresponding w_c in model A. For $w < w_c$, the system is in the *condensate phase*, with $\langle m_*^{(L)}\rangle \propto \langle M\rangle \propto L$, as evinced by the upward trend of the ratios for small w in fig. 2.8, and the finite non-zero value of $\langle m_*^{(L)}\rangle/\langle M\rangle$ (inset, fig. 2.8). For $w > w_c$, the system is in the *normal phase*, and contains no macroscopic condensate, so that $\langle m_*^{(L)}\rangle/\langle M\rangle \to 0$. Below we discuss the static and dynamical properties of the phases in more detail, with a special emphasis on the points of difference from model A.

2.4.1 Condensate Phase

We study the full probability distribution of the largest mass m_* using numerics, and find that, just as in model A, it has the non-Gaussian form $P(m_*^{(L)}) \propto (1/\tilde{M}_0)\exp(-m_*^{(L)}/\tilde{M}_0)$ for large $m_*^{(L)}$, where $\tilde{M}_0 \propto L$. Consequently, $m_*^{(L)}$ exhibits giant rms fluctuations, with $\Delta m_*^{(L)} \propto \langle m_*^{(L)}\rangle \propto L$.

We analyse various other properties of the system with the aim of addressing the following two questions. First, does this system also exhibit phase co-existence, i.e., is there a location with relative position $x_c(w)$, such that in the limit $L \to \infty$, the condensate is found with a finite probability in any macroscopic region lying to the left of this location, but with vanishing probability, if the region lies to the right. Second, how does the presence of the condensate affect the properties of the total mass M? More specifically, is the total mass characterised by giant fluctuations and temporal intermittency, as in the case of model A? Recall that in

model A, the occurrence of giant fluctuations and temporal intermittency of M have their origin in the occasional exit of the macroscopic condensate from the system boundaries. In model B, where exit occurs only at the far boundary i.e., from site L, it is reasonable to expect that if the condensate dissipates completely before reaching this boundary, then there will be no exit events involving macroscopic aggregates and hence no giant fluctuations of M or temporal intermittency in $M(t)$ [7]. Thus, phase co-existence militates against giant fluctuations in model B, though not in model A.

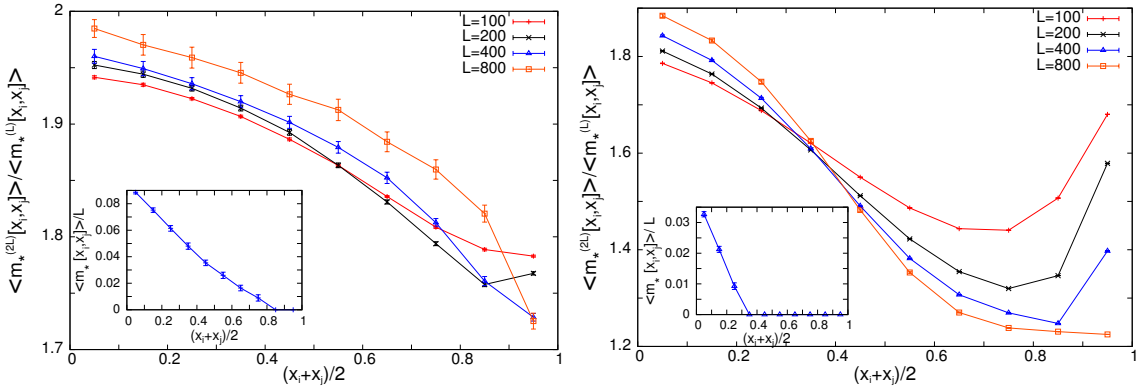


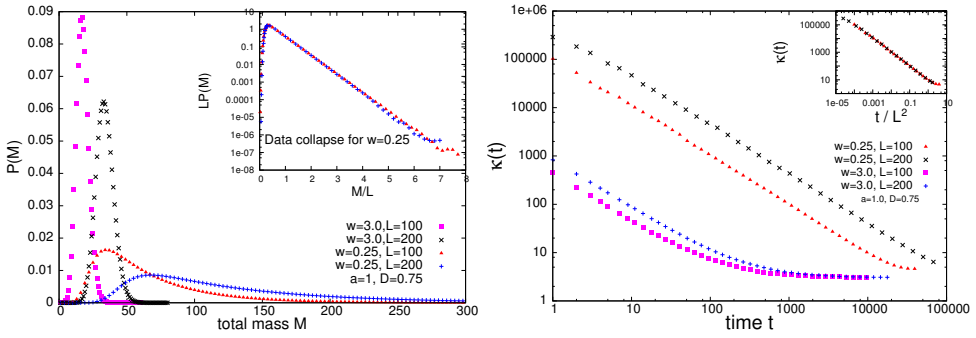
Figure 2.9: Data for average region-wise largest mass $\langle m_*[x_i, x_j]\rangle$ for (a) small w ($a = 1$, $D = 0.75$, $w = 0.5$) (b) intermediate w ($a = 1$, $D = 0.75$, $w = 1.125$).
(a) Small w: The ratios $\langle m_*^{(2L)}[x_i, x_j]\rangle / \langle m_*^{(L)}[x_i, x_j]\rangle$ show an upward trend with L for all regions. Inset: The asymptotic $L \to \infty$ value of $\langle m_*^{(L)}[x_i, x_j]\rangle / L$ goes to zero for $x > 0.8$. The evidence for phase co-existence is mixed and inconclusive.
(b) Intermediate w: The ratios $\langle m_*^{(2L)}[x_i, x_j]\rangle / \langle m_*^{(L)}[x_i, x_j]\rangle$ increase/decrease with L for left/right lying regions. Inset: $\langle m_*^{(L)}[x_i, x_j]\rangle / L$ goes to zero for $x > 0.3$. There is clear evidence for phase co-existence.

In order to probe phase co-existence, we monitor the region-wise largest mass $m_*^{(L)}[x_i, x_j]$ for different values of w and system sizes L. As before, we use the ratio method and the intercept method to analyse the L-dependence of $\langle m_*^{(L)}[x_i, x_j]\rangle$ for different regions $[x_i, x_j]$. The main conclusions of this analysis are similar to those found for model A. At intermediate values of w, within the condensate phase, both the methods suggest that there is phase co-existence, with $\langle m_*^{(L)}[x_i, x_j]\rangle \propto \langle M_{x_i,x_j}^{(L)}\rangle \propto L$ for left-lying regions and $\langle m_*^{(L)}[x_i, x_j]\rangle / \langle M_{x_i,x_j}^{(L)}\rangle \to 0$ for right-lying regions, as $L \to \infty$ (fig. 2.9a). At small values of w, the two methods seem to predict different behaviour– the intercept method provides evidence in support of phase co-existence, while the ratio method suggests that condensate-like behaviour occurs in all regions, including those close to the right boundary (fig. 2.9b). Based on the considerations outlined in sec. 2.3.1.1, it is possible that this discrepancy is a finite size effect, and in the $L \to \infty$ limit, co-existence of condensate-like and condensate-free regions occurs in the system for any non-zero w less than w_c.

[7] This is not the case for model A, where exit also takes place at the left boundary, ensuring that M shows giant fluctuations and temporal intermittency, as long as there is a condensate-bearing region in the system, however small.

We now investigate static and dynamical properties of the total mass M to see if these are consistent with the above conjecture. Numerics show that at small w, up till $L \sim 800$, the probability distribution $P(M)$ of the total mass has a broad condensate tail of the form $P(M) \sim (1/\tilde{M}_0) \exp(-M/\tilde{M}_0)$ where $\tilde{M}_0 \propto L$ (see fig. 2.10a), leading to giant rms fluctuations $\Delta M \propto \langle M \rangle \propto L$. For w that are close to but less than w_c, the distribution $P(M)$ does not appear to have a condensate tail of this sort.

We also probe the self-similarity properties of the time series $M(t)$ using dynamical structure functions $S_n(t)$, as defined in eq. (2.5). For the typical system sizes used in our simulations ($L \sim \mathcal{O}(100)$), the behaviour of $S_n(t)$ for small w, is similar to that found in the $w = 0$ limit (see chapter 3), i.e., $S_n(t) \sim L^n \mathcal{F}_n(t/L^2)$, where $\mathcal{F}_n(t/L^2) \propto t/L^2$, independent of n. Note that unlike in model A, there are no multiplicative $log(t)$ terms in the structure functions [8]. The flatness $\kappa(t)$ also diverges at small t, going as $\kappa(t) \propto (t/L^2)^{-1}$ (see fig. 2.10b). Thus, at small w, the time series $M(t)$ exhibits strong intermittency over time-scales that are proportional to L^2, at least for the relatively small L used in our simulations. At higher values of w that are close to w_c, the structure functions behave differently: $S_n(t)$ again show intermittency but at much shorter $\mathcal{O}(L)$ time scales. This behaviour is characteristic of the structure functions in the normal phase (discussed in sec. 2.4.2).



(a) (b)

Figure 2.10: Different static and dynamical properties of M in the small w ($a = 1$, $D = 0.75$, $w = 0.25$) and large w ($a = 1$, $D = 0.75$, $w = 3.0$) regimes of model B, for relatively small system sizes $L = 100$ and $L = 200$.
(a) $P(M)$ has a long 'condensate' tail for small w and a Gaussian tail for large w. Inset: Data collapse of the condensate tails for small w on plotting $LP(M)$ vs. M/L.
(b) Flatness $\kappa(t)$ as a function of t: Divergence of $\kappa(t)$ on $\mathcal{O}(L^2)$ time scales for small w and on $\mathcal{O}(L)$ time scales for large w. Inset: Data collapse for small w on plotting $\kappa(t)$ vs. t/L^2.

The numerical results for total mass M, as discussed above, lend themselves to two alternative hypotheses. The first possibility is that in the limit $L \to \infty$, the total mass is characterised by Gaussian rms fluctuations and temporal intermittency on $\mathcal{O}(L)$ time-scales, *at any non-zero* w, but this behaviour becomes evident only beyond some crossover length scale $L_*(w)$. It is possible that the system sizes we consider are much smaller than L_*, so that the behaviour of M that is observed at these sizes is effectively the $w = 0$ behaviour.

[8] This difference between the behaviour of $S_n(t)$ in models A and B can be demonstrated analytically for the $w = 0$ case (see Chapter 3).

This hypothesis is based on the analysis of the region-wise largest mass data using the intercept method, which suggests that the region near the right boundary is locally in the normal phase, even for small w, in the $L \to \infty$ limit. The second possibility is that there is indeed no phase coexistence at small w, and giant fluctuations and temporal intermittency are genuine and not merely finite size effects. This hypothesis is consistent with our numerical data for M, where there is no visible trend towards $\Delta M \propto \sqrt{L}$ as L is increased, for small values of w. Ascertaining the true asymptotic behaviour of $P(M)$ and $S_n(t)$ in the $L \to \infty$ limit has thus not been possible using the existing numerical data. Nevertheless, for the relatively small sizes ($100 \leq L \leq 1600$) considered in the simulations, the system mass M does show giant fluctuations and temporal intermittency. Even if this is just a finite size effect, it may be of some relevance to real systems which are typically characterised by a finite L.

2.4.2 Normal Phase

For $w > w_c$, the quantity $\langle m_*^{(L)} \rangle / \langle M \rangle$ goes to zero in the large L limit, implying that there is no condensate formation. The behaviour of steady state distributions is similar to that found for model A. For instance, the total mass shows normal rms fluctuations ΔM that scale as $\sqrt{L}$, and the distribution of $(M - \langle M \rangle)/\Delta M$ approaches a Gaussian form at large L (see fig. 2.10a for $P(M)$).

The dynamical properties of $M(t)$ are, however, somewhat different in this phase as compared to model A. As in model A, the 'long' time scale τ in the model is associated with the autocorrelation of total mass, and is proportional to L^2. However, in this model, there is another time scale $\tilde{t}$, which is simply the average time between successive changes in $M(t)$. It is proportional to L due to our choice of $\mathcal{O}(1/L)$ injection rates, which were needed to ensure extensivity of total mass. Numerics show that the structure functions behave differently over these two time scales: at short times $t \ll \tilde{t}$, we have $S_n(t) \propto t/\tilde{t}$, while at larger times $\tilde{t} \ll t \ll \tau$, the structure functions follow $S_n(t) \sim (t/\tilde{t})^{\gamma n}$ where $\gamma \approx 1/2$. Thus, in model B, the time series $M(t)$ is intermittent over the short time scale $\tilde{t} \propto L$ (see fig. 2.10b), but shows self-similarity on the longer time scale $\tau \propto L^2$. The intermittency over short time scales is simply a consequence of the $1/L$ injection rates.

2.4.3 Critical Point

The phase transition takes place at a critical chipping rate w_c, which is very close to the corresponding chipping rate for model A (if a/D in model A and $\tilde{a}/D$ in model B are equal). At the critical point, the mean largest mass in the system scales as $\langle m_*^L \rangle \propto \sqrt{L}$. Other properties of the system are the same as in the normal phase– the total mass shows Gaussian $\mathcal{O}(\sqrt{L})$ fluctuation, and $M(t)$ is intermittent on $\mathcal{O}(L)$ time scales due to the $1/L$ injection rates.

2.5 Model C: Biased stack hopping; outflux from right boundary

To investigate the role of bias in this model, we consider the case with fully asymmetric stack movement ($D' = 0$), but symmetric chipping (and movement) of single particles. Numerics

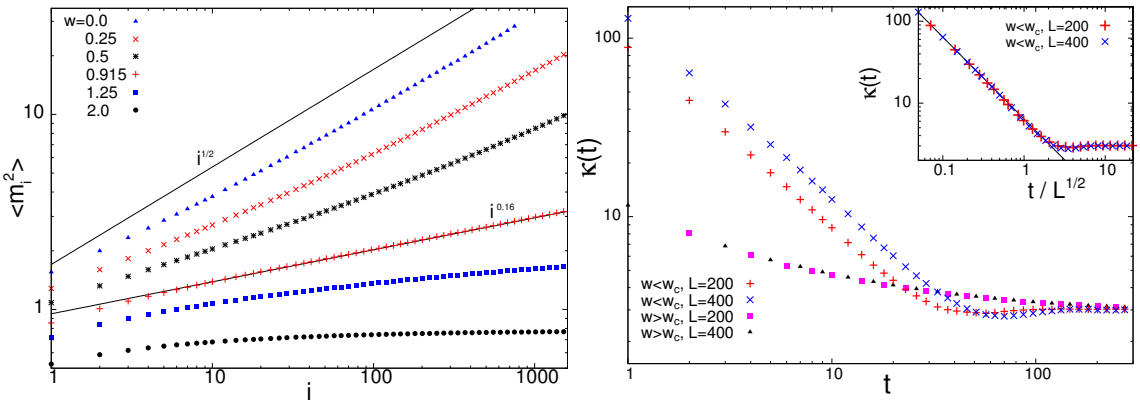
show that this system also undergoes a phase transition from an aggregation-dominated phase to a normal phase. The aggregation-dominated phase is characterised by formation and exit of $\mathcal{O}(\sqrt{L})$ mass aggregates, while no such L^α-sized aggregates form in the normal phase. This phase transition does not survive if directional bias is introduced in chipping; thus, we only consider directionally symmetric chipping.

Before studying the two phases in detail, it is useful to understand the above statements in terms of the interplay of various time scales in the model. The argument that follows is essentially the one introduced in [47] for the closed periodic case. The typical time τ_{trav} taken by an aggregate to traverse the whole system from the left to the right is proportional to L, when motion of aggregates is biased. The typical time $\tau_{form}(m)$ taken for an aggregate of mass m to form due to movement and aggregation is proportional to m^2, for both biased and unbiased stack movement. The typical time $\tau_{diss}(m)$ taken for an aggregate of mass m to dissipate by chipping alone is proportional to m^2 when chipping occurs symmetrically in both directions, and to $m^{3/2}$ for chipping with a forward bias [9]. From the above, we see that if chipping is asymmetric, then τ_{diss} is always smaller than τ_{form}, ensuring that chipping always wins over aggregation, so that no large aggregates can form. However, if chipping is symmetric, then the two time scales $\tau_{diss}(m)$ and $\tau_{form}(m)$ have the same dependence on m, leading to a competition between chipping and aggregation, and eventually to a phase transition between an aggregation-dominated phase and a chipping-dominated phase at a *non-zero* chipping rate. Further, since the traversal time of a particle is proportional to L, it follows that in the aggregation-dominated phase, $\tau_{form}(m) \propto L$, so that the typical mass m of an aggregate exiting the system is proportional to $\sqrt{L}$, rather than L as in the unbiased models.

We now discuss the distinguishing features of the aggregation-dominated and the chipping-dominated phases. Numerics show that in the aggregation-dominated phase (at small w), the typical mass of an aggregate at a distance i from the point of injection, scales as $\sqrt{i}$ (as in the limiting case $w = 0$, where this can be established analytically [32]). In the chipping-dominated phase, on the other hand, there is no i-dependence of the aggregates at sufficiently large i. This distinction is captured in the behaviour of the second moment $\langle m_i^2 \rangle$ of the site-wise mass [fig. 5(a)]. For large w, the plots of $\langle m_i^2 \rangle$ vs. i show a downward bend on the log-log plot and approach a constant value at large i, while for small w, the $\langle m_i^2 \rangle$ vs. i curves bend upwards, consistent with an approach to $\langle m_i^2 \rangle \propto \sqrt{i}$ at large i. The transition takes place at w_c (corresponding to the curve with no bending on the log-log plot), at which $\langle m_i^2 \rangle \sim i^{\alpha_c}$ with $\alpha_c \simeq 0.16$.

Distinctive signatures of the two phases also appear in the properties of the total mass M. Numerical simulations show that in both the phases, the total mass has normal rms fluctuations $\Delta M \propto \sqrt{L}$, with the probability distribution of the rescaled mass, $(M - \langle M \rangle)/\Delta M$, approaching a Gaussian form at large L. However, in the aggregation-dominated phase, this is a consequence of sudden $\mathcal{O}(\sqrt{L})$ changes occurring sporadically at $\mathcal{O}(\sqrt{L})$ time separations, whereas in the normal phase, it emerges from the addition of uncorrelated or weakly correlated

[9] This follows from the exact mapping between unbiased (biased) chipping and the symmetric (asymmetric) exclusion process. The dynamical exponents for the exclusion processes are known and are given by 2 and 3/2 respectively for the symmetric and asymmetric case. This yields the corresponding time scales $\tau_{diss}(m)$.



(a) (b)

Figure 2.11: (a) $\langle m^2(x)\rangle$ vs. x for $a = 1$, $D = 1.5$ and different w. Note the upward (downward) bending of curves on log-log plot in aggregation (chipping) dominated phase. There is no bending at w_c.
(b) $\kappa(t)$ vs. t for two different L in the aggregation-dominated and normal phases. Inset: Data collapse of $\kappa(t)$ vs. t for different L on scaling time as $t/\sqrt{L}$ in the aggregation-dominated phase. The divergence at small t roughly follows $\sim (t/\sqrt{L})^{-1}$ (solid line).

$\mathcal{O}(1)$ variations in M at every time step. This distinction is captured in the dynamical properties of $M(t)$: in the aggregate phase, the total mass M shows intermittency on $\mathcal{O}(\sqrt{L})$ time scales, corresponding to a typical time interval of $\mathcal{O}(\sqrt{L})$ duration between large exit events. Flatness and hyperflatness are functions of $t/\sqrt{L}$ and diverge as power laws as $t/\sqrt{L} \to 0$ [see fig. 2.11b]. By contrast, for large w, the time series $M(t)$ is not intermittent, and there is no L-dependent increase in flatness as $t \to 0$ [fig. 2.11b]. Thus, in this model, it is the dynamical rather than static properties of M that show imprints of aggregate formation.

2.6 Conclusions

In this chapter, we have studied the open boundary aggregation-chipping model with influx of single particles and outflux of aggregates at the boundaries. The model was studied for both unbiased and directionally biased movement of aggregates as well as different boundary conditions. The principal result of this part of the thesis is that in spite of non-conservation of the total mass, the open-boundary system (with unbiased stack movement) undergoes a *phase transition to a condensate phase*, which is quite different from the condensate phase of the closed periodic aggregation-chipping model [44, 45].

The condensate phase of the open-boundary system is characterised by the formation of a macroscopic aggregate or condensate whose mass shows giant fluctuations, and has a long-tailed distribution. Further, at least for intermediate values of w, this condensate appears to be confined to a finite left-lying fraction of the system. The system thus exhibits *co-existence* of a condensate-like and normal behaviours in different regions. However, using the existing numerical data, we have not been able to resolve whether this kind of phase co-existence occurs at all non-zero w less than w_c, or only for intermediate values of w. In general, the system shows strong finite size effects, making it difficult to ascertain the true asymptotic behaviour

in the thermodynamic limit using numerical simulations alone. An analytical treatment of this model would thus be helpful in shedding more light on this issue.

An interesting feature of the open model is that global quantities such as the total mass also show widely different properties in the two phases. In particular, in model A where exit of aggregates occurs at both boundaries, the total mass exhibits *anomalously large steady state fluctuations* and has a non-Gaussian probability distribution whose width is proportional to system size L. Another distinctive feature of the condensate phase is the non-self-similar time evolution of the total mass M. In fact, we find that the time series $M(t)$ of the total mass shows *intermittency*, which we quantify using dynamical structure functions, as used in the study of turbulence. Turbulence-like features of aggregation have been studied earlier in [31], where multi-scaling of n-point mass-mass correlation functions was demonstrated for an aggregation model. Our work focuses on a different connection with turbulence, which is related to *temporal* intermittency of mass and multi-scaling of dynamical structure functions.

We also study the case with biased movement of stacks, and find that on decreasing w, this model undergoes a related phase transition to an aggregation-dominated phase, which is characterised by formation and exit of sub-macroscopic ($\sqrt{L}$-sized) aggregates. Intermittency of total mass is an important signature of the aggregation-dominated phase in this case as well.

A broad conclusion that emerges from this study is that cluster or aggregate formation can have a strong imprint on dynamical properties of a system, and may cause breakdown of self-similarity. Using dynamical structure functions to analyse the self-similarity or its breakdown, for various temporal signals in systems with clustering, could thus be an interesting direction for future study.

Chapter 3

Analytical Study of the Only-Aggregation Limit

3.1 Introduction

In this chapter, we focus on the zero fragmentation or only-aggregation limit of the model discussed in the previous chapter. This special limit of the model involves *influx* of single particles at the first site at rate a, *movement of full stacks* with rate D (or D') to the site to the right (or left) and *aggregation* with other masses on contact, and *outflux* of aggregates with rate βD (or $\alpha D'$) from the right (or left) boundary. This work is also described in [63].

The only-aggregation limit is an instructive limit to study, as it exhibits most of the distinctive static and dynamical properties that characterise the full aggregation-dominated phase of the general model (see Chapter 2), but is also simple enough to be analytically tractable. In particular, in this limit, it becomes possible to explicitly derive the probability distribution and dynamical structure function of the total mass for both driven and diffusive movement of aggregates and different boundary conditions. These calculations establish that the total mass has a non-Gaussian probability distribution and exhibits giant fluctuations, in the case where aggregate motion is diffusive. Further, the calculation of structure functions and their scaling exponents, explicitly demonstrates that the total mass shows strong intermittency in time, for both driven and diffusive movement of aggregates.

Intermittency, which is a feature of many turbulent systems, is difficult to study analytically in contexts such as fluid turbulence. This model provides a simple example of a system where an analytical calculation of the anomalous scaling exponents of structure functions becomes possible. Moreover, unlike other turbulent systems such as the Burgers fluid [64], the structure functions considered here are *temporal* structure functions and are related to various dynamical correlation functions of the total mass in the system. Calculation of dynamical quantities in such models can be somewhat involved and has been done earlier only for the case of aggregation with spatially uniform injection [29, 30].

The analytical approach used in these calculations is based on a 'closure' property: the probability distributions of mass $M_{i,j}$ in any single continuous region $[i+1, j]$ of the system satisfy recursion relations that involve *only single-region* probability distributions. This closed

set of equations can be solved to obtain the distribution or equivalently, various moments of $M_{i,j}$. The distribution of the total mass in the system can then be recovered as a special case of this. The closure property also forms the basis of the calculation of various spatio-temporal correlation functions of mass, from which structure functions can be obtained. The only approximation involved in these calculations is the use of a spatial continuum limit for the diffusive model, so that recursion relations involving discrete lattice sites become partial differential equations in a continuous spatial variable.

This sort of closure property is a generic feature of aggregation/annihilation models in *one-dimension*. It was first utilised by Doering et al in [65] to calculate the probability of there being no particle in a particular spatial interval or stretch; these empty interval probabilities were then used to derive various other quantities such as the inter-particle distribution functions and particle concentration. This method was extended in by Takayasu et al in [25] to calculate the distribution of mass contained in an arbitrary spatial interval, and consequently the mass distribution at a single site. A similar analysis was carried out by Majumdar et al in [29] to obtain the time-dependent mass distribution for the Takayasu model. Our treatment is very close to the analysis of Takayasu et al, except that we consider a spatially inhomogeneous system. More importantly, we extend the analysis to also calculate various spatio-temporal correlation functions of mass, from which dynamical structure-functions can then be obtained.

A limitation of this approach is that it can be used only for one-dimensional systems where lattice sites have a natural ordering. More generally, while several techniques for finding exact solutions of such models exist in 1D [25, 27, 65], they cannot be generalised to higher dimensions [1] Another limitation is that this method does not work for mass-dependent diffusion of aggregates or generalisations of the model which also allow for fragmentation of particles; the closure property of the equations for $M_{i,j}$ is lost in both these cases.

The rest of this chapter is devoted to the actual calculation of various static and dynamical properties of the system for the three cases discussed in the previous chapter, namely:

A. Diffusive movement, outflux only from right boundary: $D = D'$ and $\alpha = 0, \beta = 1$.

B. Diffusive movement, outflux from both boundaries: $D = D'$ and $\alpha = 1, \beta = 1$. [2]

C. Driven movement, outflux from right boundary: $D' = 0$ and $\beta = 1$.

In section 3.2, we give a brief statement of the results derived for the three cases, and place them in the context of related aggregation models. Sections 3.3, 3.4 and 3.5 present a detailed analysis of the three cases (A), (B) and (C) respectively. Section 3.6 contains some concluding remarks. The appendices are devoted to some mathematical details: in appendix 3.7, we derive asymptotic expressions of various quantities; the solutions of various partial differential equations are discussed in appendix 3.8.

[1] The only studies in higher dimensions that we are aware of, analyse reaction-diffusion models by studying the corresponding field theory using renormalisation group techniques [31, 66, 67].

[2] Note that unlike in the previous chapter, in this chapter, we study the case with outflux only from the right boundary first as it is easier to analyse in the absence of fragmentation.

3.2 Main Results and their Relevance to Other Aggregation Models

3.2.1 A brief statement of the main results

For each of the three cases (A)-(C), we calculate the following static and dynamical quantities:

1. The probability distribution $P(M)$ of the total mass M in the system in steady state. Equivalently, one can specify the distribution by the moments $\langle M^n \rangle$.

2. Dynamical structure functions:

$$S_n(t) = \langle [M(t) - M(0)]^n \rangle \tag{3.1}$$

which monitor time correlations of the total mass $M(t)$ at time t. Here $t = 0$ is any arbitrary time instant after the system has attained steady state.

As discussed in Chapter 2, the motivation for analysing structure functions comes from studies of turbulent systems where they are used to quantify the breakdown of self-similarity due to intermittency. For self-similar signals, $S_n(t)$ scale as $S_n(t) \sim t^{u(n)}$ as $t/\tau \to 0$, where $u(n) \propto n$ and τ is a time scale characterising the lifetime of the largest structures in the system; structure functions of intermittent signals deviate from this scaling form. From these, we can also derive the flatness $\kappa(t)$, which is defined as [57]:

$$\kappa(t) = S_4(t)/S_2^2(t) \tag{3.2}$$

We now summarise the main results derived in this chapter for these static and dynamical properties of total mass M in the three cases (A)-(C):

(A). *Diffusive movement, outflux only from right boundary*:
Statics: For injection rates that scale as $1/L$ with system size L, we explicitly calculate the distribution $P(M)$ and show that it has a non-Gaussian tail of the form $P(M) \sim (1/M_0)\exp(-M/M_0)$ where $M_0 \propto L$. In keeping with this, the total mass has 'giant' root mean square (rms) fluctuations ΔM that scale as $\Delta M \propto L$, implying that $\Delta M/\langle M \rangle$ is finite even as $L \to \infty$. Further, the mass in any part of the system also shows giant rms fluctuations and a similar non-Gaussian tailed distribution with characteristic scale M_0.
Dynamics: We calculate the structure functions $S_n(t)$ [defined in eq. (3.1)] for $n = 2, 3, 4$ and find that they obey the scaling form: $S_n(t) = L^n \mathcal{F}_n(t/L^2)$. The functions $\mathcal{F}_n(x)$ are proportional to x for $x \ll 1$ independently of n, thus exhibiting the extreme anomalous scaling associated with strong intermittency. The flatness [as defined in eq. (3.2)] also diverges as $t \to 0$, with the divergence becoming stronger for larger L.

(B). *Diffusive movement, outflux from both boundaries:*
Statics: In this case, we are not able to calculate the full distribution, but instead obtain various moments of the total mass and also of mass in any fraction of the lattice. This model, with $\mathcal{O}(1)$ injection rates, shows the same scaling of mass moments with L as in case (A): in particular, the rms fluctuations are giant, scaling as L rather than $\sqrt{L}$ with system size.

Dynamics: We explicitly calculate $S_2(t)$ and show that in the limit $t \to 0$, it behaves as $S_2(t) \propto -L^2[\frac{t}{L^2}\log\left(\frac{A_1 Dt}{L^2}\right)]$ (where A_1 is a numerical constant). In this case also, structure functions show the anomalous time-dependence associated with intermittency. However, their functional form for small t is different from that in case (A) due to the presence of additional multiplicative log terms.

(C). *Driven movement, outflux from right boundary:*
Statics: In this case, the total mass has a Gaussian distribution and rms fluctuations that scale as $\Delta M \propto \sqrt{L}$.
Dynamics: We calculate dynamical structure functions $S_n(t)$ for $n = 2, 3, 4$ and find that at short times, $t \ll \sqrt{L}$, they scale as $S_n(t) \sim L^{n/2}\mathcal{G}_n\left(\frac{Dt}{\sqrt{L}}\right)$ where $\mathcal{G}_n(x) \sim x$ as $x \to 0$. Thus, this system also shows strong intermittency, but at much shorter $[\mathcal{O}(\sqrt{L})]$ time scales than the diffusive systems.

The calculations for cases (A) and (B) are performed in the continuum approximation, wherein we replace recursion relations involving discrete lattice sites with differential equations in a continuous spatial coordinate and retain only the lowest order derivatives in the differential equations. We check that this approximation is self-consistent, and also find that the resulting expressions for various quantities are in excellent agreement with numerics for the discrete lattice. In case (C), where the continuum approximation is not good, we work with exact recursion relations.

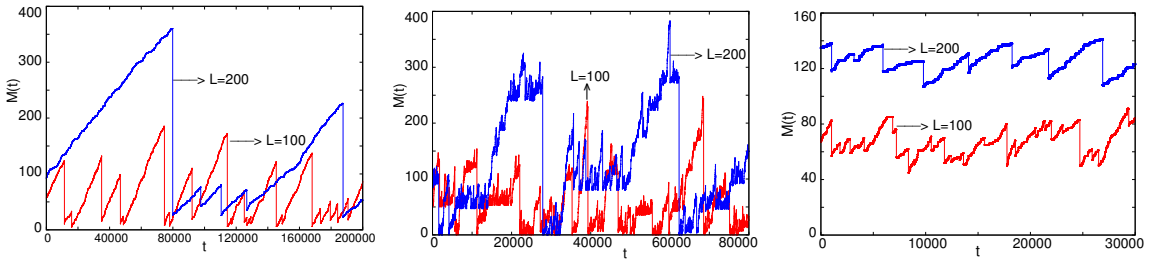


Figure 3.1: A typical realisation of $M(t)$ vs. t for two different system sizes L for (a) Case A: diffusive movement; outflux only from right boundary (b) Case B: diffusive movement; outflux from both boundaries (c) Case C: driven movement; outflux from right boundary. Sharp drops in $M(t)$ dominate the small t behaviour of structure functions in all three cases, giving rise to temporal intermittency.

While the calculations presented in sections 3.3-3.5 are somewhat involved, the physical origin of some of these anomalous features can be understood quite simply. Both large fluctuations and intermittency of the total mass have their origin in the exit of large aggregates from the system. The typical size of the exiting aggregates sets the scale of the rms fluctuations of the total mass: in the diffusive models (A) and (B), these aggregates can be macroscopic, with average mass scaling as system size L, thus giving rise to $\mathcal{O}(L)$ fluctuations of the total mass. In the driven case (C), exiting aggregates typically have $\mathcal{O}(\sqrt{L})$ mass leading to $\sqrt{L}$ fluctuations of the total mass. Temporal intermittency is also a consequence of the sudden changes in $M(t)$ due to the exit of a large aggregate. This can be seen clearly in figure 3.1 which shows a typical time series $M(t)$ for each of the three cases (A)-(C). The intervals

between large exit events sets the characteristic time scale τ which appears in the scaling of the structure functions. For the diffusive systems, this is $\mathcal{O}(L^2)$ while for driven systems, it is $\mathcal{O}(\sqrt{L})$. For t much smaller than this characteristic time τ, higher moments of $M(t) - M(0)$ are dominated by the large crashes in fig. 3.1, giving rise to the anomalous scaling of structure functions at these time scales.

3.2.2 Related analytical work

Theoretical studies of aggregation span a variety of models which may be used to address different sorts of questions. The simplest description of aggregation dynamics is provided by the mean field Smoluchowski equation which describes the time evolution of the mass distribution of aggregates, ignoring spatial distributions altogether (see [28]). It provides a useful framework for studying aggregation phenomena in high dimensions. However, in low dimensions, it is important to explicitly account for the movement of aggregates in space as fluctuations are strong, leading to non mean field behaviour. In many studies, aggregates are assumed to be point objects and the main focus is on studying the kinetics of the aggregation process and distribution of aggregate masses, both with [23, 24, 25, 26, 65, 67] and without influx [27, 68, 69]. In some of these studies, aggregation is treated as an $A + A \to A$ reaction [23, 24, 65, 68] while in others, the mass of the reacting aggregates is also taken into account [25, 26, 27, 69]. Below we present a more detailed discussion of some of these models and their relation to the present work.

(i) Takayasu et al studied diffusion and aggregation with spatially uniform injection of particles [25] and found that at large times, the interplay of injection and aggregation gives rise to a quasi-stationary state in which the mean mass at any site grows indefinitely, while the mass distribution of aggregates approaches a stationary form with a power law decay. In 1D, they calculated the power law exponent exactly and found that it is different from the mean field value. This quasi-stationary phase survives even if the model is generalised to allow for evaporation of single particles from aggregates, as long as the evaporation rate is less than a critical value [44, 70]. The model we analyse in this chapter is quite different in that it has *spatially localised* influx and outflux and constant average mass, but it too shows a broad-tailed distribution of aggregate sizes. However, this distribution is not scale free; instead it has a characteristic width which diverges with system size L.

(ii) Models of aggregation with localised injection have been studied for finite and infinite systems, and different sorts of boundaries [23, 24, 26]. These models lack translational invariance and are characterised by position-dependent distributions of mass. For example, Cheng et al considered diffusing and aggregating masses in an infinite system with a source of particles at the origin and explicitly calculated the steady state distribution of aggregate size as a function of distance from the source [26]. Further, they argued that in the presence of a sink, the distribution of masses exiting the system has an exponential tail which depends on system size. The present work also considers aggregation on a finite system with a source and a sink and finds that this exponential tail appears in the mass distribution at *any location*, however far from the sink.

(iii) In models with localised injection of particles, bias in movement of aggregates can lead to behaviour which is qualitatively different from the unbiased case. This behaviour was

mentioned in [26], and studied in detail by Jain et al in [32] and Reuveni et al in [33], where it was referred to as the Asymmetric Inclusion Process. These studies analyse various mass distributions of the system and show that their spatial dependence is quite different from that in the unbiased case. Here, we have considered this model again (see sec. 3.5), but have gone beyond earlier treatments by calculating time-dependent properties and demonstrating temporal intermittency.

(iv) Some studies have investigated the effect of outflux of large masses on aggregation by introducing models where aggregates with mass higher than some cutoff value are removed from the system [71, 72, 73]. These systems typically show non-monotonic and, in some cases, even oscillatory time evolution. By contrast, the model we study in this chapter does not have a fixed cutoff scale. Instead, the mass of the exiting clusters follows a broad distribution with a characteristic mass scale that emerges from the interplay of the time taken to form an aggregate and the typical residence time of an aggregate in the system.

(v) Connaughton et al demonstrated that the Takayasu model with aggregation and injection shows turbulence-like behaviour in the sense of multi-scaling of n-point mass-mass correlation functions [31]. They calculated these n-point functions using field theoretic techniques and found that for $d \leq 2$, these have an anomalous dependence on mass that deviates from the prediction of a Kolmogorov-like, self-similarity hypothesis. The turbulence-like behaviour that we observe in the present work is somewhat different, being related to the temporal intermittency of the total mass. It may, however, be interesting to explore whether there is a connection between these two characterisations of turbulent behaviour in the context of aggregation.

3.3 Case A: Diffusive movement; outflux only from right boundary

In this section, we analyse the case where single particles are injected onto site 1 at rate a, aggregates diffuse to the left or right with symmetric rates D, and aggregates exit from the opposite end i.e. site L, also with rate D.

As discussed in the previous chapter, if injection rate a and diffusion rate D are both $\mathcal{O}(1)$, then the mean value of the total mass in the system is super-extensive. To ensure extensivity of total mass, we need to consider rates of the sort $a = \tilde{a}/L$, where $\tilde{a}$ is $\mathcal{O}(1)$, so that the injection rate scales as the inverse of the system size. With such injection rates, the total mass becomes an extensive quantity and the average particle current $D[\langle m_i \rangle - \langle m_{i+1} \rangle]$ in the system becomes proportional to $1/L$. It also turns out that the continuum approximation needed for the analysis of this case is self-consistent for $\mathcal{O}(1/L)$ injection rates, but not $\mathcal{O}(1)$ rates.

This section is organised as follows: In sec. 3.3.1, we calculate the steady state probability distribution of mass in any region of the system and recover the total mass distribution and single site mass distributions as special cases of this. This calculation establishes that all these distributions have a broad, exponential tail with the same characteristic width,

which is proportional to system size L. Section 3.3.2 deals with the calculation of dynamical quantities, specifically temporal structure functions $S_n(t)$ of the total mass. Structure functions are obtained by finding correlation functions which relate the total mass in the system at a particular time instant to the mass in any sub-part of the system at another instant. This calculation demonstrates that structure functions show strongly intermittent behaviour, $S_n(t) \propto t$, at short times, *independent of* n.

3.3.1 Statics

To derive the probability distribution of the total mass M, we follow the approach used by Takayasu et al [25] and write the evolution equations for:

$$M_{i,j} = \sum_{l=i+1}^{j} m_l \qquad \text{and} \quad P_{i,j}(M,t) = \text{Prob}[M_{i,j} = M \text{ at time } t] \tag{3.3}$$

The aim is to solve for $P_{i,j}(M)$, which is the steady state probability distribution of mass in the region $[i+1,\, j]$. The distribution $P(M)$ of total mass M can be recovered from the solution as the special case $i = 0, j = L$, while the mass distribution $p_k(m_k)$ at the k^{th} site corresponds to the case $i = k-1, j = k$.

The evolution equations for the time-dependent probability distribution $P_{i,j}(M,t)$ are as follows:

$$\frac{\partial P_{i,j}(M,t)}{\partial t} = D\{P_{i,j+1}(M,t) + P_{i,j-1}(M,t) + P_{i-1,j}(M,t) + P_{i+1,j}(M,t) - 4P_{i,j}(M,t)\} \qquad i > 0, j > i+1, j < L \tag{3.4a}$$

$$\frac{\partial P_{0,j}(M,t)}{\partial t} = a\{[1-\delta_{M,0}]P_{0.j}(M-1,t) - P_{0,j}(M,t)\} + D\{P_{0,j+1}(M,t) + P_{0,j-1}(M,t) - 2P_{0,j}(M,t)\} \qquad i = 0, j > i+1, j < L \tag{3.4b}$$

$$\frac{\partial P_{i,L}(M,t)}{\partial t} = D\{P_{i,L-1}(M,t) + P_{i-1,L}(M,t) + P_{i+1,L}(M,t) - 3P_{i,L}(M,t)\} \qquad i > 0, j > i+1, j = L \tag{3.4c}$$

$$\frac{\partial P_{i,i+1}(M,t)}{\partial t} = D\{2\delta_{M,0} + P_{i,i+2}(M,t) + P_{i-1,i+1}(M,t) - 4P_{i,i+1}(M,t)\} \qquad i > 0, j = i+1, j < L \tag{3.4d}$$

Equation (3.4a) reflects how the mass in the region $[i+1,\, j]$ evolves by exchange of aggregates between sites i and $i+1$ at one end and j and $j+1$ at the other. If this region includes the first site [eq. (3.4b)], then the mass can also change due to injection of single particles. Equation (3.4c) takes into account the effect of the sink next to site L and eq. (3.4d) describes the special case of a single site. Note that the time evolution equations for the probabilities $P_{i,j}(M)$ form a closed set of equations.

In steady state, all the time derivatives are equal to zero. The steady state probability distributions $P_{i,j}(M)$ can be obtained by solving for the corresponding generating functions $F_{i,j}(z) = \sum_{M=0}^{\infty} P_{i,j}(M) z^M$. From eq. (3.4b) it can be seen that $P_{0,j}(M)$ does not depend on $P_{i,j}(M)$ for $i > 0$. Thus, $F_{0,j}(z)$ can be solved for independently of other values of i. The equations satisfied by $F_{0,j}(z)$ are:

$$F_{0,j+1}(z) + F_{0,j-1}(z) - 2F_{0,j}(z) = \frac{a}{D}[1-z]F_{0,j}(z) \qquad 1 < j < L \tag{3.5a}$$

$$F_{0,0}(z) = 1, \qquad F_{0,L+1}(z) = F_{0,L}(z) \tag{3.5b}$$

To solve eq. (3.5), we take the continuum limit in space: $j/L \to y$, $F_{0,j}(z) \to Q(y,z)$, and Taylor expand $F_{0,j\pm1}(z)$ such that $F_{0,j\pm1}(z) \to Q(y,z) \pm \frac{1}{L}\frac{\partial Q}{\partial y} + \frac{1}{2L^2}\frac{\partial^2 Q}{\partial y^2} + \ldots$ Retaining only the leading order term in $1/L$, we find:

$$\frac{\partial^2 Q(y,z)}{\partial y^2} = \lambda(1-z)Q(y,z) \qquad \text{where} \quad \lambda = \frac{aL^2}{D} = \frac{\tilde{a}L}{D} \tag{3.6a}$$

$$Q(y=0,z) = 1, \qquad \left.\frac{\partial Q}{\partial y}\right|_{y=1} = 0 \tag{3.6b}$$

Equation (3.6) can be solved to obtain:

$$Q(y,z) = \frac{\cosh\left[\sqrt{\lambda(1-z)}(1-y)\right]}{\cosh\left[\sqrt{\lambda(1-z)}\right]} \tag{3.7}$$

To check the self-consistency of the continuum approximation, note that the next correction to the leading order term in eq. (3.6a) would be $\sim (1/L^2)\partial^4 Q/\partial y^4$, which from eq.(3.7) is λ/L^2 times the leading order term. Thus, for injection rates of the sort $a = \tilde{a}/L$, the correction term is $1/L$ times the leading order term, whereas for $\mathcal{O}(1)$ injection rates, all higher order terms are comparable to the leading order term. The continuum limit is thus, self-consistent only for injection rates of the sort $a = \tilde{a}/L$, and all the results we derive in this section are valid only for such rates [3]

The generating function for the distribution of total mass in the system can be obtained by setting $y = 1$ in eq. (3.7). This can be inverted for large M and large L (details in appendix 3.7.1), to give the following expression for the tail of the probability distribution $P(M)$:

$$P(M) \sim \frac{\pi}{\lambda} \exp\left(-\frac{\pi^2 M}{4\lambda}\right) \tag{3.8}$$

where $\lambda \propto L$ for $\mathcal{O}(1/L)$ injection rates.

[3] Although the exact results derived in this section are not valid for $\mathcal{O}(1)$ injection rates, numerics show that the scaling of the moments of mass and the structure functions with λ, as derived here, holds in that case too.

We now turn to the calculation of $F_{i,j}(z)$ for $i > 0$. As for the $i = 0$ case, this becomes easier in the continuum limit: $i/L \to x$, $j/L \to y$, $F_{i,j}(z) \to F(x, y, z)$, $\langle M^n_{i,j} \rangle \to \langle M^n_{xy} \rangle$. Then, it follows from eq. (3.4), that in steady state, $F(x, y, z)$ satisfies:

$$\frac{\partial^2 F(x,y,z)}{\partial x^2} + \frac{\partial^2 F(x,y,z)}{\partial y^2} = 0, \qquad x < y \tag{3.9a}$$

$$F(x, x, z) = 1, \qquad \left.\frac{\partial F}{\partial y}\right|_{y=1} = 0, \qquad F(0, y, z) = Q(y, z) \tag{3.9b}$$

This is just the Laplace equation on a triangle with vertices $(0, 0)$, $(0, 1)$, $(1, 0)$, and mixed, inhomogeneous boundary conditions. It can be solved by mapping onto the Laplace equation on a square with suitably chosen boundary conditions [74, 75] (details in appendix 3.8.1), to give:

$$F(x,y,z) = 1 + 2\sum_{n=0}^{\infty} \frac{\lambda(1-z)}{\alpha_n(\alpha_n^2 + \lambda(1-z))} \left[\frac{\sin[\alpha_n x]\cosh[\alpha_n(1-y)] - \sin[\alpha_n y]\cosh[\alpha_n(1-x)]}{\cosh \alpha_n}\right] \tag{3.10}$$

where $\alpha_n = (n + \frac{1}{2})\pi$.

$F(x, y, z)$ can be inverted to give $P_{i,j}(M)$:

$$P_{i,j}(M) = 2\lambda^M \sum_{n=0}^{\infty} \frac{\alpha_n}{(\alpha_n^2 + \lambda)^{M+1}} \left[\frac{\sin[\alpha_n \frac{j}{L}]\cosh[\alpha_n(1 - \frac{i}{L})] - \sin[\alpha_n \frac{i}{L}]\cosh[\alpha_n(1 - \frac{j}{L})]}{\cosh \alpha_n}\right], \quad M > 0 \tag{3.11}$$

For large M and large L, just the first term ($n = 0$) in the series in eq. (3.11) is enough to give a good approximation for $P_{i,j}(M)$, so that:

$$P_{i,j}(M) \sim \frac{\pi}{\lambda} \left[\frac{\sin[\frac{\pi j}{2L}]\cosh[\frac{\pi}{2}(1 - \frac{i}{L})] - \sin[\frac{\pi i}{2L}]\cosh[\frac{\pi}{2}(1 - \frac{j}{L})]}{\cosh \frac{\pi}{2}}\right] \exp\left(-\frac{\pi^2 M}{4\lambda}\right) \tag{3.12}$$

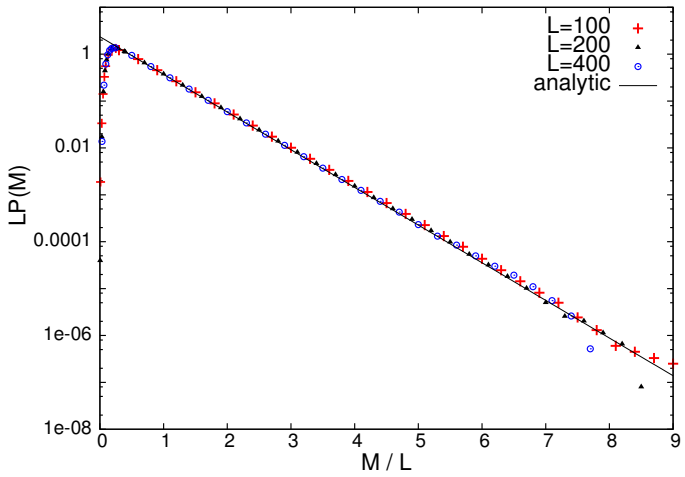
The asymptotic expression for the probability distribution $p_i(m)$ of mass at a single site i can be obtained similarly:

$$p_i(m) \sim \frac{\pi^2}{2\lambda L} \left[\frac{\sin[\frac{\pi i}{2L}]\sinh[\frac{\pi}{2}(1 - \frac{i}{L})] + \cos[\frac{\pi i}{2L}]\cosh[\frac{\pi}{2}(1 - \frac{i}{L})]}{\cosh \frac{\pi}{2}}\right] \exp\left(-\frac{\pi^2 m}{4\lambda}\right) \tag{3.13}$$

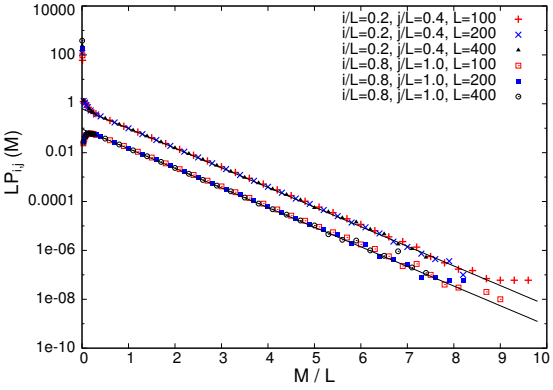
These asymptotic analytic expressions (eqs. (3.8), (3.12) and (3.13)) agree well with numerics (3.2). This calculation thus demonstrates that all mass distributions in the system have in common, a 'condensate' tail with the exponentially decaying form $\sim (1/M_0)\exp[-m/M_0]$ where $M_0 \sim (4/\pi^2)\lambda$. It also gives an expression for the pre-factor of the exponential tail, which is non-universal and depends on the particular distribution.

Moments and cumulants of the mass in any region of the system can also be obtained from the generating functions $F(x, y, z)$ and $Q(y, z)$. For example, for the total mass M, we have:

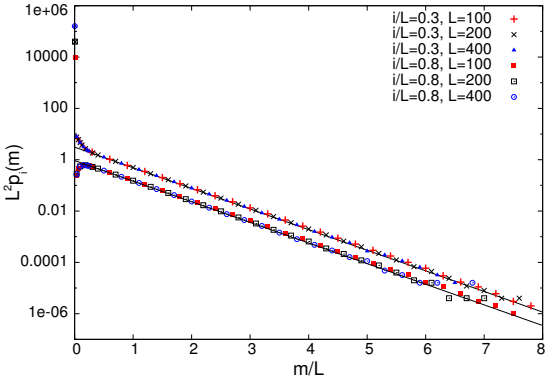
$$\langle M \rangle = \frac{\lambda}{2}, \qquad \langle M^2 \rangle = \frac{5}{12}\lambda^2 + \frac{\lambda}{2}, \qquad \langle M^3 \rangle = \frac{61}{120}\lambda^3 + \frac{5}{4}\lambda^2 + \frac{\lambda}{2} \tag{3.14a}$$



(a)



(b)



(c)

Figure 3.2: Comparison of numerics (data points) and analytical expressions (solid lines) for various steady state mass distributions with $a/D = 4/3$ fixed and different L:
(a) Total mass: Data collapse of tails of the probability distribution $P(M)$ for different L on plotting $LP(M)$ vs. M/L.
(b) Mass in a region $[i+1, j]$: Data collapse of tails on plotting $LP_{i,j}(M)$ vs. M/L for i/L, j/L fixed.
(c) Mass at a single site i: Data collapse of tails on plotting $L^2 p_i(m)$ vs. m/L for a fixed i/L.

$$\langle [M - \langle M \rangle]^2 \rangle = \frac{\lambda^2}{6} + \frac{\lambda}{2}, \qquad \langle [M - \langle M \rangle]^3 \rangle = \frac{2}{15}\lambda^3 + \frac{\lambda^2}{2} + \frac{\lambda}{2}, \qquad \lambda = \tilde{a}L/D \tag{3.14b}$$

The rms fluctuations ΔM of the total mass are anomalously large, scaling as system size L, rather than $\sqrt{L}$, implying that $\Delta M/\langle M \rangle$ remains finite even in the thermodynamic limit. Similarly, by calculating $\langle M_{i,j}^n \rangle$ from $F(x, y, z)$ it is possible to establish that the mass in any region $[i+1,\, j]$ of the system also shows giant fluctuations which scale as L for fixed i/L and j/L.

The moment-generating function $F(x, y, z)$ can be used to calculate many other quantities as well. For example, the 2-point correlation function $\langle m_i m_j \rangle$ can be obtained from the second moments using:

$$\langle m_i m_j \rangle = \frac{1}{2}[\langle M_{i,j-1}^2 \rangle + \langle M_{i-1,j}^2 \rangle - \langle M_{i,j}^2 \rangle - \langle M_{i-1,j-1}^2 \rangle] \tag{3.15}$$

The fluctuations of the particle current J_i in the i^{th} bond of the lattice can also be calculated using:

$$\langle J_i^2 \rangle = \langle [D(m_i - m_{i+1})]^2 \rangle = D^2[2\langle M_{i-1,i}^2 \rangle + 2\langle M_{i,i+1}^2 \rangle - \langle M_{i-1,i+1}^2 \rangle] \tag{3.16}$$

Finding higher moments of the current is more difficult, as it involves quantities like $\langle m_i^2 m_{i+1} \rangle$ for which we need to calculate multi-region moments like $\langle M_{i,j}^2 M_{j,k} \rangle$. In principle, this can be done by generalising the previous analysis, thus allowing for the calculation of various statistics of the particle current.

In summary, in this section we have obtained various mass distributions that characterise the system and shown that all of these have the same kind of non-Gaussian tail which depends explicitly on system size. This calculation also establishes the occurrence of *giant* fluctuations of total mass in the system.

3.3.2 Dynamics

The exit of the macroscopic aggregate from the system causes the total mass M to change sharply over short time scales (see fig. 3.1a). As a result, the time series $M(t)$ exhibits intermittency, which we characterise in terms of temporal structure functions

$$S_n(t) = \langle [M(t) - M(0)]^n \rangle \tag{3.17}$$

where $t = 0$ is any time instant after the system attains steady state. As discussed in sec. 3.2.1, in general, structure functions behave as $S_n(t) \sim t^{u(n)}$ as $t \to 0$, where $u(n)$ is typically sublinear in n for intermittent signals. The main goal of this section is to calculate structure functions for various n for this model and hence, find the n-dependence of the exponents $u(n)$.

Structure functions, as defined in (3.17) are closely related to autocorrelation functions of the total mass. For example,

$$S_2(t) = 2\langle M^2 \rangle - 2\langle M \rangle^2 - 2[\langle M(0)M(t) \rangle - \langle M(0) \rangle \langle M(t) \rangle] \tag{3.18}$$

The correlation function $\langle M(0)M(t)\rangle$ depends on the function $\langle M_{0,L}(0)M_{0,L-1}(t-1)\rangle$ which in turn also depends on $\langle M_{0,L}(0)M_{0,L-2}(t-2)\rangle$ and so on. Thus, it becomes necessary to define and solve for spatio-temporal correlation functions of the sort:

$$C_2(j,t) = \langle M_{0,L}(0)M_{0,j}(t)\rangle - \langle M_{0,L}(0)\rangle\langle M_{0,j}(t)\rangle \tag{3.19a}$$

$$C_{31}(j,t) = \langle M_{0,L}^2(0)M_{0,j}(t)\rangle - \langle M_{0,L}^2(0)\rangle\langle M_{0,j}(t)\rangle \tag{3.19b}$$

$$C_{32}(j,t) = \langle M_{0,L}(0)M_{0,j}^2(t)\rangle - \langle M_{0,L}(0)\rangle\langle M_{0,j}^2(t)\rangle \tag{3.19c}$$

$$C_{41}(j,t) = \langle M_{0,L}^3(0)M_{0,j}(t)\rangle - \langle M_{0,L}^3(0)\rangle\langle M_{0,j}(t)\rangle \tag{3.19d}$$

$$C_{42}(j,t) = \langle M_{0,L}^2(0)M_{0,j}^2(t)\rangle - \langle M_{0,L}^2(0)\rangle\langle M_{0,j}^2(t)\rangle \tag{3.19e}$$

$$C_{43}(j,t) = \langle M_{0,L}(0)M_{0,j}^3(t)\rangle - \langle M_{0,L}(0)\rangle\langle M_{0,j}^3(t)\rangle \tag{3.19f}$$

By solving for the above correlation functions and setting $j = L$, structure functions of various orders can be obtained from:

$$S_2(t) = 2\langle M_{0,L}^2\rangle - 2\langle M_{0,L}\rangle^2 - 2C_2(L,t) \tag{3.20a}$$

$$S_3(t) = 3C_{31}(L,t) - 3C_{32}(L,t) \tag{3.20b}$$

$$S_4(t) = 2\langle M_{0,L}^4\rangle - 8\langle M_{0,L}^3\rangle\langle M_{0,L}\rangle + 6\langle M_{0,L}^2\rangle^2 - 4C_{41}(L,t) + 6C_{42}(L,t) - 4C_{43}(L,t) \tag{3.20c}$$

Now we turn to the actual computation of the correlation functions defined in eq. (3.19). This becomes easier in the spatial continuum limit ($j/L \to y$, $M_{0,j} \to M_y$), as the continuum versions of the correlation functions in eq. (3.19) satisfy relatively simple partial differential equations in y and t. Motivated by eq. (3.20), we write down equations for the following functions:

$$G_2(y,t) = 2C_2(j = yL, t) \tag{3.21a}$$

$$G_3(y,t) = 3C_{31}(yL,t) - 3C_{32}(yL,t) \tag{3.21b}$$

$$G_4(y,t) = 4C_{41}(yL,t) - 6C_{42}(yL,t) + 4C_{43}(yL,t) \tag{3.21c}$$

To derive the time-evolution of the above functions, note that the time evolution of $M_{0,j}$, as captured by eq. (3.4b), can be re-expressed as follows:

$$M_{0,j}(t+\Delta t) = \begin{cases} M_{0,j+1}(t), & \text{with probability } D\Delta t \\ M_{0,j-1}(t), & \text{with probability } D\Delta t \\ M_{0,j}(t)+1, & \text{with probability } a\Delta t \\ M_{0,j}(t), & \text{with probability } 1-(a+2D)\Delta t. \end{cases} \tag{3.22}$$

The time-dependent equation satisfied by $\langle M_{0,L}(0)M_{0,j}(t)\rangle$ can be obtained by multiplying eq. (3.22) by $M_{0,L}(0)$ and then taking an average over different realisations of the stochastic time evolution. Similarly, the equation for $\langle M_{0,L}(0)M_{0,j}^2(t)\rangle$ is obtained by taking the square of eq. (3.22), multiplying it by $M_{0,L}(0)$ and taking averages. In this way, we can write time-evolution equations satisfied by each of the correlation functions in eq. (3.19). Taking the

continuum limit ($j/L \to y$) in space, finally yields the following partial differential equations for $G_n(y,t)$:

$$\begin{aligned} &\frac{\partial G_2(y,t)}{\partial t} = \frac{D}{L^2}\frac{\partial^2 G_2(y,t)}{\partial y^2} \\ &G_2(0,t) = 0, \qquad \left.\frac{\partial G_2}{\partial y}\right|_{y=1} = 0. \end{aligned} \tag{3.23a}$$

$$\begin{aligned} &\frac{\partial G_3(y,t)}{\partial t} = \frac{D}{L^2}\left[\frac{\partial^2 G_3(y,t)}{\partial y^2} - 3\lambda G_2(y,t)\right] \\ &G_3(0,t) = 0, \qquad \left.\frac{\partial G_3}{\partial y}\right|_{y=1} = 0. \end{aligned} \tag{3.23b}$$

$$\begin{aligned} &\frac{\partial G_4(y,t)}{\partial t} = \frac{D}{L^2}\left[\frac{\partial^2 G_4(y,t)}{\partial y^2} + 4\lambda G_3(y,t) - 6\lambda G_2(y,t)\right] \\ &G_4(0,t) = 0, \qquad \left.\frac{\partial G_4}{\partial y}\right|_{y=1} = 0. \end{aligned} \tag{3.23c}$$

To solve these equations, additional boundary conditions are needed, which specify the correlation functions at $t = 0$. These are given by:

$$\begin{aligned} G_2(y,0) &= 2\langle M_y M_1\rangle - 2\langle M_y\rangle\langle M_1\rangle = -[\langle (M_1 - M_y)^2\rangle - \langle M_1^2\rangle - \langle M_y^2\rangle] - 2\langle M_y\rangle\langle M_1\rangle \\ &= \langle M_1^2\rangle + \langle M_y^2\rangle - \langle M_{y1}^2\rangle - 2\langle M_y\rangle\langle M_1\rangle \end{aligned} \tag{3.24a}$$

$$\begin{aligned} G_3(y,0) &= 3\langle M_1^2 M_y\rangle - 3\langle M_1 M_y^2\rangle - 3\langle M_1^2\rangle\langle M_y\rangle + 3\langle M_1\rangle\langle M_y^2\rangle \\ &= \langle M_1^3\rangle - \langle M_y^3\rangle - \langle M_{y1}^3\rangle - 3\langle M_1^2\rangle\langle M_y\rangle + 3\langle M_1\rangle\langle M_y^2\rangle \end{aligned} \tag{3.24b}$$

$$\begin{aligned} G_4(y,0) &= 4\langle M_1^3 M_y\rangle - 6\langle M_1^2 M_y^2\rangle + 4\langle M_1 M_y^3\rangle - 4\langle M_1^3\rangle\langle M_y\rangle + 6\langle M_1^2\rangle\langle M_y^2\rangle - 4\langle M_1\rangle\langle M_y^3\rangle \\ &= \langle M_1^4\rangle + \langle M_y^4\rangle - \langle M_{y1}^4\rangle - 4\langle M_1^3\rangle\langle M_y\rangle + 6\langle M_1^2\rangle\langle M_y^2\rangle - 4\langle M_1\rangle\langle M_y^3\rangle. \end{aligned} \tag{3.24c}$$

Thus, all correlation functions at $t = 0$ can be expressed in terms of $\langle M_y^n\rangle$ and $\langle M_{y1}^n\rangle$ which can be obtained respectively from the moment generating functions $Q(y,z)$ [eq. (3.7)] and $F(x,y,z)$ [eq. (3.10)].

Each of the equations in eq. (3.23) is a heat equation in one spatial dimension with an inhomogeneous source term. It can be solved for the initial conditions given in eq. (3.24) using standard techniques (details in appendix 3.8.2) [76]. The structure functions $S_n(t)$ can be now obtained from $G_n(y,t)$ using eq. (3.20):

$$S_2(t) = \sum_{n=0}^{\infty} 4(-1)^n \left(\frac{2\lambda^2}{\alpha_n^5}\operatorname{sech}\alpha_n + \frac{\lambda}{\alpha_n^3}\right)\left(1 - \exp\left[-\alpha_n^2\frac{Dt}{L^2}\right]\right) \qquad \text{where} \quad \alpha_n = \left(n + \frac{1}{2}\right)\pi \tag{3.25a}$$

$$S_3(t) = \sum_{n=0}^{\infty} 6(-1)^n \lambda^2 \left\{\frac{4}{\alpha_n^5}\left(\left[1 - \lambda\frac{Dt}{L^2}\right]\operatorname{sech}\alpha_n - 1\right) + \frac{1}{\alpha_n^3}\left(1 - 2\frac{Dt}{L^2}\right)\right\}\exp\left[-\alpha_n^2\frac{Dt}{L^2}\right] \tag{3.25b}$$

$$S_4(t) = \Bigg\{ \sum_{n=0}^{\infty} (-1)^n \left(\frac{-96\lambda^4 \operatorname{sech} \alpha_n}{\alpha_n^9} + \lambda^3 \left[\frac{-10}{\alpha_n^3} + \frac{48}{\alpha_n^5} + \frac{-144}{\alpha_n^7} \right] - \lambda^2 \left[\frac{56 \operatorname{sech} \alpha_n}{\alpha_n^5} + \frac{12}{\alpha_n^3} \right] - \frac{4\lambda}{\alpha_n^3} \right)$$
$$\times \left(1 - \exp\left[-\alpha_n^2 \frac{Dt}{L^2} \right] \right) \Bigg\}$$
$$+ \sum_{n=0}^{\infty} 24\lambda^3 \left(\frac{2 \operatorname{sech} \alpha_n}{\alpha_n^5} - \frac{4}{\alpha_n^5} - \frac{Dt}{L^2} \left[\frac{1}{\alpha_n^3} + \frac{2\lambda \operatorname{sech} \alpha_n}{\alpha_n^5} \right] \right) \frac{Dt}{L^2} \exp\left[-\alpha_n^2 \frac{Dt}{L^2} \right] \tag{3.25c}$$

To study intermittency, we need to analyse the small t behaviour of the above expressions. Note that t always appears in the exponentially decaying terms in the combination t/τ where $\tau = L^2/D$. The time scale τ is proportional to L^2 which is roughly the time interval between successive exit events involving a macroscopic aggregate. This, thus, provides a natural time scale for the problem, so that the small t behaviour can be obtained by Taylor expanding the exponentials in eq. (3.25) in powers of t/τ and retaining only the lowest order term. This yields the following small t ($t \ll L^2$) expressions for the structure functions:

$$S_2(t) \sim \lambda \left(0.82\lambda + 2\right) \frac{Dt}{L^2} \tag{3.26a}$$

$$S_3(t) \sim -\lambda^2(\lambda + 2.46)\frac{Dt}{L^2} \tag{3.26b}$$

$$S_4(t) \sim \lambda(1.62\lambda^3 - 8.90\lambda^2 - 0.25\lambda - 2)\frac{Dt}{L^2} \tag{3.26c}$$

In the limit $L \to \infty$, only leading order terms in L are important, so that:

$$S_n(t) \sim c_n(-1)^n \left(\frac{aL^2}{D} \right)^n \left(\frac{Dt}{L^2} \right) = c_n(-1)^n (a\tau)^{n-1} at \tag{3.27}$$

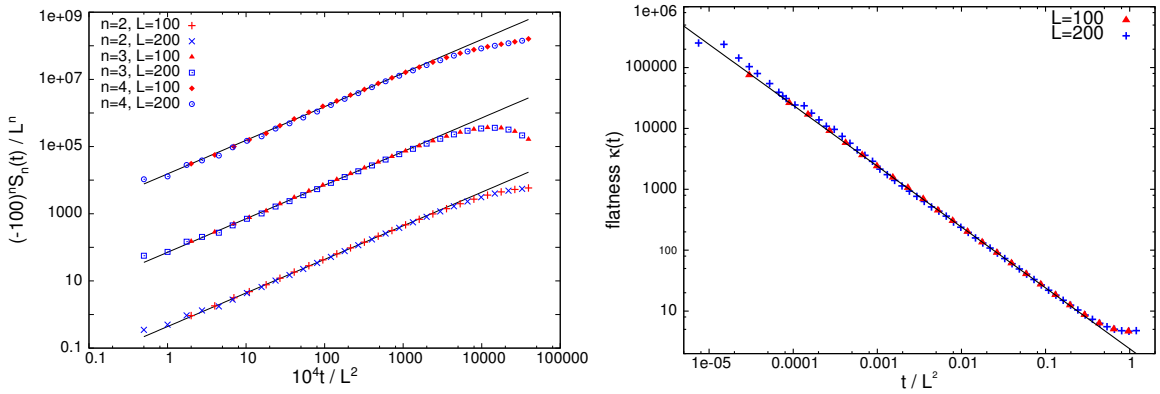
where c_n are the $\mathcal{O}(1)$ coefficients of λ^n in eq. (3.26). The above calculation, thus establishes that $S_n(t) \propto t^{u(n)}$ for small t, where the exponent $u(n)$ is actually *independent of* n instead of being linear in n, as would be expected for a self-similar signal. The anomalous time scale in the problem, which results in breakdown of self-similarity is simply $\tau \propto L^2$. Since we have assumed $\mathcal{O}(1/L)$ injection rates, $a\tau$ is however, proportional to L, resulting in the following scaling of $S_n(t)$ with L:

$$S_n(t) = L^n f_n \left(\frac{Dt}{L^2} \right) \tag{3.28}$$

From eq. (3.26), it also follows that for small t, the flatness is given by:

$$\kappa(t) = \frac{S_4(t)}{(S_2(t))^2} \sim 2.4 \left(\frac{Dt}{L^2} \right)^{-1} \tag{3.29}$$

to leading order in L. Thus, the flatness diverges as a power law as $Dt/L^2 \to 0$. Figure 3.3 shows a comparison of numerical data with the analytical expressions for structure functions and flatness derived within the continuum approximation. The two are in very good agreement.



(a) (b)

Figure 3.3: (a) Structure functions: $(-1)^n S_n(t)/\tilde{L}^n$ vs. $t/\tilde{L}^2$ where $\tilde{L} = L/100$ (rescaling of L done to display all three structure functions in the same plot clearly.) Solid lines representing the small t analytical expressions of eq. (3.26) are in very good agreement with data points obtained from numerical simulations.
(b) Flatness $\kappa(t)$ vs. Dt/L^2: Data collapse for different L. Analytical prediction for small t as given by eq. (3.29) (solid line) agrees well with numerics.

The extreme form of anomalous scaling exhibited by structure functions ($u(n)$ independent of n) is referred to as strong intermittency and is also seen in Burgers turbulence, where spatial structure functions of the velocity field exhibit similar behaviour [64]. The Burgers equation which describes the time evolution of a fully compressible fluid admits solutions with shocks (discontinuities) in the velocity field in the inviscid limit. These shocks dominate the behaviour of higher order structure functions, giving rise to anomalous scaling. This is qualitatively similar to the system we study, where the behaviour of higher order structure functions in time is dominated by the occasional large crashes of total mass [see fig. 3.1a] that occur when a macroscopic aggregate exits the system.

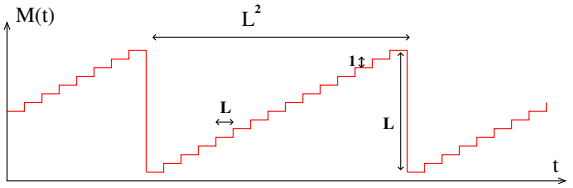


Figure 3.4: A toy model approximating the time series $M(t)$ vs. t for case (A)

Another feature of Burgers turbulence is the bifractal behaviour of the structure functions of the velocity field: these exhibit self-similar scaling for $n \leq 1$ and anomalous scaling for $n \geq 1$. In the context of the present work, this raises the question whether generalised structure functions $\tilde{S}_n(t) = \langle |M(t) - M(0)|^n \rangle$, defined for general (possibly non-integer) n, also show this kind of bifractal behaviour. To answer this question, it is necessary to calculate structure functions of non-integer order. This, however, cannot be done using the preceding analysis. To get an insight into the behaviour of these generalised structure functions, let us consider a toy model in which the time evolution of $M(t)$ takes place in the regular, deterministic

manner illustrated in fig. 3.4. The deterministic increase of M by 1 after every L steps in the figure is akin to the injection of unit mass with rate $1/L$ in the real time series $M(t)$, while the decrease by L after every L^2 steps, represents the exit of the macroscopic aggregate. For the regular pattern in fig. 3.4, we can analyse the probability distribution of $M(t) - M(0)$ for a randomly chosen $t = 0$, and from this obtain $\tilde{S}_n(t) = \langle |M(t) - M(0)|^n \rangle$. This simple analysis shows that $\tilde{S}_n(t) \propto t$ for $n \geq 1$ when $t \ll L^2$, whereas $\tilde{S}_n(t) \propto t^n$ for $n \leq 1$ and $L \ll t \ll L^2$. For $t \ll L$ however, $\tilde{S}_n(t) \propto t/L$ when $n \leq 1$. Thus, except for the $t \ll L$ behaviour when $n \leq 1$, this is analogous to the behaviour of structure functions in Burgers turbulence. The anomalous $t \ll L$ behaviour is simply a consequence of $1/L$ injection rates. The predictions of this simple sawtooth-like picture are borne out by numerical results for $\tilde{S}_n(t)$ for non-integer n.

To summarise, in this sub-section, we have obtained expressions for dynamical structure functions $S_n(t)$ for n upto 4, by calculating various spatio-temporal correlation functions of mass. These calculations yield the anomalous scaling exponents of structure functions, and demonstrate that the total mass exhibits strong intermittency in time.

3.4 Case B: Diffusive movement; outflux from both boundaries

In this section, we consider diffusive movement of masses with equal rates D to the left and right, injection of single particles at site 1 at rate a and exit of aggregates from both boundaries, i.e., from site 1 as well as site L at rate D. In this case, the total mass is extensive and the average particle current through the system is proportional to $1/L$, even with $\mathcal{O}(1)$ injection rates. This is in contrast to model A, where $\mathcal{O}(1/L)$ injection rates were required to ensure extensivity. As discussed in chapter 2, like model A, this system also shows giant fluctuations and temporal intermittency of total mass. However, there are some subtle differences between the two models, especially in the dynamics, which emerge from a detailed analysis.

To analyse this case, we follow the general approach used in the previous section. First, in sec. 3.4.1, we write recursion relations satisfied by the steady state distributions $P_{i,j}(M)$ and solve for various moments $\langle M_{i,j}^n \rangle$ of the distributions. This also allows us to obtain the moments $\langle M^n \rangle$ of the total mass and demonstrate the occurrence of giant fluctuations $\Delta M \propto L$. The behaviour of dynamical structure functions of the total mass is discussed in sec. 3.4.2; we explicitly calculate $S_2(t)$ and indicate how $S_n(t)$ can be obtained for higher n. The structure functions in this case also exhibit the anomalous scaling associated with intermittency, but differ in their functional form case A. As in sec. 3.3, the analysis is simplified by taking a continuum limit in space. Details of these calculations are presented below.

3.4.1 Statics

As in section 3.3.1, it is possible to write time evolution equations for the distributions $P_{i,j}(M, t)$. These yield the following recursion relations for the steady state distributions

$P_{i,j}(M)$:

$$P_{i,j+1}(M) + P_{i,j-1}(M) + P_{i-1,j}(M) + P_{i+1,j}(M) - 4P_{i,j}(M) = 0 \qquad i > 0, j > i+1, j < L \tag{3.30a}$$

$$a[(1-\delta_{M,0})P_{0,j}(M-1) - P_{0,j}(M)] + D[P_{0,j+1}(M) + P_{0,j-1}(M) + P_{1,j}(M) - 3P_{0,j}(M)] = 0 \qquad i = 0, j > i+1, j < L \tag{3.30b}$$

$$P_{i,L-1}(M) + P_{i-1,L}(M) + P_{i+1,L}(M) - 3P_{i,L}(M) = 0 \qquad i > 0, j > i+1, j = L \tag{3.30c}$$

$$2\delta_{M,0} + P_{i,i+2}(M) + P_{i-1,i+1}(M) - 4P_{i,i+1}(M) = 0 \qquad i > 0, j = i+1, j < L \tag{3.30d}$$

Like eq. (3.4), these recursion relations also form a closed set of equations. However, unlike eq. (3.4b) in section 3, the $i = 0$ equation in this case does not decouple from the other equations. Hence, we need to solve the full system of equations in eq. (3.30) to obtain the distribution of total mass. To do this, we define the generating function $G_{i,j}(z) = \sum_{M=1}^{\infty} P_{i,j}(M) z^M$ and take the continuum limit: $i/L \to x$, $j/L \to y$, $G_{i,j}(z) \to G(x,y,z)$. The function $G(x,y,z)$ then satisfies:

$$\frac{\partial^2 G(x,y,z)}{\partial x^2} + \frac{\partial^2 G(x,y,z)}{\partial y^2} = 0 \qquad x < y \tag{3.31a}$$

$$G(x,x,z) = 0, \qquad \left.\frac{\partial G}{\partial y}\right|_{y=1} = 0, \qquad \left.\frac{\partial G}{\partial x}\right|_{x=0} = \frac{aL}{D}\left[(1-z)G(0,y,z) - zP_{0,j=yL}(0)\right] \tag{3.31b}$$

The boundary condition at $x = 0$ in this case is a Robin boundary condition, for which we have not been able to solve the corresponding Laplace equation on a square. However, the functions $g_n(x,y) = \partial^n G/\partial z^n|_{z=1}$ satisfy equations that are easier to solve. The functions $g_n(x,y)$, in turn give $\langle M_{xy}^n \rangle$ which are just the continuum limit versions of the moments $\langle M_{i,j}^n \rangle$. The equations for $g_n(x,y)$ [obtained by differentiating eq. (3.31)] are:

$$\frac{\partial^2 g_n(x,y)}{\partial x^2} + \frac{\partial^2 g_n(x,y)}{\partial y^2} = 0 \qquad x < y \tag{3.32a}$$

$$g_n(x,x) = 0, \qquad \left.\frac{\partial g_n(x,y)}{\partial y}\right|_{y=1} = 0, \qquad \left.\frac{\partial g_n(x,y)}{\partial x}\right|_{x=0} = -\zeta n g_{n-1}(0,y)$$

$$\text{where} \qquad \zeta = \frac{aL}{D} \qquad \text{and} \quad g_0(0,y) = 1 \tag{3.32b}$$

Note that unlike the generating function $G(x,y,z)$, the functions $g_n(x,y)$ obey Neumann boundary conditions at $x = 0$ and $y = 1$. We solve the above system of equations recursively, with the solution of $g_n(x,y)$ providing the boundary conditions for $g_{n+1}(x,y)$ and so on. The expressions for $\langle M_{xy}^n \rangle$, thus obtained (details in appendix 3.8.1), retaining only leading order terms in $\zeta = aL/D$ are:

$$\langle M_{xy} \rangle = \frac{\zeta}{2}(x-y)(x+y-2) \tag{3.33a}$$

$$\langle M_{xy}^2\rangle = \frac{\zeta^2}{3}(x-y)(x+y-2)+ 4\zeta^2\sum_{n=1}^{\infty}\frac{1}{(n\pi)^3}\left[\frac{\cos[n\pi x]\cosh[n\pi(1-y)]-\cos[n\pi y]\cosh[n\pi(1-x)]}{\sinh[n\pi]}\right] \tag{3.33b}$$

$$\langle M_{xy}^3\rangle = \frac{2\zeta^3}{5}(x-y)(x+y-2)+ 12\zeta^3\sum_{n=1}^{\infty}\left[\frac{\mathrm{cosech}[n\pi]}{(n\pi)^4}\left(\coth[n\pi]-\frac{1}{n\pi}\right)\right]\left[\frac{\cos[n\pi x]\cosh[n\pi(1-y)]-\cos[n\pi y]\cosh[n\pi(1-x)]}{\sinh[n\pi]}\right] \tag{3.33c}$$

$$\langle M_{xy}^4\rangle = 8\zeta^4\left[\frac{1}{15}-3\sum_{n=1}^{\infty}\left(\frac{1}{(n\pi)^6}-\frac{2\coth[n\pi]}{(n\pi)^5}\right)\right](x-y)(x+y-2) + 16\zeta^4\sum_{n=1}^{\infty}\left\{\left[\frac{\mathrm{cosech}[n\pi]}{(n\pi)^3}\left(12\sum_{m=1}^{\infty}\frac{\coth[m\pi]}{m\pi(m^2\pi^2+n^2\pi^2)}+\frac{6\coth^2[n\pi]-1}{(n\pi)^2}-\frac{3}{(n\pi)^4}\right)\right]\times \left[\frac{\cos[n\pi x]\cosh[n\pi(1-y)]-\cos[n\pi y]\cosh[n\pi(1-x)]}{\sinh[n\pi]}\right]\right\} \tag{3.33d}$$

Equation (3.33) also gives the moments of the total mass $M = M_{01}$ in the system. To leading order in $\zeta = aL/D$, these are:

$$\langle M\rangle = \zeta/2, \qquad \langle M^2\rangle \sim 0.46\zeta^2, \qquad \langle M^3\rangle \sim 0.62\zeta^3, \qquad \langle M^4\rangle \sim 1.07\zeta^4 \tag{3.34a}$$

$$\langle [M-\langle M\rangle]^2\rangle \sim 0.21\zeta^2, \qquad \langle [M-\langle M\rangle]^3\rangle \sim 0.17\zeta^3, \qquad \langle [M-\langle M\rangle]^4\rangle \sim 0.35\zeta^4 \tag{3.34b}$$

From eq. (3.34), it is evident that the various moments and cumulants of the total mass in this case scale as the same power of L as in model A in sec. 3.3 (with $1/L$ injection rates). In particular, the rms fluctuations of the total mass are again anomalously large, scaling as L rather than $\sqrt{L}$, so that $\Delta M/\langle M\rangle$ is finite, even in the limit $L \to \infty$. A comparison with numerics (fig. 3.5) shows that the numerical prefactors obtained in eq. (3.34) are close to exact. Note that these are different from the corresponding pre-factors in eq. (3.14) for model A in sec. 3.3.1.

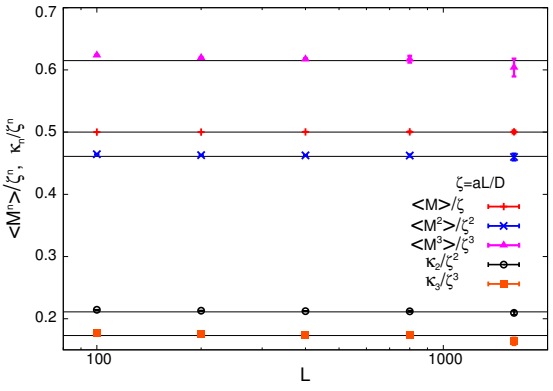


Figure 3.5: $\langle M^n\rangle/\zeta^n$ vs. L for $n = 1, 2, 3$. Cumulants $\kappa_2 = \langle [M-\langle M\rangle]^2\rangle$ and $\kappa_3 = \langle [M-\langle M\rangle]^3\rangle$ plotted as κ_n/ζ^n vs. L. Solid lines are the analytical predictions given by eq. (3.34)

The method that we have used to derive $\langle M_{xy}^n \rangle$ can also be generalised to calculate moments of the r-sector probability $P_{x_1,x_2,\ldots x_{r+1}}(M_1, M_2, ..M_r)$ which is the joint probability of finding masses M_1, M_2, .. M_r in r contiguous stretches x_1 to x_2, x_2 to x_3, ... x_r to x_{r+1} of the lattice respectively. Moments of such r-sector probability distributions will be required in sec. 3.4.2 for the computation of structure functions of order 3 or more. For example, we need $\langle M_{xy}^2 M_{y1} \rangle$ in order to calculate $S_3(t)$ and moments such as $\langle M_{xy}^2 M_{yz} M_{z1} \rangle$ to calculate $S_4(t)$ and so on.

We indicate below how such multi-sector moments can be calculated. For simplicity, consider the two-sector distribution $P_{i,j,k}(M_1, M_2)$. This evolves in time by exchange of mass at sites i, j and k, and in steady state, satisfies the discrete Laplace equation in three variables. We can, as before, use the generating function approach, take the continuum limit in space, and obtain the differential equation satisfied by the moment $\langle M_{xy}^2 M_{yz} \rangle$ where $0 \leq x \leq y \leq z \leq 1$. This equation is just the three dimensional Laplace equation on a pyramidal region with mixed boundary conditions (Neumann and Dirichlet) on various faces of the pyramid. It can be solved in the same way as the 2D equation considered so far, i.e. by mapping it onto a Laplace equation on a cube with appropriately chosen boundary conditions (see appendix 3.8.1 for details). This method can also be generalised to higher r: in general, to obtain moments of $P_{x_1,x_2,\ldots x_{r+1}}(M_1, M_2, .., M_r)$, it is necessary to solve the $r+1$ dimensional Laplace equation (on an $r+1$ dimensional hyper-pyramid). However, while the generalisation to $r \geq 2$ is conceptually straightforward, the actual calculations become very cumbersome and we do not pursue this further here.

3.4.2 Dynamics

We now turn to the computation of dynamical structure functions $S_n(t)$ for this model. We will explicitly calculate only $S_2(t)$, and indicate how higher order structure functions can be obtained.

As discussed in sec. 3.3.2, the structure function of order n can be obtained from the corresponding autocorrelation functions of the total mass. Thus, to calculate $S_2(t)$, we need $\langle M(0)M(t) \rangle - \langle M(0) \rangle \langle M(t) \rangle$, which is a special case of the general spatio-temporal correlation function $\langle M_{0,L}(0) M_{i,j}(t) \rangle - \langle M_{0,L}(0) \rangle \langle M_{i,j}(t) \rangle$ or its continuum limit version: $H_2(x, y, t) = \langle M_{01}(0) M_{xy}(t) \rangle - \langle M_{01}(0) \rangle \langle M_{xy}(t) \rangle$.

By following a procedure similar to that followed in section 3.3.2, we obtain the following time-dependent equation for $H_2(x, y, t)$:

$$\frac{\partial H_2(x,y,t)}{\partial t} = \frac{D}{L^2} \left[\frac{\partial^2 H_2(x,y,t)}{\partial x^2} + \frac{\partial^2 H_2(x,y,t)}{\partial y^2} \right] \tag{3.35a}$$

$$\left. \frac{\partial H_2}{\partial x} \right|_{x=0} = 0, \qquad \left. \frac{\partial H_2}{\partial y} \right|_{y=1} = 0, \qquad H_2(x,x,t) = 0. \tag{3.35b}$$

$$\begin{aligned} H_2(x,y,0) &= 2\langle M_{01} M_{xy} \rangle - 2\langle M_{01} \rangle \langle M_{xy} \rangle \\ &= 2\langle M_{0x} M_{xy} \rangle + 2\langle M_{xy}^2 \rangle + 2\langle M_{xy} M_{y1} \rangle - 2\langle M_{01} \rangle \langle M_{xy} \rangle \\ &= \langle M_{0y}^2 \rangle - \langle M_{0x}^2 \rangle + \langle M_{x1}^2 \rangle - \langle M_{y1}^2 \rangle - 2\langle M_{01} \rangle \langle M_{xy} \rangle \end{aligned} \tag{3.35c}$$

This is the 2D heat equation on a right triangle with mixed boundary conditions and a specified initial state $H_2(x,y,0)$ which is expressible in terms of single-sector mass moments $\langle M_{xy}\rangle$ and $\langle M_{xy}^2\rangle$ [eq. (3.33)]. Equation (3.35) can be solved (details in appendix 3.8.3); the $x=0$, $y=1$ solution then gives $\langle M(0)M(t)\rangle - \langle M(0)\rangle\langle M(t)\rangle$, from which we obtain the following expression for $S_2(t)$:

$$S_2(t) = \sum_{n=1,3,5\ldots}^{\infty} \left\{ \frac{16\zeta^2}{(n\pi)^4}\left(n\pi\coth\left[\frac{n\pi}{2}\right]-1\right) + \frac{4\zeta}{(n\pi)^2}\right\}\left\{1-\exp\left[-\frac{D\pi^2n^2t}{L^2}\right]\right\}, \qquad \zeta = \frac{aL}{D} \tag{3.36}$$

To study intermittency, we need to to extract the small t behaviour of $S_2(t)$ from eq. (3.36). This cannot be done by simply expanding $1-\exp\left[-\frac{D\pi^2n^2t}{2L^2}\right]$ as $\sim \frac{D\pi^2n^2t}{L^2}$, as the resultant sum over n diverges due to the presence of the coth term. Instead, we keep the exponential term as it is, and approximate the discrete sum in eq. (3.36) by an integral over n. On taking the limit $Dt/L^2 \to 0$ of this integral (details in appendix 3.7.2), we get the following small t form of $S_2(t)$:

$$S_2(t) \sim -\frac{4}{\pi}\left(\frac{aL}{D}\right)^2 \frac{Dt}{L^2}\log\left[A_1\frac{Dt}{L^2}\right] \tag{3.37}$$

where A_1 is a constant which can be derived approximately using our analysis (see appendix 3.7.2) but is more accurately obtained by fitting to numerical data.

Higher order structure functions can be calculated using a similar procedure but the calculations become progressively more cumbersome. To calculate $S_3(t)$, we need to calculate the function $H_3(x,y,t) = 3\langle M_{01}^2(0)M_{xy}(t)\rangle - 3\langle M_{01}^2\rangle\langle M_{xy}\rangle - 3\langle M_{01}(0)M_{xy}^2(t)\rangle + 3\langle M_{01}\rangle\langle M_{xy}^2\rangle$ which satisfies the following time evolution equation:

$$\frac{\partial H_3(x,y,t)}{\partial t} = \frac{D}{L^2}\left[\frac{\partial^2 H_3(x,y,t)}{\partial x^2} + \frac{\partial^2 H_3(x,y,t)}{\partial y^2}\right] \tag{3.38a}$$

$$\left.\frac{\partial H_3}{\partial x}\right|_{x=0} = 3\zeta H_2(0,y,t), \qquad \left.\frac{\partial H_3}{\partial y}\right|_{y=1} = 0, \qquad H_3(x,x,t) = 0. \tag{3.38b}$$

$$\begin{aligned} H_3(x,y,0) &= 3\langle M_{01}^2 M_{xy}\rangle - 3\langle M_{01}\rangle\langle M_{xy}\rangle - 3\langle M_{01}M_{xy}^2\rangle + 3\langle M_{01}\rangle\langle M_{xy}^2\rangle \\ &= \langle M_{0y}^3\rangle - \langle M_{0x}^3\rangle + \langle M_{x1}^3\rangle - \langle M_{y1}^3\rangle - 2\langle M_{xy}^3\rangle - 3\langle M_{01}^2\rangle\langle M_{xy}\rangle + 3\langle M_{01}\rangle\langle M_{xy}^2\rangle \\ &\quad + 3\left[\langle M_{0x}^2 M_{xy}\rangle + \langle M_{0y}^2 M_{y1}\rangle - \langle M_{0x}^2 M_{x1}\rangle - \langle M_{xy}^2 M_{y1}\rangle\right] \end{aligned} \tag{3.38c}$$

An important difference between equations (3.35) and (3.38) is in the nature of the initial conditions: while $H_2(x,y,0)$ is expressible in terms of the moments $\langle M_{xy}^2\rangle$ and $\langle M_{xy}\rangle$ of mass in a single sector, $H_3(x,y,0)$ also involves terms like $\langle M_{xy}^2 M_{y1}\rangle$ which require knowledge of two-sector probabilities i.e. the joint probability of finding mass M_1 and M_2 in two contiguous stretches of the lattice. Similarly, the expression for $H_4(x,y,0)$ involves terms like $\langle M_{xy}^2 M_{yz} M_{z1}\rangle$ for which we need to calculate joint probabilities of mass in three contiguous stretches. As discussed at the end of section 3.4.1, the calculation of the moments of these

multi-region joint probabilities is rather cumbersome. However, in principle, it can be done, thus giving the $t = 0$ value of correlation functions such as $H_3(x, y, t)$. Once this $t = 0$ value is known, eq. (3.38) can be solved (details in appendix 3.8.3) to obtain $S_3(t)$.

As discussed in chapter 2, higher order structure functions obey the scaling: $S_n(t) = L^n \mathcal{F}_n(Dt/L^2)$, which is consistent with eq. (3.37). Motivated by the small t expression for $S_2(t)$ [eq. (3.37)], we try to fit $\mathcal{F}_n$ to the form $\mathcal{F}_n(y) \sim y\mathcal{G}_n[\log(y)]$ for small y and obtain good fits for $S_3(t)$ and $S_4(t)$ by choosing a polynomial form for the function $\mathcal{G}_n$.

As before, structure functions $\tilde{S_n}(t) = \langle |M(t) - M(0)|^n \rangle$ for non-integer n cannot be calculated analytically by this approach and must be studied using numerical simulations. Numerics show that unlike the case with spatially separated influx and outflux [sec. 3.3], in this model, structure functions show self-similar scaling $\tilde{S_n}(t) \sim t^{n/2}$ for $n \leq 2$ and anomalous scaling for $n \geq 2$. This behaviour can be explained heuristically by considering a toy model [similar to fig. 3.4 in sec. 3.3.2] in which $M(t)$ grows as $\sqrt{t}$ instead of $\sim t$ between successive crashes. This is a plausible approximation to the real time series $M(t)$, as the real $M(t)$ evolves through a large number of $\mathcal{O}(1)$ changes, which may be positive or negative, thus resulting in an effective random walk behaviour.

To summarise, in this sub-section we have explicitly calculated the structure function $S_2(t)$ and indicated how higher order structure functions can be obtained. A comparison of this model [model (B)] with the model studied in sec. 3.3 [model (A)] shows that while the static properties of the total mass in both cases are qualitatively similar, significant differences appear in the dynamical properties. This is apparent even in a typical time series $M(t)$ in the two cases [figs. 3.1a and 3.1b], where the time series corresponding to case (B) shows sharp drops of $M(t)$ at many scales. This difference manifests itself in the multiplicative $log(t)$ terms in the structure functions. Moreover, the anomalous, intermittency-associated scaling of $S_n(t)$ in model (B) sets in only for $n \geq 2$, rather than for $n \geq 1$ as in model (A).

3.5 Case C: Driven movement; outflux from right boundary

In this section, we analyse the case where single particles are injected onto site 1 at rate a, aggregates move unidirectionally to the right with rate D, and exit from site L, also with rate D. Note that this is the same as monitoring only the first L sites of a semi-infinite lattice with injection at the origin, and unidirectional movement of aggregates everywhere.

The analysis of this section shows that the properties of the aggregation model with localised influx are qualitatively different in the presence of bias: the model with biased stack movement is characterised by formation and exit of aggregates of $\mathcal{O}(\sqrt{L})$ mass, rather than $\mathcal{O}(L)$ mass, leading to static and dynamical properties that are different from the unbiased case. Some of these static properties, in particular the single site mass distribution, have already been calculated by Jain et al in [32]. In this section, we extend the analysis in [32] to calculate the distribution of mass in an arbitrary region of the system. We then use these results to compute the dynamical structure functions $S_n(t)$ and show that this system also exhibits strong temporal intermittency of the total mass, though at much shorter time scales than

the unbiased system. Unlike the previous two sections, in this case, we work with recursion relations directly, without taking a continuum limit in space.

3.5.1 Statics

For the calculations in this section, it turns out to be more convenient to define $M_{i,j}$ and $P_{i,j}(M)$ in the following way:

$$M_{i,j} = \sum_{l=i+1}^{i+j} m_l \qquad P_{i,j}(M) = \text{Prob}(M_{i,j} = M) \tag{3.39}$$

Thus, $P_{i,j}(M)$ is now the probability of finding mass M in the region $[i+1,\, i+j]$. As in the unbiased case with outflux only from site L (section 3.3), the equations for $P_{0,j}(M)$ decouple from $P_{i,j}(M)$ with $i > 0$. In steady state, $P_{0,j}(M)$ satisfy the following recursion relations:

$$a[1-\delta_{M,0}]P_{0,j}(M-1) - aP_{0,j}(M) + D[P_{0,j-1}(M) - P_{0,j}(M)] = 0 \qquad j > 1 \tag{3.40a}$$

$$a[1-\delta_{M,0}]P_{0,1}(M-1) - aP_{0,1}(M) + D[\delta_{M,0} - P_{0,1}(M)] = 0 \qquad j = 1 \tag{3.40b}$$

From eq. (3.40), it follows that the generating function $Q_j(z) = \sum_{M=0}^{\infty} P_{0,j}(M) z^M$ satisfies:

$$Q_j(z) = \frac{Q_{j-1}(z)}{1+\eta(1-z)}, \qquad Q_0(z) = 1, \qquad \eta = a/D \tag{3.41}$$

From eq. (3.41), we can obtain the generating function $Q_L(z)$ of the probability distribution of mass in a system of L sites:

$$Q_L(z) = \frac{1}{[1+\eta(1-z)]^L} \tag{3.42}$$

Inverting $Q_L(z)$ yields the following expression for the distribution $P(M)$ of mass M in a system of L sites:

$$P(M) = \frac{\eta^M}{(1+\eta)^{L+M}} \binom{L+M-1}{M} \tag{3.43}$$

Various moments and cumulants of M can also be worked out from eq. (3.42). They are:

$$\langle M \rangle = \eta L \qquad \langle M^2 \rangle = \eta^2 L^2 + \eta(1+\eta)L \qquad \langle M^3 \rangle = \eta^3 L^3 + 3\eta^2(1+\eta)L^2 + \eta(1+\eta)(1+2\eta)L \tag{3.44a}$$

$$\Delta M^2 = \langle [M - \langle M \rangle]^2 \rangle = \eta(1+\eta)L, \qquad \langle [M - \langle M \rangle]^3 \rangle = \eta(1+\eta)(1+2\eta)L \tag{3.44b}$$

In fact, it follows from eq. (3.42) that all cumulants of the total mass scale as L. This calculation, thus demonstrates an important difference between the driven and diffusive cases: the total mass in the driven case *does not have a broad distribution*; in the limit $L \to \infty$, the distribution $P(M)$ in eq. (3.53) approaches a Gaussian distribution for the variable $(M - \langle M \rangle)/\Delta M$. Further, the rms fluctuations of the total mass scale as $\sqrt{L}$, and not L. Nevertheless, as demonstrated in sec. 3.5.2, the total mass is intermittent in time, as in the diffusive case.

Another point of difference from the diffusive case is that the recursion relation in eq. (3.41) cannot be solved by taking a continuum limit in space and replacing $Q_j - Q_{j-1}$ by $\partial Q/\partial x$. This follows from a self-consistency argument similar to that employed in sec. 3.3.1. [4] Thus, in the analysis of the driven case, we need to necessarily work with recursion relations rather than differential equations.

We now turn to the calculation of $P_{i,j}(M)$ for $i > 0$. Most of this analysis has already been carried out in [32] but we briefly summarise it here in the interest of continuity. In steady state, the probabilities $P_{i,j}(M)$ satisfy the following recursion relations:

$$P_{i-1,j+1}(M) + P_{i,j-1}(M) - 2P_{i,j}(M) = 0 \qquad i > 0, j > 1 \tag{3.45a}$$

$$P_{i-1,2}(M) + \delta_{M,0} - 2P_{i,1}(M) = 0 \qquad j = 1 \tag{3.45b}$$

The equation for the case $i = 0$ [eq. (3.40)] decouples from these and has already been solved. From eq. (3.45) and (3.41), it follows that the generating functions $F_{i,j}(M) = \sum_{M=0}^{\infty} P_{i,j}(M) z^M$ satisfy:

$$F_{i-1,j+1}(z) + F_{i,j-1}(z) - 2F_{i,j}(z) = 0, \qquad i > 0, j > 1 \tag{3.46a}$$

$$F_{i,0}(z) = 1, \qquad F_{0,j}(z) = Q_j(z) = \frac{1}{[1 + \eta(1-z)]^j}, \quad \eta = a/D \tag{3.46b}$$

The above set of equations can be solved using the functions:

$$G_{u,j}(z) = \sum_{i=1}^{\infty} F_{i,j}(z) u^i \qquad H_{u,v}(z) = \sum_{j=1}^{\infty} G_{u,j}(z) v^j \tag{3.47}$$

By performing the sums over both i and j on eq. (3.46), the following equation for $H_{u,v}(z)$ is obtained:

$$H_{u,v}(z) = \frac{uvG_{u,1}(z) - \frac{uv^2}{1-u} - \frac{uv^2}{[1+\eta(1-z)][1-v+\eta(1-z)]}}{v^2 - 2v + u} \tag{3.48}$$

Equation (3.48) can be solved to get a closed form expression for $H_{u,v}(z)$ in the same way as in [32]: the denominator of the R.H.S of eq. (3.48) can be expressed as $v^2 - 2v + u = (v_- - v)(v_+ - v)$ where $v_\pm = 1 \pm \sqrt{1-u}$. Since the denominator has a zero at $v = v_-$ which is less than 1, it follows that the numerator must also have a zero at $v = v_-$, to ensure that the function $H_{u,v}(z)$ is analytic for $v < 1$. This condition allows us to solve for $G_{u,1}(z)$:

$$G_{u,1}(z) = \frac{v_-}{1-u} + \frac{v_-}{[1+\eta(1-z)][1-v_-+\eta(1-z)]} \qquad \text{where } v_- = 1 - \sqrt{1-u} \tag{3.49}$$

The single site mass distribution $p_i(m)$ can be obtained by inverting $G_{u,1}(z)$ w.r.t u and z [32].

[4] A continuum approximation would be self-consistent if $\eta \propto L^{-\alpha}$, but this would result in sub-extensive total mass

We now extend the above analysis to find $F_{i,j}(z)$. Substituting from eq. (3.49) into eq. (3.48) allows for the determination of $H_{u,v}(z)$:

$$H_{u,v}(z) = \frac{uv}{1-v+\sqrt{1-u}}\left[\frac{1}{1-u} + \frac{1}{[1-v+\eta(1-z)][\sqrt{1-u}+\eta(1-z)]}\right] \tag{3.50}$$

$H_{u,v}(z)$ can be Taylor expanded in powers of v to recover $G_{u,j}(z)$:

$$G_{u,j}(z) = \left(\frac{u}{1-u}\right)\left(\frac{\eta^2(1-z)^2}{\eta^2(1-z)^2-(1-u)}\right)\left[[1+\sqrt{1-u}]^{-j} - \left(\frac{1-u}{\eta^2(1-z)^2}\right)[1+\eta(1-z)]^{-j}\right] \tag{3.51}$$

Using $[1+\sqrt{1-u}]^{-j} = \sum_{k=0}^{\infty} 2^{-2k-j}\frac{j}{j+2k}\binom{j+2k}{k}u^k$, it is possible to invert $G_{u,j}(z)$ and get a formal expression for $F_{i,j}(z)$:

$$F_{i,j}(z) = [1+\eta(1-z)]^{-j-i}[1-\eta(1-z)]^{-i} + j2^{-j}\sum_{k=0}^{i-1}\left[\frac{2^{-2k}}{j+2k}\binom{j+2k}{k}\left\{1-[1-\eta^2(1-z)^2]^{k-i}\right\}\right] \tag{3.52}$$

Formal expressions for the moments $\langle M_{i,j}^n\rangle$ of the mass in the region $[i+1, i+j]$ can be obtained by differentiating eq. (3.52). In the limit $i\to\infty$, $j\to\infty$ with $j/\sqrt{i}$ finite, however, simpler asymptotic expressions for $\langle M_{i,j}^n\rangle$ emerge. An analysis of eq. (3.51) [details in appendix 3.7.3] shows that in this limit, $\langle M_{i,j}^n\rangle$ satisfy the simple scaling form $\langle M_{i,j}^n\rangle \sim i^{n/2}\mathcal{R}_n\left(\frac{j}{\sqrt{i}}\right)$:

$$\langle M_{i,j}\rangle = \eta j \tag{3.53a}$$

$$\langle M_{i,j}^2\rangle \sim i\mathcal{R}_2\left(\frac{j}{\sqrt{i}}\right) \sim \frac{4\eta^2 i}{\sqrt{\pi}}\left[\left(\frac{j}{\sqrt{i}}\right) + \frac{1}{12}\left(\frac{j}{\sqrt{i}}\right)^3 + ...\right] \tag{3.53b}$$

$$\langle M_{i,j}^3\rangle \sim i^{3/2}\mathcal{R}_3\left(\frac{j}{\sqrt{i}}\right) \sim 6\eta^3 i^{3/2}\left[\left(\frac{j}{\sqrt{i}}\right) + \frac{1}{6}\left(\frac{j}{\sqrt{i}}\right)^3 + ...\right] \tag{3.53c}$$

$$\langle M_{i,j}^4\rangle \sim i^2\mathcal{R}_4\left(\frac{j}{\sqrt{i}}\right) \sim \frac{32\eta^4 i^2}{\sqrt{\pi}}\left[\left(\frac{j}{\sqrt{i}}\right) + \frac{1}{4}\left(\frac{j}{\sqrt{i}}\right)^3 + ...\right] \tag{3.53d}$$

As discussed in the next section, this scaling limit also turns out to be crucial for the analysis of the intermittency properties of the system.

3.5.2 Dynamics

To calculate structure functions, we follow the general approach of sections 3.3.2 and 3.4.2, i.e. we first find the correlation functions defined in eq. (3.19) and then relate them to structure functions using eq. (3.20). Let $f_j(t)$, $g_j(t)$ and $h_j(t)$ denote the following correlation functions:

$$f_j(t) = 2C_2(j,t) \tag{3.54a}$$

$$g_j(t) = 3C_{31}(j,t) - 3C_{32}(j,t) \tag{3.54b}$$

$$h_j(t) = 4C_{41}(j,t) - 6C_{42}(j,t) + 4C_{43}(j,t) \tag{3.54c}$$

where $C_2(j,t)$, $C_{31}(j,t)$ etc. are defined in eq. (3.19). By solving for these correlation functions and using the $j = L$ solution, various structure functions can be obtained from:

$$S_2(t) = 2\langle M_{0,L}^2\rangle - 2\langle M_{0,L}\rangle^2 - f_L(t) \tag{3.55a}$$

$$S_3(t) = g_L(t) \tag{3.55b}$$

$$S_4(t) = 2\langle M_{0,L}^4\rangle - 8\langle M_{0,L}^3\rangle\langle M_{0,L}\rangle + 6\langle M_{0,L}^2\rangle^2 - h_L(t) \tag{3.55c}$$

The time evolution equations satisfied by $f_j(t)$, $g_j(t)$ and $h_j(t)$ can be derived from the time evolution equation for $M_{0,j}(t)$ by following the same procedure as in sec. 3.3.2. These equations are:

$$f_j(t+1) = (1-D)f_j(t) + Df_{j-1}(t) \tag{3.56a}$$

$$g_j(t+1) = (1-D)g_j(t) + Dg_{j-1}(t) - 3af_j(t) \tag{3.56b}$$

$$h_j(t+1) = (1-D)h_j(t) + Dh_{j-1}(t) - 4ag_j(t) + 6af_j(t) \tag{3.56c}$$

These equations can be solved for $t < j$ by induction to obtain:

$$f_j(t) = \sum_{k=0}^{t} \binom{t}{k} (1-D)^{t-k} D^k f_{j-k}(0) \tag{3.57a}$$

$$g_j(t) = \sum_{k=0}^{t} \left[\binom{t}{k} (1-D)^{t-k} D^k g_{j-k}(0) \right] - 3atf_j(t-1) \tag{3.57b}$$

$$h_j(t) = \sum_{k=0}^{t} \left[\binom{t}{k} (1-D)^{t-k} D^k h_{j-k}(0) \right] - at[4g_j(t-1) - 6f_j(t-1)] - 12a^2 \frac{t(t-1)}{2} f_j(t-2) \tag{3.57c}$$

Equation (3.57) expresses $f_j(t)$, $g_j(t)$ and $h_j(t)$ in terms of their $t = 0$ values. The $t = 0$ values $f_j(0)$, $g_j(0)$ and $h_j(0)$ can be expressed in terms of $\langle M_{0,j}^n\rangle$ and $\langle M_{j,L-j}^n\rangle$ as follows:

$$f_j(0) = \langle M_{0,L}^2\rangle + \langle M_{0,j}^2\rangle - \langle M_{j,L-j}^2\rangle - 2\langle M_{0,j}\rangle\langle M_{0,L}\rangle \tag{3.58a}$$

$$g_j(0) = \langle M_{0,L}^3\rangle - \langle M_{0,j}^3\rangle - \langle M_{j,L-j}^3\rangle - 3\langle M_{0,L}^2\rangle\langle M_{0,j}\rangle + 3\langle M_{0,L}\rangle\langle M_{0,j}^2\rangle \tag{3.58b}$$

$$h_j(0) = \langle M_{0,L}^4\rangle + \langle M_{0,j}^4\rangle - \langle M_{j,L-j}^4\rangle - 4\langle M_{0,L}^3\rangle\langle M_{0,j}\rangle + 6\langle M_{0,L}^2\rangle\langle M_{0,j}^2\rangle - 4\langle M_{0,L}\rangle\langle M_{0,j}^3\rangle. \tag{3.58c}$$

Note that this is just eq. (3.24), if we take into account the difference in the way $M_{i,j}$ is defined in sections 3.3 and 3.5. It is now possible to obtain the structure functions by using eqs. (3.57) and (3.58). For example, $S_2(t)$ can be expressed as:

$$\begin{aligned} S_2(t) &= 2[\langle M_L^2\rangle - \langle M_L\rangle^2] - 2\sum_{k=0}^{t} \binom{t}{k} (1-D)^{t-k} D^k [\langle M_{L-k}M_L\rangle - \langle M_{L-k}\rangle\langle M_L\rangle] \\ &= 2[\langle M_L^2\rangle - \langle M_L\rangle^2] - \sum_{k=0}^{t} \binom{t}{k} (1-D)^{t-k} D^k [\langle M_L^2\rangle + \langle M_{L-k}^2\rangle - \langle M_{L-k,k}^2\rangle - 2\langle M_{L-k}\rangle\langle M_L\rangle] \end{aligned} \tag{3.59}$$

Since the time interval between successive exit events is $\mathcal{O}(\sqrt{L})$, we expect the system to show temporal intermittency in the limit where $t/\sqrt{L}$ is small. Thus, we need to investigate the behaviour of $S_2(t)$ in the scaling limit $L \to \infty$, $t \to \infty$, $t/\sqrt{L}$ finite but small. In this limit, the dominant contribution to the sum over k in eq. (3.59) must come from $k \sim \mathcal{O}(\sqrt{L})$. Thus, term such as $\langle M^2_{L-k,k} \rangle$ can be approximated as $\langle M^2_{L,k} \rangle$ for which we can use the asymptotic ($k \to \infty$, $L \to \infty$, $k/\sqrt{L}$ finite) expressions in eq. (3.53). Using the exact expressions for $\langle M^2_L \rangle$, $\langle M^2_{L-k} \rangle$ etc. and the asymptotic expression for $\langle M^2_{L-k,k} \rangle$ [eq. (3.53)] in eq. (3.59) gives:

$$S_2(t) \sim \sum_{k=0}^{t} \binom{t}{k} (1-D)^{t-k} D^k \left[\frac{4\eta^2 k\sqrt{L}}{\sqrt{\pi}} - \eta^2 k^2 + \frac{\eta^2 k^3}{3\sqrt{\pi}\sqrt{L}} + ... \right] \tag{3.60}$$

On performing the sum over k, an expression in powers of t is obtained. Since we are interested in the thermodynamic limit $L \to \infty$, we need to retain only the leading order term in L in the coefficient of each power of t. This finally gives:

$$S_2(t) \sim \frac{4\eta^2 L}{\sqrt{\pi}} \left[\left(\frac{Dt}{\sqrt{L}} \right) - \frac{\sqrt{\pi}}{4} \left(\frac{Dt}{\sqrt{L}} \right)^2 + \frac{1}{12} \left(\frac{Dt}{\sqrt{L}} \right)^3 + ... \right] \tag{3.61}$$

It is possible to compute higher order structure functions in exactly the same way, the only additional feature being that now the calculation of $g_j(t)$ involves expressions for $f_j(t)$ and so on [see eq. (3.57)].

In the scaling limit $t \to \infty$, $L \to \infty$, $t/\sqrt{L} \to 0$, we finally obtain:

$$S_n(t) \equiv \langle [M_L(t) - M_L(0)]^n \rangle \sim L^{n/2} \mathcal{G}_n \left(\frac{Dt}{\sqrt{L}} \right) \tag{3.62}$$

where:

$$\mathcal{G}_2(q) \sim \frac{4\eta^2}{\sqrt{\pi}} q \left[1 - \frac{\sqrt{\pi}}{4} q + \frac{1}{12} q^2 + ... \right] \tag{3.63a}$$

$$\mathcal{G}_3(q) \sim -6\eta^3 q \left[1 - \frac{2}{\sqrt{\pi}} q + \frac{1}{2} q^2 + ... \right] \tag{3.63b}$$

$$\mathcal{G}_4(q) \sim \frac{32\eta^4}{\sqrt{\pi}} q \left[1 - \frac{3\sqrt{\pi}}{4} q + q^2 + ... \right] \tag{3.63c}$$

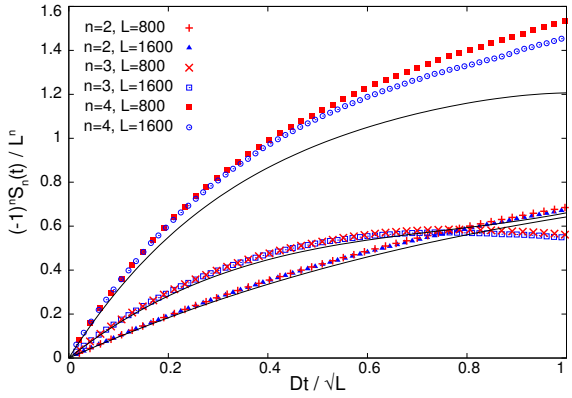


Figure 3.6: Structure functions: $(-1)^n S_n(t)/L^{n/2}$ vs. $Dt/\sqrt{L}$ for $n = 2, 3, 4$ and two different L. Solid lines represent the small t analytical expressions in eq. (3.63). Deviation of analytical predictions from numerics due to corrections to scaling which decay very slowly as $1/\sqrt{L}$ with system size L

Thus, this system also shows strong intermittency, but at time scales that are $\mathcal{O}(\sqrt{L})$, rather than $\mathcal{O}(L^2)$. Figure 3.6 shows $S_2(t)$, $S_3(t)$ and $S_4(t)$, as obtained from numerics along with the analytical predictions of eq. (3.63). The deviation of the analytical curves from the numerical data is due to corrections to scaling which decay very slowly as $1/\sqrt{L}$ with system size L. If we keep track of these sub-leading terms in our analysis, then the resultant expressions for $S_n(t)$ show good agreement with numerics. The flatness $\kappa(t)$ in the limit $t/\sqrt{L} \to 0$, as obtained from eq. (3.63) is:

$$\kappa(t) \sim 2\sqrt{\pi} \left(\frac{Dt}{\sqrt{L}}\right)^{-1} \left[1 - \frac{\sqrt{\pi}}{4}\left(\frac{Dt}{\sqrt{L}}\right) + \left(\frac{40-9\pi}{48}\right)\left(\frac{Dt}{\sqrt{L}}\right)^2 + ...\right] \tag{3.64}$$

This also shows an L-dependent divergence as $t \to 0$, but with a weaker dependence on L than in the diffusive case.

In conclusion, in this section we showed that the steady state distribution of total mass M has a Gaussian form, with rms fluctuations that scale as $\sqrt{L}$ with system size L. Further, by calculating dynamical structure functions, we demonstrate that the total mass shows strong intermittency in time in the limit $Dt/\sqrt{L} \to 0$.

3.6 Conclusions

In this chapter, we studied an extreme limit of the single species aggregation-fragmentation model discussed in Chapter 2, one in which the chipping rate is set to zero. We have analytically calculated various static and dynamical properties, focusing, in particular, on the computation of dynamical structure functions, in order to probe turbulence-like behaviour in the system. These calculations demonstrate that the system shows giant number fluctuations and temporal intermittency of the total mass in the case where aggregates undergo unbiased diffusion. We rationalise these results in terms of the formation and exit of a macroscopic aggregate containing a finite fraction of the total mass. Further, we show that this system is sensitive to boundary conditions: subtle differences arise if both influx and outflux are allowed at the same boundary, as opposed to opposite ends of the 1D system. We also analyse the case with fully biased movement. In this case, the typical mass of aggregates exiting the system scales as $\sqrt{L}$, which gives rise to normal, Gaussian fluctuations of the total mass. Nevertheless, the passage of such aggregates through the system is captured by the behaviour of dynamical structure functions which still exhibit intermittency, but at a different time scale.

While the analytic techniques presented in this chapter allow for the calculation of both stationary mass distributions and dynamical correlation functions, there are other important questions that we are not able to address using this approach. One such question is related to the properties of the macroscopic aggregate in the system. While we infer its presence in the unbiased case from the form of the tail of various probability distributions, we are not able to directly calculate its properties, for instance, by analysing the properties of the largest aggregate in the system. Apart from yielding more detailed information about this system, such a line of inquiry may also reveal interesting connections with earlier work which relates the properties of macroscopic aggregates to extreme value statistics [77].

Another important open question is related to the behaviour of the system in higher dimensions. Aggregation models typically have an upper critical dimension equal to two, above which the role of fluctuations becomes unimportant, and a mean field description suffices [28]. It would be interesting to investigate how the intermittency properties of the present model change with dimension, and whether this model also has the same upper critical dimension. As with other reaction-diffusion models in higher dimensions [31, 66, 67], field theoretic approaches may provide some insight into these questions.

Many of these results derived for the zero-fragmentation case, hold more generally for the aggregation-dominated phase of the full aggregation-fragmentation model that was discussed in Chapter 2. This general model, however, cannot be studied using the methods of this chapter, as the equations for $M_{i,j}$ in a single stretch are no longer closed in the presence of fragmentation. Finding analytical techniques that work for the full aggregation-fragmentation model, and using these to understand the position-dependent properties of the condensate phase, thus remains an interesting open problem.

An interesting direction for future study would be to explore whether other systems with cluster or aggregate formation also show temporal intermittency of mass or particle number. A broader question is whether there is a more general connection between temporal intermittency of particle number and giant number fluctuations, which often arise due to a 'clustering' tendency in a system. This question could be relevant for some of the aggregation models discussed in sec. 3.2.2, or other systems with giant fluctuations such as self-propelled particles [78, 79] and sliding particles on fluctuating interfaces [61, 80]. Dynamical structure functions of particle number and the time dependence of flatness could be useful probes in these systems as well.

3.7 Appendix A: Asymptotics

3.7.1 $P(M)$ for large M and large L [eq. (3.8), sec. 3.3.1]:

The generating function $Q(z) = \sum_{M=0}^{\infty} P(M) z^M$ of the probability distribution $P(M)$ in sec. 3.3.1 can be obtained by setting $y = 1$ in eq. (3.7). $Q(z)$ can be inverted to obtain $P(M)$ as follows:

$$\begin{aligned} Q(z) &= \operatorname{sech}\left[\sqrt{\lambda(1-z)}\right] \qquad \text{where} \qquad \lambda = \frac{2\tilde{a}L}{D} \\ &= \sum_{n=0}^{\infty} \frac{E_{2n}}{(2n)!}(1-z)^n \lambda^n \qquad \text{where } E_{2n} \text{ are Euler numbers} \\ &= \sum_{n=0}^{\infty} \frac{E_{2n}}{(2n)!}\lambda^n \sum_{M=0}^{n} \binom{n}{M} (-1)^M z^M \\ &= \sum_{M=0}^{\infty} \sum_{n=M}^{\infty} (-1)^M z^M \frac{E_{2n}}{(2n)!}\lambda^n \binom{n}{M} \end{aligned} \tag{3.65}$$

Thus, $P(M)$ is given by:

$$P(M) = (-1)^M \sum_{n=M}^{\infty} \frac{E_{2n}}{(2n)!} \lambda^n \binom{n}{M} \tag{3.66}$$

This is exact so far, but can be further simplified for large M by using the large n form of E_{2n}:

$$E_{2n} \sim 8(-1)^n \sqrt{\frac{n}{\pi}} \left(\frac{4n}{\pi e}\right)^{2n} \sim (-1)^n \frac{4}{\pi} \left(\frac{4}{\pi^2}\right)^n (2n)! \qquad \text{using Stirling's approximation} \tag{3.67}$$

Substituting from eq. (3.67) into eq. (3.66), we get:

$$\begin{aligned} P(M) &\sim \frac{4}{\pi}(-1)^M \sum_{n=M}^{\infty} \binom{n}{M} \left(\frac{-4\lambda}{\pi^2}\right)^n \\ &= \frac{4}{\pi} \left(\frac{4\lambda}{\pi^2}\right)^M \sum_{n=0}^{\infty} \binom{n+M}{n} \left(\frac{-4\lambda}{\pi^2}\right)^n \\ &= \frac{\frac{4}{\pi}\left(\frac{4\lambda}{\pi^2}\right)^M}{\left(1+\frac{4\lambda}{\pi^2}\right)^{M+1}} \\ &\sim \frac{\pi}{\lambda}\left(1-\frac{\pi^2}{4\lambda}\right)^M \qquad \text{for } M \gg 1 \text{ and } \frac{\pi^2}{4\lambda} \ll 1 \\ &\sim \frac{\pi}{\lambda} \exp\left(-\frac{\pi^2 M}{4\lambda}\right) \end{aligned} \tag{3.68}$$

which is the same as eq. (3.8).

3.7.2 Small t behaviour of $S_2(t)$ from eq. (3.36) [sec. 3.4.2]:

Equation (3.36) expresses the structure function $S_2(t)$ as the following infinite sum (ignoring the sub-leading terms in L):

$$S_2(t) = 16\zeta^2 \sum_{n=1,3,5\ldots}^{\infty} \left\{ \frac{1}{(n\pi)^4} \left(n\pi \coth\left[\frac{n\pi}{2}\right] - 1 \right) \right\} \left\{ 1 - \exp\left[-\frac{D\pi^2 n^2 t}{L^2} \right] \right\} \qquad \zeta = aL/D \tag{3.69}$$

To extract the small t ($Dt/L^2 \ll 1$) behaviour of this expression, we approximate the sum over n by an integral, and then take the $Dt/L^2 \to 0$ limit of this integral, so that:

$$\begin{aligned} S_2(t) &\sim 16\zeta^2 \int_0^{\infty} \frac{1-e^{-\pi^2(2l+1)^2\tau}}{\pi^3(2l+1)^3} \left\{ \coth\left[\frac{\pi}{2}(2l+1)\right] - \frac{1}{\pi(2l+1)} \right\} dl \qquad \text{where} \quad \tau = \frac{Dt}{L^2} \\ &= \frac{8\zeta^2\tau}{\pi} \int_{\pi\sqrt{\tau}}^{\infty} \left[\frac{1-e^{-x^2}}{x^3}\right] \left\{ \coth\left[\frac{x}{2\sqrt{\tau}}\right] - \frac{\sqrt{\tau}}{x} \right\} dx \qquad x = (2l+1)\pi\sqrt{\tau} \end{aligned} \tag{3.70}$$

In the limit $\tau \to 0$, we have $\coth[x/2\sqrt{\tau}] \sim 1$, so that $S_2(t)$ becomes:

$$S_2(t) \sim \frac{8\zeta^2\tau}{\pi} \int_{\pi\sqrt{\tau}}^{\infty} \left[\frac{1-e^{-x^2}}{x^3}\right]\left[1-\frac{\sqrt{\tau}}{x}\right] dx$$

$$= \frac{8\zeta^2\tau}{\pi}\left\{\left[\frac{1}{2\pi^2\tau} - \frac{e^{-\pi^2\tau}}{2\pi^2\tau} - \frac{1}{2}Ei[-\pi^2\tau]\right] + \left[\frac{e^{-\pi^2\tau}(1-2\pi^2\tau) - (1-2\pi^{7/2}\tau^{3/2}\,\mathrm{Erfc}[\pi\sqrt{\tau}])}{3\pi^3\tau}\right]\right\}$$

$$\text{where} \quad Ei[-\pi^2\tau] = -\int_{\pi^2\tau}^{\infty} \frac{e^{-u}}{u} du \quad \text{and} \quad \mathrm{Erfc}[\pi\sqrt{\tau}] = (2/\sqrt{\pi}) \int_{\pi\sqrt{\tau}}^{\infty} e^{-u^2} du \tag{3.71}$$

In the limit $\tau \to 0$, the complementary error function behaves as $\mathrm{Erfc}[\pi\sqrt{\tau}] \sim 1 - 2\sqrt{\pi\tau}$ and the exponential integral has the asymptotic form $Ei[-\pi^2\tau] \sim \gamma + \log[\pi^2\tau]$ where γ is the Euler-Mascheroni constant. Thus, as $\tau \to 0$, the expression in eq. (3.71) tends to:

$$S_2(t) \sim \frac{4\zeta^2\tau}{\pi}\left[1 - \gamma - \frac{2}{\pi} - \log[\pi^2\tau]\right] \tag{3.72}$$

This expression is not in very good agreement with numerics because of the significant corrections that appear while approximating the discrete sum with an integral. By taking these correction terms into account using the Euler-Maclaurin formula, better agreement with numerics is obtained. These correction terms basically modify the constants inside the square bracket in eq. (3.72), so that the small t form of $S_2(t)$ is still given by:

$$S_2(t) \sim \frac{4\zeta^2\tau}{\pi}(A_0 - \log[\tau]) = -\frac{4}{\pi}\left(\frac{aL}{D}\right)^2 \frac{Dt}{L^2} \log\left[A_1 \frac{Dt}{L^2}\right] \tag{3.73}$$

where the simplest way of obtaining the constant A_1 is by fitting to numerical data.

3.7.3 Asymptotic expressions for $\langle M_{i,j}^n \rangle$ [eq. (3.53), sec. 3.5.1]:

We start with the generating function $G_{u,j}(z)$ in eq. (3.51),

$$G_{u,j}(z) = \sum_{i=1}^{\infty} F_{i,j}(z) u^i = \sum_{i=1}^{\infty}\left[\sum_{M=0}^{\infty} P_{i,j}(M) z^M\right] u^i$$

$$= \left(\frac{u}{1-u}\right)\left(\frac{\eta^2(1-z)^2}{\eta^2(1-z)^2 - (1-u)}\right)\left[[1+\sqrt{1-u}]^{-j} - \left(\frac{1-u}{\eta^2(1-z)^2}\right)[1+\eta(1-z)]^{-j}\right] \tag{3.74}$$

The generating function $\sum_{i=1}^{\infty} \langle M_{i,j}^2 \rangle u^i$ can be obtained by differentiating the above expression w.r.t. z and then setting $z = 1$

$$\sum_{i=1}^{\infty} \langle M_{i,j}^2 \rangle u^i = \frac{\eta^2 u}{(1-u)^2}\left[2 - 2\{1+\sqrt{1-u}\}^{-j} + j(1+j)(1-u)\right] + \frac{\eta u j}{1-u} \tag{3.75}$$

Since we are interested in $\langle M_{i,j}^2 \rangle$ in the limit $i \to \infty$, we consider the $u \to 1$ limit of the above equation. This can be obtained by Taylor expanding eq. (3.75) in powers of the small

parameter $1-u$ and retaining only the first few terms (terms that become asymptotically large) in $1-u$. This gives:

$$\sum_{i=1}^{\infty}\langle M_{i,j}^2\rangle u^i \sim u\left[\frac{2\eta^2 j}{(1-u)^{3/2}}+\frac{\eta j}{1-u}+\frac{\eta^2 j(j^2+3j+2)}{3(1-u)^{1/2}}+...\right] \tag{3.76}$$

Each of the above terms can be now Taylor expanded about $u=0$ to give:

$$\langle M_{i,j}^2\rangle \sim 2\eta^2 j(2i)\left(\frac{(2i)!}{2^{2i}(i!)^2}\right)+\eta j+\frac{\eta^2 j(j^2+3j+2)}{3}\left(\frac{2i}{2i-1}\right)\left(\frac{(2i)!}{2^{2i}(i!)^2}\right)+.. \tag{3.77}$$

By taking the limit $i\to\infty$ and using Stirling's approximation for $i!$, eq. (3.77) becomes:

$$\langle M_{i,j}^2\rangle \sim \frac{4\eta^2 j\sqrt{i}}{\sqrt{\pi}}+\eta j+\frac{\eta^2 j(j^2+3j+2)}{3\sqrt{\pi}\sqrt{i}}+.... \tag{3.78}$$

The same procedure can be followed to compute $\langle M_{i,j}^3\rangle$ and $\langle M_{i,j}^4\rangle$ in the $i\to\infty$ limit: first, $\sum_{i=1}^{\infty}\langle M_{i,j}^3\rangle u^i$ and $\sum_{i=1}^{\infty}\langle M_{i,j}^4\rangle u^i$ are expanded in powers of $1-u$ and next, each of the $(1-u)^\alpha$ terms in this expansion is further expanded about $u=0$. Finally, by using Stirling approximation for $i!$ etc. in the $i\to\infty$ limit, we get:

$$\langle M_{i,j}^3\rangle \sim 6\eta^3 ji+\frac{12\eta^2}{\sqrt{\pi}}j\sqrt{i}+\eta j+\eta^3 j(j^2+3j+2)+\frac{\eta^2 j(j^2+3j+2)}{\sqrt{\pi}\sqrt{i}}+... \tag{3.79a}$$

$$\langle M_{i,j}^4\rangle \sim \frac{32\eta^4}{\sqrt{\pi}}(ji^{3/2})+36\eta^3(ji)+\frac{28\eta^2 j+8\eta^4 j(j^2+3j+2)}{\sqrt{\pi}}\sqrt{i}+6\eta^3 j(j^2+3j+2)+\eta j+... \tag{3.79b}$$

The expressions for $\langle M_{i,j}^n\rangle$ simplify further in the limit $j\to\infty$, $i\to\infty$ with $j/\sqrt{i}$ finite. In this limit, terms that are $\mathcal{O}(j/i)$ etc. go to zero, so that $\langle M_{i,j}^n\rangle$ are given by:

$$\langle M_{i,j}^2\rangle \sim i\mathcal{R}_2\left(\frac{j}{\sqrt{i}}\right)\sim\frac{4\eta^2 i}{\sqrt{\pi}}\left[\left(\frac{j}{\sqrt{i}}\right)+\frac{1}{12}\left(\frac{j}{\sqrt{i}}\right)^3+....\right] \tag{3.80a}$$

$$\langle M_{i,j}^3\rangle \sim i^{3/2}\mathcal{R}_3\left(\frac{j}{\sqrt{i}}\right)\sim 6\eta^3 i^{3/2}\left[\left(\frac{j}{\sqrt{i}}\right)+\frac{1}{6}\left(\frac{j}{\sqrt{i}}\right)^3+....\right] \tag{3.80b}$$

$$\langle M_{i,j}^4\rangle \sim i^2\mathcal{R}_4\left(\frac{j}{\sqrt{i}}\right)\sim\frac{32\eta^4 i^2}{\sqrt{\pi}}\left[\left(\frac{j}{\sqrt{i}}\right)+\frac{1}{4}\left(\frac{j}{\sqrt{i}}\right)^3+....\right] \tag{3.80c}$$

3.8 Appendix B: Solution of various partial differential equations:

3.8.1 Laplace equation on a right isosceles triangle [eqs. (3.9) and (3.32)]:

Consider a function $p(x,y)$ which:

i. satisfies the Laplace equation on a right isosceles triangle with vertices $(0,0)$, $(0,1)$ and $(1,0)$.

ii. is equal to zero uniformly on the hypotenuse. [5]

iii. satisfies some specified boundary conditions [mixed B.C. for eq. (3.9) and Neumann B.C. for eq. (3.32)] on the other two sides of the triangle.

The key step in solving this equation is the folding transformation introduced in [75] for the Laplace equation on the equilateral triangle, and adapted to the right isosceles triangle in [74]. This transformation, as applied to eqs. (3.9) and (3.32), is described below and also shown schematically in fig. 3.7.

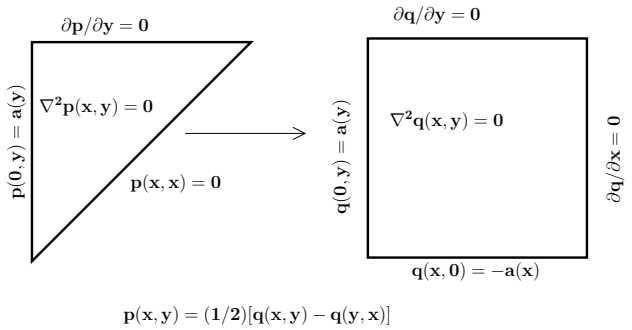


Figure 3.7: Folding transformation used to solve Laplace equation on the right isosceles triangle

We first consider the case, where the boundary conditions at $x = 0$ and $y = 1$ are of the mixed kind, as in eq. (3.9):

$$\frac{\partial^2 p(x,y)}{\partial x^2} + \frac{\partial^2 p(x,y)}{\partial y^2} = 0, \qquad 0 \le x \le y \le 1 \tag{3.81a}$$

$$p(x, y = x) = 0, \qquad \left.\frac{\partial p}{\partial y}\right|_{y=1} = 0, \qquad p(0,y) = a(y) \tag{3.81b}$$

Consider another function $q(x,y)$ on the square $[0 \le x \le 1,\ 0 \le y \le 1]$ which satisfies:

$$\frac{\partial^2 q(x,y)}{\partial x^2} + \frac{\partial^2 q(x,y)}{\partial y^2} = 0 \tag{3.82a}$$

$$q(0,y) = a(y) \qquad q(x,0) = -a(x) \qquad \left.\frac{\partial q}{\partial y}\right|_{y=1} = \left.\frac{\partial q}{\partial x}\right|_{x=1} = 0 \tag{3.82b}$$

Then, it can be seen that the function $w(x,y) = \frac{1}{2}[q(x,y) - q(y,x)]$ is the required solution of eq. (3.81) in the triangular region as:

$$\nabla^2 w(x,y) = \frac{1}{2}[\nabla^2 q(x,y) - \nabla^2 q(y,x)] = 0 \tag{3.83a}$$

[5] If the function is equal to some constant c on the hypotenuse, then we can define a new function $p(x,y) - c$ which satisfies all the three conditions (i)-(iii) and can be solved for using the method described subsequently.

$$w(x, y = x) = \frac{1}{2}[q(x, x) - q(x, x)] = 0$$
$$\left.\frac{\partial w}{\partial y}\right|_{y=1} = \frac{1}{2}\left(\frac{\partial q(x,y)}{\partial y} - \frac{\partial q(y,x)}{\partial y}\right)\Bigg|_{y=1} = \frac{1}{2}\left(\left.\frac{\partial q(x,y)}{\partial y}\right|_{y=1} - \left.\frac{\partial q(x,y)}{\partial x}\right|_{x=1}\right) = 0 \quad (3.83b)$$
$$w(0, y) = \frac{1}{2}[q(0, y) - q(y, 0)] = \frac{1}{2}[a(y) - (-a(y))] = a(y)$$

Equation (3.82) can be solved by a standard application of the superposition method [76] i.e. by decomposing $q(x, y)$ as $q(x, y) = u(x, y) + v(x, y)$, such that:

$$\nabla^2 u(x, y) = 0 \qquad \left.\frac{\partial u}{\partial y}\right|_{y=1} = \left.\frac{\partial u}{\partial x}\right|_{x=1} = 0 \qquad u(x, 0) = 0 \qquad u(0, y) = a(y) \qquad (3.84a)$$

$$\nabla^2 v(x, y) = 0 \qquad \left.\frac{\partial v}{\partial y}\right|_{y=1} = \left.\frac{\partial v}{\partial x}\right|_{x=1} = 0 \qquad v(x, 0) = -a(x) \qquad v(0, y) = 0 \qquad (3.84b)$$

Each of the functions $u(x, y)$ and $v(x, y)$ can be solved for by separation of variables. Their sum gives the function $q(x, y)$, which in turn gives $p(x, y)$.

The case where both the $x = 0$ and $y = 1$ boundary conditions are of the Neumann kind [as in eq. (3.32)] can be dealt with similarly using the folding transformation. Suppose, the boundary condition at $x = 0$ in eq. (3.81b) is given by $\left.\frac{\partial p}{\partial x}\right|_{x=0} = b(y)$, then the corresponding Laplace equation on the square that needs to be solved is:

$$\nabla^2 q(x, y) = 0, \qquad \left.\frac{\partial q}{\partial y}\right|_{y=1} = \left.\frac{\partial q}{\partial x}\right|_{x=1} = 0, \qquad \left.\frac{\partial q}{\partial x}\right|_{x=0} = b(y), \qquad \left.\frac{\partial q}{\partial y}\right|_{y=0} = -b(x) \quad (3.85)$$

Note that *all* the boundary conditions for this equation are of the Neumann kind. The Laplace equation with Neumann boundary conditions has a solution only if $\int_0^1 dy \left.\frac{\partial q}{\partial x}\right|_{x=1} - \int_0^1 dy \left.\frac{\partial q}{\partial x}\right|_{x=0} + \int_0^1 dx \left.\frac{\partial q}{\partial y}\right|_{y=1} - \int_0^1 dx \left.\frac{\partial q}{\partial y}\right|_{y=0} = 0$ i.e. there is no net flux through the boundaries [76, 81]. This is simply because the solution of the Laplace equation is the steady state solution of a heat equation and for a steady state solution to exist, the net heat flux through the boundaries must be zero. Thus, while decomposing $q(x, y)$ as $q(x, y) = u(x, y) + v(x, y)$, both $u(x, y)$ and $v(x, y)$ must individually satisfy the condition of no net flux through the boundaries. This can be done [81] by defining a new function:

$$s(x, y) = q(x, y) + (K/2)\left[(1 - x)^2 - (1 - y)^2\right] \qquad (3.86)$$

The function $s(x, y)$ now satisfies:

$$\nabla^2 s(x, y) = 0, \qquad \left.\frac{\partial s}{\partial y}\right|_{y=1} = \left.\frac{\partial s}{\partial x}\right|_{x=1} = 0, \qquad \left.\frac{\partial s}{\partial x}\right|_{x=0} = b(y) - K, \qquad \left.\frac{\partial s}{\partial y}\right|_{y=0} = -b(x) + K \qquad (3.87)$$

Choosing $K = \int_0^1 b(y)dy$ ensures that the net flux of $s(x,y)$ through each side of the square is zero. Then $s(x,y)$ can be decomposed as $s(x,y) = u(x,y) + v(x,y)$ such that:

$$\nabla^2 u(x,y) = 0, \qquad \left.\frac{\partial u}{\partial y}\right|_{y=1} = \left.\frac{\partial u}{\partial x}\right|_{x=1} = \left.\frac{\partial u}{\partial y}\right|_{y=0} = 0, \qquad \left.\frac{\partial u}{\partial x}\right|_{x=0} = b(y) - K \tag{3.88a}$$

$$\nabla^2 v(x,y) = 0, \qquad \left.\frac{\partial v}{\partial y}\right|_{y=1} = \left.\frac{\partial v}{\partial x}\right|_{x=1} = \left.\frac{\partial v}{\partial x}\right|_{x=0} = 0, \qquad \left.\frac{\partial v}{\partial y}\right|_{y=0} = -b(x) + K \tag{3.88b}$$

The functions $u(x,y)$ and $v(x,y)$ both satisfy the condition of zero net flux at the boundaries. Thus, solutions to eqs. (3.88a) and (3.88b) exist and can be found by separation of variables.

The folding transformation described above, can also be used to solve the 3D Laplace equation on the region $0 \le x \le y \le z \le 1$, which comes up in the calculation of multi-sector moments such as $\langle M_{xy}^2 M_{yz} \rangle$ [see sec. 3.4.1]. For example, suppose $g(x,y,z)$ satisfies the Laplace equation on the region $0 \le x \le y \le z \le 1$ with the boundary conditions $g(x,x,z) = g(x,y,y) = 0$ and some specified Dirichlet or Neumann boundary conditions on the other two surfaces. Then, we can find $g(x,y,z)$ by solving for a function $h(x,y,z)$ which satisfies the Laplace equation inside a cube $[0 \le x \le 1,\ 0 \le y \le 1,\ 0 \le z \le 1]$ and appropriately chosen boundary conditions on the six faces. If $h(x,y,z)$ is known, then $g(x,y,z)$ can be obtained as the antisymmetric combination:

$$g(x,y,z) = (1/6)[h(x,y,z) - h(y,x,z) + h(y,z,x) - h(z,y,x) + h(z,x,y) - h(x,z,y)]$$

This procedure can be generalised to solve Laplace equations on higher dimensional regions as well.

3.8.2 Inhomogeneous heat equation in 1 spatial dimension [eqs. (3.23) and (3.24)]

Equations (3.23) and (3.24) constitute a set of equations of the type:

$$\begin{aligned} &\frac{\partial p(x,t)}{\partial t} = \gamma \frac{\partial^2 p(x,t)}{\partial x^2} + a(x,t) \\ &p(0,t) = 0, \qquad \left.\frac{\partial p}{\partial x}\right|_{x=1} = 0, \qquad p(x,0) = b(x) \end{aligned} \tag{3.89}$$

The inhomogeneous heat equation can be solved [76] by expressing $p(x,t)$ as the sum of the complementary and particular solutions. The complementary solution, which is the solution of the homogeneous equation corresponding to eq. (3.89), has a variable separable form given by:

$$p_c(x,t) = \sum_{n=0}^{\infty} u_n \exp(-\gamma \alpha_n^2 t) \sin[\alpha_n x] \qquad \text{where } \alpha_n = \left(n + \frac{1}{2}\right)\pi \tag{3.90}$$

By assuming the particular solution to be of the form $\sum_{n=0}^{\infty} B_n(t) \sin[\alpha_n x]$, the general solution can be expressed as $p(x,t) = \sum_{n=0}^{\infty} [B_n(t) + u_n \exp(-\gamma \alpha_n^2 t)] \sin[\alpha_n x] = \sum_{n=0}^{\infty} C_n(t) \sin[\alpha_n x]$. The

inhomogeneous source term $a(x,t)$ in eq. (3.89) can also be written in the same eigenbasis as $a(x,t) = \sum_{n=0}^{\infty} A_n(t)\sin[\alpha_n x]$. Then, it follows that $C_n(t)$ satisfies:

$$\dot{C}_n(t) = -\gamma\alpha_n^2 C_n(t) + A_n(t) \tag{3.91}$$

This can be solved to give $C_n(t) = C_n(0)\exp[-\gamma\alpha_n^2 t] + \int_0^t A_n(t')\exp[-\gamma\alpha_n^2(t-t')]dt'$ where $C_n(0)$ can be obtained from the initial condition $p(x,0) = \sum_{n=0}^{\infty} C_n(0)\sin[\alpha_n x] = b(x)$.

3.8.3 Heat equation in 2 spatial dimensions on a right isosceles triangle with Neumann boundary conditions [eq. (3.35) and eq. (3.38)]:

Consider a function $p(x,y,t)$ which satisfies the heat equation on the triangular region $0 \le x \le y \le 1$:

$$\frac{\partial p(x,y,t)}{\partial t} = \gamma\left[\frac{\partial^2 p(x,y,t)}{\partial x^2} + \frac{\partial^2 p(x,y,t)}{\partial y^2}\right] \tag{3.92a}$$

$$p(x,y=x,t) = 0, \qquad \left.\frac{\partial p}{\partial y}\right|_{y=1} = 0, \qquad \left.\frac{\partial p}{\partial x}\right|_{x=0} = a(y,t) \qquad p(x,y,0) = b(x,y) \tag{3.92b}$$

As in the case of the Laplace equation, this can be solved by first solving the corresponding heat equation on a square:

$$\frac{\partial q(x,y,t)}{\partial t} = \gamma\left[\frac{\partial^2 q(x,y,t)}{\partial x^2} + \frac{\partial^2 q(x,y,t)}{\partial y^2}\right] \tag{3.93a}$$

$$\left.\frac{\partial q}{\partial y}\right|_{y=1} = \left.\frac{\partial q}{\partial x}\right|_{x=1} = 0, \qquad \left.\frac{\partial q}{\partial x}\right|_{x=0} = a(y,t) \qquad \left.\frac{\partial q}{\partial y}\right|_{y=0} = -a(x,t) \tag{3.93b}$$

where the initial condition at $t = 0$ is unspecified. If we can solve eq. (3.92) (upto undetermined constant coefficients corresponding to the unknown initial condition), then the function $p(x,y,t)$ can be obtained from this solution using $p(x,y,t) = \frac{1}{2}[q(x,y,t) - q(y,x,t)]$. The constant coefficients can now be determined from the initial condition $p(x,y,0) = b(x,y)$ by using appropriate Fourier transforms.

If $a(x,t) = 0$ [as in eq. (3.35)], then the boundary conditions are homogeneous and eq. (3.93) can be solved simply by separation of variables. If however, $a(x,t)$ is non-zero [as in eq. (3.38)], then an additional transformation is required. This involves defining a new function:

$$s(x,y,t) = q(x,y,t) + \frac{(1-x)^2}{2}a(y,t) - \frac{(1-y)^2}{2}a(x,t) \tag{3.94}$$

It can be checked that the function $s(x,y,t)$ now satisfies an inhomogeneous heat equation (with a source term) and homogeneous boundary conditions:

$$\begin{aligned}\frac{\partial s(x,y,t)}{\partial t} =& \gamma\left[\frac{\partial^2 s(x,y,t)}{\partial x^2} + \frac{\partial^2 s(x,y,t)}{\partial y^2}\right] \\ &+ \frac{(1-x)^2}{2}\left[\frac{\partial a(y,t)}{\partial t} - \gamma\frac{\partial^2 a(y,t)}{\partial y^2}\right] - \frac{(1-y)^2}{2}\left[\frac{\partial a(x,t)}{\partial t} - \gamma\frac{\partial^2 a(x,t)}{\partial x^2}\right] + \gamma[a(x,t) - a(y,t)]\end{aligned} \tag{3.95a}$$

$$\left.\frac{\partial s}{\partial y}\right|_{y=1} = \left.\frac{\partial s}{\partial x}\right|_{x=1} = \left.\frac{\partial s}{\partial y}\right|_{y=0} = \left.\frac{\partial s}{\partial x}\right|_{x=0} = 0 \tag{3.95b}$$

The transformation in eq. (3.94) is not a general prescription for solving eq. (3.93) for an arbitrary function $a(x,t)$. It works only when $\left.\frac{\partial a(x,t)}{\partial x}\right|_{x=1} = \left.\frac{\partial a(x,t)}{\partial x}\right|_{x=0} = 0$, which is the case for eq. (3.38).

Equation (3.95) is the heat equation with an inhomogeneous term. It can be solved, as in the 1D case, by first obtaining the complementary solution of the corresponding homogeneous equation and then obtaining the particular solution by appropriate Fourier transforms of the inhomogeneous source terms.

Chapter 4

The Multi-species Model with Interconversion

4.1 Introduction

Having analysed the single species open-boundary aggregation-fragmentation model in the previous chapters, we now come back to the full multi-species model, as defined in Chapter 1. To recapitulate, the two-species version of this model allows for influx of single particles of type 'A', conversion of particles of type A to type B and vice versa, chipping of single particles from stacks with species-dependent rates, diffusion and aggregation of multi-particle stacks, and outflux of single particles or whole stacks from the boundaries.

Due to the large number of processes in this system, it becomes necessary to consider only a subset of these processes at a time and analyse the effects that arise due to their interplay. In the previous chapters, we focused primarily on aggregation-fragmentation dynamics in the system by analysing the model with only one type of particles. In this chapter, we take the two extreme limits of the single-species model, namely the only-chipping limit and the only-stack-movement limit and generalise each of these separately to include multiple, interconverting species of particles. In particular, for both these limits, we ask the following question: how is the multi-species model different from the single species version and what new effects arise due to the interconversion between species? Based on our understanding of these two limits, we comment on the full multi-species model at the end of this chapter.

Our main finding is that in the only-chipping limit, i.e., in the absence of stack movement, the multi-species model undergoes a phase transition from a stationary phase characterised by a constant average number of particles, to a growing phase in which the total number of particles increases indefinitely in time. The growing phases have interesting spatial structure, with unbounded growth of one or more species of particles occurring in spatial regions whose width and location can be tuned (by varying parameters) , while steady state is attained in the rest of the system. Most of this chapter is devoted to a detailed analysis of the phases that arise in this limit. This analysis is carried out by solving equations for particle currents at each site. This work is also described in [82].

It is also possible to study the multi-species version of the only-stack-movement limit. If the rate of stack movement is taken to be a constant, independent of the relative amounts of A, B in the stack, then it is easy to see that the multi-species version is not very different from the single species model studied in Chapters 2 and 3. In fact, the only effect of having multiple, interconverting species is that the composition of aggregates becomes position dependent; the number, typical mass and spatial distribution of aggregates remain unaltered from the single-species case. If, on the other hand, we consider stack movement at a rate that depends on the composition of the stack, then the position-dependence of various quantities, including the mass at each site, is expected to be different from the single species case. However, numerics show that the main qualitative features are not affected. Thus, for simplicity, we will take the rate of stack movement to be independent of the composition of the stack.

The chapter is organised as follows. Sections 4.2-4.6 deal with the only-chipping model. In section 4.2, we define the only-chipping limit of the multi-species model precisely, state the main results for this limit and discuss these in the light of related models. Section 4.3 contains an analysis of a system containing just one site; this simple system also exhibits a phase transition between a stationary and a growing phase. In Sec. 4.4, we present a detailed study of the case of fully asymmetric (biased) single particle movement on a 1D lattice by solving a continuity equation at each site. In Sec. 4.5, we take a continuum limit in space and analyse the system for movement with some arbitrary asymmetry. In Sec. 4.6, several generalizations of the only-chipping model are discussed. In section 4.7, we return to the full multi-species model with chipping *and* stack movement, and comment briefly on some qualitative features of this model. Some details of the analysis of occupation probabilities are presented in appendices 4.8 and 4.9.

4.2 The Only-Chipping Model: Main Results

The main focus of this chapter is the only-chipping model, obtained by setting the rate of stack movement to zero. The model is defined on a 1D lattice. At any time t, a lattice site i has m_i^A particles of type A and m_i^B particles of type B. We start with an empty lattice with L sites at $t = 0$. At each instant, a site i is chosen at random and one of the following moves (illustrated in Fig. 4.1) occurs

i. injection of a single A particle at site 1 with rate a

ii. conversion of one of the A particles on site i to a B particle with rate u or conversion of an A particle to a B particle with rate v

iii. chipping (fragmentation) of a single A particle from site i and movement to the right (left) neighbouring site with rate w (w'); chipping and movement of a single B particle to the right (left) with rate q (q')

iv. outflux (exit) of single A (B) particles from site L with rate w (q) or from site 1 with rate w' (q'), i.e., with the same rate as in the bulk.

In general, A particles and B particles can move with a different degree of bias. However, for simplicity, we take $w/w' = q/q'$. Then we can redefine the chipping rates in terms of a common asymmetry factor γ as follows: $w \to \gamma w$, $w' \to (1-\gamma)w$; $q \to \gamma q$, $q' \to (1-\gamma)q$.

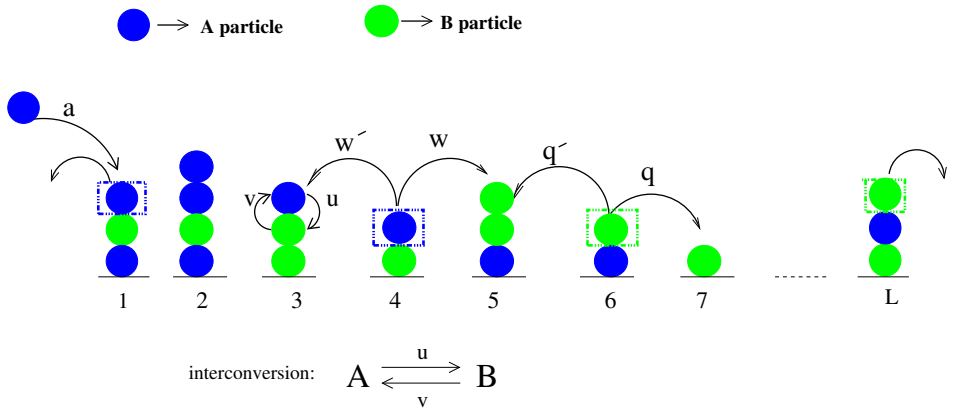


Figure 4.1: Illustration of the allowed moves in the only-chipping model: injection of single particles at site 1 at rate A; interconversion of single particles $A \to B$ at rate u and $B \to A$ at rate v; chipping of an A (B) particle and rightward hopping with rate w (q) or leftward hopping with rate w' (q'); outflux of single particles from left and/or right boundary with same rate as in the bulk

Here, γ is the measure of the asymmetry (or bias) in movement and takes on values between 0 and 1. The model is thus parametrized by the rates a, w, q, u, v and the asymmetry factor γ, where $\gamma = 1/2$ corresponds to symmetric and $\gamma = 1$ to fully asymmetric rightward movement. Note that chipping and interconversion processes take place at rates that are independent of the number of particles at a site.

The model defined above exists in qualitatively different phases in different regions of parameter space. In order to discuss these, it is first useful to define the occupation probability s_i^A (s_i^B) as the probability of finding at least one A (B) particle at a site i. Formally,

$$s_i^A(t) = 1 - \sum_{m^B=0}^{\infty} P_i(0, m^B, t)$$
$$s_i^B(t) = 1 - \sum_{m^A=0}^{\infty} P_i(m^A, 0, t) \tag{4.1}$$

where $P_i(m^A, m^B, t)dt$ is the probability of finding m^A A particles and m^B B particles on site i between time t and $t+dt$. Similarly, we define $\langle m_i^A(t)\rangle$ ($\langle m_i^B(t)\rangle$) as the average number of A (B) particles at site i at time t. To analyse the model, we write the time evolution equations for the average masses $\langle m_i^A(t)\rangle$ and $\langle m_i^B(t)\rangle$ at each site in terms of s_i^A and s_i^B. In solving these equations, we sometimes take the continuum limit $i/L \to x$, such that $\langle m_i^A\rangle \to \langle m^A(x)\rangle$, $\langle m_i^B\rangle \to \langle m^B(x)\rangle$, $s_i^A \to s^A(x)$ and $s_i^B \to s^B(x)$.

The main results that emerge from this analysis can be summarised as follows:

i. At long times, depending on the values of various parameters, the system may exist either in a *steady* phase, in which the total mass in the system (and at each site) reaches a finite time-independent value, or in a *growing phase* characterised by indefinitely growing mass of A and/or B particles in some regions of the system.

ii. In the *steady phase*, s_i^A and s_i^B reach time independent values *less than 1*, at each site i. Also, the particle currents due to chipping, interconversion etc. balance out at each site, resulting in finite, time-independent values of $\langle m_i^A\rangle$ and $\langle m_i^B\rangle$ for all i.

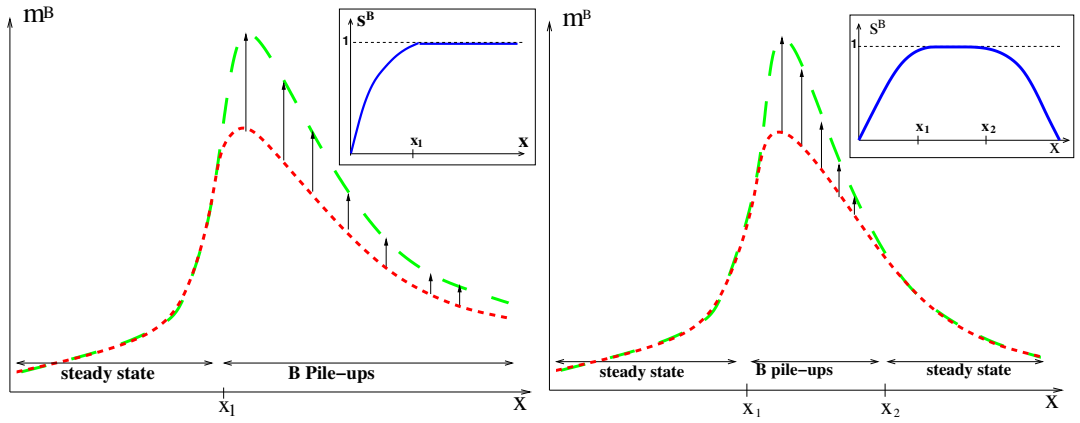


Figure 4.2: Schematic depiction of a phase with pile-ups of B for (a) fully asymmetric ($\gamma = 1$) (b) fully symmetric ($\gamma = 1/2$) hopping. Dashed lines show $\langle m^B(x,t)\rangle$ as a function of x at two time instants t_1 (small dashes) and t_2 (big dashes) respectively such that $0 \ll t_1 < t_2$. In regions of steady state, $\langle m^B(x,t_2)\rangle = \langle m^B(x,t_1)\rangle$, while $\langle m^B(x,t_2)\rangle > \langle m^B(x,t_1)\rangle$ in regions of B pile-ups. Insets: Occupation probability s^B as a function of x. While $s^B = 1$ in regions of B pile-ups, $s^B < 1$ in regions of steady state. For $\gamma = 1/2$, the region with pile-ups ($x_1 < x < x_2$) lies between two regions in steady state; for $\gamma = 1$, the region with pile-ups ($x > x_1$) lies entirely to the right of the steady state region.

iii. In the *growing phase*, there exist sites where the 'out' current always falls short of the 'in' current, causing the average A (B) mass at the site to grow indefinitely in time. We refer to this unbounded growth as an A (B) *pile-up* at the site. At the site with an A (B) pile-up, the mean occupancy $s^A(s^B)$ approaches 1 at long times and the average mass $\langle m^A\rangle$ and $\langle m^B\rangle$ increases at a finite rate indefinitely.

iv. In the growing phase, a system may have pile-ups of A, of B or of both species. While an A pile-up can be found only on site 1 (except in the special case $\gamma = 1$, $q = 0$), B pile-ups can also occur in the bulk. In a phase with B pile-ups, an interesting situation can arise where the lattice has a *region in steady state (with finite mean mass) coexisting with a region of B pile-ups (where the mean mass grows indefinitely)*. For fully asymmetric movement, the region with pile-ups always lie to the right of a region in steady state (shown schematically in Fig 4.2a), while for symmetric or partially asymmetric movement, the region with pile-ups lies between two regions in steady state (fig. 4.2b). For B pile-ups to occur a macroscopic (i.e. proportional to L) distance away from the left boundary, interconversion rates should be much smaller than chipping rates: specifically, we require $u/w + v/q$ to be $\mathcal{O}(1/L)$ (or $\mathcal{O}(1/L^2)$) for asymmetric (or symmetric) chipping.

In the rest of the chapter, we will adopt the following terminology. If the average mass of both A and B reaches a finite time-independent value, we will refer to the system as being in a steady phase. If one region of the system attains constant average mass and another region shows unbounded growth, we will refer to the two regions as a region in steady state and a region with pile-ups respectively. The system as a whole will be referred to as being in a growing phase.

The occurrence of pile-ups in this model is a simple consequence of the fact that the chipping and interconversion rates are independent of m_i^A and m_i^B. For such rates, the particle currents are constrained by an upper bound. and hence, may fail to balance in certain regions of the system for some parameters [1]. Pile-ups are also known to occur in the single species version of this model with generalised rates [43], for a similar reason. However, in the single-species model, linearly growing pile-ups can be found only at the boundary sites. The occurrence of B pile-ups anywhere in the bulk, in our model, is a direct consequence of the interconversion between different species. Interconversion causes the net A (B) current in any bond of the lattice to show non-trivial spatial variation. Thus, a situation can arise where currents fail to balance only in a certain region of the system. A similar effect is also seen in models with hard core particles such as the two lane Asymmetric Exclusion Process (ASEP) with switching of particles between lanes [55, 56]. Lane switching, which is the equivalent of interconversion in this model, gives rise to phase co-existence in the ASEP: the system develops a localised shock separating a low density region and a high density region. Our model too shows co-existence of two phases of one species (a steady phase and a growing phase) in different spatial regions separated by a domain wall with a tunable location.

4.3 Analysis of a single site system

Before studying this model for the 1D lattice, we analyse the behaviour of a system consisting of just *one* site; this analysis yields considerable insight into the full problem. It turns out that the single site system can exhibit four possible behaviours: steady state; pile-up of A but not B; pile-up of B but not A; pile-ups of both A and B. The results derived here also hold for the first site of a system with fully asymmetric (rightward) hopping ($\gamma = 1$), as in that case, a site is not affected by any sites to the right of it.

The master equation for the probability distribution $\mathcal{P}(m^A, m^B)$ (as defined in section 4.2) at the first site can be written as follows:

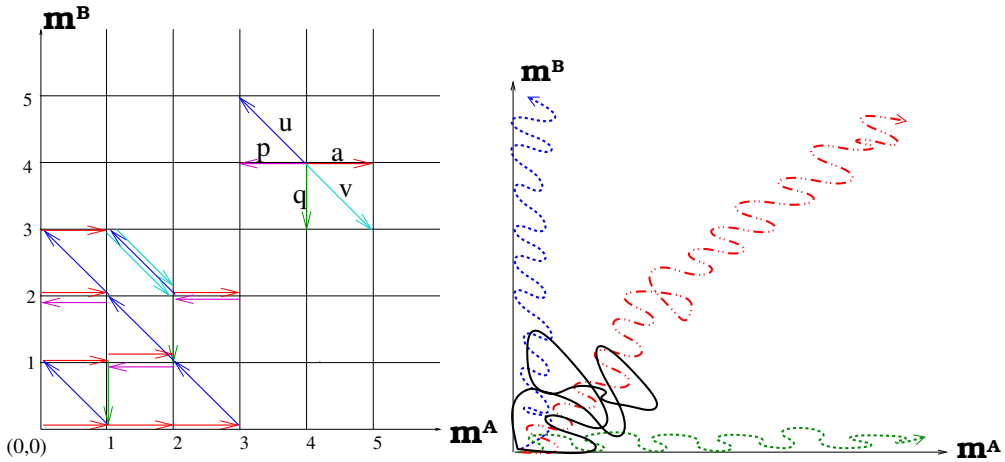
$$\begin{aligned}\frac{\partial \mathcal{P}(m^A, m^B)}{\partial t} =& w\mathcal{P}(m^A+1, m^B) + q\mathcal{P}(m^A, m^B+1) + a\mathcal{P}(m^A-1, m^B) + u\mathcal{P}(m^A+1, m^B-1) \\ &+ v\mathcal{P}(m^A-1, m^B+1) - (a+u+v+w+q)\mathcal{P}(m^A, m^B) \\ & \qquad\qquad\qquad\qquad\qquad\qquad m^A \geq 1 \text{ and } m^B \geq 1\end{aligned} \tag{4.2a}$$

$$\frac{\partial P(0, m^B)}{\partial t} = wP(1, m^B) + qP(0, m^B+1) + uP(1, m^B-1) - (a+v+q)P(0, m^B) \qquad m^B \geq 1 \tag{4.2b}$$

$$\begin{aligned}\frac{\partial \mathcal{P}(m^A, 0)}{\partial t} =& w\mathcal{P}(m^A+1, 0) + q\mathcal{P}(m^A, 1) + a\mathcal{P}(m^A-1, 0) + v\mathcal{P}(m^A-1, 1) \\ &- (a+u+w)\mathcal{P}(m^A, 0) \qquad\qquad m^A \geq 1\end{aligned} \tag{4.2c}$$

$$\frac{\partial P(0,0)}{\partial t} = wP(1,0) + qP(0,1) - aP(0,0) \tag{4.2d}$$

[1] Note that the particle currents themselves attain stationary values at long enough times. An indefinitely growing pile-up arises at a given site simply because the particle currents, though stationary, are not balanced at that site.



(a) (b)

Figure 4.3: Time evolution on the first site can be mapped to random walk in 2D (m^A, m^B) space. (a) A possible random walk in (m^A, m^B) space. Allowed moves are enumerated in the upper right corner of the figure along with the corresponding rates at which they occur.
(b) Trajectories in (m^A, m^B) space corresponding to various scenarios. Trajectory in black corresponds to steady state, the green one directed along m^A axes to pile-up of A, the blue one along m^B axes to pile-up of B, and the fourth (red) trajectory to pile-up of both A and B.

We are not able to solve the above master equation, but instead work with the time-evolution equations for $\langle m^A \rangle$ and $\langle m^B \rangle$:

$$\frac{d\langle m^A \rangle}{dt} = a + vs^B - us^A - ws^A \tag{4.3a}$$

$$\frac{d\langle m^B \rangle}{dt} = us^A - vs^B - qs^B \tag{4.3b}$$

Here $\langle m^A \rangle$ and $\langle m^B \rangle$ denote the average masses of A and B respectively on site 1 averaged over time histories and s^A and s^B the occupation probabilities of A and B (defined in eq. (4.1)); all of these are functions of time. Evidently, eq. (4.3) is just a continuity equation for the average masses.

The time-evolution of the first site can also be mapped to a random walk in two dimensions. The two-dimensional (2D) space here is the (m^A, m^B) space. The random walk starts at the origin at $t = 0$, moves in both positive and negative directions along the m^A axis (with rates a and w respectively) but moves only in the negative direction along the m^B axis (with a rate q) as there is no injection of B particles. Diagonal moves, corresponding to interconversion, are also allowed *but only in the direction of increasing* m^A *and decreasing* m^B *or vice versa*. The m^A and m^B axes are reflecting boundaries. A possible random walk, with all allowed moves in (m^A, m^B) space, is illustrated in Fig. 4.3a. Clearly, the random walk cannot be decomposed into independent walks along the m^A and m^B axes because of the diagonal moves.

The random walk picture is useful in the following way. To begin with, let us consider a simple 1D random walk on the positive half of the x-axis with a reflecting barrier at the

origin. This walk has a drift velocity c and a diffusion constant D. After a length $\sim D/|c|$, drift becomes more important than diffusion. If $c > 0$, then beyond this length, the mean displacement will grow linearly with time, while if $c < 0$, the mean displacement reaches a constant value $\sim D/|c|$. If $c = 0$, then the motion is always diffusive, and the mean displacement grows as $\sqrt{t}$.

The 2D walk illustrated in fig. 4.3a is more complicated because motion along m^A and m^B axes is coupled due to the diagonal moves (corresponding to interconversion). The diagonal moves effectively make c and D in the m^A direction dependent on the displacement of the walker along the m^B direction. If the effective c in the m^A direction is greater than or equal to zero everywhere, then we get a pile-up of A on the first site (with $c = 0$ implying $\sqrt{t}$ growth of the pile-up and $c > 0$, growth that is linear in time). Also, as expected from the theory of random walks, the fluctuation about the mean mass, i.e. $\Delta m^A = \sqrt{\langle [m^A]^2 \rangle - \langle m^A \rangle^2}$ grows as $\sqrt{t}$ (for $c \geq 0$). If $c < 0$, then the mean mass $\langle m^A \rangle$ and the rms fluctuations Δm^A reach time-independent values. The above can also be recast in terms of the occupation probability: $s^A < 1$ corresponds to finite $\langle m^A \rangle$, and $s^A = 1$ to $\langle m^A \rangle$ that is infinite at long times. To see this, note that the return from $m^A = M$ to $m^A = 0$ requires M steps, implying that $\mathcal{P}(m^A = 0)$ falls as $e^{-\alpha M}$, which is non-vanishing only for finite M. Thus, if $s^A = 1 - \mathcal{P}(m^A = 0)$ is less than 1 at long times, $\langle m^A \rangle$ must be finite; if s^A is equal to 1 at long times, $\langle m^A \rangle$ approaches infinity, i.e. it keeps increasing indefinitely, even at long times.

From the above discussion it is clear that depending on the rates, the effective c in either direction (i.e. along m^A and m^B axes) can be negative or non-negative, resulting in four possibilities altogether. Figure 4.3b shows schematically, possible random walks corresponding to the four scenarios – steady state, pile-up of A, pile-up of B and pile-up of both A and B.

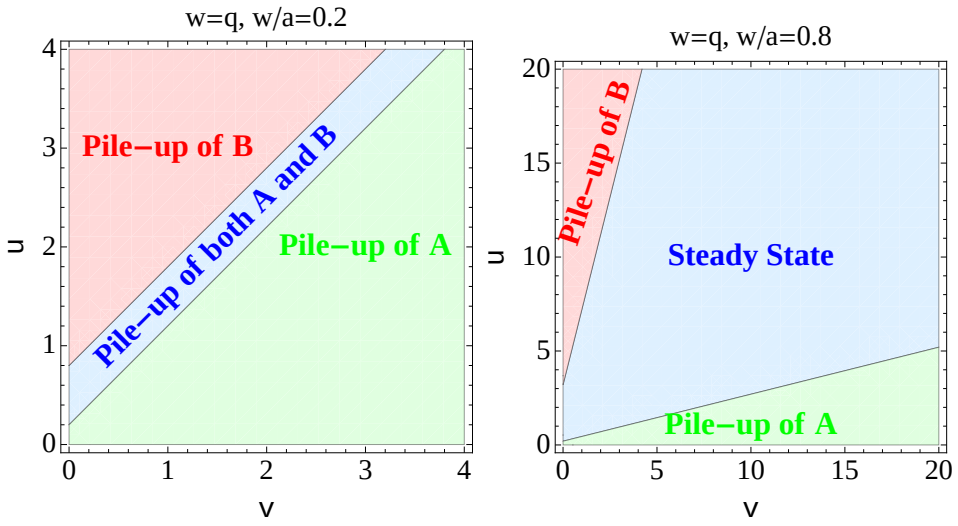
From eq. (4.3), we can determine the regions in parameter space that correspond to each of the four scenarios. For example, the occurrence of a pile-up of B but not of A at site 1 implies at long times, $d\langle m^B \rangle/dt \geq 0$ with $s^B = 1$ and $d\langle m^A \rangle/dt = 0$ with $s^A < 1$. Substituting these into eq. (4.3) gives the two inequalities in parameter space that must be satisfied in this case. The conditions corresponding to the other scenarios can be derived similarly and are stated below:

i. *Steady state* attained for $\frac{a(v+q)}{qw+qu+wv} < 1$ and $\frac{au}{qw+qu+wv} < 1$.
At long times; $s^A = \frac{a(v+q)}{qw+qu+wv}$ and $s^B = \frac{au}{qw+qu+wv}$; $\langle m^A \rangle$ and $\langle m^B \rangle$ are finite.

ii. *Pile-up of B but not A* for $\frac{au}{qw+qu+wv} \geq 1$ and $\frac{a+v}{u+w} < 1$
At long times: $s^A = \frac{a+v}{u+w}$ and $s^B = 1$; $\langle m^A \rangle$ is finite and $\frac{d\langle m^B \rangle}{dt} = \frac{ua-vw}{u+w} - q$

iii. *Pile-up of B but not A* for $\frac{a(v+q)}{qw+qu+wv} \geq 1$ and $\frac{u}{v+q} < 1$,
At long times: $s^B = \frac{u}{v+q}$ and $s^A = 1$; $\frac{d\langle m^A \rangle}{dt} = a - w - \frac{uq}{v+q}$ and $\langle m^B \rangle$ is finite.

iv. *Pile-ups of both A and B* for $\frac{a+v}{u+w} \geq 1$ and $\frac{u}{v+q} \geq 1$.
At long times: $s^B = s^A = 1$, $\frac{d\langle m^A \rangle}{dt} = a + v - u - w$ and $\frac{d\langle m^B \rangle}{dt} = u - v - q$



(a) (b)

Figure 4.4: (v, u) phase diagram for the first site when A and B chipping rates are equal (a) $w = q$, $w/a = 0.2$ (b) $w = q$, $w/a = 0.8$

The equalities corresponding to the inequalities in (i)-(iv) above specify the phase boundaries in parameter space between the four different 'phases'. These can be visualised more easily by considering a reduced parameter space obtained by setting $w = q$, in which case, particles of both species A and B chip with the same rate. The total mass at the site then evolves according to:

$$\frac{d[\langle m^A \rangle + \langle m^B \rangle]}{dt} = a - w(s^A + s^B) \tag{4.4}$$

Clearly, if $a < w$, then s^A, $s^B < 1$ and the system always attains steady state. If $a \geq 2w$, the system can never attain steady state and must show pile-ups (as $d[\langle m^A \rangle + \langle m^B \rangle]/dt \geq 0$ even when s^A and s^B take on the maximum possible value i.e. 1). In this case, if $u \sim v$, i.e. the interconversion rates are similar, both A and B pile up. However, if one of the conversion rates is much higher than the other, then the species which converts fast is able to stabilize, so that particles of only one species piles up. Figure 4.4a shows the (v, u) phase diagram for the case $a \geq 2w$. Finally, in the case where $w \leq a < 2w$, the occupation probabilities s^A and s^B can take on values such that steady state is attained. However, as before, if one of the species converts much faster than the other, it drives the latter out of steady state, causing it to pile up. Figure 4.4b shows the (v, u) phase diagram for the case $w \leq a < 2w$.

In conclusion, the single site is an instructive case to study as it exhibits all the different kinds of behaviour that occur in the full system, namely, steady state or growing phase, pile-up of one or both species. It also demonstrates how the analysis of simple continuity equations can yield information about the phase of the system and also predict phase boundaries; this is the approach we follow for the analysis of the full 1D system as well. A limitation of this kind of analysis is that it cannot be used to calculate the average mass $\langle m^A \rangle$ and $\langle m^B \rangle$. In the steady state, we know that the average mass is finite, while in the growing phase it grows linearly (where our analysis yields the rate of growth). Numerics show that the transition

from the steady phase to a pile-up is associated with a divergence of the mean mass, of the form $m^{A/B} \propto (1 - s^{A/B})^{-1}$ as $s^{A/B} \to 1$ [2]. Thus, close to the phase boundary, the system supports a large but finite aggregate.

4.4 Fully Asymmetric Chipping on a 1D lattice

We now consider the full 1D lattice with L sites, with chipping and interconversion in the bulk and influx and outflux at the boundary sites. As this system is open at the boundaries, it lacks translational invariance and can in fact, exhibit nontrivial position-dependent properties. Here, we study these for the case with fully asymmetric movement of particles ($\gamma = 1$). In this limit, the equations involving s_i^A and s_i^B turn out to have a simple solution. Also, this case illustrates in a clear way, how interconversion drives spatial gradients in A and B particle currents, leading to 'phase-coexistence' in the growing phase.

We begin by writing time-evolution equations for $\langle m_i^A \rangle$ and $\langle m_i^B \rangle$, similar to eq. (4.3), at each site in the system:

$$\frac{d\langle m_i^A \rangle}{dt} = w s_{i-1}^A + v s_i^B - u s_i^A - w s_i^A \tag{4.5a}$$

$$\frac{d\langle m_i^B \rangle}{dt} = q s_{i-1}^B + u s_i^A - v s_i^B - q s_i^B \qquad i \neq 1 \tag{4.5b}$$

$$\frac{d\langle m_1^A \rangle}{dt} = a + v s_1^B - u s_1^A - w s_1^A \tag{4.5c}$$

$$\frac{d\langle m_1^B \rangle}{dt} = u s_1^A - v s_1^B - q s_1^B \tag{4.5d}$$

The above equations are exact, being just the continuity equations for the average masses. They can also be obtained by writing a master equation for the total probability $\mathcal{P}(\{m_i^A\}, \{m_i^B\})$, multiplying it by m_i^A (or m_i^B) and averaging over all configurations.

4.4.1 Steady Phase:

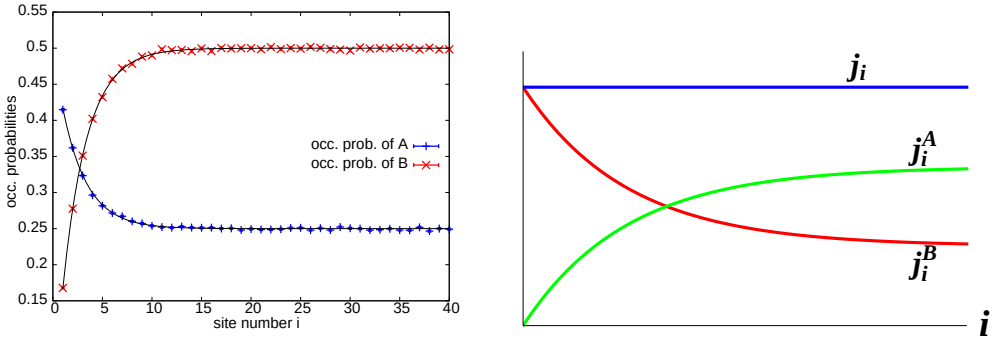
In the steady phase, $\langle m_i^A \rangle$ and $\langle m_i^B \rangle$ reach finite time-independent values, so that all time derivatives in equations (4.5) can be set to zero. This gives recursion relations relating s_i^A and s_i^B to s_{i-1}^A and s_{i-1}^B. Iterating these recursion relations yields the following steady phase occupation probabilities:

$$\begin{aligned} s_i^A &= \frac{av}{uq + vw}\left[1 + \left(\frac{uq}{vw}\right)\lambda^i\right] \\ s_i^B &= \frac{au}{uq + vw}\left[1 - \lambda^i\right] \end{aligned} \tag{4.6}$$

where $\lambda = (wq)/(uq + wv + wq)$.

Figure 4.5a shows s_i^A and s_i^B as a function of lattice site i for a system in the steady phase, as obtained from Monte Carlo simulations, along with the analytical results [eq. (4.6)].

[2] This can also be shown exactly for the single species case without interconversion



(a) (b)

Figure 4.5: (a) Occupation probabilities s_i^A and s_i^B as a function of site number i in a steady phase for $\gamma = 1$, $a = 1$, $u = 0.5$, $v = 0.25$, $w = 2$, $q = 1$. Data points are obtained from Monte Carlo simulations, and solid lines are plots of the analytical results in eq. (4.6). s^A and s^B approach their asymptotic values over a length l_D defined in eq. (4.7).
(b) Schematic diagram showing variation of the particle currents j_i^A, j_i^B and j_i with distance i from left boundary for a system in steady phase. While j_i^A and j_i^B change with i, the total particle current j_i remains constant.

Several features of the steady phase can be deduced from the expressions in eq. (4.6):

i. A necessary and sufficient condition for the system to be in steady phase is that the following two inequalities are satisfied: $a(v+q)/(uq+wv+wq) < 1$ and $au/(uq+wv) < 1$. To see this, note that if the system is in the steady phase, the occupation probabilities s_i^A, s_i^B, as given by (4.6), must be less than 1 for all i. This is ensured if the maxima of the two expressions in (4.6) are less than 1, thus yielding the two inequalities given above.

ii. Away from the left boundary, s_i^A and s_i^B approach asymptotic values that satisfy the relation $s_i^A/s_i^B = v/u$ which corresponds to zero interconversion current at site i.

iii. The characteristic length l_D associated with this (exponential) approach is given by:

$$l_D = \left[\ln\left(\frac{1}{\lambda}\right)\right]^{-1} \tag{4.7}$$

which is $(v/q + u/w)^{-1}$ to first order in interconversion rates.

The above can also be rationalised in terms of the behaviour of the net particle currents j_i^A, j_i^B and $j_i = j_i^A + j_i^B$ in the i^{th} bond of the lattice (i.e. from site i to $i+1$) and their dependence on i (depicted schematically in fig. 4.5b). In the absence of interconversion, the particle current j_i^A is the same in each bond, and so is j_i^B. Thus, s_i^A and s_i^B are constant across the lattice, being a/w and 0 respectively. Interconversion changes j_i^A and j_i^B and consequently s_i^A and s_i^B. The particle current into site 1 (j_0) is entirely of type A. At site 1, some of the A particles convert to type B, resulting in a small probability that site 1 also has B particles. Hence, j_1 is different in composition, having a small B component as well. At site 2, there is still a much larger concentration of A particles than B particles, and hence a

net conversion of A to B particles. Thus, the composition of j_2 shifts further towards B. This process continues until we reach a site at which there is *no net conversion*. Beyond this site j_i^A and j_i^B do not change with i.

If interconversions are faster than chipping (i.e. $u, v \gtrsim w, q$), then at a given site, the relative amounts of A and B can alter significantly through interconversion before chipping occurs. This is also reflected in the expression for the characteristic length l_D that governs the spatial gradients (see eq. (4.7)). In particular, the first order term has the following interpretation: $1/w$ is the typical time scale over which a site loses a single A particle due to chipping; the ratio u/w is thus proportional to the average number of conversion events (of A particles) that take place over this time scale. The term v/q has a similar interpretation. Thus, $(v/q + u/w)$ is an estimate of the average number of interconversions taking place in the system in a time interval over which a single chipping event occurs. When this number is large, then j_i^A and j_i^B, and hence s_i^A and s_i^B change sharply and reach their asymptotic values over just a few sites from the left boundary. Conversely, when this number is proportional to $1/L$, the length l_D is macroscopic and the gradients in s^A, s^B extend over the bulk of the lattice. (Interconversion rates that are $\mathcal{O}(1/L)$ were also considered in [55, 56].)

4.4.2 Growing Phases:

As discussed in section 4.4.1, the system is in steady phase only if the parameters satisfy certain inequalities. If these inequalities are not satisfied, then the s_i^A and s_i^B, as derived in eq. (4.6) will be greater than 1, for some i, which is clearly unphysical. This implies that the steady state assumption used in deriving eq. (4.6) is not valid, and $d\langle m_i^A\rangle/dt$ and/or $d\langle m_i^B\rangle/dt$ must be non-zero for at least some i at long times. Thus, for such parameters, the system exists in a growing phase characterised by pile-ups of A and/or B at certain sites.

Many features of these growing phases can be deduced from the time-evolution equations (4.5) which are represented diagrammatically in fig. 4.6a. The figure depicts each site in the system as having an A and B compartment, with the arrows representing particle currents between the various compartments. Vertical arrows correspond to intra-site interconversion currents and horizontal arrows to the inter-site chipping currents j_i^A and j_i^B discussed earlier. The currents into and out of any compartment can either balance each other (resulting in a steady state for the compartment) or the in-current can be more than the out-current (resulting in a pile-up). In steady state regions, both j_i^A and j_i^B vary with site i, but in such a way that the total particle current $j_i = j_i^A + j_i^B$ remains independent of i. In contrast, in a region with pile-ups, the net in-current at each site is more than the net out-current; thus, j_i is no longer independent of i but decreases with increasing i (as depicted schematically in fig. 4.6b). In fact, steady and growing phases can also be distinguished by the qualitatively different spatial profile of j_i in the two phases (figs. 4.5b, 4.6b).

Evidently, the out-current cannot exceed the in-current at any site as this would eventually lead to negative mass. This simple observation can be used to establish the following result (see appendix 4.8 for proof):
In both steady and growing phases, if $q \neq 0$, *then* $s_{i+1}^A < s_i^A$ *and* $s_{i+1}^B \geq s_i^B \ \forall i$.
Consequently, $us_{i+1}^A - vs_{i+1}^B < us_i^A - vs_i^B$, i.e. interconversion current keeps decreasing as we move right, asymptotically approaching zero.

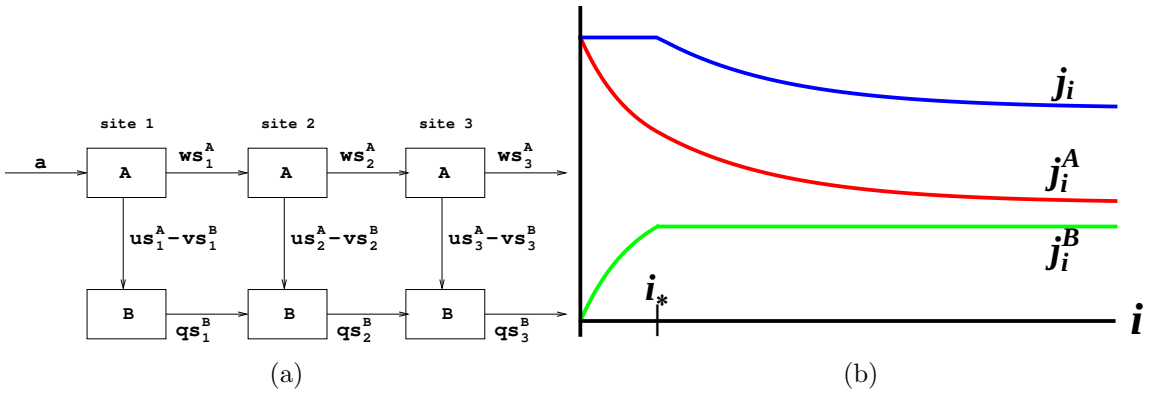


Figure 4.6: (a) Each site can be thought of as having A and B compartments. Exchange of particles takes place between compartments at a give site, i.e. through interconversion (represented by vertical arrows). Exchange of particles takes place between sites through chipping (represented by horizontal arrows).
(b) Schematic diagram showing variation of the particle currents j_i^A, j_i^B and j_i with i for a system in a growing phase. For the steady state region in the system ($i < i_*$), the total current j_i is independent of i, while in the region with pile-ups ($i > i_*$), j_i decreases with i.

The above inequalities for the occupation probabilities have the following implications (for $q \neq 0$):

i. An A pile-up can only occur at the first site, and not at any other site. This follows from the fact that s_i^A is a strictly decreasing function of i; thus, even if $s_1^A = 1$, i.e., there is an A pile-up at the first site, s_i^A must be less than 1 for $i > 1$, implying that there cannot be an A pile-up at subsequent sites.

ii. If there is a pile-up of B at some site k, then all sites i to the right of k also show B pile-ups. to see this, note that if $s_k^B = 1$, then $s_i^B = 1$ for $i > k$ as s_i^B is an increasing function of i.

Thus, if k is the leftmost site with a B pile-up, then all sites to the left of k have a finite mass of B, while all sites to the right of k show B pile-ups. It also follows from eq. (4.5) that at long times, $\langle m_i^B \rangle$ grows linearly with time for $i > k$ (at the k^{th} site, it could also grow as $\sqrt{t}$). Moreover, if $m_{k+1}^B \sim ct$, then $m_{k+2}^B \sim [w/(w+u)]\, ct$, $m_{k+3}^B \sim [w/(w+u)]^2\, ct$ etc. Thus, after sufficiently long times, a snapshot of the lattice at any instant would show an exponential decay in space in the mass profile of B, site $k+1$ onwards, with a decay length given by $[\ln((w+u)/w)]^{-1}$. The location of the site k depends on the parameters and can even be tuned to be at site 1. Left of site k, the system exists in steady state (except for a possible A pile-up on site 1) while right of it, the system exists in a growing phase.

When $q = 0$, i.e. B particles cannot move, there are two important differences. Steady state on site 1 implies a steady state behaviour for the whole system, and it is not possible to have a region of steady state followed by a region of pile-ups. Secondly, $q = 0$ is the only surface in parameter space on which *pile-ups of A need not be localised at site 1.* If there exists a pile-up of A on site 1, then as long as $u \leq v$, there are pile-ups of A on all subsequent sites. These pile-ups grow as $\sqrt{t}$ with time.

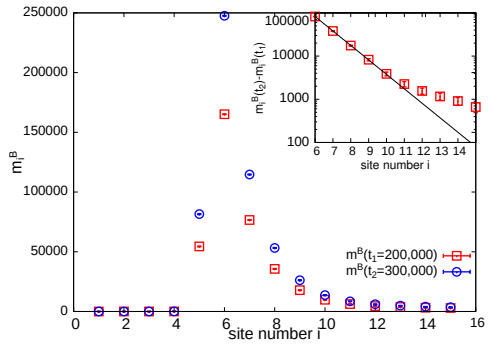


Figure 4.7: A system with pile-ups of B ($a = 1.3$, $u = 1$, $v = 0.33$, $w = 0.86$, $q = 1$): $\langle m_i^B \rangle$ profile for two different time instants $t_1 = 10^8$ (squares) and $t_2 = 1.5 \times 10^8$ (circles). Inset: Semi-log plot of $m_i^B(t_2) - m_i^B(t_1)$ vs i. Solid line shows the analytical prediction.

We also study the system numerically using Monte Carlo simulations for small system sizes, typically $L = 20$. Since later sites cannot affect earlier sites, this is the same as doing a simulation with large L but keeping track of only the first few sites. Figure 4.7 shows numerical results for a choice of parameters for which the theory predicts the existence of steady state at the first four sites, and pile-ups of B site 5 onwards. The $\langle m_i^B \rangle$ profile at two time instants $t_1 = 10^8$ and $t_2 = 1.5 \times 10^8$ (where t is in units of Monte Carlo (MC) steps). In agreement with the analytical prediction, the system shows co-existence of a region with finite time-independent $\langle m^B \rangle$ (upto site 4), and a region with pile-ups of B (site 5 onwards). The rate of growth of $\langle m_i^B(t) \rangle$ with t is highest at site 6 and decreases on moving rightward from site 6. The inset of Fig. 4.7 shows $\langle m^B(t_2) \rangle - \langle m^B(t_1) \rangle$ as a function of site number in a semi-log plot. The solid line is the analytically expected exponential decay.

To summarise, for purely rightward chipping and movement of single particles ($\gamma = 1$), with injection of A particles at site 1, the occupation probability s_i^A decreases monotonically while s_i^B increases monotonically with site number i. The system may attain stationarity or it may exist in a growing phase characterised by 'pile-ups'. An A pile-up can occur only at the first (boundary) site (for $q \neq 0$). B pile-ups occur in the bulk; in fact, it is possible to have steady state existing near the left boundary and pile-ups of B further on (as in Fig 4.7). By making interconversion much slower than chipping ($u/w + v/q \sim \mathcal{O}(1/L)$), both the steady state and the pile-up region can be made macroscopically large ($\mathcal{O}(L)$). If a site has a pile-up of B, then all sites right of it also have B pile-ups. These pile-ups grow as $\sim t$. In the pile-up region, at any instant, the spatial profile of m^B shows an exponential decay.

4.5 Analysis of the System in the Continuum Limit

For fully asymmetric chipping ($\gamma = 1$), steady phase occupation probabilities were obtained by solving eqs. (4.5). For any other value of γ, however, a given site is affected by both its left and right neighbours and the resulting time evolution equations for $\langle m_i^A \rangle$ and $\langle m_i^B \rangle$, (similar to Eqs. (4.5)) yield complicated recursion relations for the occupation probabilities which are difficult to solve. Thus, to analyse the system for any general asymmetry, we assume the total number of sites L in the system to be large, and take a continuum limit for the lattice, thereby going from the integer valued site number i to the real valued position co-ordinate x, defined as $x = i/L$. By this definition, x can take on real values between 0 and 1. In the

continuum limit,

$$s_{i\pm1}^{A,B} = s^{A,B}(x) \pm \frac{1}{L}\frac{\partial s^{A,B}}{\partial x} + \frac{1}{2L^2}\frac{\partial^2 s^{A,B}}{\partial x^2} + ... \tag{4.8}$$

4.5.1 $\gamma = 1$

First, we check whether the steady phase occupation probabilities given by eq. (4.6) can be recovered using a continuum limit analysis. Substituting from eq. (4.8) into eqs. (4.5), setting time derivatives equal to zero and retaining only leading order terms in $1/L$, we get:

$$-w\frac{ds^A}{dx} + \tilde{v}s^B - \tilde{u}s^A = 0 \tag{4.9a}$$

$$-q\frac{ds^B}{dx} + \tilde{u}s^A - \tilde{v}s^B = 0 \tag{4.9b}$$

along with the boundary conditions $s^A(x=0) = a/w$ and $s^B(x=0) = 0$. Here, $\tilde{u}$ and $\tilde{v}$ refer to the rescaled rates $\tilde{u} = uL$ and $\tilde{v} = vL$. In eq. (4.9) the full derivative with respect to x has been used instead of the partial derivative as the steady state occupation probabilities are functions of x only. The solution of eq. (4.9) is given by:

$$s^A(x) = \frac{a\tilde{v}}{\tilde{u}q + \tilde{v}w}\left[1 + \frac{\tilde{u}q}{\tilde{v}w}\exp\left(-\frac{x}{\tilde{l}}\right)\right] \tag{4.10a}$$

$$s^B(x) = \frac{a\tilde{u}}{\tilde{u}q + \tilde{v}w}\left[1 - \exp\left(-\frac{x}{\tilde{l}}\right)\right] \tag{4.10b}$$

with $\tilde{l} = (\tilde{v}/q + \tilde{u}/w)^{-1} = (1/L)\,(v/q + u/w)^{-1}$. A comparison with the decay length computed for the discrete lattice (in eq. (4.7)) shows that apart from the length rescaling factor $1/L$, $\tilde{l}$ is just the leading order term in the expansion of the exact decay length l_D in terms of $(v/q + u/w)$. This is consistent with the fact that in taking the continuum limit, we retain only the leading order term in eq. (4.8). Thus, we conclude that the continuum limit is valid for $\gamma = 1$ when interconversion is much slower than chipping, i.e. $(v/q + u/w) << 1$.

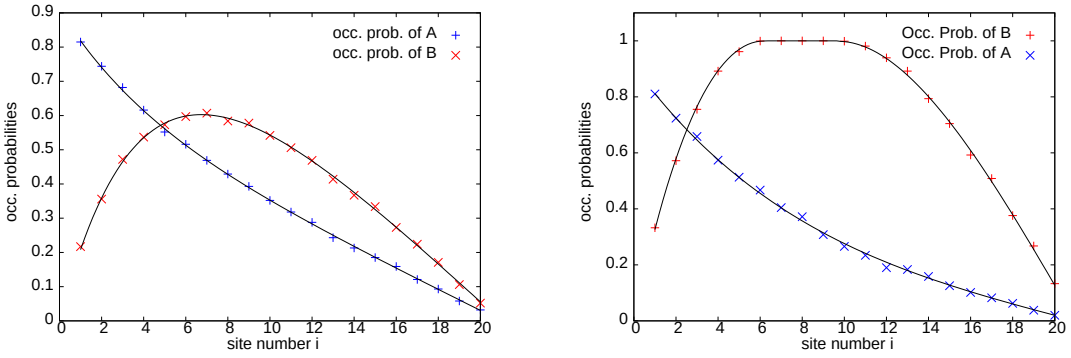
4.5.2 $\gamma = 1/2$

The other limiting case that we consider is that of perfectly symmetric chipping, corresponding to $\gamma = 1/2$. As before, continuity equations similar to eqs. (4.5) can be written for the discrete lattice. On taking the continuum limit in space, these yield the following two coupled differential equations:

$$\frac{\partial\langle m^A(x,t)\rangle}{\partial t} = \frac{w}{2}\frac{\partial^2 s^A}{\partial x^2} + \tilde{v}s^B - \tilde{u}s^A \tag{4.11a}$$

$$\frac{\partial\langle m^B(x,t)\rangle}{\partial t} = \frac{q}{2}\frac{\partial^2 s^B}{\partial x^2} + \tilde{u}s^A - \tilde{v}s^B \tag{4.11b}$$

Here the rescaled interconversion rates are $\tilde{u} = uL^2$ and $\tilde{v} = vL^2$. The boundary conditions are given by $s^A(0) = 2a/w$, $s^B(0) = 0$ and $s^A(1) = s^B(1) = 0$. The boundary conditions at $x = 1$ just reflects the fact that the right boundary acts as a sink, as particles cannot return to the system once they leave from site L.



(a) (b)

Figure 4.8: Occupation probabilities s_i^A and s_i^B as a function of site number i
(a) in a steady phase (with $a = 1.3$, $u = 0.02$, $v = 0.01$, $w = 2.2$, $q = 0.5$).
(b) in a growing phase (with $a = 1.0$, $u = 0.0136$, $v = 0$, $w = 2.2$, $q = 0.3$).
Data points obtained from Monte Carlo simulations seem to agree well with analytical predictions (solid lines).

Steady Phase: Equation (4.11) can be solved for the steady phase by setting time derivatives equal to zero, taking second derivative of eq. (4.11a) with respect to x and eliminating the $s^B(x)$ terms by using eq. (4.11b) and the original eq. (4.11a). This yields an equation which is the second derivative with respect to x of a second order differential equation involving $s^A(x)$. This equation can be solved to obtain the following steady phase occupation probabilities:

$$s^A(x) = \frac{2a\tilde{v}}{\tilde{u}q + \tilde{v}w}\left[(1-x) + \frac{\tilde{u}q}{\tilde{v}w}\left(\frac{\sinh\left(\sqrt{\tilde{\eta}}(1-x)\right)}{\sinh\left(\sqrt{\tilde{\eta}}\right)}\right)\right] \tag{4.12a}$$

$$s^B(x) = \frac{2a\tilde{u}}{\tilde{u}q + \tilde{v}w}\left[(1-x) - \left(\frac{\sinh\left(\sqrt{\tilde{\eta}}(1-x)\right)}{\sinh\left(\sqrt{\tilde{\eta}}\right)}\right)\right] \tag{4.12b}$$

where $\tilde{\eta} = 2\left(\tilde{v}/q + \tilde{u}/w\right) = 2\left(v/q + u/w\right)L^2$

From eq. 4.12, it can be seen that the length scale governing the spatial gradients of $s^A(x)$ and $s^B(x)$ is proportional to $1/\sqrt{\tilde{\eta}}$. Thus, for these gradients to extend a macroscopic distance into the system, the interconversion rates must be much smaller, i.e., $\mathcal{O}\left(1/L^2\right)$ compared to the chipping rates. Figure 4.8a shows the spatial profile of s^A and s^B in steady phase for a certain choice of parameters. The analytic expressions (eq. (4.12)) obtained using the continuum approximation agree well with the Monte Carlo results.

Growing Phase: As before, if the steady phase occupation probabilities, as given by eqs. (4.12), exceed 1 for any x, then eqs. (4.12) cannot be true, indicating that the system cannot attain stationarity and must exist in a growing phase. Nevertheless, the steady phase spatial profiles of s^A and s^B (Fig. 4.8a) provide some indications as to the spatial structure of the growing phase. As in the asymmetric case with $q \neq 0$, since s^A is always a decreasing function

of x, it follows that an A pile-up can exist only on the first site [3]. Further, since the maximum of $s_B(x)$ lies in the bulk of the system, away from the boundaries, it is reasonable to expect that pile-ups of B would typically occur in some central region of the lattice flanked by regions of steady state on either side. To test this prediction, simulations were done with parameters for which $s^B(x)$, as obtained from eq. (4.12), crosses 1 in some range of x values. Figure 4.8b shows the occupation probability profiles for such a choice of parameters. As expected, the region of pile-ups (with $s^B = 1$) exists in the middle of two steady state regions. Appendix 4.9 presents the details of how we analytically calculate the spatial boundaries of the pile-up region as well as the occupation probability profiles which are shown as solid lines in Fig. 4.8b. The main assumption involved in this calculation is that the functions $s^A(x)$ and $s_B(x)$ as well as the derivatives $\partial s^A/\partial x$ and $\partial s^B/\partial x$ are continuous at the boundaries of the pile-up region.

4.5.3 $1/2 < \gamma < 1$

After studying the two limiting cases, namely, pure drift ($\gamma = 1$) and pure diffusion ($\gamma = 1/2$), we now consider partially asymmetric chipping and movement ($1/2 < \gamma < 1$). Site-wise continuity equations similar to eq. (4.11) can be written for this case as well. These equations involve both first and second order derivatives of s^A and s^B with respect to x, and can be solved in the same manner as eq. (4.11), giving the following expressions for the steady phase occupation probabilities $s^A(x)$ and $s^B(x)$:

$$s^A(x) = \frac{a\tilde{v}}{\gamma(\tilde{u}q + \tilde{v}w)} \left[\frac{1 - \exp(-2\tilde{\xi}(1-x))}{1 - \exp(-2\tilde{\xi})} + \frac{\tilde{u}q}{\tilde{v}w} \exp(\tilde{\xi}x) \left(\frac{\sinh\left(\sqrt{\tilde{\eta} + \tilde{\xi}^2}(1-x)\right)}{\sinh\left(\sqrt{\tilde{\eta} + \tilde{\xi}^2}\right)} \right) \right] \tag{4.13a}$$

$$s^B(x) = \frac{a\tilde{u}}{\gamma(\tilde{u}q + \tilde{v}w)} \left[\frac{1 - \exp(-2\tilde{\xi}(1-x))}{1 - \exp(-2\tilde{\xi})} - \exp(\tilde{\xi}x) \left(\frac{\sinh\left(\sqrt{\tilde{\eta} + \tilde{\xi}^2}(1-x)\right)}{\sinh\left(\sqrt{\tilde{\eta} + \tilde{\xi}^2}\right)} \right) \right] \tag{4.13b}$$

where $\tilde{u} = uL^2$, $\tilde{v} = vL^2$, $\tilde{\eta} = 2\left(\tilde{v}/q + \tilde{u}/w\right)$ and $\tilde{\xi} = (2\gamma - 1)L$

Equation (4.13) is also valid for $0 < \gamma < 1/2$. However, if particles have a net drift back towards the source, only a small region close to $x = 0$ is occupied. Thus, we confine our analysis to the more interesting case $1/2 < \gamma < 1$. Strictly at the two limits, $\gamma = 1/2$ and $\gamma = 1$, eq. (4.13) is not valid. However, taking the limit $\gamma \to 1/2$, or equivalently $\tilde{\xi} \to 0$ for eq. (4.13) gives back the profiles of eq. (4.12). At $\gamma = 1$, a second order description cannot be used as there is only one boundary condition for s^A (or s^B).

The spatial variation of $s^A(x)$ and $s^B(x)$ is governed by the rescaled variables $\tilde{\eta}$ and $\tilde{\xi}$. If $\tilde{\eta} \sim \mathcal{O}(L^2)$ and $\tilde{\xi} \sim \mathcal{O}(L)$, this variation is confined to the boundary regions, eventually becoming a boundary discontinuity in the limit $L \to \infty$. If $\tilde{\eta} \sim \mathcal{O}(1)$ (corresponding to

[3] If an A pile-up exists at site 1, then we use the boundary condition $s^A(x = 1/L) = 1$ instead of the boundary condition for $s^A(x = 0)$ in eq. (4.11) to solve for the steady state of the rest of the system.

$\mathcal{O}(1/L^2)$ interconversion rates) and $\tilde{\xi} \sim \mathcal{O}(1)$ (corresponding to weak $\mathcal{O}(1/L)$ asymmetry), the gradients due to the two boundaries extend into the whole lattice and together, determine the behaviour of the system in a complex way. For large but not infinite values of $\tilde{\eta}$ and $\tilde{\xi}$ ($\tilde{\eta} \sim \mathcal{O}(100)$, $\tilde{\xi} \sim \mathcal{O}(10)$), gradients near the two boundaries can be treated independently and yet are not mere boundary discontinuities. This situation is shown in Fig. 4.9. For such values of $\tilde{\eta}$ and $\tilde{\xi}$ and sufficiently large L, two distinct length scales emerge, one characterizing the variation near $x = 0$ and the other near $x = 1$. These length scales are, respectively:

$$\tilde{l}_1 = \left[\sqrt{\tilde{\eta} + \tilde{\xi}^2} - \tilde{\xi}\right]^{-1}, \qquad \tilde{l}_2 = \left[2\tilde{\xi}\right]^{-1} \tag{4.14}$$

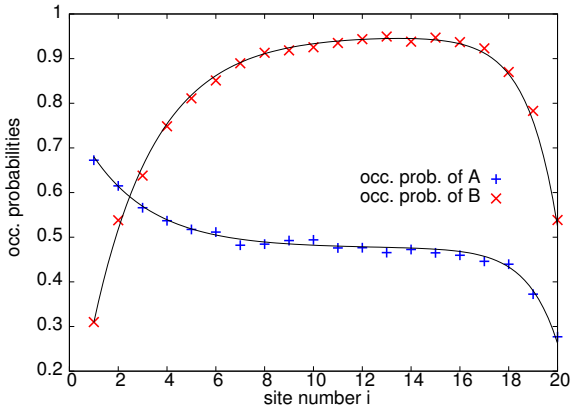


Figure 4.9: s_i^A and s_i^B for $a = 1.2$, $u = 0.2$, $v = 0.1$, $w = 2.2$, $q = 0.7$ and $\gamma = 0.7$. For these parameters, $\tilde{\eta} = 103$ and $\tilde{\xi} = 8.4$. Solid lines are the analytical predictions of eq. (4.13).

The length $\tilde{l}_1$ is the distance from $x = 0$ beyond which A and B are effectively decoupled (no net interconversion). It is thus, the length scale that characterises the spatial gradient of A and B concentrations near the left boundary. As discussed earlier, $\tilde{l}_1$ is small for fast interconversions (high values of $\tilde{\eta}$). It also decreases as $\tilde{\xi}$ decreases, i.e. with decrease in movement asymmetry.

The effect of the absorbing right boundary (at $x = 1$) extends over a length which is given by $\tilde{l}_2$. This can be seen in the following way. Since A and B are effectively decoupled beyond $\tilde{l}_1$, the spatial variation of $s^A(x)$ (and similarly $s^B(x)$) beyond this length is described by a drift diffusion equation $(w/2(\partial^2 s^A/\partial x^2) - w\tilde{\xi}(\partial s^A/\partial x) \simeq 0)$ along with the boundary condition $s^A(x = 1) = 0$. The diffusion length associated with this equation is $[2\tilde{\xi}]^{-1}$ which is just $\tilde{l}_2$. As $\tilde{l}_1$ and $\tilde{l}_2$ increase, the variations near the two boundaries can no longer be treated independently and the middle region of nearly constant $s^A(x)$ and $s^B(x)$ seen in Fig. 4.9 disappears. Perfectly symmetric chipping, for which $\tilde{l}_2$ diverges, is an extreme case of this.

As discussed earlier, if the steady state solutions $s^A(x)$ and $s^B(x)$ (in eq. (4.13)) exceed 1 for any x, then this implies that the system exists in a growing phase. As before, pile-ups of A can occur only at the boundary site, while pile-ups of B are typically found in the bulk of the system.

4.6 Generalisations of the Only-Chipping Model

To summarise, so far we have studied a stochastic two-species model on a 1D lattice with boundary injection of type A, interconversion between types A and B, and transport of both species through the lattice (in general, in a driven diffusive way) by chipping of single particles. The rates of chipping and interconversion at any site were assumed to be constants, independent of the number of particles at that site. The main finding of this study is that depending on the rates of various processes, the system may either eventually attain steady state or show unbounded growth of mass at all times. Pile-ups (as defined in the text) may be composed of A particles or B particles or both. Unlike in translationally invariant systems, the phases in our model have interesting spatial structure. Generically, in growing phases, a part of the lattice attains steady state while other regions show unbounded growth of mass (pile-ups).

The occurrence of pile-ups in this model is a simple consequence of the chipping rates at site i being independent of m_i^A and m_i^B. For such rates, the particle currents are constrained by an upper bound and hence, may fail to balance in certain regions of the system if the rate of injection of particles is sufficiently high. The emergence of pile-ups in specific portions of the system arises due to the interconversion between different species. Interconversion causes the net A (B) current in any bond of the lattice to show non-trivial spatial variation. Thus, a situation can arise where currents of a particular species fail to balance only in a certain region of the system, while current balance and steady state is attained in the rest of the system.

Since the physical picture underlying pile-ups is fairly simple and involves somewhat general effects, it is reasonable to expect that the main results derived so far, i.e., the occurrence of stationary and 'growing' phases, and the co-existence of steady-state regions and regions with pile-ups, will be robust to small variations in the rules of the model. To test this, we look at the following extensions of the model: (i) multi-species model with more than two species of particles (ii) model in higher number of spatial dimensions (iii) model with generalised chipping and interconversion rates that depend on number of particles on the site. Some of these generalised versions are also more realistic from the point of view of the Golgi, as discussed in the next chapter.

4.6.1 Model with Many Species of Particles

The two-species model can be extended quite easily to include more species of particles. For example, consider the three-species version with A, B and C particles. Particles of type A are injected at the left boundary. An A particle can convert to B and vice versa; in addition, B particles can convert to C (at rate k) and C to B (at rate l). Also, as before, a single particle of any species can chip off a site (with rates w, q, p for species A,B,C respectively) and move, in general, in a driven diffusive way.

The model can be analysed in exactly the same way as the two-species model by writing equations involving particle currents and solving for the occupation probabilities at each site. These calculations show that while $s^A(x)$ decreases monotonically as a function of x, the probabilities $s^B(x)$ and $s^C(x)$ both increase monotonically with x in the driven case and

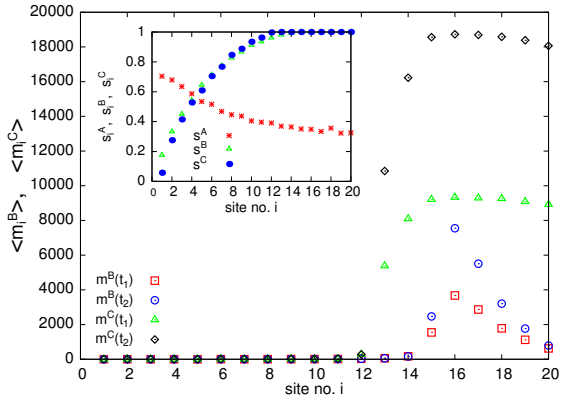


Figure 4.10: Results of Monte Carlo simulations for a phase with pile-ups of both B and C ($\gamma = 1$). Parameters: $a = 5$, $w = 6$, $q = 2$, $p = 0.3$, $u = 0.7$, $v = 0.2$, $k = 0.65$, $l = 0.62$. Data points show $\langle m_i^B \rangle$ and $\langle m_i^C \rangle$ at two time intervals $t_1 = 5 \times 10^6$ and $t_2 = 10^7$. Pile-ups of C occur earlier in the lattice (site 13 onwards) than pile-ups of B (site 15 onwards). Inset: s_i^A, s_i^B, s_i^C at long times.

show a single maximum in the diffusive case. Using arguments similar to the two-species case, we conclude that this model also shows stationary and growing phases. The growing phases may exhibit pile-ups of A and/or B and/or C in overlapping or disjoint spatial regions, depending on various parameters and the asymmetry of chipping. While pile-ups of B or C can form in the bulk of the lattice, a pile-up of A occurs only at the first site. Interestingly, although interconversion takes place in the order $A \rightleftharpoons B \rightleftharpoons C$, pile-ups need not appear in the same order in space. For example, C pile-ups may be found to the left of B pile-ups (i.e. closer to the left boundary) as shown in Fig 4.10.

The generalisation to a higher number of species is straightforward, though the model becomes somewhat unwieldy due to the large number of parameters. We expect that these multi-species models will also exhibit pile-ups of one or a few or all species of particles.

4.6.2 Generalised chipping and Interconversion Rates

So far, we have studied the case where the chipping and interconversion processes taking place at a site do not depend on the number of particles at that site. However, this is not true in many physical situations. Below, we consider the more general scenario where a (chipping or interconversion) process involving particles of species 'Z' at a site with m^Z number of Z particles occurs with a rate that is proportional to some function $f\left(m^Z\right)$ of m^Z [4]. The natural question to ask in this context is whether there is a general class of functions $f\left(m^Z\right)$, for which this model exhibits pile-ups.

It follows from arguments related to the boundedness of currents, that this generalised model also shows pile-ups in specific regions of the system, as long as the function $f\left(m^Z\right)$ is bounded for large m^Z, i.e. $f\left(m^Z\right) < F_0$ for all m^Z. To see this, note that this model can also be analysed by writing continuity equations at each site, similar to eqs. (4.5), as below:

$$\frac{\partial \langle m_i^A(t) \rangle}{\partial t} = w[\gamma \langle f(m_{i-1}^A) \rangle + (1-\gamma)\langle f(m_{i+1}^A) \rangle - \langle f(m_i^A) \rangle] - u\langle f(m_i^A) \rangle \tag{4.15a}$$

$$\frac{\partial \langle m_i^B(t) \rangle}{\partial t} = q[\gamma \langle f(m_{i-1}^B) \rangle + (1-\gamma)\langle f(m_{i+1}^B) \rangle - \langle f(m_i^B) \rangle] + u\langle f(m_i^A) \rangle \tag{4.15b}$$

$$\langle f(m_{L+1}^A) \rangle = \langle f(m_{L+1}^B) \rangle = \langle f(m_0^B) \rangle = 0; \qquad w\gamma \langle f(m_0^A) \rangle = a \tag{4.15c}$$

[4] In general, the function $f\left(m^Z\right)$ can be different for chipping and interconversion processes, but here, for simplicity, we take the rates of both processes to be governed by the same function

These equations, however, involve the quantities $\langle f\left(m_i^Z\right)\rangle$ rather than s_i^Z. As discussed in sec. 4.4.2, if the steady state solution $\langle f\left(m_i^Z\right)\rangle$ exceeds the bound F_0 at any i, then this signifies that the assumption of stationary $\langle m_i^Z\rangle$ is not valid, and the system must exhibit pile-ups, i.e. unbounded growth of mass at all times. Also, since eqs. 4.15 involving $\langle f\left(m_i^Z\right)\rangle$, have exactly the same structure as in the m-independent case, it follows that this model also shows co-existence of regions in steady state with regions that support pile-ups.

An important example of such generalised rates that allow for pile-ups is provided by rates of the Michelis-Menten sort [83], which have been observed for many enzyme-mediated processes. For Michelis-Menten rates, the function $f\left(m^Z\right)$ has the form

$$f(m) = Rm/(T+m) \tag{4.16}$$

where R and T are positive constants. Numerical simulations performed for chipping and interconversion rates of this sort, confirm that this model also exhibits a growing phase which is qualitatively similar to that in the m-independent case.

4.6.3 Higher Dimensions

This model can also be defined on a d-dimensional lattice, with injection of single particles of type A at some $d-1$ dimensional surface and exit of particles of all species at other surfaces. As in the 1D case, we can write continuity equations at each site. Since the argument we have made for the emergence of pile-ups depends only on the boundedness of particle currents in different bonds of the lattice and not on the dimensionality of the lattice, it follows that a growing phase with pile-ups should occur in higher dimensions as well. The detailed spatial structure of the growing phase will depend on the boundary conditions and the degree of spatial bias in movement of different particles. However, as before, we expect the growing phase to be characterised by co-existence of regions in steady state with regions with pile-ups.

In addition to the generalisations discussed above, other minor variations can also be introduced in this model. One such variation involves allowing for injection of B particles at the left boundary, at a rate b. Now, depending on the value of the ratio auq/bwv, the system either has net interconversion from A to B (and behaviour similar to what has been discussed in the preceding sections) or net interconversion from B to A (resulting in A pile-ups in the bulk etc.) or no net interconversion, i.e. no effective coupling between the two species. Another possibility is to allow single particles to exit from the right boundary at a rate that is different from the chipping rate in the bulk. This produces just a boundary effect unless the exit rate is $\mathcal{O}(1/L)$ times the chipping rate. If the exit rate is small in the sense defined above, then this has an effect on the $s^A(x)$ and $s^B(x)$ spatial profiles even at a macroscopic distance away from the boundaries. However, the system still supports growing phases with B pile-ups in the bulk.

4.7 The Multi-Species Model with Chipping and Stack Movement

In this section, we comment briefly on the full multi-species model which allows for both chipping of single particles and movement of entire stacks. This model is difficult to study

because of the large number of parameters and competing effects that are involved. However, some insight can be gained by looking at the two limiting cases– the only-chipping model and the only-stack-movement model, and considering the effect of small perturbations about these limits.

For example, a natural question to ask in the context of the only-chipping model is whether pile-ups survive if the model also allows for stack movement at some small rate D. In the presence of stack movement, the average current of 'Z' type particles out of a site acquires a component that is proportional to $D\langle m^Z\rangle$. As a result, the current is no longer bounded above, and at sufficiently large but finite $\langle m^Z\rangle$, the out-current will be able to balance the in-current. Thus, in the presence of stack movement, there is no growing phase, and the average mass at each site always attains a time-independent value [5]. This value diverges as D goes to zero, but is finite for a finite non-zero D. If we switch on stack movement as a small perturbation to the only-chipping model in a growing phase, then this eliminates pile-ups, leading instead to the emergence of a stationary mass profile with a sharply-peaked (small width) maximum in the same region where the pile-up occurred in the corresponding $D = 0$ model.

The other limiting case of the model is one where transport occurs solely via movement of stacks. As mentioned in section 4.1, the multi-species model differs from the single species version discussed in Chapters 2 and 3, only in that the A/B composition of aggregates varies with position, with $\langle m_i^A\rangle$ falling and $\langle m_i^B\rangle$ rising with increasing site number. As before, it is important to consider the effect of perturbations about this limit, such as the introduction of chipping at a small rate. The discussion of the single-species case in Chapter 2 suggests that as long as the chipping rate is sufficiently small, the basic features of the model such as the formation of large aggregates with $\mathcal{O}(L)$ mass (or $\mathcal{O}(\sqrt{L})$ mass in the case with biased stack movement), remain unaltered. Numerics show that this is the case with the multi-species model as well. However, in the multi-species model, if the chipping rates for different species of particles are very different, then more complex, position-dependent behaviour can arise.

In conclusion, it is possible to obtain large aggregates in two different limits of the full model introduced in Chapter 1. In the aggregation-dominated regime where transport occurs predominantly by stack movement, the system is characterised by the formation of macroscopic condensates (or $\mathcal{O}(\sqrt{L})$ aggregates in the biased case). These emerge due to the tendency of a moving aggregate to sweep the surrounding mass into itself. In the only-chipping limit discussed in this chapter, the system supports pile-ups or indefinitely growing aggregates at *fixed* locations in the system. These emerge for chipping (and interconversion) rates for which particle currents between sites are bounded above, so that there are certain regions in the system, where the in-current and out-current fail to balance. If stack movement is introduced at a small rate to a system with pile-ups, then the unbounded growth is curtailed, leading instead to the formation of a large but finite aggregates. Thus, in this way, it is possible to obtain large, stationary aggregates in the chipping-dominated regime. Note that these two limits, the chipping-dominated regime and the stack-movement-dominated regime are also

[5] If the rate of stack movement is mass dependent i.e. $D(m) \propto m^{-\alpha}$, then the out-current which now has a $D\langle m^{1-\alpha}\rangle$ component, will be bounded above for $\alpha \geq 1$. In this case, pile-ups can occur and the behaviour of the system would be expected to be similar to that of the only-chipping model.

the two limits most relevant to Golgi traffic, as they correspond respectively to the vesicular transport and cisternal movement scenarios. In the next chapter, we study these two limits from the point of view of Golgi structure and dynamics, with the aim of making a closer connection between the model and the biological system.

4.8 Appendix C: Proof of inequalities satisfied by s_i^A and s_i^B for $\gamma = 1$, $q \neq 0$

In sec. 4.4.2, the following inequalities was stated without proof:
For purely rightward movement ($\gamma = 1$), and $q \neq 0$, the occupation probabilities satisfy $s_{i+1}^A < s_i^A$ and $s_{i+1}^B \geq s_i^B$ $\forall i$ in *both steady and growing phases*

This result is based on the observation that the average particle current out of a 'compartment' (as depicted in fig. 4.6a) cannot exceed the current into the compartment, as this would eventually lead to negative number of particles in that compartment. Thus for $q \neq 0$,at site 1, the interconversion current, being the only 'in' current for compartment B, is always positive, i.e. from A to B. This implies that $s_2^A < s_1^A$ and $s_2^B \geq s_1^B$. We can prove this by ruling out the other three possibilities which are:
(i) $s_2^A \geq s_1^A$, $s_2^B \geq s_1^B$: This would imply that either the net current out of site 2 is more than the net current into it which is not possible; or if $s_1^A = s_2^A$ and $s_1^B = s_2^B$, then $us_2^A - vs_2^B = us_1^A - vs_1^B > 0$, i.e. there is a net interconversion current out of A. Since the in and out hopping currents for A are equal ($s_1^A = s_2^A$), this means there is a net out-current from compartment A which is not possible.
(ii) $s_2^A < s_1^A$, $s_2^B < s_1^B$: Since s_1^A and s_1^B can at most be equal to 1, this implies that $s_2^A, s_2^B < 1$, i.e. site 2 is in steady state and currents for both compartments balance. To satisfy current balance for A with $s_2^A < s_1^A$, interconversion current must be from A to B (see Fig. 4.6a) but for current balance for B with $s_2^B < s_1^B$, it must be in the opposite direction. This leads to a contradiction.
(iii) $s_2^A \geq s_1^A$, $s_2^B < s_1^B$: This implies that $us_2^A - vs_2^B > us_1^A - vs_1^B > 0$. However (from Fig. 4.6a), to have $s_2^A \geq s_1^A$, the interconversion current has to be from B to A or at least has to be zero. This leads to an inconsistency.
Thus, the only possibility is that $s_2^A < s_1^A$ and $s_2^B \geq s_1^B$. In the same way, the proof can be extended to all subsequent sites. It may be noted that we have not assumed the existence of steady state i.e. $d\langle m^{A,B}\rangle/dt = 0$, anywhere in this argument. Thus, the inequalities hold for the growing phase also. The $q = 0$ case has many sub cases and can be analysed in a similar way.

4.9 Appendix D: Method for calculating $s^A(x)$ and $s^B(x)$ for $\gamma = 1/2$ when pile-ups of B occur

Let x_1 and x_2 divide the lattice into three regions such that in:
Region I ($0 \leq x < x_1$): Steady state ($s^A(x) < 1$, $s^B(x) < 1$);
Region II ($x_1 \leq x \leq x_2$): Steady state of A, pile-ups of B ($s^A(x) < 1$, $s^B(x) = 1$);
Region III ($x_2 < x \leq 1$): Steady state ($s^A(x) < 1$, $s^B(x) < 1$).

In regions I and III, since steady state exists, $s^A(x)$ and $s^B(x)$ are obtained by setting the L.H.S. equal to 0 in both of Eqs. (4.11a) and (4.11b) and solving them with the following boundary conditions for:
Region I: $s^A(0) = 2a/w$, $s^B(0) = 0$, $s^B(x_1) = 1$, $\partial s^B/\partial x\big|_{x=x_1} = 0$
Region III: $s^B(x_2) = 1$, $\partial s^B/\partial x\big|_{x=x_2} = 0$, $s^A(1) = s^B(1) = 0$

In region II, substituting $s^B(x) = 1$ into eq. (4.11a), and setting the left-hand-side equal to zero, we get $s^A(x)$ upto two unknowns. We can find $s^A(x_1)$ and $s^A(x_2)$ from the solutions in regions I and III respectively. Since $s^A(x)$ must be a continuous function, these provide the boundary conditions for the solution of $s^A(x)$ in region II. Thus, we have $s^A(x)$ and $s^B(x)$ for all three regions upto the two unknowns x_1 and x_2. To determine x_1 and x_2, we make the assumption that $\partial s^A/\partial x$ is continuous at x_1 and x_2, i.e. $\partial s^A/\partial x\big|_{x_1}^{I} = \partial s^A/\partial x\big|_{x_1}^{II}$ and $\partial s^A/\partial x\big|_{x_2}^{II} = \partial s^A/\partial x\big|_{x_2}^{III}$ which yields two transcendental equations in x_1 and x_2. For a given set of parameters, these can be solved numerically to get x_1 and x_2, Since we have expressions for $s^A(x)$ and $s^B(x)$ in terms of x_1 and x_2, this also gives the occupation probabilities in all three regions.

This calculation involves two *ad hoc* assumptions, namely, that $\partial s^B/\partial x$ vanishes and $\partial s^A/\partial x$ is continuous at the two boundaries of the pile-up region. These assumptions give results which match well with numerical data (see Fig. 4.8b), but they need to be investigated and adequately justified.

Chapter 5

Relevance of the Model to Golgi Structure and Traffic

5.1 Golgi Apparatus: Structure and Dynamics

The Golgi apparatus is an organelle found in most eukaryotic (nucleus-containing) cells [13]. Its main function is the processing of proteins synthesised in the endoplasmic reticulum (ER), typically by addition of carbohydrate chains, and subsequent sorting and export of these processed proteins to various cellular locations. Morphologically, the Golgi consists of multiple disc-like structures known as cisternae, each with a distinct chemical composition. The cisternae close to the ER constitute the cis Golgi, the cisternae on the other end close to cell membrane form the trans Golgi, and the intermediate cisternae are referred to as the medial Golgi.

Golgi apparatus across different cell types share some common features, but also show certain variations. For example, the Golgi apparatus in almost all cell types shows a biochemical polarity in the cis to trans direction– the cis, medial and trans cisternae each have distinct resident enzymes which mediate different processing events [13]. A feature which shows some variation across cells, however, is the number, shape and spatial spread of the Golgi cisternae. While cisternae are tightly stacked in most mammalian cells, in saccharomyces cerevisiae (budding yeast), they can be completely unstacked, making it possible to track individual cisternae through live imaging [15]. In some cells, the Golgi apparatus also shows intercisternal connections [84].

Different types of cargo, mainly protein and lipid molecules, enter the Golgi through the cis face, traverse the Golgi in the cis to trans direction, get sequentially modified by different enzymes at each cisterna and finally exit from the trans face. As discussed in Chapter 1, there has been a lot of interest in understanding the exact mechanism by which cargo molecules move through the Golgi. The two main competing models [14, 17] for intra-Golgi transport (schematic illustration in fig. 5.1) are:
Vesicular Transport: In this view, the cisternae are stable structures with a fixed spatial location within the Golgi and a fixed set of resident enzymes. Molecules shuttle from one *stationary* cisterna to the next in small sacs called vesicles, and get chemically modified by the resident enzymes.

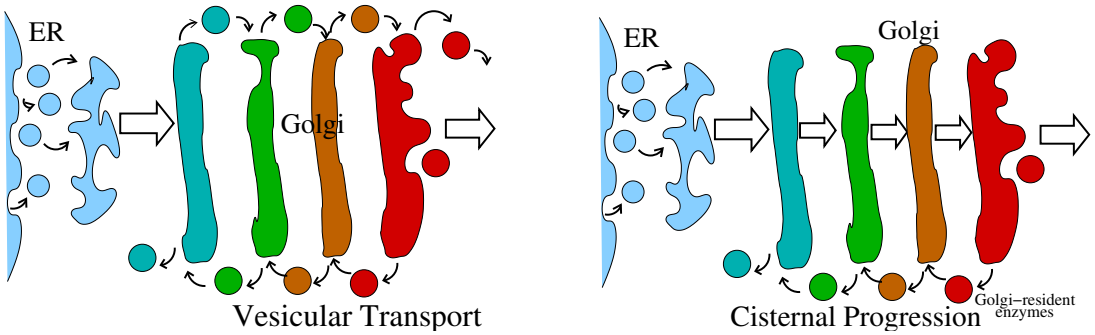


Figure 5.1: Schematic depiction of the two models of intra-Golgi transport– (a) Vesicular Transport (vesicles shuttle between stationary cisternae) and (b) Cisternal Maturation (cisternae move through Golgi and undergo maturation during movement.)

Cisternal Progression and Maturation: This model considers the cisternae to be transient structures that are formed by fusion of incoming vesicles. In this scenario, it is the cisternae that move as a whole through the Golgi apparatus, carrying cargo molecules with them. Specific enzymes get attached to a cisterna in different stages of its progression, and modify its contents. The final cisterna eventually breaks up, releasing processed biomolecules.
The main difference between these two mechanisms of transport is thus, whether the primary carriers of molecules are the large aggregates (cisternae) in the system or small fragments (vesicles) of the large aggregates.

Experimental evidence in support of the two hypotheses, namely, vesicular transport and cisternal progression, is mixed. Cisternal maturation has been directly visualised in yeast cells [15], where the cisternae are unstacked, making it possible to track the individual cisternae in time. However, there are other experiments measuring the exit statistics of molecules through the Golgi [16], which appear to refute the maturation hypothesis. It has also been suggested that these two mechanisms need not be mutually exclusive, and transport may be occurring via some combination of both [17].

How can the two mechanisms be distinguished from each other? In most cells, due to the closely stacked nature of the Golgi cisternae, it has not been possible to monitor the detailed dynamics at the level of the cisterna. However, different signatures of the two mechanisms may appear in dynamical or transport properties of cargo molecules, e.g. in the distribution of transit times of cargo molecules through the Golgi. Devising indirect measurements of this sort thus becomes an important step in probing Golgi dynamics. In this context, an important task of theory is to interpret transport measurements and work out whether these are consistent with a particular mode of transport [85].

In addition, different transport mechanisms may give rise to Golgi structures that are qualitatively different– this is because the spatial structure of the Golgi and its chemical polarity must arise naturally from and be maintained by the same basic processes that govern transport. Thus, a second (theoretical) approach to studying the Golgi is to build simple models incorporating the basic traffic processes described above, and analyse the conditions

under which these models gives rise to Golgi-like structures. A question which emerges naturally in this framework is whether variations as well as similarities in Golgi structures across cell types can be explained by varying some parameter or a particular rule within the model.

Motivated by these questions, in this chapter, we focus on the two limiting cases of the general multi-species model– the only chipping limit and the only stack hopping limit, which correspond respectively to vesicular transport and cisternal movement. For each limit, we identify the conditions under which the system exhibits multiple, spatially resolvable aggregates (cisternae) with position-dependent chemical composition. Further, we ask (a) whether the steady state structures in the two limits are qualitatively different (b) what other diagnostics can be used to distinguish between the different transport processes in the two limits of the model.

The rest of the chapter is organised as follows. In section 2, we define a slightly generalised version of the model studied so far, and also discuss some finer points of the correspondence between the model and the Golgi. Sections 3 and 4 consider in detail the only-chipping and only-stack-movement limits respectively and the conditions under which these yield multiple spatially and chemically distinct aggregates. In section 5, we contrast the structures generated in the two limits, and enumerate key differences in the static properties of these structures. Section 6 deals with some dynamical measurements that can be used to distinguish between the structures in the two limits. Finally, in section 7, we present some broad conclusions of the study and also comment on the relevance of this kind of a model to the study of Golgi dynamics.

5.2 The Model and its Correspondence with the Golgi

In this chapter we consider a slightly generalised version of the multi-species model defined in Chapter 1. This model incorporates three species, 'A', 'B' and 'C', of particles, and allows for the following processes to occur:

i. Injection of single particles of type A at the first site at rate a

ii. Conversion of a single A particle at a given site i to a B particle with rate $uf_A\left(m_i^A\right)$ which now also *depends on the number m_i^A of A particles at site i* through the function $f_A\left(m_i^A\right)$. Similarly, a B particle at a given site can convert to a C particle with rate $kf_B\left(m_i^B\right)$ [1].

iii. Chipping and movement of single particles with species-dependent and number-dependent rates, e.g., an A particle on a site i chips off a site with rate $wf_A\left(m_i^A\right)$ and moves to the right (or left) with probability γ_A (or $1-\gamma_A$). Similarly a B particle chips off with rate $qf_B\left(m_i^B\right)$ and a C particle with rate $pf_C\left(m_i^C\right)$ (with the corresponding asymmetry factors being γ_B and γ_C).

[1] In principle, we can also have $B \to A$ and $C \to B$ conversions, but to limit the number of parameters, we do not allow for these processes. The basic features of the model are not altered in the presence of these reverse conversions

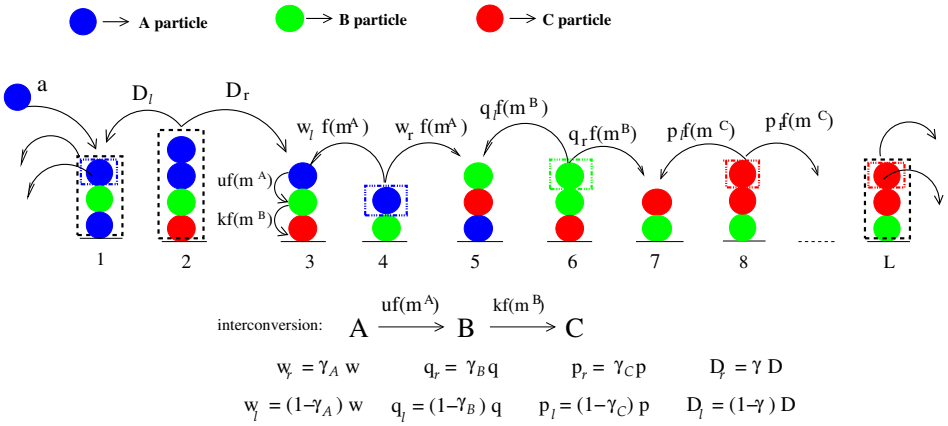


Figure 5.2: Illustration of three-species model with 'A', 'B' and 'C' particles. The model allows for: injection of single particles of type A at site 1; $A \to B$ and $B \to C$ conversions of single particles with a rate that depends on number of A (or B) particles at the site (see site 3); chipping of a single A/B/C particle with a rate that depends on the number of particles of that species at the site and movement to neighbouring site (depicted on sites 4,5,6); movement of full stack to the right or left neighbour and aggregation with the mass there (see site 2); outflux of full stacks or single particles from left and/or right boundary (site 1 and/or site L). Rates of various processes are depicted in the figure.

iv. Movement of the entire stack (all the particles at a site as a whole) to the right or left with rates γD and $(1-\gamma)D$ respectively.

v. Exit of the full stack or a single 'chipped' particle from either boundary with the same rates as those for movement in the bulk [2].

This model is different from the one defined in Chapter 1 only in that the chipping and interconversion events at a site occur with generalised rates which also depend on the number of particles at the site, through the function f. In principle, the function f can be different for chipping and conversion processes, but to keep the model analytically tractable, we take f to be the same for both. The parameter space of this model is quite large, encompassing the injection rate a, the interconversion rates u, k, the chipping rates w, q, p, and the corresponding spatial asymmetry factors γ_A, γ_B, γ_C, the stack movement rate D and the corresponding asymmetry factor γ, and finally the functions f_A, f_B, f_C, along with any parameters that may be needed to define these.

The broad correspondence between this model and the Golgi has already been outlined in chapter 1: injection of A particles at site 1 corresponds to influx of vesicles containing unprocessed proteins from the ER to the cis-Golgi; the $A \to B \to C$ interconversion is a highly simplified representation of the sequential processing of proteins as they pass through the Golgi; single particle movement corresponds to vesicle exchange and collective motion of the entire stack of particles at a site to cisternal movement. Below, we discuss some finer points of this correspondence as well as certain issues that arise in making it:

1. A,B,C particles as analogues of molecules in different stages of processing in the Golgi: In constructing this model, we imagine an A particle to be the equivalent of the unpro-

[2]If exit is allowed to occur only from the right boundary, then we need to consider $\mathcal{O}(1/L)$ injection rates.

cessed protein molecules arriving at the cis-Golgi from the ER, and particles of type B, C, D as representing processed proteins in different (successive) stages of processing. Thus, an aggregate which is primarily of type B is analogous to a cisterna with a large fraction of semi-processed molecules in an early stage of processing. In the Golgi, however, cisternae are characterised by the enzymes they carry and not by the processing level of the cargo molecules present in them. These two different characterisations can be reconciled by treating this model as an effective version of a more detailed model incorporating enzymes, where each enzyme preferentially attaches to cargo molecule in a specific stage of processing. With this assumption, it is reasonable to expect that a region with a large concentration of 'B' particles would also have a high concentration of the enzyme with an affinity for 'B' and so on. Thus, in this sense, an 'A','B','C' particle can be considered as composite of a cargo molecule at a particular stage of processing along with the associated enzyme which has affinity for molecules in that stage.

2. Analogues of vesicles and cisternae in the model:
   In the model, we consider single particles as vesicles and any *multi-particle* aggregate as a cisterna [3]. The single particles move with a rate that depends on the species of the particle– this reflects the fact that vesicles with different cargo and/or enzymes may have different fission and fusion properties. Further, in the interest of simplicity, we assume that a multi-particle aggregate or stack moves at a rate which is independent of both its composition and mass. (Preliminary simulations with mass-dependent and composition-dependent rates suggest that such rates do not cause a qualitative change in the behaviour of the system.)

3. Modeling enzyme-mediated processes as simple rate processes:
   The $A \to B$ and $B \to C$ conversions which we have treated as Poissonian rate processes in the model are actually catalysed by enzymes in the real system. Thus, a more realistic model of the Golgi would involve both A, B, C 'cargo particles' and E_A, E_B 'enzyme particles'. In this kind of a detailed model, not only would the distribution of an enzyme (say E_B) depend on the distribution of the corresponding species (say B), but the distribution of B particles would also be influenced by the enzyme E_B in turn, as the E_B molecule activates the $B \to C$ conversion. Thus, in general, the distribution of enzymes may not be mirrored in just the distribution of A,B,C molecules as argued in point (i) above; instead the enzyme plus cargo model could show more complex effects than the model with effective rates that we study. Nevertheless, the effective model does incorporate the sequential interconversion process using just a few parameters, and has the advantage that it is more amenable to analysis than a complex model.

4. Parameters as composite of many rates:
   The various chipping, interconversion rates in our model are effective or composite rates. For example, the $B \to C$ conversion is influenced by the rates of attachment and detachment of enzymes with a B molecule, the concentration of enzymes, the reaction rate between the enzyme and the B molecule etc. Similarly, the chipping rate combines

[3] Some of these aggregates may consist of only a few particles, and thus are too small to qualify as cisterna. However, we find such small aggregates primarily near the left boundary (see sec. 5.4), which may also representative of the situation in the Golgi, where (in the cisternal progression scenario), smaller 'pre-cisternae' could be forming near the ER, and fusing together eventually to form the cis-cisterna.

both the rate at which a vesicle buds from an aggregate, as well as the rate at which it fuses with the next aggregate. Because of the composite nature of the rates, it is not straightforward to find the corresponding parameter in the real system or vary these rates in a controlled manner in experiments. However, such a model with a small number of composite parameters is more suited for gaining an insight into broad qualitative effects at play in the system.

5. Representing the three-dimensional Golgi by a one-dimensional model:
   Studying a 1D model can be justified on the grounds that there is a net direction of transport of molecules through the Golgi; the spatial variation of enzymes across the Golgi also occurs along this (cis to trans) direction. The 1D model can thus be considered an effective model obtained by integrating over the two directions perpendicular to this cis to trans direction. However, a 1D model of this sort cannot be used to address questions related to the shape of individual cisternae.

In short, we study a simplified 1D model which includes the basic processes involved in Golgi dynamics. This model takes into account the sequential modification of molecules in the Golgi, and treats molecules in different stages of processing as distinct entities. However, it does not explicitly incorporate the enzymes responsible for various biochemical modifications, and considers these modification processes as simple Poissonian processes that occur with specified rates. The various parameters of the model are thus effective or composite parameters. While such parameters cannot be connected easily to experimentally measurable rates, they lend themselves to an economical description of the system, which makes it possible to analyse in detail the various effects that can arise in such a system, simply due to self-organisation. As argued in the ensuing sections of this chapter, an understanding of these effects could be relevant to experiments as well.

5.3 Modeling Vesicular Transport by the Only-Chipping Model

From the correspondence outlined above, it is clear that the analogue of the vesicular transport scenario in the Golgi is the only-chipping-model, obtained by setting $D_L = D_R = 0$ in the general model (fig. 5.2). The only-chipping-model allows for movement of only single particles between neighbouring spatial locations with rates that depend on the type of particle, and the number of particles of that type present at the departure site.

This limiting case was analysed in detail in Chapter 4. The main results of this analysis may be stated as follows. If $f(m)$ is bounded for large m, then the system admits two qualitatively different phases– the steady and the growing phase. In the steady phase, the mean mass at any location has a finite value. The growing phase is characterised by pile-ups or unbounded growth of mass of one or more species of particles occurring in specific regions with tunable extent and location, even as steady state is attained in the rest of the system. The transition from constant mass to a pile-up at any site is associated with divergence of the mean mass at the site. Both steady and growing phases exhibit spatial gradients of A, B, C particles, where the length scales governing the various gradients are determined by the ratio of the interconversion to the chipping rates, and the degree of asymmetry in the movement of A,B,C particles.

The aim of this section is to identify the conditions under which the only-chipping-model generates Golgi-like structures. More specifically, we ask: in what regions of parameter space, does the steady state of the model exhibit basic features of the Golgi, such as the presence of multiple, spatially and biochemically distinct cisternae that are stable in time, and robust to small variations in the rates governing various transport processes. The above requirements translate into certain conditions on the mass profile, which are stated below:

i. The mass profile should have multiple maxima, corresponding to multiple cisternae in the Golgi.

ii. Each maximum in the mass profile should have a different composition, i.e., the relative amounts of A, B, C should be different at each maximum; this corresponds to having cisternae with position-varying chemical composition.

iii. The peaks of the mass profile should be well-separated, i.e., the distance between successive maxima should be more than the width of the maxima; this corresponds to cisternae that are spatially distinct or resolvable.

iv. The rms fluctuations Δm of the peak height should be much smaller than the average height $\langle m \rangle$ of the peak, i.e., $\Delta m / \langle m \rangle \to 0$ for $\langle m \rangle >> 1$; this corresponds to having cisternae that are stable in time.

v. The steady state mass profile should not change qualitatively due to a small perturbation in rates: this ensures that structures are robust to small fluctuations in the underlying transport processes.

What functional form $f(m^Z)$ and what regions of parameter space would admit steady states with the above properties? To gain some insight into this, we study three kinds of $f(m^Z)$ (represented schematically in figure 5.3), where Z can stand for any of the three species A, B, C.

I. Mass-independent rates, i.e.

$$f(m_i^Z) = \begin{cases} 0, & \text{if} \quad m_i^Z = 0 \\ 1, & \text{if} \quad m_i^Z > 0 \end{cases} \tag{5.1}$$

II. Rates proportional to the number m_i^Z of Z particles, or some power of m_i^Z, e.g., $f(m^Z) \propto \sqrt{m^Z}$.

III. Rates that increases with m_i^Z when m_i^Z is small but saturate to a constant value at large m_i^Z. An example of $f(m_i^Z)$ with this kind of crossover is given by Michelis-Menten rates [83]:
$f(m_i^Z) = \frac{\sqrt{m_{sat}^Z}\sqrt{m_i^Z}}{\sqrt{m_{sat}^Z}+\sqrt{m_i^Z}}$
where m_{sat}^Z is the typical mass scale over which the crossover occurs.

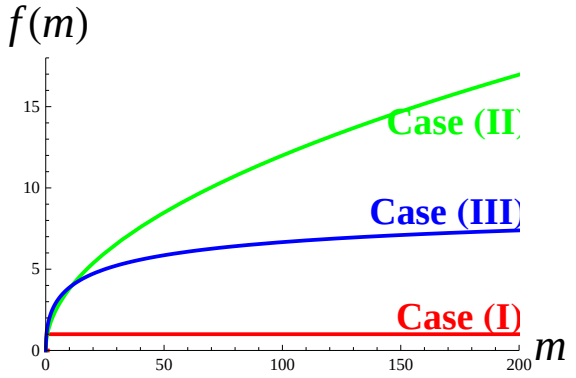


Figure 5.3: Schematic representation of the three functional forms $f(m)$ considered for the only-chipping model. Case (I): $f(m)$ is constant. Case (II): $f(m) \propto \sqrt{m}$. Case (III): $f(m) \propto \sqrt{m}\sqrt{m_{sat}}/(\sqrt{m}+\sqrt{m_{sat}})$.

The above forms of $f(m^Z)$ emerge naturally for processes that are mediated by enzymes, with the enzyme concentration playing a crucial role in determining the kinetics of the process, and hence $f(m^Z)$. For instance, if the number of enzyme molecules is much lower than the number of A, B, C particles available for chipping or conversion, then the reaction rate is limited by the enzyme concentration and is more or less independent of the number of particles at a site: this corresponds to case (I). If the enzyme concentration is in excess of the reactant particles, then the rate is proportional to the number of A, B, C particles available for reaction at a site i: this could in general be $[m_i^Z]^\alpha$ where $\alpha > 0$ [4] [see case (II)]. For intermediate enzyme concentrations there could be a crossover from a reaction rate that goes as $\sqrt{m^Z}$ for small m^Z and becomes independent of m^Z at large m^Z. This is encapsulated in case (III) above. This is just a heuristic justification of the three kinds of rates. A more detailed model would explicitly take enzymes into account.

Note that cases (I) and (III) admit pile-ups, as the function $f(m^Z)$ is bounded for large m^Z. Thus, in order to generate large but *finite* aggregates, in these cases, we must consider only that region in parameter space that corresponds to the steady phase. In the steady phase, it is possible to obtain $\langle f(m_i^Z)\rangle$ as a function of i, by solving for the various particle currents in the system (see section 4.6.2, Chapter 4). The average mass profiles $\langle m_i^Z\rangle$ can then be obtained approximately from $\langle f[m_i^Z]\rangle$, at least in the limit $\langle m_i^Z\rangle \gg 1$. As discussed in Chapter 4, in case (I), $\langle m^Z\rangle \propto (1 - \langle f[m^Z]\rangle)^{-1}$ as $\langle f[m^Z]\rangle \to 1$, since $\langle f[m_i^Z]\rangle$ is just the occupation probability s_i^Z of Z particles at site i. For the other two cases, if we assume that m_i^Z has a narrow distribution centred about $\langle m_i^Z\rangle$, then we can use the approximation $\langle f[m_i^Z]\rangle \sim f[\langle m_i^Z\rangle]$. This assumption turns out to be quite good for $\langle m_i^Z\rangle \gg 1$, and yields:

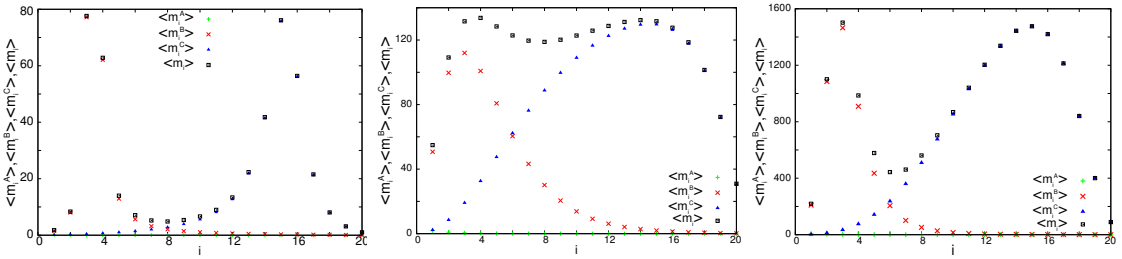
$$\langle m_i^Z\rangle \sim \langle f[m_i^Z]\rangle^2 \text{ for case (II), and}$$
$$\langle m^Z\rangle \sim m_{sat}^Z \left[\left(\frac{\sqrt{m_{sat}^Z}}{\langle f(m^Z)\rangle}\right)^2 - 1\right]^{-2} \text{ for case (III).}$$

Equipped with the above approximate expressions for $\langle m_i^Z\rangle$ (for large $\langle m_i^Z\rangle$), we can now examine which of the conditions (i)-(v) stated at the beginning of the section are satisfied by the mass profiles generated in the three cases (I)-(III). From the solutions for $\langle f[m_i^Z]\rangle$ (see chapter 4), we know that in the model with $A \to B \to C$ conversion and left-right symmetric chipping, $\langle f[m_i^A]\rangle$ decreases monotonically with i, while $\langle f[m_i^B]\rangle$ and $\langle f[m_i^C]\rangle$ increases with i, reach a maximum somewhere in the bulk, and then decrease towards 0 on approaching the right boundary. As a result, the $\langle m_i^B\rangle$ and $\langle m_i^C\rangle$ profiles also have maxima lying somewhere in the bulk of the system, in all three cases (I)-(III). If these maxima are sufficiently well-separated and narrow, then the total mass profile $\langle m_i\rangle$ will have multiple maxima, where each maximum is automatically *associated with a different species of particles.*

As discussed earlier, for the mass profile to have distinct maxima, the distance between successive maxima should be more than the width of each maximum. To achieve this, we should be able to control both the position of the maxima and their width.

[4] $\sqrt{m^Z}$ rates emerge if we assume that the mass at any location in the 1D model is obtained by integrating over the 2D disc-like structure. A 2D structure with m particles typically has $\sqrt{m}$ molecules at its perimeter. Assuming that only the perimeter molecules are available for reaction, we get reaction rates that are proportional to $\sqrt{m^Z}$

Position of maxima: The analysis of the two-species model in Chapter 4 suggests that there are two effects that determine the position of the maximum of $\langle f(m_i^Z)\rangle$, and consequently of $\langle m_i^Z\rangle$ [5]. First, the ratio of the interconversion rates to the chipping rates controls the length scale over which a typical $\langle f[m_i^Z]\rangle$ profile rises and second, the asymmetry factors γ_B, γ_C control the length scale associated with the falling part of the profile. By choosing interconversion rates that are $\mathcal{O}(1/L^2)$ times the chipping rates, and weak $\mathcal{O}(1/L)$ asymmetry, it is possible to tune the maximum to lie in the bulk of the system, i.e., a macroscopic distance away from both boundaries. Moreover, by allowing different asymmetries for transport of different species, greater tuning of spatial positions becomes possible. For example, a wide separation between maxima of $\langle f[m_i^B]\rangle$ and $\langle f[m_i^C]\rangle$ profiles can be obtained quite easily by choosing $\gamma_C > \gamma_B$. Thus, the maxima of the average mass profiles $\langle f[m_i^Z]\rangle$ for different species 'Z' can be made sufficiently far apart in all three cases (I)-(III), by tuning interconversion and chipping rates and the asymmetry factors.



(a) (b) (c)

Figure 5.4: The average mass profiles $\langle m_i^A\rangle$, $\langle m_i^B\rangle$, $\langle m_i^C\rangle$ and $\langle m_i\rangle = \langle m_i^A\rangle + \langle m_i^B\rangle + \langle m_i^C\rangle$ in (a) case (I) (b) case (II) (c) Case III (The parameters used for the three cases are given in appendix). The total mass profile in all three cases has two maxima. which can be resolved for cases (I) and (III) but not in (II). The maxima in case (III) typically have a much larger width than in case (I).

Width of maxima In order to obtain well-separated peaks in the total mass profile, it is also necessary to tune the width of the maxima to be much smaller than the distance between the maxima. Recall that with rates of type (I) or (III), the system can undergo a phase transition from a steady to a growing phase. From our earlier discussion, it follows that $\langle m_i^Z\rangle$ diverges if $\langle f[m_i^Z]\rangle$ is close to its maximum value (which is 1 for rates of type A, and $\sqrt{m_{sat}^Z}$ for rates of type (III)), i.e., if the system is poised close to the phase boundary between the steady and growing phases. Thus, for rates of types (I) and (III), the peaks of the various mass profiles $\langle m_i^Z\rangle$ can be made arbitrarily narrow by tuning the maximum of the corresponding $\langle f[m_i^Z]\rangle$ profile to be arbitrarily close to its highest allowed value. In case (II), where there is no phase transition to a growing phase, and hence no such divergence of mass, the maxima are generically wide, making it very difficult to resolve adjacent maxima.

In summary, boundedness of particle currents is a pre-requisite for getting maxima with small width in the only-chipping-model; sharp peaks are obtained in cases (I) and (III) by tuning the system to be close to the phase boundary between the stationary and the pile-up

[5]For cases (I)-(III), we discussed how $\langle m_i^Z\rangle$ can be obtained from $\langle f[m_i^Z]\rangle$ for $\langle m_i^Z\rangle \gg 1$. From this relationship, it follows that $d\langle m_i^Z\rangle/di \propto d\langle f[m_i^Z]\rangle/di$ in all three cases, implying that the maximum of $\langle m_i^Z\rangle$ and $\langle f[m_i^Z]\rangle$ must coincide.

phase. Fig. 5.4 depicts average mass profiles $\langle m_i^A \rangle$, $\langle m_i^B \rangle$, $\langle m_i^C \rangle$ and the total mass profile $\langle m_i \rangle$ in all three cases, for parameters which have been chosen so as obtain spatially resolvable maxima (to the extent possible).

Stability of mass profile: To estimate the extent to which the mass profile varies over time, we monitor the rms fluctuations Δm^Z about the average number $\langle m^Z \rangle$ of Z particles at each site. We study the three cases (I)-(III) numerically and find the following behaviour of the fluctuations for large $\langle m^Z \rangle$:
Case (I): $\Delta m^Z \sim \langle m^Z \rangle$;
Case (II): $\Delta m^Z \sim \sqrt{\langle m^Z \rangle}$;
Case (III), $\Delta m^Z \sim \langle m^Z \rangle^\alpha$ where $\alpha \sim 0.6$.
Figure 5.5 shows a plot of $\Delta m^Z / \langle m^Z \rangle$ as a function of $\langle m^Z \rangle$ for the three cases. Thus, case (I) is characterised by giant fluctuations about the mass profile, which can even be of the order of the average mass (see fig. 5.6a, 5.6b), whereas for rates of type (II) and (III), the mass profiles are relatively stable (see fig. 5.6c, 5.6d for case (II) and fig. 5.6e, 5.6f for case (III)). We conclude that for stability of aggregates, chipping rates should increase with mass of the parent aggregate. This ensures a kind of stabilising negative feedback– when the aggregate becomes too large, the number of particles breaking off from it increases and vice versa, thus tending to restore the aggregate to its average size.

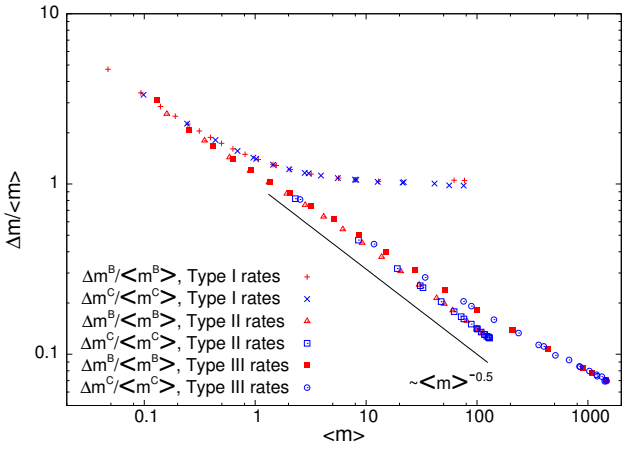


Figure 5.5: Log-log plot showing $\Delta m_i^Z / \langle m_i^Z \rangle$ vs. $\langle m_i^Z \rangle$ for B and C particles for rates of type (I), (II) and (III). For case I, $\Delta m_i^Z / \langle m_i^Z \rangle$ approaches a constant value at large $\langle m_i^Z \rangle$. For case II, $\Delta m_i^Z / \langle m_i^Z \rangle$ decreases as $1/\sqrt{\langle m_i^Z \rangle}$, while for case III, the decrease is slightly milder than this.

Robustness of mass profile to small variations in rates: A general consideration for any such model is that the cellular environment is intrinsically noisy and any mechanism that maintains differentiated cisternae must be robust to small perturbations in the underlying transport processes. Thus, it is important to quantify how the mass profiles generated in the three cases, change in response to small variations in the parameters of the model. A closely related issue is the extent of fine tuning required to obtain mass profiles with multiple, resolvable peaks.

Varying any parameter of the model typically results in a shift in both the position and the height of the maximum of any $\langle m_i^Z \rangle$ profile. However, in the regime of $\mathcal{O}(1/L^2)$ interconversion rates, the position is relatively insensitive to small changes in parameters. The main effect of any perturbation is to change the height of the maximum of $\langle m^Z \rangle$ for some or all Z. For example, decreasing the rate of chipping of A particles increases the heights of the maxima of $\langle m_i^A \rangle$, $\langle m_i^B \rangle$, $\langle m_i^C \rangle$ profiles. Similarly, a small decrease in the chipping rate of B particles causes an increase in the heights of the maxima of $\langle m_i^B \rangle$ and $\langle m_i^C \rangle$, and so on.

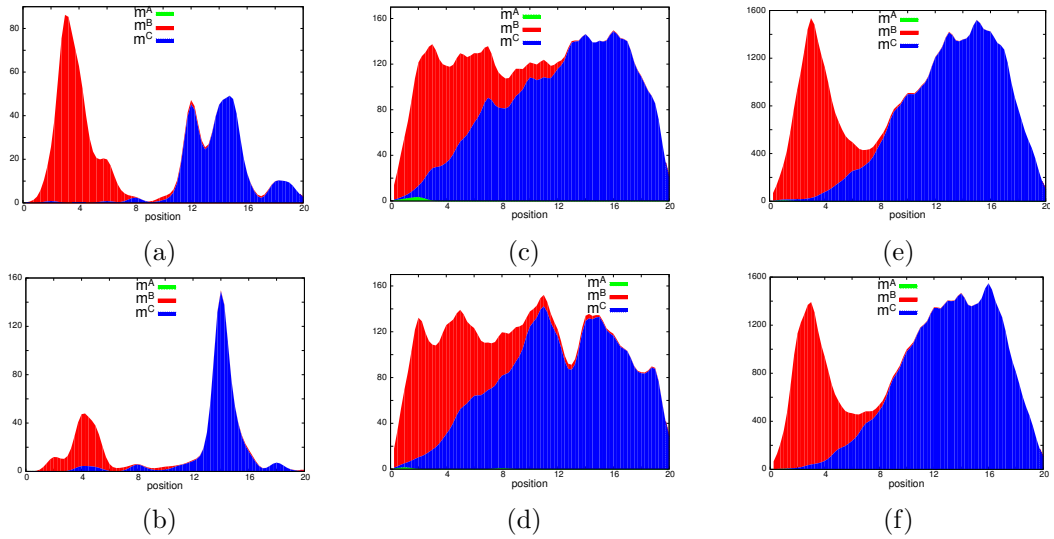


Figure 5.6: Two snapshots taken at two different time instants for (a)-(b) a system with rates of type (I) (c)-(d) a system with rates of type (II) (e)-(f) a system with rates of type (III). The parameters used are the same as for fig. 5.4 and are given in appendix. Note the extreme variation over time in case (I), and the milder fluctuations in cases (II) and (III).

We now analyse in some detail how the mass profiles respond to a change in influx, for each of the three case (I)-(III). This response can be quantified using the relative change $\delta m^Z/\langle m^Z\rangle$ that occurs at a location with average mass $\langle m^Z\rangle$, if the influx rate a is increased by δa [6]. From the relationship between $\langle m^Z\rangle$ and $\langle f\left(m^Z\right)\rangle$ discussed earlier, it follows that for for $\langle m^Z\rangle \gg 1$ and $\delta a/a \to 0$, we have:
Case (I): $\delta m^Z/\langle m^Z\rangle \propto \langle m^Z\rangle(\delta a/a)$
Case (II): $\delta m^Z/\langle m^Z\rangle \propto 2(\delta a/a)$
Case (III): $\delta m^Z/\langle m^Z\rangle \propto 2(1+\sqrt{\langle m^Z\rangle/m^Z_{sat}})(\delta a/a)$

From the above, we can see that the strongest response occurs for rates of type (I), followed by rates of type (III), and the mildest for type (II) rates. Further, in case (II), the proportionate change is the same for different portions of the mass profile, so that the shape of the mass profile is preserved even as it shifts upwards due to an increase in influx. By contrast, for rates of type (I) and (III), the response depends on the region, with the maxima of the mass profile showing much greater proportionate change, as compared to regions with small mass. Thus, in case (I) and (III), the mass profile changes shape as influx is increased, becoming more

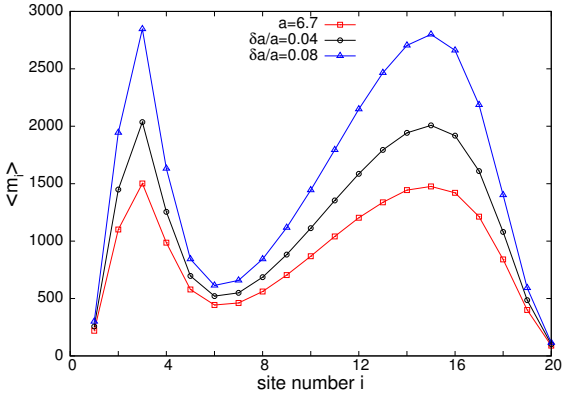


Figure 5.7: $\langle m_i\rangle$ vs. i for 3 different values of influx rate a for a system with rates of type (III): mass profile changes shape (becomes more peaked) as a is increased.

[6] Variation in rate of influx is an important perturbation to consider, as it affects $\langle m^Z\rangle$ for all species Z in a similar way. Thus, the response of the total mass profile is simpler to analyse.

and more sharply peaked (see fig. 5.7 for a representative example). Moreover, for these types of rates, if the increase δa in influx is sufficiently large, then the system can undergo a phase transition, wherein large aggregates undergo an instability, and start growing without bound. Since this is clearly unphysical in the context of the Golgi, an additional mechanism is needed to stabilise the system and curtail runaway growth. In sec. 5.3.1, we discuss one such mechanism, namely the introduction of stack movement at a small rate. Another conclusion from the above discussion is that a higher degree of fine tuning is required for rates of type (I) as compared to type (III), to obtain sharp peaks of a certain height in the mass profile.

All the main results derived above also hold for generalised versions of the model with a higher number of species. Fig. 5.8 shows the average mass profile for a particular choice of parameters for the four-species model with $A \to B \to C \to D$ conversion, and chipping and interconversion rates of the Michelis-Menten sort (type (III)). Parameters can be chosen such that the maxima of the $\langle m_i^B \rangle$, $\langle m_i^C \rangle$ and $\langle m_i^D \rangle$ profiles lie in the bulk and are sufficiently well separated. Further, by tuning the $\langle f(m_i^Z) \rangle$ corresponding to each maximum to be close to its saturation value, the maxima can be made sharp, so that the total mass profile has three distinct peaks, corresponding to particles of type B, C and D.

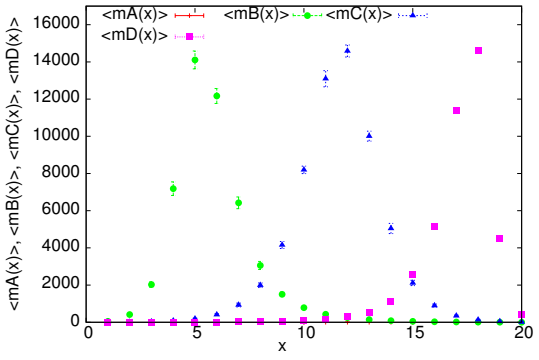


Figure 5.8: Average mass profiles $\langle m_i^A \rangle$, $\langle m_i^B \rangle$, $\langle m_i^C \rangle$ and $\langle m_i^D \rangle$ for the 4-species model with $A \to B \to C \to D$ model. Parameters chosen (see appendix) to obtain multiple, distinct peaks.

To summarise, in this section, we have analysed the only-chipping model for three different kinds of dependence of chipping and interconversion rates on particle number. In all three cases, the mass profiles exhibit multiple peaks where each peak corresponds to a different species of particles. For individual peaks to be distinct and resolvable, the width of the maximum of each $\langle m^Z \rangle$ profile should be small. However, in our model there appears to be a trade-off between the sharpness (smallness of half-width) and stability of maxima in the mass profiles. Sharp maxima can be obtained with (type I) rates that are independent of the number of particles at a site, but such maxima are highly unstable, with the fluctuations about the maximum being almost equal to the value of the maximum itself (such large fluctuations reflect the fact that a large aggregate can disappear completely once in a while). Stable maxima with small fluctuations in height, are obtained for (type II) chipping rates that increase with number of particles at a site, but with such rates, it is not possible to tune the width of the maxima to be small. The best compromise between the competing requirements of stability and sharpness is obtained with (type III) 'crossover rates' which increase with the number of particles at a site when the number is small, but saturate to a constant value for large numbers. With such rates, it is possible to get aggregates that are somewhat sharply peaked and stable in time. However, even with such rates, the steady state structures are quite sensitive to perturbations– the mass profiles exhibit a strong shape-altering response to a small change in parameters (though the response is weaker than in case I), and can

even undergo an instability, leading to runaway growth if the change is sufficiently large. In subsequent sections, we will only consider rates of type III, when comparing the only-chipping model with the only-stack-movement model.

5.3.1 Perturbation about the only-chipping limit: modeling mixed transport

The discussion of the only-chipping limit suggests that some additional process is needed to ensure stability of aggregates in the face of small variations in parameters and to also prevent runaway growth. Here, we briefly comment on the effect of introducing perturbations such as stack movement to the only-chipping model, and their effect on the robustness of the mass profile. Allowing for stack movement at a small rate is also relevant in view of the theory that Golgi transport is not mediated either purely by vesicles or purely by cisternae, but involves both to a different extent.

As discussed in Chapter 4, the inclusion of stack movement in the only-chipping model (for instance, in the forward direction) at any non-zero rate D is sufficient to eliminate runaway growth. If the rate D is comparable to the chipping rates, then stack movement tends to homogenize the system, leading to mild gradients rather than sharply localised spatial distributions of various species of particles (see sec. 5.4). On the other hand, if stack movement is introduced at a very small rate D to a system which in the $D = 0$ shows pile-ups, then this leads to the emergence of sharply peaked but *finite* aggregates in the same region where the pile-up occurred in the corresponding $D = 0$ model. Aggregates obtained in this way are thus, more robust as they always remain finite and do not require as much fine tuning of parameters. However, the mass profile now shows significant stochastic fluctuations over time scales proportional to $1/D$, as an entire stack can move away once in a while.

Another strategy for improving the robustness of the only-chipping model is to allow chunks containing a finite fraction α of the mass (rather than the entire mass) at a location to break off at a small constant rate, and move to a neighbouring location to join with the mass present there. As before, if the rate of breakage is very small, it does not alter the basic multi-peak nature of the mass profile, but ensures greater robustness, by ensuring that there is no runaway growth [7], even if there is a substantial change in other rates. Moreover, stochastic fluctuations about the average mass profile in this case are smaller (with $\Delta m/m \sim \alpha$) than in the case where entire stacks are allowed to move, where $\Delta m/m \sim 1$.

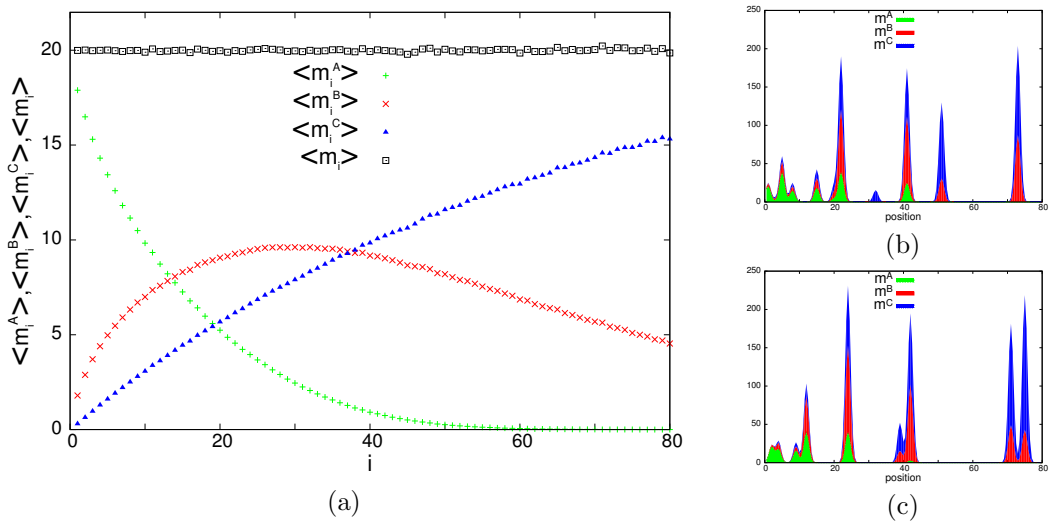
5.4 Modeling Cisternal Progression by the Only-Stack-Movement Model

We now consider the alternative scenario for intra-Golgi transport, namely cisternal progression and maturation. As discussed in sec. 5.1, in this view, a Golgi cisterna forms at the cis end due to fusion of incoming vesicles, progresses through the Golgi in the cis to trans direction, carrying cargo molecules with it, and disintegrates at the trans end. The contents of the cisterna also undergo maturation or processing as the cisterna moves, with

[7] As discussed in Chapter 4, any perturbation which makes the particle current flowing out of a site with m particles proportional to m, eliminates pile-ups.

different enzymes getting attached to a cisterna in different stages of its progression. In our simplified model, cisternal progression can be represented by the *uni-directional* movement of stacks from left to right, with stacks aggregating when they meet to form larger stacks [8]. The analogue of cisternal maturation is the sequential $A \to B \to C \to$ conversion of particles within each stack. Thus, in effect, we use the only-stack-movement model with interconversion between species to model cisternal progression and maturation.

In chapter 4, we outlined the basic features of the steady state of the multi-species only-stack-movement model. To recapitulate, the mass in the system is contained in a few moving aggregates. The typical mass of an aggregate increases with increasing distance x from the point of injection, while the probability of finding an aggregate decreases with x. Thus, aggregates become larger but sparser from left to right. In the case where rate of stack movement is independent of the mass of stack, the typical mass at a site scales as $\sqrt{x}$, the probability that the site is occupied by an aggregate scales as $1/\sqrt{x}$, so that $\langle m(x) \rangle$ is constant and $\langle m^2(x) \rangle \propto \sqrt{x}$, and so on for higher moments. The system also exhibits spatial gradients in the abundance of different species, with the composition of aggregates changing from A-rich to B-rich to C-rich on moving from left to right. Another distinctive feature of this system is that the total mass in a system of size L exhibits temporal intermittency on time scales that are proportional to $\sqrt{L}$ (due to the exit of aggregates of typical size $\sqrt{L}$, from the system at $\mathcal{O}(\sqrt{L})$ time intervals).



(a) (b) (c)

Figure 5.9: (a) The average mass profiles $\langle m_i^A \rangle$, $\langle m_i^B \rangle$, $\langle m_i^C \rangle$ and $\langle m_i \rangle = \langle m_i^A \rangle + \langle m_i^B \rangle + \langle m_i^C \rangle$ for a system in the only-stack-movement limit (parameters specified in appendix).
(b)-(c) Two snapshots of the same system taken at two different time instants. Note how the two snapshots are quite different from each other, and very different from the average mass profiles.

[8] In the Golgi, most fusion events involving vesicles are believed to occur near the ER, leading to the formation of the cis cisternae. In our model, we assume that two stacks fuse to form a larger stack upon meeting *anywhere* in the system. However, in the model, the particles naturally organise themselves into a state where the probability of two stacks encountering each other decreases with distance from source, so that most fusion events happen close to the source.

As before, in order to replicate Golgi-like structures, the mass profile must possess multiple, spatially resolvable maxima with a position-dependent chemical composition. Further, if we monitor a particular aggregate in time, its position, composition and possibly size should change with time: this is the analogue of cisternal maturation in our model. Also, even though individual cisternae are dynamic, the system should be stationary in the sense that the probability distributions of various quantities, obtained by averaging over a large number of configurations, are stationary at long times. Finally, these features should be robust to small variations in the parameters of the model.

From the discussion above, it is clear that most of these requirements are automatically satisfied in the only-stack-movement limit of the model. For instance, the snapshots of the system depicted in figs. 5.9b, 5.9c, show that a typical configuration of the system contains multiple resolvable peaks. The emergence of distinct aggregates happens naturally in this limit, without any fine tuning, because of the tendency of a moving aggregate to gather particles it encounters in its neighborhood, thus leading to the concentration of mass. Spatial gradients of chemical species also emerge naturally due to the sequential $A \to B \to C$ conversion, as evident in the average mass profiles $\langle m_i^A \rangle$, $\langle m_i^B \rangle$, $\langle m_i^C \rangle$ in fig. 5.9a. The length scales governing the gradients depend on the interconversion rates and the function $f(m)$. For slowly varying or system-wide gradients, the interconversion rates must be much smaller than the rate of stack movement.

We can also track an individual stack (cisterna) in time in simulations; for a typical stack, the position and mass change as t and $\sqrt{t}$ respectively with time t, while the composition changes from A-rich to B-rich to C-rich. Thus, each stack 'matures' as it moves through the system. Finally, we note that the system is robust in the sense that very little fine tuning of rates is required to generate such mass profiles. Another way of assessing robustness is to monitor the response of the system to small variations in parameters, e.g., the influx rate a. If the rate of stack movement is mass-independent then the system shows a shape-preserving response, with $\delta m_i^Z / \langle m_i^z \rangle \propto \delta a / a$.

A noteworthy feature of the structures generated in the two limits is that the average mass profile (fig. 5.9a) and typical configurations (figs. 5.9b, 5.9c) are dramatically different. In fact, while multiple maxima appear in any *typical* configuration of the system, the *average* mass profile, which is obtained by averaging over many such configurations, is flat. This is just a manifestation of the large fluctuations that arise in the system, due to movement of entire stacks. This is also very different from the only-chipping limit (with rates of type III) discussed in sec. 5.3, where the average and typical mass profiles were quite similar. In sections 5.5 and 5.6, we explore whether the qualitatively different behaviour of fluctuations in the two limits can be used to deduce the transport processes at work in a system.

To investigate whether the qualitative features described above depend on our choice of the rule for stack movement, we also study two variants of this model. First, we consider the case where the rate D of stack movement depends on the mass m of the stack, going as $D(m) \sim m^{-\alpha}$ (with $0 < \alpha < 1$), such that larger stacks move more slowly than smaller stacks. The second variant we study involves rates of stack movement that depends on the proportion of A, B and C particles in the stack. Preliminary numerics suggest that in both

these cases, the system mass is concentrated in a few aggregates, which become larger and sparser with distance from the point of injection. Further, as in the case of mass-independent and composition-independent D, the composition of the stacks changes with position. Thus, these generalised versions are qualitatively similar to the constant D model we have studied, and differ only in the specific form of the x-dependence of various quantities.

5.4.1 Perturbations about the only-stack-movement limit: a different regime of mixed transport

We also investigate whether the basic properties of the only-stack-movement model remain unaltered in the presence of single particle chipping at a small rate. This would represent a 'mixed' scenario of intra-Golgi transport, in which bulk of the transport is via cisternal movement, but with vesicle exchange between cisternae also taking place. As discussed in Chapter 2, if the chipping rate is small and m-independent, then the system continues to be in an aggregation-dominated phase, and supports large aggregates, as in the zero chipping limit. Numerics show that this is true even if chipping occurs with Michelis-Menten rates, as long as the maximum (saturation) rate is very small. Thus, we conclude, that if the chipping rates for A,B,C particles are all sufficiently small, then inclusion of chipping does not qualitatively alter the main features of the only-stack-movement model.

An interesting situation can arise if chipping rates are very different for the three species. For example if the chipping rate for particles of type C is large, while the corresponding rates for A and B are small, then large aggregates form close to cis end but disintegrate into smaller fragments near the trans end. This scenario is illustrated in fig. 5.10, and is also relevant for the Golgi, where cisternae are believed to disintegrate near the trans end into smaller vesicles which then leave the Golgi for different locations.

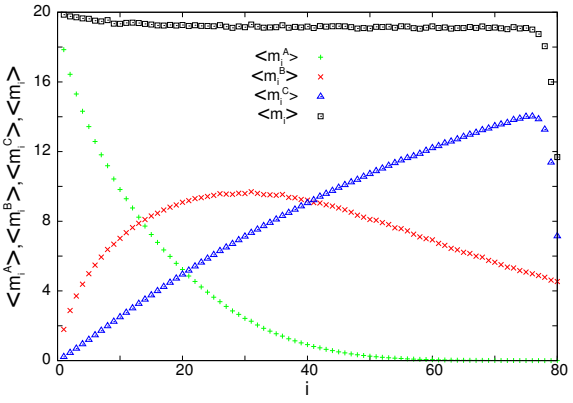
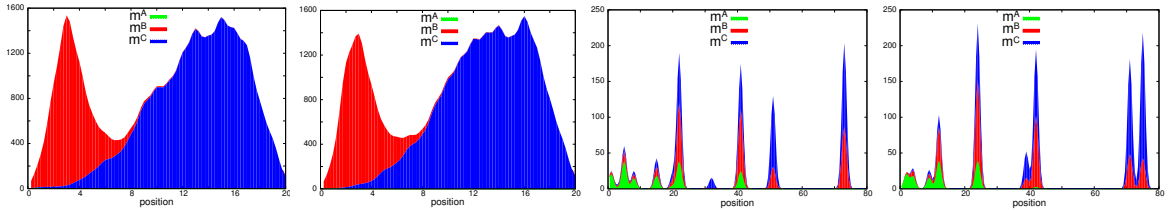


Figure 5.10: Average mass profiles $\langle m_i^A \rangle$, $\langle m_i^B \rangle$, $\langle m_i^C \rangle$ and $\langle m_i \rangle$ for a model with movement of stacks and chipping of only C particles. Average mass dips close to trans end.

5.5 Comparing the two limits: Static features

So far, we have studied two limits of the model that correspond to vesicular transport and cisternal progression, with the aim of identifying regions in parameter space, for which these yield Golgi-like structures. While the structures obtained in this way in the two limits share some similarities such as the presence of multiple, resolvable maxima in the mass profile, with a position-dependent composition, there are also significant differences. In this section, we list some points of difference in the static properties of the structures in the two cases. Many of these are evident from a comparison of typical 'snapshots' of the system (fig. 5.11). In



(a) (b) (c) (d)

Figure 5.11: Two snapshots of the system at two different instants for (a)-(b) the only-chipping model with Michelis-Menten rates of chipping and interconversion (c)-(d) the only-stack-movement model with Michelis-Menten rates of interconversion. Parameters same as in fig. 5.4c and fig. 5.9

the next section, we focus on dynamical measurements that bear signatures of the transport mechanism and can hence, be used to distinguish between the two scenarios (at least within our simplified model). We also comment on whether the distinctive features of the structures in each limit depend on fine details of the model or are more general, and hence have greater relevance to the Golgi.

Some salient differences in the properties of the structures in the two limits are as follows:

1. The only-chipping model with Michelis-Menten rates generates *stable* aggregates [figs. 5.11a,5.11b] whose position, size and composition does not change much over time. By contrast, in the only-stack-movement limit, aggregates are highly *dynamic* [figs. 5.11c,5.11d]– their positions, composition and even number can show significant variation over time. The stability of aggregates is a fairly general feature of the only-chipping model, and is ensured if $f(m)$ is an increasing function of m at least for small m (see sec. 5.3). [9]
   Relevance for Golgi: In addition to the local, stochastic processes incorporated in our model, there could also be other long-range communication and control mechanisms in the real system, which suppress variability. Nevertheless, the main conclusion that movement of large cisternae would be associated with large fluctuations and hence, significant variations over time, appears to be a reasonable one. *Thus, monitoring the extent of variability in static images of the Golgi cisternae taken at different time instants could provide clues to the mechanism of transport.*

2. In the only-chipping model with Michelis-Menten rates, the mass profile is *continuous*, with a significant amount of mass in the region between large aggregates. In the only-stack-hopping model, on the other hand, the aggregates have a more *discrete* character, with the region between adjacent aggregates containing little or no mass. While this distinction between the two kinds of structures emerges quite naturally in our model, it is not obvious whether it would hold if the details are altered, for instance, if the small m behaviour of $f(m)$ in the only-chipping model is chosen to be of a from which is different from types (I)-(III) considered in this study. Similarly, the separated character of the aggregates in the stack-movement limit may depend on the extent to which a

[9] If there are fluctuations in the rates of underlying processes in the system, then the mass profile can show significant fluctuations even in the only-chipping limit, due to the high sensitivity to parameters of the model. Such a high degree of sensitivity is, however, not very realistic in a biological system, and as discussed earlier, there is likely to be a compensatory mechanism which makes the system more robust.

moving aggregate collects the mass in its vicinity.
Relevance for Golgi: The present model predicts that interconnected cisternae are correlated with vesicular transport and more discrete cisternae with cisternal movement. While the exact nature of this prediction may hinge on specific details of the model, *the broader conclusion that there is a connection between the degree of interconnectedness of cisternae and the mode of mass transport in the system, could be relevant to the Golgi.* This also suggests that an interesting direction for future study would be to compare Golgi morphologies (especially inter-cisternal continuities) across different cell types, and investigate whether there is a correlation between these and the transport mechanism in the cell (as probed by other measurements).

3. In the only-chipping model, particles of a particular type are *sharply localised*, i.e., most are concentrated at the location of a single peak (or aggregate), and not spread across multiple aggregates (see figs. 5.11a, 5.11b). In the only-stack-movement model, the spatial distribution of particles of any species shows a *gentle gradation*; the system typically contains many aggregates of mixed composition, with the relative abundance of different species in the aggregate changing gradually in the cis to trans direction (see figs. 5.11c, 5.11d). This distinction between the structures in the two limits does not appear to depend on the fine details of the model.
Relevance for Golgi: Measuring the spatial profiles of various Golgi-resident enzymes and other markers associated with the processing level of cargo molecules, could be a fruitful direction for experiments. As before, *by simultaneously carrying out measurements that probe transport properties more directly, it may be possible to establish whether there is a general connection between the mechanism of transport in the Golgi, and the the spatial localisation/delocalisation of enzymes and markers.*

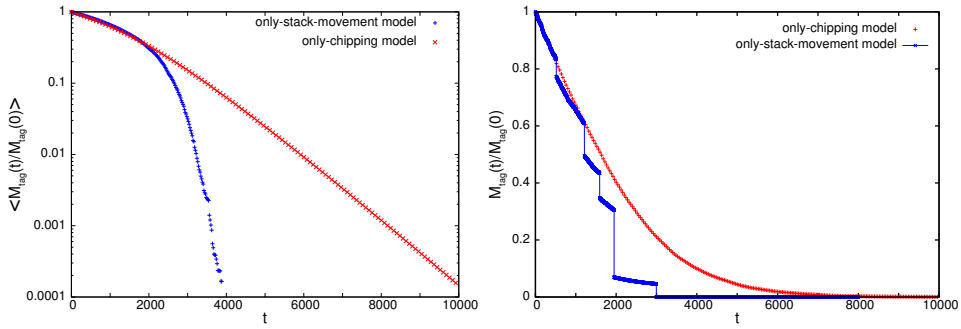
4. The structures generated in the two limits respond differently to an increase in influx. This cannot be assessed from snapshots, and involves perturbing the system. In the only-chipping model with Michelis Menten rates, the mass profile is very sensitive to small changes in the influx, with large aggregates showing a much larger proportionate change as compared to smaller aggregates, so that the overall *shape of the mass profile changes* and becomes more sharply peaked. In the only-stack-movement limit, the response is *shape preserving* ($(\delta m/m)/(\delta a/a)$ independent of m). Within the parameters of our model at least, the above features are fairly robust. The strong, shape-altering response of the mass profile in the only-chipping limit hinges only on the boundedness of $f(m)$ for large m (which is a pre-requisite for obtaining distinct aggregates in this limit).
Relevance for Golgi: This suggests that *treating the Golgi with reagents which alter budding from the ER and hence influx into the cis Golgi, and monitoring the resultant change in the size of Golgi cisternae, could be a relevant direction for experiments.*

5.6 Comparing the two limits: Dynamical properties

We describe below two dynamical measurements that can be used to distinguish between the structures generated in the two limits.

Tagged particle dynamics: Motivated by fluorescence experiments on the Golgi in mammalian cells [16], we perform numerical simulations in which all the particles present in the

system at a particular time instant $t = 0$ are tagged, and the number $M_{tag}(t)$ of tagged particles in the system is monitored as a function of time t. Numerics show that for the model with stack movement, the average concentration $\langle M_{tag}(t)/M_{tag}(0)\rangle$ (averaged over many realisations of the dynamics) decays linearly ($\sim 1 - t/\tau_A$), whereas for the only-chipping model, the decay is nearly exponential ($\sim \exp[-t/\tau_B]$) [see fig 5.12a]. Experiments on certain cell types have also reported an exponential decay [16].



(a) (b)

Figure 5.12: (a) Average fraction $\langle M_{tag}(t)/M_{tag}(0)\rangle$ of tagged particles in the system as a function of time t for (i) only-chipping model (exponential decay) and (ii) only-stack-movement model (linear decay). The curves are obtained by averaging over ~ 500 realisations of the dynamics. Unit of time for the two curves has been chosen differently in order to display both clearly in the same plot.
(b) A single realisation of $M_{tag}(t)/M_{tag}(0)$ for the only-chipping model and the only-stack movement model. Distinctive sharp drops in the latter case arise due to exit of large aggregates from the region.

We also observe $M_{tag}(t)/M_{tag}(0)$ during a *single* realisation of the dynamics, and find that it almost coincides with the average $\langle M_{tag}(t)\rangle/\langle M_{tag}(0)\rangle$ for the only-chipping case, but deviates significantly from the averaged curve for the only-stack-movement case. In fact, a single realisation of $M_{tag}(t)/M_{tag}(0)$ for the only-stack-movement model is characterised by strong fluctuations, and distinctive sharp drops (see fig 5.12b) This suggests that fluctuations, which are often treated only as noise to be averaged over, may actually provide valuable information about the system, and should be carefully analysed in such experiments .

The above procedure can also be performed for a portion of the system, by tagging all the particles present in a particular region, and monitoring the decay in the number of tagged particles in the same region [10]. As before, the decay is exponential or linear depending on whether mass exits the region in the form of single particles or stacks. Carrying out this measurement separately for the cis, medial and trans regions of the Golgi is important in a scenario where cisternal movement occurs in the cis and medial Golgi, but the cisternae disintegrate into vesicles near the trans end before exit from the Golgi.

In [16], the observation of exponential decay was taken as evidence against cisternal progression. However, as argued by Dmitrieff et al in [85], such an observation simply implies

[10]This would be different from the experiment performed in [16], where a pool of particles localised typically in the cis region is tagged fluorescently, and the decay in fluorescence in the entire Golgi region is monitored

that there is some single particle (vesicle) exchange and efflux, which could even be occurring in conjunction with cisternal movement.

Dynamics of the total number of particles: Another possible measurement involves tracking the total mass $M(t)$, i.e., the total number of particles in the system as a function of time t. Figs. 5.13a, 5.13b show a representative time series $M(t)$ for each of the two cases, as obtained from numerical simulations. While the time series in the only-chipping model has a smoothly varying, nearly Brownian character, the time series in the only-stack-movement model is 'intermittent' and displays sporadic sharp drops, corresponding to the exit of large aggregates. As discussed in Chapter 2, this difference between the time series is captured by a statistical measure known as flatness, which is given by $\kappa(t) = \langle [M(t) - M(0)]^4 \rangle / \langle [M(t) - M(0)]^2 \rangle^2$ [11]. The behaviour of $\kappa(t)$ in the two cases is depicted in fig. 5.13c. In the only-stack-movement model, $\kappa(t)$ diverges as $t \to 0$, where the time scale over which the divergence occurs, is the typical time interval between successive exit events. In the only-chipping model, there is no such divergence of flatness.

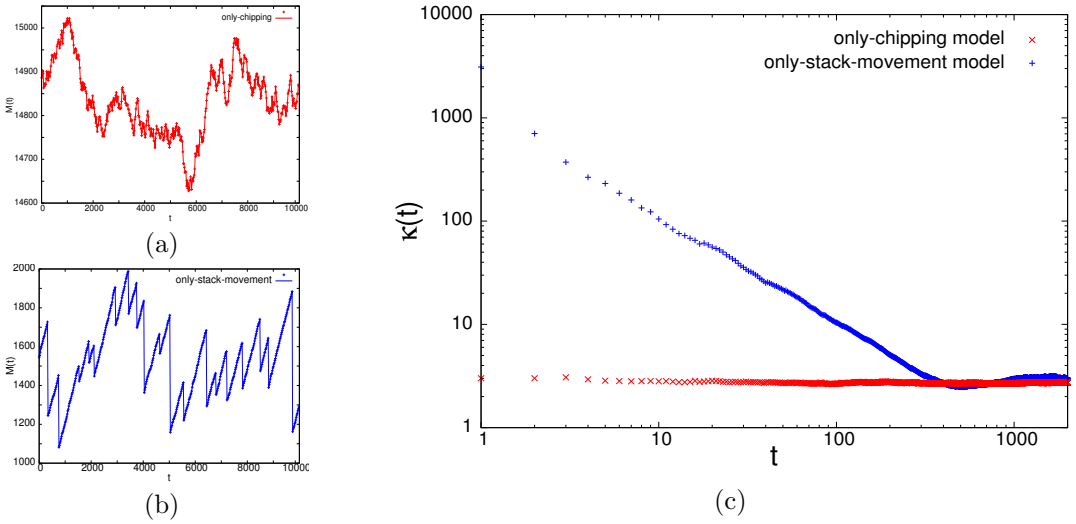


Figure 5.13: A typical time series $M(t)$ vs. t in the (a) only-chipping model (b) only-stack movement model. Note that the y-axis in (a) and (b) has a different scale and origin.
(c) Flatness $\kappa(t)$ vs. t in the two cases on a log-log plot: The occurrence of sharp drops in $M(t)$ is the only stack movement case is captured by the small t divergence of $\kappa(t)$. No such divergence in the only-chipping case.

In the context of the Golgi, such a measurement would involve using a fluorescent molecule which is present in the Golgi in abundant quantities (possibly after over-expression), and measuring the fluorescence signal from the Golgi region as a function of time. As before, this measurement could also be carried out for different parts of the Golgi separately. Note that this is different from the previous measurement where only the proteins present in the Golgi at $t = 0$ were made fluorescent, so that there was a monotonic decay in the fluorescence from the Golgi in time. In the present measurement, the fluorescence from the Golgi region

[11] For convenience, we will define a unit of time as the typical time interval over which an injection event takes place

is not expected to show an overall change, but fluctuate about a steady state value. As discussed above, it is these fluctuations that have very different characteristics in the only-chipping and only-stack movement limits, which are the analogues of the vesicular transport and cisternal progression scenarios respectively. Thus, *a systematic analysis of the statistics of these fluctuations using measures such as flatness, could be an effective way of probing the transport mechanism.* However, measuring steady state fluctuations of the fluorescence signal requires very high resolution.

5.7 Conclusions and Discussion

In this thesis, we have defined and studied a simple stochastic model with local rules inspired by Golgi dynamics. The aim of the present chapter was to harness our understanding of various aspects of the model into identifying conditions under which the particles in the system self-organise into Golgi-like structures, i.e., into multiple, spatially resolvable aggregates with a position-dependent chemical composition. For this purpose, we considered two limits of the model, namely the only-chipping limit and the only-stack-movement limit, which correspond respectively to the vesicular transport and cisternal progression (and maturation) scenarios for intra-Golgi traffic.

Obtaining sharp, spatially resolved aggregates which are also stable in time requires some fine-tuning of parameters in the only-chipping model, where there is a trade-off between the sharpness and stability of maxima in the mass profile. The competing requirements of stability and sharpness of aggregates are best satisfied in the only-chipping model with Michelis-Menten kind of chipping rates which increase with increasing size of an aggregate for small aggregates but saturate to a constant value for large aggregates. In the only-stack-movement limit, by contrast, the particle in the system generically organise into a few spatially separated aggregates with position-dependent composition, without much fine tuning of rates. We also consider small perturbations about the two limits, for instance by introducing stack movement at a very small rate in the only-chipping model (and vice versa) and find that many of the features of the pure limits remain qualitatively unaltered, as long as the perturbation is sufficiently small.

Further, we contrast the structures obtained in the two limits and find several interesting differences. For example, simple snapshots of the system in the two limits reveal that particles of a particular species or type are far more spatially localised in aggregates obtained in the only-chipping model as compared to the only-stack-movement model. We also perform numerical simulations of dynamical measurements, and find that the time-evolution of various quantities is strikingly different in the two cases, with the only-stack-movement model, in particular, being characterised by temporal intermittency. The two biologically relevant limits of the model can also serve as a starting point for simulation studies of other experimentally relevant perturbations/protocols such as blockade of outflux at the trans end, and Golgi disassembly/re-assembly dynamics. We have performed some preliminary simulations for these measurements (which are not reported in the thesis), but have not been able to find clear signatures of the underlying transport process in these. Also, for any such measurement, it is important to assess whether the observed features are typical of that limit of the model or depend on the specific choice of parameters. In the absence of an intuitive understanding,

this requires an exploration of the entire high-dimensional parameter space, which is difficult to carry out.

We now outline some general conclusions and directions suggested by our study. (i) Whether and how the Golgi can form de novo (for example, after treatment with reagents that cause Golgi cisternae to dismantle) has been an important question in cell biology [86]. In our work, we show that a system with Golgi-like traffic processes naturally self-organises into structures that are qualitatively similar to Golgi cisternae under certain conditions. Further, this kind of self-organisation can happen even in the limit corresponding to vesicular transport [12]. (ii) A second broad conclusion that emerges from this work is that the transport mechanism could have a qualitative effect on the structural features of the cisternae such as inter-cisternal continuities or the degree of localisation of Golgi-resident enzymes. Thus, it may be useful to carry out transport measurements in the Golgi in conjunction with a detailed observation of morphological features of the Golgi cisternae across a variety of cells, to investigate whether there are some correlations between the two. (iii) Finally, the study suggests that fluctuations can provide valuable insight into Golgi dynamics; the appearance of sporadic, relatively large fluctuations in particular, may be a manifestation of cisternal movement. Analysing the statistics of fluctuations, for example in fluorescence experiments, could thus be an important direction for experiments.

A significant limitation of this study is that it does not explicitly take into account the role of enzymes in the processing of cargo molecules in the Golgi, and the complex dynamics of the enzymes themselves. Thus, extending the present model to include enzymes as separate entities would lead to a more realistic model, the parameters of which could also be connected to experimentally relevant quantities. However, the increase in the number of parameters also makes it difficult to analyse such an extended model and extract the relevant regions of parameter space. An abstract model of the sort studied in this chapter, is more suited to gaining insight into the qualitative effects at play in the system. Nevertheless, it would be useful to consider a 2-species model with enzyme-mediated $A \to B$ conversion, and investigate whether biophysically relevant rules for enzyme dynamics and enzyme-cargo interactions lead to the more abstract model with effective rates that we have studied here.

In conclusion, the present model which incorporates the basic processes involved in Golgi traffic, can serve as a useful starting point for more detailed models of the Golgi. More importantly, in comparing different scenarios for Golgi traffic or envisioning new ones, and also in designing experiments and interpreting observations, it is useful to understand the qualitative effects that can arise simply by self-organisation in a minimal model of this sort.

[12] In the cell biology literature, de novo formation has been associated with the cisternal progression scenario (where the formation of cisternae can be understood intuitively) or with the existence of a Golgi matrix or 'template' on which cisternae can reassemble, even with vesicular transport. In our model, distinct cisternae emerge in the only-chipping (pure vesicle transport) limit simply by the interplay of various processes, even in the absence of a template.

5.8 Appendix E: Parameters used for various figures in the chapter

Fig. 5.4a, Fig. 5.5 (type I rates), Figs. 5.6a, 5.6b: $a = 0.574$, $u = 0.8$, $k = 0.00972$, $w = 6.0$, $q = 0.433$, $p = 0.28$, $\gamma_A = 0.5$, $\gamma_B = 0.54$, $\gamma_C = 0.65$, $D = 0$

Fig. 5.4b, Fig. 5.5 (type II rates), Figs. 5.6c, 5.6d: $a = 5.0$, $u = 0.8$, $k = 0.01$, $w = 5.0$, $q = 0.4$, $p = 0.215$, $\gamma_A = 0.53$, $\gamma_B = 0.5$, $\gamma_C = 0.65$, $D = 0$

Fig. 5.4c, Fig. 5.5 (type III rates), Figs. 5.6e, 5.6f, Figs. 5.11a, 5.11b, plots corresponding to the only-chipping limit in Fig. 5.12 and Fig. 5.13: $a = 6.7$, $u = 0.18$, $k = 0.0155$, $w = 5.0$, $q = 0.1$, $p = 0.29$, $\gamma_A = 0.5$, $\gamma_B = 0.54$, $\gamma_C = 0.68$, $D = 0$;
$f_Z\left(m^Z\right) = \sqrt{m}\sqrt{m^Z_{sat}}/(\sqrt{m} + \sqrt{m^Z_{sat}})$ where $m^A_{sat} = m^B_{sat} = m^C_{sat} = 200$

Fig. 5.7: Parameters for red curve same as in fig. 5.4c above. For black and blue curves, $a = 6.968$ and $a = 7.236$ respectively, while all other parameters are held fixed.

Fig. 5.8: $a = 1.17$, $u = 0.03$, $k = 0.000093$, $g = 0.00000835$, $w = 0.4$, $q = 0.1505$, $p = 0.0059$, $r = 0.0002118$, $\gamma_A = \gamma_B = 0.5$, $\gamma_C = 0.54$, $\gamma_D = 0.75$, $D = 0$;
$f_Z\left(m^Z\right) = \sqrt{m}\sqrt{m^Z_{sat}}/(\sqrt{m} + \sqrt{m^Z_{sat}})$ where $m^A_{sat} = m^B_{sat} = m^C_{sat} = m^C_{sat} = 100$
Here, r denotes the chipping rate for 'D' type particles and g is the rate for $C \to D$ conversion.

Fig. 5.9, Fig. 5.11c, 5.11d, plots corresponding to the only-stack-movement limit in Fig. 5.12 and Fig. 5.13: $a = 20$, $u = 0.75$, $k = 0.375$, $w = q = p = 0$, $D = 1.0$, $\gamma = 1.0$;
$f_Z\left(m^Z\right) = \sqrt{m}\sqrt{m^Z_{sat}}/(\sqrt{m} + \sqrt{m^Z_{sat}})$ where $m^A_{sat} = m^B_{sat} = m^C_{sat} = 200$

Fig. 5.10: $a = 20$, $u = 0.75$, $k = 0.375$, $w = q = 0$, $p = 10.0$, $D = 1.0$, $\gamma = 1.0$ $\gamma_C = 0.5$;
$f_Z\left(m^Z\right) = \sqrt{m}\sqrt{m^Z_{sat}}/(\sqrt{m} + \sqrt{m^Z_{sat}})$ where $m^A_{sat} = m^B_{sat} = m^C_{sat} = 200$

Chapter 6

Conclusions

In this thesis, we have defined and studied a stochastic model inspired by the phenomenology of structure formation and transport in the Golgi apparatus of cells. The model is a simple lattice model in which different kinds of molecules are represented by different species of 'particles', and the main processes involved in Golgi dynamics are represented as elementary moves of these particles. Thus, we deal with a model with multiple species of particles that undergo processes such as diffusion and aggregation, fragmentation, interconversion between different species, as well as influx and outflux at the boundaries. While the model ignores molecular details and is necessarily very simple, it lends itself to an economical description of the biological system, which makes it easier to identify and study broad qualitative effects at play in the system.

The elementary processes in the model play an important role in several other physical phenomena, and have been studied in different contexts in statistical physics. One of the main questions we have explored in this thesis is whether the *interplay* of these processes can give rise to new kinds of non-equilibrium phases. Identifying and characterising the static and dynamical features of these phases formed a major part of this work, and has been described in detail in Chapters 2, 3 and 4 of the thesis.

A second aspect of this work involved identifying the limits of the model in which the system *self-organises* into a state that exhibits the rudimentary features of Golgi organisation. The steady states generated in these limits were examined in detail, to assess whether they bear (experimentally) observable signatures of the underlying transport process. This work has been described in Chapter 5.

Below we discuss some broad conclusions that emerge from our study.

1. The open boundary aggregation-fragementation model undergoes a phase transition to an aggregation-dominated phase characterised by the formation of a macroscopic aggregate. The presence of large mobile aggregates or clusters can have a strong imprint on the dynamical properties of the system. In our work, we monitor the time series of the total mass in the system, and find that it is non-self-similar and exhibits intermittency in phases with formation and exit of large aggregates. The breakdown of self-similarity can be quantified using dynamical structure functions, which we define in analogy with spatial structure functions of the velocity field, as used in fluid turbulence [57]. For

intermittent time series, the structure functions are characterised by anomalous scaling exponents that cannot be predicted from a Kolmogorov-like self-similarity hypothesis [57]. We were able to calculate these anomalous exponents analytically in one limit of the model. While we have focused on temporal intermittency of the total number of particles in an open system, the basic results are more general, and could also hold for the time evolution of particles in any *part* of a closed system with aggregate formation. An interesting direction for future work could be to analyse time series of mass or particle number in other systems with giant number fluctuations, such as collections of self-propelled particles [12], and investigate whether these also exhibit temporal intermittency. Dynamical structure functions, as used in this study, could be useful probes in these systems as well.

2. Systems with clustering or aggregation typically show strong density inhomogeneities and anomalously large number fluctuations [62]. However, the length scale that governs the scale of these fluctuations depends crucially on whether or not the systems allows for hard core constraints between particles. In cluster-forming systems with hard core constraints [60, 61], the fluctuations in a region of linear size Δl scale as $(\Delta l)^d$, and do not depend on the system size L, at least to leading order. By contrast, in the absence of hard core constraints, the fluctuations in any region of the system scale with system size L rather than with the size of the region in question [45, 80]. In fact, for the condensate phase of our model, the rms fluctuations of mass in any region, from the full system, down to a single site, show non-Gaussian behaviour, and depend directly on L. Thus, the system size L enters in a central way into the properties of the system, in contrast to the usual finite size corrections which become increasingly irrelevant for large systems.

3. In our work, we have studied several limits of the model analytically. In particular, we were able to calculate exactly various static and dynamical quantities for the aggregation-only limit of the single species model, in the continuum limit. This calculation is based on solving master equations for the mass contained in different spatial intervals of the system. While this kind of interval-based approach has been widely used in the study of reaction-diffusion systems [25, 29, 65], to our knowledge, this work constitutes the first application of the method to the calculation of dynamical quantities such as structure functions, which help demonstrate the occurrence of intermittency. Further, using this method, we were able to calculate structure functions of several integer orders and explicitly demonstrate their anomalous scaling behaviour. The method used here may have wider applicability and could be useful in the calculation of dynamical quantities for other 1D systems as well.

4. In the presence of open boundaries, non-equilibrium systems can develop interesting position-dependent behaviour. In the models that we study, there is in fact evidence for co-existence of different spatial regions in *qualitatively different* 'phases' within the same system. One example of this kind of 'phase co-existence' arises in the condensate phase of the single species model, where the condensate can only be found in a finite fraction of the system, while the remaining part of the system approaches a normal or condensate-free phase in the thermodynamic limit. Another example appears in the context of the multi-species chipping model, where indefinitely growing pile-ups occur in specific regions of the system, even as the rest of the system attains steady state.

5. Interconversion between different types of particles leads to non-conservation of particles of each type in the bulk of the system (even though the total number of particles is conserved). Non-conservation in the bulk, in conjunction with movement and boundary fluxes has earlier been associated with non-trivial spatial effects and phase co-existence in the exclusion process [55]. The multi-species chipping model we study also shows a similar effect, namely the co-existence of two phases of one species (a steady phase and a growing phase) in different spatial regions. This sort of space-dependent behaviour is a consequence of interconversion and is not found in the single species version of the chipping model [43].

6. While models of this sort cannot be connected directly to experiments, they shed light on the effects that can arise due to the interplay of various processes in the biological system, and also help in identifying features that can emerge simply from self-organisation. In the context of the Golgi, our model makes some qualitative predictions. It also suggests some new directions for experiments, most notably, the possibility of extracting information about the transport processes from the statistics of fluctuations in fluorescence experiments.

References

[1] P. C. Nelson, *Biological physics: Energy, Information, Life* (W.H. Freeman, 2003).

[2] M. C. Marchetti, J. F. Joanny, S. Ramaswamy, T. B. Liverpool, J. Prost, M. Rao and R. A. Simha, Rev. Mod. Phys. **85**, 1143 (2013).

[3] T. Vicsek and A. Zafeiris, Phys. Rep. **517**, 71 (2012).

[4] P. Bak and K. Sneppen, Phys. Rev. Lett. **71**, 4083 (1993).

[5] C. T. MacDonald, J. H. Gibbs and A. C. Pipkin, Biopolymers **6**, 1 (1968); C. T. MacDonald and J. H. Gibbs, Biopolymers **7**, 707 (1969).

[6] F. Spitzer, Adv. Math. **5**, 246 (1970).

[7] G. M. Schütz, *Phase Transitions and Critical Phenomena*, edited by C. Domb and J. Lebowitz (Academic Press, San Diego, 2001), Vol. 19, pp. 1–251.

[8] *Scaling and Fluctuations in Biology*, edited by T. Vicsek (Oxford University Press, Oxford, 2001).

[9] R. Milo, S. Shen-Orr, S. Itzkovitz, N. Kashtan, D. Chklovskii and U. Alon, Science **298**, 824 (2002).

[10] N. Barkai and S. Leibler, Nature, **387**(6636), 913 (1997).

[11] G. B. West, J. H. Brown, and B. J. Enquist, *Scaling in Biology*, edited by J. H. Brown and G. B. West (Oxford University Press, New York, 2000).

[12] S. Ramaswamy, Annu. Rev. Condens. Matter Phys. **1**, 323 (2010).

[13] B. Alberts et al, *Molecular Biology of the Cell*, 4/e (New York: Garland, 2003).

[14] B.S. Glick and A. Nakano, Annu Rev Cell Dev Biol **25**:113–132 (2009).

[15] E. Losev, C.A. Reinke, J. Jellen, D.E. Strongin, B.J. Bevis and B.S. Glick, Nature **441**, 1002 (2006); K. Matsuura, M. Takeuchi, A. Ichihara, K. Mikuriya and A. Nakano, Nature **441**, 1007 (2006).

[16] G.H. Patterson, K. Hirschberg, R.S. Polishchuk, D. Gerlich, R.D. Phair and J. Lippincott-Schwartz, Cell **133**, 1055 (2008).

[17] H. R. Pelham, J. E. Rothman, Cell **102**, 713 (2000).

[18] J. Krug, Phys. Rev. Lett. **67**, 1882 (1991); B. Derrida, E. Domany and D. Mukamel, J. Stat. Phys. **69**, 667-687 (1992).

[19] N. Brilliantov, P. Krapivsky, H. Hayakawa, A. Bodrova, F. Spahn and J. Schmidt, arxiv:1302.4097v3.

[20] S. K. Friedlander, *Smoke, Dust and Haze* (Wiley Interscience, New York, 1977).

[21] R. M. Ziff, J. Stat. Phys. **23**, 241 (1980).

[22] T.A. Witten, Jr. and L.M. Sander, Phys. Rev. Lett. **47**, 1400 (1981).

[23] B. Derrida, V. Hakim and V. Pasquier, Phys. Rev. Lett. **75**, 751 (1995); B. Derrida, J. Phys. A **28**:1481 (1995).

[24] H. Hinrichsen, V. Rittenberg and H. Simon, J. Stat. Phys. **86**, 1203 (1997).

[25] H. Takayasu, Phys. Rev. Lett. **63**, 2563 (1989); H. Takayasu, I. Nishikawa and H. Tasaki, Phys. Rev. A **37**, 3110 (1988).

[26] Z. Cheng, S. Redner and F. Leyvraz, Phys. Rev. Lett. **62**, 2321 (1989).

[27] J. L. Spouge, Phys. Rev. Lett. **60**, 871 (1988); J. Phys. A **21**, 4183 (1988).

[28] F. Leyvraz, Phys. Rep. **383**, 95 (2003).

[29] S.N. Majumdar and C. Sire, Phys. Rev. Lett. **71**, 3729 (1993).

[30] R. Rajesh and S.N. Majumdar, Phys. Rev. E **62**, 3186 (2000).

[31] C. Connaughton, R. Rajesh, and O. Zaboronski, Phys. Rev. Lett. **94**, 194503 (2005); Physica D **222**, 97 (2006).

[32] K. Jain and M. Barma, Phys. Rev. E **64**, 016107 (2001).

[33] S. Reuveni, I. Eliazar, U. Yechiali, Phys. Rev. E **84**, 041101 (2011); Phys. Rev. Lett. **109**, 020603 (2012); S. Reuveni, O. Hirschberg, I. Eliazar, U. Yechiali, arxiv:1309.2894v1.

[34] R. Rajesh and S. N. Majumdar, J. Stat. Phys. **99**, 943 (2000).

[35] M. R. Evans and T. Hanney, J. Phys. A **38**, R195 (2005).

[36] Y. Kafri, E. Levine, D. Mukamel, G. M. Schütz and J. Török, Phys. Rev. Lett. **89**, 035702-1 (2002).

[37] J. Krug and P. A. Ferrari, J. Phys. A **29**, L465 (1996).

[38] M. R. Evans, Europhys. Lett. **36**, 13 (1996).

[39] K. Jain and M. Barma, Phys. Rev. Lett. **91**, 135701 (2003)

[40] T. Hanney and M. R. Evans, Phys. Rev. E **69**, 016107 (2004).

[41] M.R. Evans and T. Hanney, J. Phys. A **36**, L441 (2003).

[42] M. R. Evans, T. Hanney and S. N. Majumdar, Phys. Rev. Lett. **97**, 010602 (2006).

[43] E. Levine, D. Mukamel and G. M. Schütz, J. Stat. Phys. **120**, 759 (2005).

[44] S. N. Majumdar, S. Krishnamurthy, and M. Barma, Phys. Rev. Lett. **81**, 3691 (1998); J. Stat. Phys. **99**, 1 (2000).

[45] R. Rajesh and S.N. Majumdar, Phys. Rev. E **63**, 036114 (2001).

[46] R. Rajesh, D. Das, B. Chakraborty and M. Barma, Phys. Rev. E **66**, 056104 (2002)

[47] R. Rajesh and S. Krishnamurthy, Phys. Rev. E **66**, 046132 (2002).

[48] R. Lahiri, M. Barma and S. Ramaswamy, Phys. Rev. E **61**, 1648 (2000).

[49] M.R. Evans, Y. Kafri, M. Koduvely and D. Mukamel, Phys. Rev. Lett. **80**, 425 (1998).

[50] P.F. Arndt, T. Heinzeland V. Rittenberg, J. Phys. A: Math. Gen. **31** L45 (1998); J. Stat. Phys. **97** 1 (1999).

[51] E. Levine and R. D. Willmann, J. Phys. A **37**, 3333 (2004).

[52] T. Chou, J. Phys. A **35**, 4515 (2002).

[53] S. Muhuri and I. Pagonabarraga, Europhys. Lett. **84**, 58009 (2008).

[54] E. Pronina and A. B. Kolomeisky, J. Phys. A **37**, 9907 (2004).

[55] T. Reichenbach, T. Franosch, and E. Frey, Phys. Rev. Lett. **97**, 050603 (2006); T. Reichenbach, E. Frey, and T. Franosch, New J. Phys. **9** 159 (2007).

[56] R. Juhász, Phys. Rev. E **76**, 021117 (2007).

[57] U. Frisch, *Turbulence: The Legacy of A.N. Kolmogorov*, (Cambridge Univ. Press, Cambridge, 1995).

[58] H. Sachdeva, M. Barma and M. Rao, Phys. Rev. Lett. **110**, 150601 (2013).

[59] S.N. Majumdar, Les Houches (2008) lecture notes, arXiv:0904:4097.

[60] V. Narayan, S. Ramaswamy and N. Menon, Science **317**, 105 (2007).

[61] D. Das and M. Barma, Phys. Rev. Lett. **85**, 1602 (2000); D. Das, M. Barma and S. N. Majumdar, Phys. Rev. E **64**, 046126 (2001).

[62] S. Dey, D. Das, and R. Rajesh, Phys. Rev. Lett **108**, 238001 (2012).

[63] H. Sachdeva and M. Barma, J. Stat. Phys. **154**, 950 (2014).

[64] J. Bec, and K. Khanin, Phys. Rep. **447**, 1 (2007).

[65] C. R. Doering, and D. ben-Avraham, Phys. Rev. Lett. **62**, 2563 (1989).

[66] M. Droz, and L. Sasvári, Phys. Rev. E **48** R2343 (1993); L. Peliti, J. Phys. A **19**, L365 (1986); O. Zaboronski, Phys. Lett. A **281**, 119 (2001)

[67] M. Howard, and C. Godrèche, J. Phys. A **31**, L209 (1998).

[68] C. R. Doering, and D. ben-Avraham, Phys. Rev. A **38**, 3035 (1988).

[69] K. Kang, and S. Redner, Phys. Rev. A **30**, 2833 (1984).

[70] S. N. Majumdar, S. Krishnamurthy, and M. Barma, Phys. Rev. E **61**, 6337 (2000).

[71] Z. Rácz, Z, Phys. Rev. Lett. **55**, 1707 (1985).

[72] T. Vicsek, P. Meakin, and F. Family, Phys. Rev. A **32**, 1122 (1985).

[73] R. C. Ball, C. Connaughton, P. P. Jones, R. Rajesh, and O. Zaboronski, Phys. Rev. Lett. **109**, 168304 (2012)

[74] A. Damle, and G. C. Peterson, SIAM Undergraduate Research Online, Volume 3, Issue 1, http://www.siam.org/students/siuro/vol3/S01061.pdf, 2010, pp. 187-208.

[75] M. Práger, Appl. Math. **43**(4), 311 (1998).

[76] R. Haberman, *Applied partial differential equations: with Fourier series and boundary value problems*, (Pearson Prentice Hall, 2004).

[77] M. R. Evans, and S. N. Majumdar, J. Stat. Mech: Theory Exp. **P05004** (2008).

[78] S. Ramaswamy, R. A. Simha, and J. Toner, Europhys. Lett. **62**, 196 (2003).

[79] H. Chaté, F. Ginelli, G. Grégoire, and F. Raynaud, Phys. Rev. E **77**, 046113 (2008).

[80] A. Nagar, M. Barma and S. N. Majumdar, Phys. Rev. Lett. **94**, 240601 (2005); A. Nagar, S. N. Majumdar and M. Barma, Phys. Rev. E **74**, 021124 (2006).

[81] Y. Pinchover, and J. Rubinstein, *An Introduction to Partial Differential Equations*, (Cambridge Univ. Press, Cambridge, 2005).

[82] H. Sachdeva, M. Barma and M. Rao, Phys. Rev. E **84**, 031106 (2011).

[83] D. L. Nelson, M. M. Cox, *Lehninger—Principles of Biochemistry*. 4/e (New York: W. H. Freeman and Company, 2005).

[84] A. Trucco et al, Nat Cell Biol **6**(11):1071–1081 (2004).

[85] S. Dmitrieff, M. Rao and P. Sens, Proc. Natl. Acad. Sci USA **110**, 15692-15697 (2013).

[86] B. S. Glick, Nat. Rev. Mol. Cell Biol. **3**, 615 (2002).