The role of mode switching in a population of actin polymers with constraints

François Robin, Anne van Gorp, Amandine Véber

The role of mode switching in a population of actin polymers with constraints

F. Robin*, A. Van Gorp†, and A. Véber‡

September 30, 2019

Abstract: In this paper, we introduce a stochastic model for the dynamics of *actin* polymers and their interactions with other proteins in the cellular envelop. Each polymer elongates and shortens, and can switch between several *modes* depending on whether it is bound to accessory proteins that modulate its behaviour as, for example, elongation-promoting factors. Our main aim is to understand the dynamics of a large population of polymers, assuming that the only limiting quantity is the total amount of monomers, set to be constant to some large N.

We first focus on the evolution of a very long polymer, of size $\mathcal{O}(N)$, with a rapid switch between modes (compared to the timescale over which the macroscopic fluctuations in the polymer size appear). Letting N tend to infinity, we obtain a *fluid limit* in which the effect of the switching appears only through the fraction of time spent in each mode at equilibrium. We show in particular that, in our situation where the number of monomers is limiting, a rapid binding-unbinding dynamics may lead to an increased elongation rate compared to the case where the polymer is trapped in any of the modes.

Next, we consider a large population of polymers and complexes, represented by a random measure on some appropriate type space. We show that as N tends to infinity, the stochastic system converges to a deterministic limit in which the switching appears as a flow between two categories of polymers. We exhibit some numerical examples in which the limiting behaviour of a single polymer differs from that of a population of competing (shorter) polymers for equivalent model parameters.

Taken together, our results demonstrate that under conditions where the total number of monomers is limiting, the study of a single polymer is not sufficient to understand the behaviour of an ensemble of competing polymers.

Key words and phrases: modelling of actin polymers, measure-valued Markov processes, convergence theorems, homogenisation theorem.

AMS 2010 subject classifications: 60G57, 60J25, 60J75, 60F05.

1 Introduction

1.1 Actin cytoskeleton and cell mechanics

During the embryonic development, cell mechanics plays a central role in many cell behaviours, from cell shape changes to cell division, cell migration and cell rearrangements [15]. In animals, these mechanical properties are largely determined by the *cortical actomyosin cytoskeleton*, a

*Sorbonne Université, Centre National de la Recherche Scientifique (CNRS) UMR7622, Institut national pour la Santé et la Recherche Médicale U1156, Institut de Biologie Paris Seine-Laboratoire de Biologie du Développement (IBPS-LBD), 75005 Paris, France

†*Corresponding author* – CMAP, CNRS, École Polytechnique, I.P. Paris, 91128 Palaiseau, France

‡CMAP, CNRS, École Polytechnique, I.P. Paris, 91128 Palaiseau, France

200-500nm thick network of actin and myosin polymers assembled beneath the cell membrane. Several parameters play a key role in determining the properties of this 2d active gel, including its turnover rate, density, crosslinking and its fine architecture. In particular, the distribution of polymer lengths in the cytoskeleton has a strong impact on the mechanical properties of the gel and the cell. In this paper, we focus on the dynamics of actin polymer assembly in the presence of two classes of proteins that modulate the assembly dynamics, formins and profilins. Our main assumption is that the total number of monomers in the system is fixed to some (large) N, reflecting the fact that this quantity is a limiting factor in the evolution of the population of polymers.

In cells, actin exists in two forms: monomers (or G-actin) and polymers (or F-actin). The G-actin monomer addition is coupled to an ATP hydrolysis. Along with the structural polar organisation of the F-actin filament, this causes both extremities of a polymer to display distinct kinetic properties: one with fast kinetics, or *plus* end, and one with slow kinetics, or *minus* end. *In vitro* actin assembly can reach a steady state where polymerisation, occurring at the plus end, is balanced by depolymerisation, occurring at the minus end. Note that in this work we are not interested in the spatial conformation of the polymers or their spatial organisation in the cellular cortex: provided there are available free monomers, a polymer can grow irrespective of its size.

A large number of accessory proteins interact with actin polymers and modify their dynamics, thus contributing to the shaping of the actin network in a cell- and tissue-specific manner. Here we consider only two such interactors: formins and profilins. Formins form a large family of dimeric proteins with a wide range of biochemical properties associated with their function in the cell. These proteins bind to the plus end of actin filaments, drive a *processive* filament elongation and modulate the actin elongation rate [13, 14]. Profilin is an abundant protein which regulates the actin dynamics by promoting ADP-to-ATP exchanges on G-Actin, renewing the ATP-actin pool. Importantly, profilin has been shown to affect differently distinct actin sub-networks, increasing the elongation rate by formins up to 15-fold, while decreasing the branching of branched networks [7, 22, 23]. In order to disentangle the different mechanisms underlying the composition in distinct polymer structures of the whole population of actin polymers, in this work we do not consider the sub-network of branched filaments (*i.e.*, tree-like structures induced by another protein called Arp2/3) and instead focus of the positive effect of formins and profilins on the elongation rate of the sub-network of simple polymers. Taking branching into account is significatively more difficult and will be the object of future work. Each formin can thus bind to or unbind from a polymer at a given rate. We shall say that a polymer is in "fast" mode if it is associated to a formin and therefore polymerises faster. When it is not bound with a formin, we say that it is in "normal" mode. To account for the role of profilin balancing the two network populations, we formally assembled filaments in "normal" mode from a G-actin complexes, while filaments in "fast" mode were assembled from G-actin/profilin (Figure 1 and [24]). Finally, note that our simple model, we also ignore the effect of the formins on actin nucleation.

Our main interest is twofold: understand the consequences of the fact that monomers are in limited numbers (being free, in a complex with a profilin or inserted in a polymer), and characterise the impact of the switching between two modes on the distribution of polymer lengths and on the fraction of monomers incorporated in a polymer. Several models for actin assembly have already been developped [6, 9, 10, 25, 26]. The majority of these models are either deterministic and focus on a steady-state equilibrium, or are stochastic and are studied by Monte-Carlo simulations [5, 19]. Most of the models taking into account a population of actin polymers rely on very specific assumptions and are not developed in a general context. Indeed, many of them concentrate on cell motility and therefore focus on the dynamics of actin in protrusions

[6, 16] or in the geometrical organisation of the actin network [8]. More general models like [9, 10] do not take into account the faster elongation of polymers bound with a formin, and neither do they consider the effect of the switching between spontaneous elongation and elongation with a formin. The works [12] and [17] use the same approach as ours to study the average length of an actin polymer and the distribution of the polymer lengths. However, [12] does not consider the large population limit of their model and [17] focuses more on the importance of having two types of monomers (ATP and ADP) than on the effects of accessory proteins. Apart from the population approaches listed above, most of the existing literature tends to focus on a single polymer, or on two competing polymers, with the idea that very large populations of polymers will then be well-described by the law of an "average" polymer. Importantly, we wanted to explore the effect of the competition between different sub-populations of polymers, elongated by distinct factors, for a unique, limited monomer pool.

To this end, we first consider a single polymer and its interactions with formins and profilins in Section 1.2. Taking a particular limit as the total number of monomers in the system tends to infinity, we show that if we allow this polymer to "live" for a very long time (by taking its initial length to be of order $\mathcal{O}(N)$), then its interactions with formins average out and the resulting asymptotic polymerisation rate is a weighted average between the polymerisation rate in fast and in normal mode (*i.e.*, when bound or not with a formin). In Section 1.3, we instead consider a large population of short competing polymers and show that, because short polymers never reach such a long lifetime, the intuitive homogenisation result of Section 1.2 does not apply and the competition for free monomers or G-actin/profilin complexes has a strong impact on the size distribution within the population of polymers. In Section 4, we show by simulation that the characteristics of a single polymer considered on its own can differ significantly from the distribution in polymer lengths in a large population even for equivalent model parameters.

1.2 Switching dynamics of a single actin polymer

1.2.1 The model

We first focus on the effect of the switching between different modes on the evolution of a single polymer. We restrict our attention to two modes (fast and normal), but our result would easily carry over to the case of finitely many modes under some natural ergodicity conditions.

The model we propose below takes into account the spontaneous elongation of the polymer as well as the effects of formins and profilins. If there are no formins attached to the plus end of the polymer, elongation is spontaneous and consumes free monomers. In contrast, if there is a formin on the polymer, elongation is induced by the formin and G-actin/profilin complexes are integrated into the polymer instead of free monomers. Only one formin can be bound to the polymer at a time. Recall that in the first case we say that the polymer is in "normal" mode, while the second case is called the "fast" mode. For every $t \geq 0$, we write $M(t) = 0$ (*resp.*, $M(t) = 1$) if the polymer is in normal (*resp.*, fast) mode at time t.

As we want monomers to be a limiting factor, we assume that the total number of monomers in the system is fixed through time to some $N \in \mathbb{N}$. These monomers can be in three states: free, in the polymer or held in a G-actin/profilin complex (each involving only one monomer). At any given time $t \geq 0$, the current number of free monomers is denoted by $L_0^N(t)$, the length of the polymer by $L_1^N(t)$ and the number of G-actin/profilin complexes by $L_2^N(t)$. We also define

$$L^N(t) \stackrel{def}{=} \big(L_0^N(t), L_1^N(t), L_2^N(t)\big). \tag{1}$$

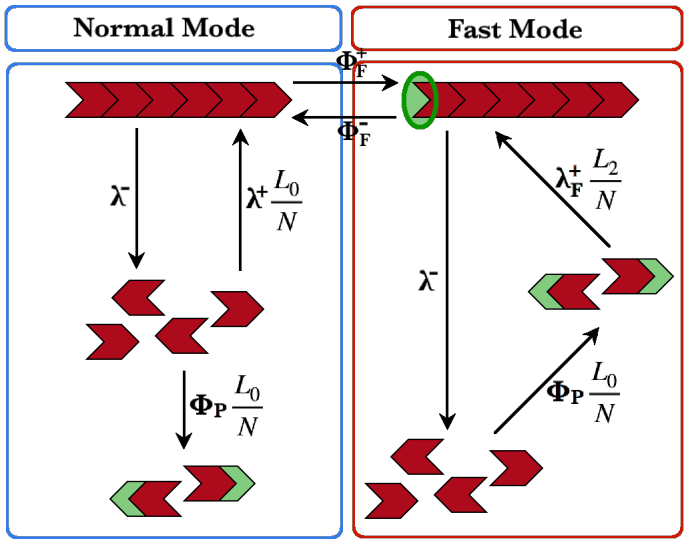


Figure 1: Schematic representation of the model for the dynamics of a single actin polymer and its interaction with formins and profilins. Red single units represent monomers, red-green units represent G-actin/profilin complexes and the strand of red units at the top of each box represents the polymer. The green oval depicts a formin bound to the polymer.

Writing $[N]$ for the set $\{0, 1, 2, \ldots, N\}$, we thus have for any $i \in \{0, 1, 2\}$ and any $t \in \mathbb{R}_+$: $L_i^N(t) \in [N]$ and

$$L_0^N(t) + L_1^N(t) + L_2^N(t) = N. \tag{2}$$

The model incorporates six types of events, illustrated in Figure 1:

1. *Spontaneous elongation.* If the polymer has positive length and is in normal mode, elongation (at the plus end) is spontaneous, and we assume that the rate at which it occurs is proportional to the current density of free monomers. That is, the instantaneous rate of spontaneous elongation a time t is given by $\mathbf{1}_{\{M(t-)=0\}}\mathbf{1}_{\{L_1^N(t-)>0\}}\lambda^+ L_0^N(t-)/N$, for some constant $\lambda^+ \geq 0$. The new value of the vector describing the system is then

$$L^N(t) = \big(L_0^N(t-) - 1, L_1^N(t-) + 1, L_2^N(t-)\big).$$

2. *Elongation with a formin.* If the system is in fast mode and the polymer has positive length, elongation (also at the plus end) is induced by a formin and consumes G-actin/profilin complexes. Therefore, in this case we suppose that the instantaneous rate of polymerisation is proportional to the density of available complexes, *i.e.* is equal to $\mathbf{1}_{\{M(t-)=1\}}\mathbf{1}_{\{L_1^N(t-)>0\}}\lambda_F^+ L_2^N(t-)/N$ for some constant $\lambda_F^+ \geq 0$. The vector L^N then jumps to

$$L^N(t) = \big(L_0^N(t-), L_1^N(t-) + 1, L_2^N(t-) - 1\big).$$

3. *Depolymerisation.* In both modes, the polymer can release a monomer (from the minus end) if its length has not already reached 0. We write $\mathbf{1}_{\{L_1^N(t-)>0\}}\lambda^-$ for rate of depolymerisation, where λ^- is a nonnegative constant, and the new value of L^N is then

$$L^N(t) = \big(L_0^N(t-) + 1, L_1^N(t-) - 1, L_2^N(t-)\big).$$

N	Total number of monomers
λ^+	Rate of spontaneous elongation
λ_F^+	Rate of elongation with formins
λ^-	Rate of depolymerisation
Φ_P	Rate of production of G-actin/profilin complexes
Φ_F^+	Rate of binding of a formin to the polymer (switch from normal to fast mode)
Φ_F^-	Rate of release of a formin by the polymer (switch from fast to normal mode)
$L_0^N(t)$	Number of free monomers at time t
$L_1^N(t)$	Length of the polymer at time t (in number of monomers)
$L_2^N(t)$	Number of G-actin/profilin complexes available at time t
$M(t)$	Mode (fast/normal) in which the polymer is at time t

Table 1: Parameters and variables in the single-polymer model

4. *Production of complexes.* Independently of the mode of the polymer, a free monomer and a profilin can react and produce a G-actin/profilin complex. Assuming that profilin is not a limiting factor, the instantaneous rate of production of G-actin/profilin complexes at time t is given by $\Phi_P L_0^N(t-)/N$, for some constant Φ_P that takes into account the concentration in profilin in the cellular cortex. The new value of L^N after such an event is

$$L^N(t) = \big(L_0^N(t-) - 1, L_1^N(t-), L_2^N(t-) + 1\big).$$

5. *Binding of a formin.* If the polymer is in normal mode, a formin can bind to its plus end and the polymer switches to the fast mode. We write Φ_F^+ for the constant rate of occurrence such a switch from normal to fast mode. This event does not modify the value of L^N, but corresponds to a jump from 0 to 1 for $M(t)$.

6. *Release of a formin.* If the polymer is in fast mode, the formin bound to it can be released and the polymer switches to the normal mode. We denote Φ_F^- the constant rate of such a switch. Again, this event only affects the process M, which jumps from 1 to 0.

Table 1 summarises all the notation introduced here. Observe that monomers cannot spontaneously detach from a G-actin/profilin complex. However, when a complex is integrated into a polymer, its depolymerisation results in the separation of the complex into a free monomer and a (free) profilin. Once the length of the polymer has reached 0, for mathematical convenience we suppose that $(M(t))_{t\geq 0}$ keeps on jumping from 0 to 1 at rate Φ_F^+ and from 1 to 0 at rate Φ_F^- for all times, even if this assumption has no biological meaning.

The system is fully described by the Markov process $(L^N(t), M(t))_{t\geq 0}$. From the above, we see that $(M(t))_{t\geq 0}$ is a Markov jump process with values in $\{0, 1\}$, switching from 0 to 1 at rate Φ_F^+ and from 1 to 0 at rate Φ_F^-. It evolves independently of the process $(L^N(t))_{t\geq 0}$, while the dynamics of $(L^N(t))_{t\geq 0}$ depends on $(M(t))_{t\geq 0}$. Since $(M(t))_{t\geq 0}$ is an irreducible Markov process on a finite state space, it has a unique stationary distribution π_M given by

$$\big(\pi_M(0)\,,\,\pi_M(1)\big) = \left(\frac{\Phi_F^-}{\Phi_F^+ + \Phi_F^-}\,,\,\frac{\Phi_F^+}{\Phi_F^+ + \Phi_F^-}\right). \tag{3}$$

This stationary distribution will appear in the homogenisation result presented in the next section.

1.2.2 Large-N limit

In order to understand the first order fluctuations of the system, we let the total number of monomers tend to infinity and look for a *fluid limit* as N tends to infinity. That is, we consider the scaled (non-Markovian) process

$$\overline{L^N}(t) = \left(\overline{L_0^N}(t), \overline{L_1^N}(t), \overline{L_2^N}(t)\right) \stackrel{def}{=} \left(\frac{L_0^N(Nt)}{N}, \frac{L_1^N(Nt)}{N}, \frac{L_2^N(Nt)}{N}\right), \tag{4}$$

and we want to prove that $\overline{L^N}$ converges to a deterministic limit as $N \to \infty$.

Let us write

$$\Delta \stackrel{def}{=} \left\{(x_1, x_2, x_3) \in [0,1]^3 \,:\, x_1 + x_2 + x_3 = 1\right\} \tag{5}$$

for the compact set in which $\overline{L^N}$ takes its values for every N and $\mathcal{D}_\Delta[0,\infty)$ for the space of all càdlàg Δ-valued paths, which we endow with the standard Skorokhod topology. Let

$$C_0 \stackrel{def}{=} \pi_M(0)\lambda^+ + \Phi_P \qquad \text{and} \qquad C_2 \stackrel{def}{=} \pi_M(1)\lambda_F^+, \tag{6}$$

where $\pi_M(0)$ and $\pi_M(1)$ are defined in (3). The constants C_0 and C_2 are respectively the rate of consumption of free monomers and of G-actin/profilin complexes when the process M is at stationarity. To avoid trivialities, we suppose that these two quantities are positive.

We can now formulate the main result of this section.

Theorem 1.1. *Suppose that the initial condition $\overline{L^N}(0) \in \Delta$ converges in distribution to some deterministic vector $(l_0(0), l_1(0), l_2(0)) \in \Delta$. Then as $N \to \infty$, the process $\overline{L^N}$ converges in distribution in $D_\Delta[0,\infty)$ to the deterministic process $(l(t))_{t\in\mathbb{R}_+} = (l_0(t), l_1(t), l_2(t))_{t\in\mathbb{R}_+}$ satisfying: for every $t \le t_0$,*

$$\begin{cases} l_0(t) &= \left(l_0(0) - \frac{\lambda^-}{C_0}\right)e^{-C_0 t} + \frac{\lambda^-}{C_0} \\ l_1(t) &= 1 - l_0(t) - l_2(t) \\ l_2(t) &= l_2(0)e^{-C_2 t} + \frac{\Phi_P}{C_2 - C_0}\left(l_0(0) - \frac{\lambda^-}{C_0}\right)(e^{-C_0 t} - e^{-C_2 t}) + \frac{\Phi_P \lambda^-}{C_0 C_2}(1 - e^{-C_2 t}), \end{cases}$$

and for every $t > t_0$,

$$\begin{cases} l_0(t) &= l_0(t_0)e^{-\Phi_P(t-t_0)} \\ l_1(t) &= 0 \\ l_2(t) &= 1 - l_0(t) \end{cases}$$

where $t_0 = \min\{t \in \mathbb{R}_+ \,:\, l_1(t) = 0\}$ is the time at which the polymer "size" cancels (by convention, $t_0 = +\infty$ if l_1 does not reach 0).

We easily see from Theorem 1.1 that the system obtained in the limit has a unique equilibrium, which depends on whether $t_0 < +\infty$ or not. If $t_0 < +\infty$ then the system converges to $(0,0,1)$ and every monomer ends up associated with a profilin. In contrast, if $t_0 = +\infty$, the limiting distribution of monomers into the three different categories is given by

$$\left(\frac{\lambda^-}{C_0}\,,\, 1 - \frac{\lambda^-}{C_0} - \frac{\lambda^-\Phi_P}{C_0 C_2}\,,\, \frac{\lambda^-\Phi_P}{C_0 C_2}\right). \tag{7}$$

Observe that the limiting amount of free monomers corresponds to the ratio of the asymptotic rate of increase of the free monomer pool, λ^-, by the asymptotic rate $C_0 = \pi_M(0)\lambda^+ + \Phi_P$ at which free monomers are consumed by the polymer elongation in normal mode and the creation of G-actin/profilin complexes. Likewise, since complexes can only be created by the reaction of a free monomer with a profilin, the limiting amount of complexes can be read as the ratio of their creation rate from the pool of free monomers at equilibrium, $(\lambda^-/C_0) \times \Phi_P$, by the rate of consumption of complexes through polymerisation in fast mode, C_2.

A natural question, on which we shall not dwell, is of course to ask when t_0 is finite. An obvious sufficient (but not necessary) condition for $t_0 < \infty$ is

$$1 - \frac{\lambda^-}{C_0} - \frac{\lambda^- \Phi_P}{C_0 C_2} < 0. \tag{8}$$

Theorem 1.1 pertains to the family of stochastic homogenisation results such as [18]. Indeed, although most reactions occur on a similar timescale, the time it takes to the polymer to see its length change by $\mathcal{O}(N)$ units (so that this change may be visible after our rescaling of L_1^N by N) is of the order of $\mathcal{O}(N)$ original time units, and is therefore much larger than the time of order $\mathcal{O}(1)$ that the binding and release of formins take. What our result shows is that, in the limit as N tends to infinity, the fast dynamics of mode switching equilibrates and only its stationary distribution π_M impacts the slow dynamics of the polymer elongation and shortening. Recalling that $\pi_M(0)$ (*resp.*, $\pi_M(1)$) can be interpreted as the fraction of time spent in the normal mode (*resp.*, fast mode) over the very long timescale considered here (by ergodicity), we see that, as mentioned earlier, C_0 (*resp.*, C_2) corresponds to a global rate of consumption of free monomers (*resp.*, of G-actin/profilin complexes) when M is at stationarity, and the average behaviour described by the equations stated in Theorem 1.1 corresponds to the barycenter between the behaviour of the system when it is in normal mode and the behaviour of the system when it is in fast mode.

Notice that our approach is similar to that of [18, 21], in which each individual switches state (or spatial location) independently, but it differs from many stochastic homogenisation results such as [2] or [3], in which it is the environment where all individuals evolve which switches fast, affecting the population as a whole. More precisely, in our study fluctuations are not on the scale of the population but on the scale of the individuals since each polymer can be in several modes independently of the others. This kind of random individual state switching seems less studied than the long-term averaging effect of environmental switches.

The limiting process $(l(t))_{t\geq 0}$ allows us to understand the impact of the restriction on monomers and of the interaction with formins and profilins on the dynamics of a single polymer. However, our model does not take into account the competition between the numerous polymers which are present in the cellular cortex. As a consequence, we now build on our model for a single polymer to construct and analyse a more complex model for a *population* of actin polymers.

1.3 Switching dynamics for a population of actin polymers

1.3.1 The model

We now consider a population of actin polymers undergoing the same reactions as in the previous section. Recall that we assume that the total number of monomers in the system is fixed, equal to $N \in \mathbb{N}$. Again, these monomers can be free, in a G-actin/profilin complex, in a polymer bound with a formin or in a simple polymer (with no formin attached). As earlier, if an actin polymer is simple its elongation consumes free actin monomers, and if the polymer is associated

to a formin, its elongation consumes only G-actin/profilin complexes. To ease the exposition, we shall consider that free monomers and G-actin/profilin complexes are two types of actin polymers.

Since we want to understand the distribution of polymer lengths within the population as well as the amounts of monomers contained in each category of polymers, we describe every element of the population by its type and its length. More precisely, free monomers are denoted by M, G-actin/profilin complexes by P, simple actin polymers by S and actin polymers bound with a formin by F. Polymers of type M and P can only be of size 1, while polymers of type S and F can be of any length in $\{1, 2, \ldots, N\}$. By convention, we distinguish between polymers of length one and free monomers as polymers made of a single monomer can grow again. Other conventions are possible and similar results to the ones presented below can be obtained using the same arguments as in the proof of Theorem 1.2.

Later it will be more convenient to have a state space independent of N, we thus let

$$\mathcal{I} \stackrel{def}{=} \{(M,1),(P,1)\} \cup \big(\{S,F\} \times \mathbb{N}\big), \tag{9}$$

and for every pair $(T,l) \in \mathcal{I}$ and every $t \geq 0$, we define $K_t^N(T,l)$ as the number of polymers of type T and length l present in the system at time t. Here we use the convention $\mathbb{N} = \{1,2,\ldots\}$. The system at every time t is described by its empirical distribution $\mathcal{Z}_t^N$:

$$\mathcal{Z}_t^N \stackrel{def}{=} K_t^N(M,1)\delta_{(M,1)} + K_t^N(P,1)\delta_{(P,1)} + \sum_{l=1}^{N} K_t^N(S,l)\delta_{(S,l)} + \sum_{l=1}^{N} K_t^N(F,l)\delta_{(F,l)}, \tag{10}$$

where δ_x denotes a Dirac mass at x. Each $\mathcal{Z}_t^N$ is thus a random measure on $\mathcal{I}$. We write $\mathcal{M}_{\mathcal{I}}$ for the space of all finite measures on $\mathcal{I}$, endowed with the topology of weak convergence.

Let us make a few remarks. Since the total number of monomers is N, for every $t \geq 0$ we have $K_t^N(\cdot, l) = 0$ if $l > N$ and

$$\langle \mathcal{Z}_t^N, l \rangle = K_t^N(M,1) + K_t^N(P,1) + \sum_{l=1}^{N} l K_t^N(S,l) + \sum_{l=1}^{N} l K_t^N(F,l) = N. \tag{11}$$

The total number of "real" polymers (*i.e.*, that are not free monomers or complexes) is denoted by $F^N(t)$:

$$F^N(t) \stackrel{def}{=} \sum_{l=1}^{N} \big(K_t^N(S,l) + K_t^N(F,l)\big) = \langle \mathcal{Z}_t^N, \mathbf{1}_{(S,\cdot)} + \mathbf{1}_{(F,\cdot)} \rangle. \tag{12}$$

Observe that the quantity $K^N(M,1)$ corresponds to what we called L_0^N in the model with one polymer and $K^N(P,1)$ corresponds to what we called L_2^N. The other quantities are not quite comparable to those introduced in the previous model, but an analogue of L_1 would be the total number of monomers inserted in a polymer:

$$M_{inF}^N(t) \stackrel{def}{=} \sum_{l=1}^{N} l\,\big(K_t^N(S,l) + K_t^N(F,l)\big) = \langle \mathcal{Z}_t^N, l\big(\mathbf{1}_{(S,\cdot)} + \mathbf{1}_{(F,\cdot)}\big) \rangle. \tag{13}$$

The process $(\mathcal{Z}_t^N)_{t\geq 0}$ evolves according to the same six types of transitions as in the single polymer model.

1. *Spontaneous elongation.* Each simple polymer can elongate by catching a free monomer, independently of the other polymers. As in the first model, $\lambda^+ K_{t-}^N(M,1)/N$ is the constant rate at which a given polymer captures a free monomer. Consequently, the total of rate at which a spontaneous elongation happens at time t is $\lambda^+(K_{t-}^N(M,1)/N)\sum_{l\geq 1} K_{t-}^N(S,l)$ and conditionally on this polymer being of size ℓ just before the event (which happens with probability $K_{t-}^N(S,\ell)/\sum_l K_{t-}^N(S,l)$), this corresponds to a transition to
$$\mathcal{Z}_t^N = \mathcal{Z}_{t-}^N - \delta_{(S,\ell)} - \delta_{(M,1)} + \delta_{(S,\ell+1)},$$
or equivalently to
$$K_t^N(S,\ell) = K_{t-}^N(S,\ell) - 1 \ ; \ K_t^N(S,\ell+1) = K_{t-}^N(S,\ell+1) + 1 \ ; \ K_t^N(M,1) = K_{t-}^N(M,1) - 1.$$

2. *Elongation with a formin.* Each polymer associated to a formin can elongate by absorbing a G-actin/profilin complex (again, independently of the other polymers). As in the previous type of transition, the total of rate at which an elongation with a formin happens at time t is $\lambda_F^+(K_{t-}^N(P,1)/N)\sum_l K_{t-}^N(F,l)$ and conditionally on this polymer being of size ℓ just before the event (which happens with probability $K_{t-}^N(F,\ell)/\sum_l K_{t-}^N(F,l)$), this corresponds to a transition to
$$\mathcal{Z}_t^N = \mathcal{Z}_{t-}^N - \delta_{(F,\ell)} - \delta_{(P,1)} + \delta_{(F,\ell+1)},$$
or equivalently to
$$K_t^N(F,\ell) = K_{t-}^N(F,\ell) - 1 \ ; \ K_t^N(F,\ell+1) = K_{t-}^N(F,\ell+1) + 1 \ ; \ K_t^N(P,1) = K_{t-}^N(P,1) - 1.$$

3. *Depolymerisation.* Each polymer of length $l \geq 1$ (bound or not with a formin) can release a monomer at rate λ^- and this monomer becomes "free". The (total) instantaneous rate at which depolymerisation occurs to a polymer of length ℓ at time t is thus $\lambda^-(K_{t-}^N(S,\ell) + K_{t-}^N(F,\ell))$, and if $T \in \{S,F\}$ is the type of this polymer, the new value of the empirical measure at time t is
$$\mathcal{Z}_t^N = \mathcal{Z}_{t-}^N - \delta_{(T,\ell)} + \delta_{(M,1)} + \delta_{(T,\ell-1)}\mathbf{1}_{\{\ell\geq 2\}}.$$

4. *Production of complexes.* Each free monomer can bind with a profilin and create a G-actin/profilin complex at rate Φ_P, independently of the others. Therefore, the total rate of creation of complexes at time t is $\Phi_P K_{t-}^N(M,1)$ and the new state of the system after such an event is
$$\mathcal{Z}_t^N = \mathcal{Z}_{t-}^N - \delta_{(M,1)} + \delta_{(P,1)}.$$

5. *Binding of a formin.* Each simple polymer of length l can bind with a formin and therefore change category (or "mode"), independently of the other polymers. If we write again Φ_F^+ for the rate at which a given simple polymer becomes bound with a formin, the total rate at which such an event occurs is $\Phi_F^+ K_{t-}^N(S,l)$ and conditionally on the length of the polymer concerned being ℓ, the new state of the system is
$$\mathcal{Z}_t^N = \mathcal{Z}_{t-}^N - \delta_{(S,\ell)} + \delta_{(F,\ell)}.$$

6. *Release of a formin.* Each polymer associated to a formin can have this association broken at rate Φ_F^-, independently of the other polymers. Hence, the total rate of release of a formin by a polymer of length ℓ at time t is $\Phi_F^- K_{t-}^N(F,\ell)$, and the new state of the system after such an event is
$$\mathcal{Z}_t^N = \mathcal{Z}_{t-}^N - \delta_{(F,\ell)} + \delta_{(S,\ell)}.$$

Figure 2 summarises all the events described above.

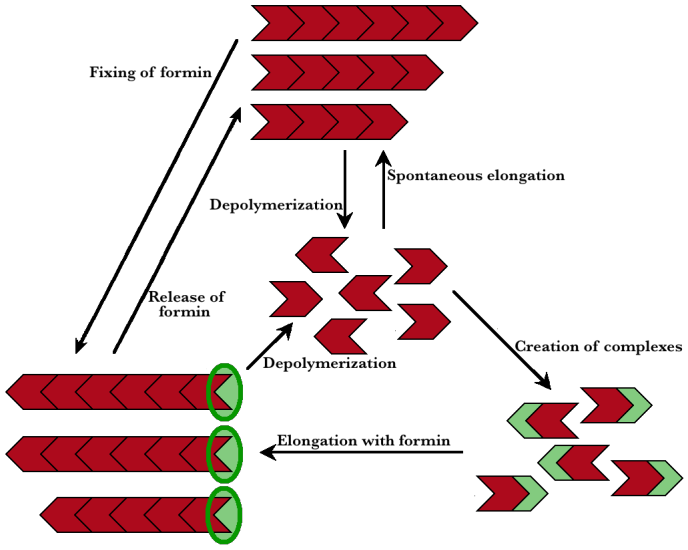


Figure 2: Transitions in the model for a population of polymers. Simple polymers can elongate by catching a free monomer, release a monomer that becomes free, or switch to the "fast" mode by binding with a formin. Polymers bound with a formin can elongate by catching a G-actin/profilin complex, release a monomer that becomes free, or switch to the "normal" mode by releasing the attached formin. Finally, free monomers and profilins can bind to form G-actin/profilin complexes.

1.3.2 Large-N limit

We now look for a large population limit, describing the first order behaviour of our population of polymers. More precisely, for every $N \in \mathbb{N}$ and $t \in \mathbb{R}_+$, we set

$$\overline{\mathcal{Z}_t^N} \stackrel{def}{=} \frac{1}{N} \mathcal{Z}_t^N. \tag{14}$$

By construction, each $\overline{\mathcal{Z}_t^N}$ takes its values in the set $\mathcal{M}_{\mathcal{I}}^*$ defined by

$$\mathcal{M}_{\mathcal{I}}^* = \big\{ Z \in \mathcal{M}_{\mathcal{I}} \,:\, \langle Z, l \rangle = 1 \big\}, \tag{15}$$

where we recall that $\mathcal{M}_{\mathcal{I}}$ is the set of all finite measures on $\mathcal{I}$ (see (9)), and $\langle Z, l \rangle$ is defined as in (11). In particular, since the length l of every monomer/polymer is at least one, we have

$$\mathcal{M}_{\mathcal{I}}^* \subset \mathcal{M}_{\mathcal{I}}^{(1)} \stackrel{def}{=} \big\{ Z \in \mathcal{M}_{\mathcal{I}} \,:\, \langle Z, 1 \rangle \leq 1 \big\}. \tag{16}$$

We now let N tend to infinity, assuming that the sequence of initial values $\overline{\mathcal{Z}_0^N}$ converges to some measure in $\mathcal{M}_{\mathcal{I}}^{(1)}$ and that the support of $\overline{\mathcal{Z}_0^N}$ is bounded uniformly in N. Note that we do not scale time by a factor N as we did to obtain the fluid limit for a single polymer in Section 1.2. Our main result is the following. We denote the space of all càdlàg trajectories with values in $\mathcal{M}_{\mathcal{I}}^{(1)}$ by $D_{\mathcal{M}_{\mathcal{I}}^{(1)}}[0, \infty)$ and we endow it with the standard Skorokhod topology.

Theorem 1.2. *Suppose that the initial condition $\overline{\mathcal{Z}_0^N}$ converges in distribution to a deterministic limit $\mathcal{Z}_0 \in \mathcal{M}_{\mathcal{I}}^{(1)}$ as $N \to +\infty$, and that there exists $K_0 \in \mathbb{N}$ such that for every N sufficiently*

large,

$$\langle \overline{\mathcal{Z}_0^N}, \mathbb{1}_{\{l \geq K_0\}} \rangle = 0. \tag{17}$$

Then the process $\left(\overline{\mathcal{Z}_t^N}\right)_{t\in\mathbb{R}_+}$ *converges in distribution in* $D_{\mathcal{M}_{\mathcal{I}}^{(1)}}[0,\infty)$ *as* $N \to \infty$*, to a deterministic limit* $(\mathcal{Z}_t)_{t\geq 0}$ *which can be characterised as follows. If for every* $(T,l) \in \mathcal{I}$ *we write* $k_t(T,l)$ *for the mass* $\mathcal{Z}_t(\{(T,l)\})$ *given by* $\mathcal{Z}_t$ *to the point* (T,l)*, then*

$$k_0(T,l) = \mathcal{Z}_0(\{(T,l)\}),$$

and

$$\begin{cases}
\frac{d}{dt}k_t(M,1) &= \lambda^- \sum_{l=1}^{+\infty}(k_t(S,l)+k_t(F,l)) - \lambda^+ k_t(M,1)\sum_{l=1}^{+\infty} k_t(S,l) - \Phi_P k_t(M,1) \\
\frac{d}{dt}k_t(P,1) &= \Phi_P k_t(M,1) - \lambda_F^+ k_t(P,1)\sum_{l=1}^{+\infty} k_t(F,l) \\
\frac{d}{dt}k_t(S,1) &= \Phi_F^- k_t(F,1) + \lambda^- k_t(S,2) - (\lambda^+ k_t(M,1) + \lambda^- + \Phi_F^+)k_t(S,1) \\
\frac{d}{dt}k_t(F,1) &= \Phi_F^+ k_t(S,1) + \lambda^- k_t(F,2) - (\lambda_F^+ k_t(P,1) + \lambda^- + \Phi_F^-)k_t(F,1) \\
\text{and } \forall l \geq 2, & \\
\frac{d}{dt}k_t(S,l) &= \lambda^+ k_t(M,1)k_t(S,l-1) + \Phi_F^- k_t(F,l) + \lambda^- k_t(S,l+1) \\
&\qquad -(\lambda^+ k_t(M,1) + \lambda^- + \Phi_F^+)k_t(S,l) \\
\frac{d}{dt}k_t(F,l) &= \lambda_F^+ k_t(P,1)k_t(F,l-1) + \Phi_F^+ k_t(S,l) + \lambda^- k_t(F,l+1) \\
&\qquad -(\lambda_F^+ k_t(P,1) + \lambda^- + \Phi_F^-)k_t(F,l).
\end{cases}$$

Remark 1.3. **(a)** *Ideally, we would like to replace the assumption that* $\mathcal{Z}_0 \in \mathcal{M}_{\mathcal{I}}^{(1)}$ *by* $\mathcal{Z}_0 \in \mathcal{M}_{\mathcal{I}}^*$ *and obtain the convergence of the sequence* $(\overline{\mathcal{Z}^N})_{N\geq 1}$ *in* $D_{\mathcal{M}_{\mathcal{I}}^*}[0,\infty)$*. However, the fact that* $\mathcal{Z}_t \in \mathcal{M}_{\mathcal{I}}^*$ *for all* $t \geq 0$ *is not guaranteed by the convergence in distribution of* $\overline{\mathcal{Z}_t^N}$ *seen as a finite measure on* $\mathcal{I}$*, since the function* $(T,l) \mapsto l$ *is not bounded continuous on* $\mathcal{I}$ *and so the property that* $\langle \mathcal{Z}_t, l\rangle = 1$ *needs not hold in the limit. Unfortunately, we were not able to prove that the solution to the limiting system in Theorem 1.2, when started in* $\mathcal{M}_{\mathcal{I}}^*$*, remains in* $\mathcal{M}_{\mathcal{I}}^*$ *for all times. Consequently, we had to relax our assumption and obtain convergence in* $D_{\mathcal{M}_{\mathcal{I}}^{(1)}}[0,\infty)$ *instead. In particular, although the total mass of the limiting system remains bounded (by* 1*), we have no guarantee that the "total length" of the system does not explode, although this seems very unrealistic.*

(b) *Notice that the assumption that the measures* $\overline{\mathcal{Z}_0^N}$ *should have uniformly bounded supports imposes that* $\overline{\mathcal{Z}_0^N}$ *should asymptotically give no weight to polymers whose sizes tend to infinity with* N*, so that no mass is lost on infinitely long polymers in the limit. If for instance the initial population was made of a few polymers of size* $\mathcal{O}(N)$ *(and free monomers, so that* (11) *would be satisfied), we would rather have to scale time and polymer lengths by* N *as in Section 1.2 to obtain a nontrivial limit. As this is not what we expect from the distribution of polymer lengths in the cellular cortex, we instead impose a bound on the polymer lengths in our model.*

The rest of the paper is organised as follows. In Section 2, we prove Theorem 1.1 on the asymptotic behaviour of a single (long) polymer. In Section 3, we prove Theorem 1.2 on the asymptotic behaviour of a population of actin polymers. Corollary 3.2 allows to compare these two results and highlights the effects of the competition between polymers for G-actin on the evolution of the population. Finally, some simulations are discussed in Section 4.

2 Homogenisation result for a single actin polymer

In this section, we prove Theorem 1.1. To this end, in Section 2.1 we prove that the sequence of processes $(\overline{L^N})_{N\geq 1}$ is C-tight, and in Section 2.2 we show that any limit point satisfies the deterministic system of equations stated in Theorem 1.1. Before proceeding to these proofs, we establish some useful preliminary properties.

Let us consider the evolution of L^N for some fixed N. Each of its coordinates takes its values in the finite set $[N]$, and so $(L^N(t), M(t))_{t\geq 0}$ is a Markov jump process with values in a finite state space. This allows us to write that for every $t \geq 0$,

$$L_0^N(t) = L_0^N(0) + \int_0^t \mathbb{1}_{\{L_1^N(s)>0\}}\lambda^- ds - \int_0^t \mathbb{1}_{\{M(s)=0\}}\mathbb{1}_{\{L_1^N(s)>0\}}\lambda^+ \frac{L_0^N(s)}{N} ds - \int_0^t \Phi_P \frac{L_0^N(s)}{N} ds + X_0^N(t), \tag{18a}$$

$$L_1^N(t) = L_1^N(0) + \int_0^t \mathbb{1}_{\{M(s)=0\}}\mathbb{1}_{\{L_1^N(s)>0\}}\lambda^+ \frac{L_0^N(s)}{N} ds - \int_0^t \mathbb{1}_{\{L_1^N(s)>0\}}\lambda^- ds + \int_0^t \mathbb{1}_{\{M(s)=1\}}\mathbb{1}_{\{L_1^N(s)>0\}}\lambda_F^+ \frac{L_2^N(s)}{N} ds + X_1^N(t), \tag{18b}$$

$$L_2^N(t) = L_2^N(0) + \int_0^t \Phi_P \frac{L_0^N(s)}{N} ds - \int_0^t \mathbb{1}_{\{M(s)=1\}}\mathbb{1}_{\{L_1^N(s)>0\}}\lambda_F^+ \frac{L_2^N(s)}{N} ds + X_2^N(t), \tag{18c}$$

where X_0^N, X_1^N and X_2^N are mean-zero martingales with respective predictable quadratic variations

$$\langle X_0^N\rangle(t) = \int_0^t \mathbb{1}_{\{L_1^N(s)>0\}}\lambda^- ds + \int_0^t \mathbb{1}_{\{M(s)=0\}}\mathbb{1}_{\{L_1^N(s)>0\}}\lambda^+ \frac{L_0^N(s)}{N} ds + \int_0^t \Phi_P \frac{L_0^N(s)}{N} ds, \tag{19a}$$

$$\langle X_1^N\rangle(t) = \int_0^t \mathbb{1}_{\{M(s)=0\}}\mathbb{1}_{\{L_1^N(s)>0\}}\lambda^+ \frac{L_0^N(s)}{N} ds + \int_0^t \mathbb{1}_{\{L_1^N(s)>0\}}\lambda^- ds + \int_0^t \mathbb{1}_{\{M(s)=1\}}\mathbb{1}_{\{L_1^N(s)>0\}}\lambda_F^+ \frac{L_2^N(s)}{N} ds, \tag{19b}$$

$$\langle X_2^N\rangle(t) = \int_0^t \Phi_P \frac{L_0^N(s)}{N} ds + \int_0^t \mathbb{1}_{\{M(s)=1\}}\mathbb{1}_{\{L_1^N(s)>0\}}\lambda_F^+ \frac{L_2^N(s)}{N} ds. \tag{19c}$$

Remark 2.1. *Notice that, for simplicity, we use the same process $(M(t))_{t\geq 0}$ to define the evolution of every $(L^N(t))_{t\geq 0}$. It has no impact on the result of Theorem 1.1 since the convergence result we obtain for $(\overline{L^N})_{N\in\mathbb{N}}$ is in distribution.*

If we now accelerate time by a factor N and scale down the coordinates of L^N by the same

quantity (see (4)), a simple change of variables gives us that

$$\overline{L_0^N}(t) = \overline{L_0^N}(0) + \int_0^t \mathbb{1}_{\{\overline{L_1^N}(u)>0\}}\lambda^- du - \int_0^t \mathbb{1}_{\{M(Nu)=0\}}\mathbb{1}_{\{\overline{L_1^N}(u)>0\}}\lambda^+\overline{L_0^N}(u)du - \int_0^t \Phi_P \overline{L_0^N}(u)du + \overline{X_0^N}(t), \tag{20a}$$

$$\overline{L_1^N}(t) = \overline{L_1^N}(0) + \int_0^t \mathbb{1}_{\{M(Nu)=0\}}\mathbb{1}_{\{\overline{L_1^N}(u)>0\}}\lambda^+\overline{L_0^N}(u)du - \int_0^t \mathbb{1}_{\{\overline{L_1^N}(u)>0\}}\lambda^- du + \int_0^t \mathbb{1}_{\{M(Nu)=1\}}\mathbb{1}_{\{\overline{L_1^N}(u)>0\}}\lambda_F^+\overline{L_2^N}(u)du + \overline{X_1^N}(t), \tag{20b}$$

$$\overline{L_2^N}(t) = \overline{L_2^N}(0) + \int_0^t \Phi_P\overline{L_0^N}(u)du - \int_0^t \mathbb{1}_{\{M(Nu)=1\}}\mathbb{1}_{\{\overline{L_1^N}(u)>0\}}\lambda_F^+\overline{L_2^N}(u)du + \overline{X_2^N}(t), \tag{20c}$$

where

$$\overline{X_i^N}(t) = \frac{X_i^N(Nt)}{N}, \qquad i \in \{0,1,2\},\ t \geq 0.$$

We now have all the ingredients to prove Theorem 1.1.

2.1 Tightness

The precise result that we prove in this section is the following.

Proposition 2.2 (Tightness). *The sequence of processes $(\overline{L^N})_{N\geq 1}$ is C-tight in $D_\Delta[0,\infty)$.*

Proof of Proposition 2.2. The proof is rather standard. Since Δ is compact, all we have to check is that for every $T > 0$ and every $\varepsilon > 0$, there exists $\delta > 0$ such that

$$\limsup_{N\to\infty} \mathbb{P}\big(\omega\big(\overline{L^N}, \delta, T\big) > \varepsilon\big) \leq \varepsilon, \tag{21}$$

where

$$\omega(\overline{L^N}, \delta, T) = \sup_{\substack{0\leq s<t\leq T\\ |t-s|\leq\delta}} \big\|\overline{L^N}(t) - \overline{L^N}(s)\big\|.$$

Theorem 15.1 in [4] will then enable us to conclude.

We proceed coordinate by coordinate, starting with $\overline{L_0^N}$. Let thus $T > 0$, $\varepsilon > 0$ and $\delta > 0$. Using (20a) and the fact that $\overline{L_0^N}(t) \leq 1$ for all $t \geq 0$, it is straightforward to show that for every $0 \leq s < t \leq T$ such that $|t - s| \leq \delta$, we have

$$\big|\overline{L_0^N}(t) - \overline{L_0^N}(s)\big| \leq (\lambda^- + \lambda^+ + \Phi_P)\delta + \big|\overline{X_0^N}(t)\big| + \big|\overline{X_0^N}(s)\big|. \tag{22}$$

Choosing $\delta < \delta_0 := \varepsilon/(9(\lambda^- + \lambda^+ + \Phi_P))$, we obtain that the first term on the r.h.s. of (22) is less than $\varepsilon/9$ with probability one. Next, by the Markov inequality we have

$$\mathbb{P}\bigg(\sup_{u\in[0,T]} \big|\overline{X_0^N}(u)\big| > \varepsilon/9\bigg) \leq \frac{81}{\varepsilon^2 N^2}\,\mathbb{E}\left[\sup_{u\in[0,T]} \big(X_0^N(Nu)\big)^2\right]. \tag{23}$$

Using Doob's maximal inequality and (19a), we can then write that

$$\mathbb{E}\left[\sup_{u\in[0,T]}(X_0^N(Nu))^2\right] \leq 2\mathbb{E}\big[(X_0^N(NT))^2\big] = 2\mathbb{E}\big[\langle X_0^N\rangle(NT)\big] \leq 2(\lambda^- + \lambda^+ + \Phi_P)NT.$$

Combining the above, we obtain that

$$\mathbb{P}\bigg(\sup_{u\in[0,T]}\big|\overline{X_0^N}(u)\big| > \varepsilon/9\bigg) \leq \frac{162(\lambda^- + \lambda^+ + \Phi_P)T}{\varepsilon^2 N}, \tag{24}$$

which converges to 0 as $N \to \infty$. Consequently, with our choice of $\delta < \delta_0$, we obtain that

$$\limsup_{N\to\infty}\mathbb{P}\left(\sup_{\substack{0\leq s<t\leq T\\ |t-s|\leq\delta}}\big|\overline{L_0^N}(t) - \overline{L_0^N}(s)\big| > \frac{\varepsilon}{3}\right) = 0.$$

The same reasoning applies to the two other coordinates $\overline{L_1^N}$ and $\overline{L_2^N}$ and yields similar thresholds $\delta_1, \delta_2 > 0$, which is more than enough to prove that if we take $\delta < \min\{\delta_0, \delta_1, \delta_2\}$,

$$\limsup_{N\to\infty}\mathbb{P}\big(\omega\big(\overline{L^N}, \delta, T\big) > \varepsilon\big) = 0.$$

This shows (21) and the proof of Proposition 2.2 is therefore complete. □

2.2 Uniqueness of the limit

Now that we have proved that the sequence of normalised processes is tight, it remains to show that the deterministic system described in Theorem 1.1 is its only possible limit as $N \to \infty$. This is what we do in this section. To do so, we shall use the following lemmas, whose proofs are postponed until the end of the section.

Lemma 2.3. *For every $t > 0$, we have*

$$\int_0^t (\mathbb{1}_{\{M(Ns)=0\}} - \pi_M(0))\overline{L_0^N}(s)ds \xrightarrow[N\to+\infty]{\mathbb{P}} 0.$$

Lemma 2.4. *For every $t > 0$, we have*

$$\int_0^t (\mathbb{1}_{\{M(Ns)=1\}} - \pi_M(1))\overline{L_2^N}(s)ds \xrightarrow[N\to+\infty]{\mathbb{P}} 0.$$

Suppose that the sub-sequence $(\overline{L^{N_k}})_{k\in\mathbb{N}}$ converges weakly to a process $(l_0(t), l_1(t), l_2(t))_{t\geq 0}$. By Proposition 2.2, we know that this process has continuous paths a.s. Furthermore, the estimates (24) (and their analogues for $\overline{X_1^N}$ and $\overline{X_2^N}$) guarantee that the limit of the quadratic variation process of each $\overline{L_i^N}$ is zero and $(l_0(t), l_1(t), l_2(t))_{t\geq 0}$ is therefore deterministic.

For every $N \in \mathbb{N}$, let t_0^N be the first time at which the polymer has length 0:

$$t_0^N = \min\big\{t \in \mathbb{R}_+ \,:\, \overline{L_1^N}(t) = 0\big\}.$$

Similarly, let us define t_0 as the first time at which $l_1(t)$ takes the value 0. In both cases, if the process considered does not reach 0, we set the corresponding $t_0^{(N)}$ to $+\infty$. Since 0 is an absorbing state for the process $\left(\overline{L_1^N}(t)\right)_{t\in\mathbb{R}_+}$, we have

$$\begin{cases} \overline{L_1^N}(t) > 0 \ \text{ if } \ t < t_0^N, \\ \overline{L_1^N}(t) = 0 \ \text{ if } \ t \geq t_0^N. \end{cases}$$

Therefore, the system of equations (20) can be rewritten as follows: for all $t < t_0^N$

$$\begin{aligned}
\overline{L_0^N}(t) &= \overline{L_0^N}(0) + \lambda^- t - \lambda^+ \int_0^t \mathbb{1}_{\{M(Nu)=0\}} \overline{L_0^N}(u)du - \Phi_P \int_0^t \overline{L_0^N}(u)du + \overline{X_0^N}(t), \\
\overline{L_1^N}(t) &= \overline{L_1^N}(0) - \lambda^- t + \lambda^+ \int_0^t \mathbb{1}_{\{M(Nu)=0\}} \overline{L_0^N}(u)du + \lambda_F^+ \int_0^t \mathbb{1}_{\{M(Nu)=1\}} \overline{L_2^N}(u)du + \overline{X_1^N}(t), \\
\overline{L_2^N}(t) &= \overline{L_2^N}(0) + \Phi_P \int_0^t \overline{L_0^N}(u)du - \lambda_F^+ \int_0^t \mathbb{1}_{\{M(Nu)=1\}} \overline{L_2^N}(u)du + \overline{X_2^N}(t),
\end{aligned}$$

and for all $t \geq t_0^N$

$$\begin{aligned}
\overline{L_0^N}(t) &= \overline{L_0^N}(0) + \lambda^- t_0^N - \lambda^+ \int_0^{t_0^N} \mathbb{1}_{\{M(Nu)=0\}} \overline{L_0^N}(u)du - \Phi_P \int_0^t \overline{L_0^N}(u)du + \overline{X_0^N}(t), \\
\overline{L_1^N}(t) &= 0, \\
\overline{L_2^N}(t) &= \overline{L_2^N}(0) + \Phi_P \int_0^t \overline{L_0^N}(u)du - \lambda_F^+ \int_0^{t_0^N} \mathbb{1}_{\{M(Nu)=1\}} \overline{L_2^N}(u)du + \overline{X_2^N}(t).
\end{aligned}$$

Splitting the integral $\int_0^{t\wedge t_0^{N_k}}$ into $\int_0^t - \int_{t\wedge t_0^{N_k}}^t$ and decomposing $\mathbb{1}_{\{M(N_k u)=0\}} = \pi_M(0) + (\mathbb{1}_{\{M(N_k u)=0\}} - \pi_M(0))$, we can rewrite the above system of equations for $t < t_0$ as

$$\begin{aligned}
\overline{L_0^{N_k}}(t) &= \overline{L_0^{N_k}}(0) + \lambda^- t - (\lambda^+ \pi_M(0) + \Phi_P) \int_0^t \overline{L_0^{N_k}}(u)du + \overline{X_0^{N_k}}(t) - \lambda^-(t - t\wedge t_0^{N_k}) \\
&\quad + \lambda^+ \pi_M(0) \int_{t\wedge t_0^{N_k}}^t \overline{L_0^{N_k}}(u)du - \lambda^+ \int_0^{t\wedge t_0^{N_k}} (\mathbb{1}_{\{M(N_k u)=0\}} - \pi_M(0)) \overline{L_0^{N_k}}(u)du, \\
\overline{L_1^{N_k}}(t) &= \overline{L_1^{N_k}}(0) - \lambda^- t + \lambda^+ \pi_M(0) \int_0^t \overline{L_0^{N_k}}(u)du + \lambda_F^+ \pi_M(1) \int_0^t \overline{L_2^{N_k}}(u)du + \overline{X_1^{N_k}}(t) \\
&\quad + \lambda^-(t - t\wedge t_0^{N_k}) + \lambda^+ \int_0^{t\wedge t_0^{N_k}} (\mathbb{1}_{\{M(N_k u)=0\}} - \pi_M(0)) \overline{L_0^{N_k}}(u)du \\
&\quad + \lambda_F^+ \int_0^{t\wedge t_0^{N_k}} (\mathbb{1}_{\{M(N_k u)=1\}} - \pi_M(1)) \overline{L_2^{N_k}}(u)du - \lambda^+ \pi_M(0) \int_{t\wedge t_0^{N_k}}^t \overline{L_0^{N_k}}(u)du \\
&\quad - \lambda_F^+ \pi_M(1) \int_{t\wedge t_0^{N_k}}^t \overline{L_2^{N_k}}(u)du, \\
\overline{L_2^{N_k}}(t) &= \overline{L_2^{N_k}}(0) + \Phi_P \int_0^t \overline{L_0^{N_k}}(u)du - \lambda_F^+ \pi_M(1) \int_0^t \overline{L_2^{N_k}}(u)du + \overline{X_2^{N_k}}(t) \\
&\quad + \lambda_F^+ \pi_M(1) \int_{t\wedge t_0^{N_k}}^t \overline{L_2^{N_k}}(u)du - \lambda_F^+ \int_0^{t\wedge t_0^{N_k}} (\mathbb{1}_{\{M(N_k u)=1\}} - \pi_M(1)) \overline{L_2^{N_k}}(u)du.
\end{aligned}$$

First, for every $t < t_0$ we have the convergence in probability

$$t - (t \wedge t_0^{N_k}) \xrightarrow[k\to+\infty]{\mathbb{P}} 0, \tag{25}$$

and

$$\int_{t\wedge t_0^{N_k}}^t \overline{L_0^{N_k}}(u)du \xrightarrow[k\to+\infty]{\mathbb{P}} 0 \qquad \text{and} \qquad \int_{t\wedge t_0^{N_k}}^t \overline{L_2^{N_k}}(u)du \xrightarrow[k\to+\infty]{\mathbb{P}} 0. \tag{26}$$

Indeed, since for any $i \in \{0,2\}$ and any $u \le t$ we have $\overline{L_i^{N_k}}(u) \le 1$, then

$$0 \le \mathbb{E}\left[\int_{t\wedge t_0^{N_k}}^t \overline{L_i^{N_k}}(u)du\right] \le t\mathbb{P}\big[t_0^{N_k} < t\big].$$

But

$$\limsup_{k\to\infty} \mathbb{P}\big[t_0^{N_k} < t\big] = \limsup_{k\to\infty} \mathbb{P}\left[\overline{L_1^{N_k}}(t) = 0\right] \le \inf_{\varepsilon>0} \mathbb{P}\big[l_1(t) \le \varepsilon\big] = 0,$$

where the last but first inequality comes from the fact that $(\overline{L_1^{N_k}}(t))_{k\ge1}$ converges in distribution to $l_1(t)$ and the last equality uses our assumption that $t < t_0$. This last result proves (25). The Markov inequality then gives us the two convergence results stated in (26).

Second, by assumption we have the convergence in distribution of $(\overline{L^{N_k}})_{k\ge1}$ to the deterministic process $(l_0(t), l_1(t), l_2(t))_{t\ge0}$ and therefore

$$\big(\overline{L_0^{N_k}}(0), \overline{L_1^{N_k}}(0), \overline{L_2^{N_k}}(0)\big) \xrightarrow[k\to+\infty]{\mathbb{P}} \big(l_0(0), l_1(0), l_2(0)\big),$$

and

$$\left(\int_0^t \overline{L_0^{N_k}}(u)du\,,\, \int_0^t \overline{L_2^{N_k}}(u)du\right) \xrightarrow[k\to+\infty]{\mathbb{P}} \left(\int_0^t l_0(u)du\,,\, \int_0^t l_2(u)du\right). \tag{27}$$

Thirdly, by (24) the sequence $(\overline{X_0^N}(t))_{N\ge1}$ converges in probability to 0 (in fact this convergence is uniform over $[0,t]$) and equivalent results hold for $(\overline{X_1^N}(t))_{N\ge1}$ and $(\overline{X_2^N}(t))_{N\ge1}$.

Finally, Lemmas 2.3 and 2.4 ensure the convergence in probability

$$\int_0^t (\mathbb{1}_{\{M(Ns)=0\}} - \pi_M(0))\overline{L_0^N}(s)ds \xrightarrow[N\to+\infty]{\mathbb{P}} 0$$

and

$$\int_0^t (\mathbb{1}_{\{M(Ns)=1\}} - \pi_M(1))\overline{L_2^N}(s)ds \xrightarrow[N\to+\infty]{\mathbb{P}} 0.$$

Combining the above and using Slutsky's theorem, we can conclude that for every $t < t_0$,

$$\begin{aligned}
l_0(t) &= l_0(0) + \lambda^- t - (\lambda^+\pi_M(0) + \Phi_P)\int_0^t l_0(u)du,\\
l_1(t) &= l_1(0) - \lambda^- t + \lambda^+\pi_M(0)\int_0^t l_0(u)du + \lambda_F^+\pi_M(1)\int_0^t l_2(u)du,\\
l_2(t) &= l_2(0) + \Phi_P\int_0^t l_0(u)du - \lambda_F^+\pi_M(1)\int_0^t l_2(u)du.
\end{aligned}$$

By the same type of arguments, we can prove that for every $t > t_0$

$$
\begin{aligned}
l_0(t) &= l_0(0) + \lambda^- t_0 - \lambda^+ \pi_M(0) \int_0^{t_0} l_0(u)du - \Phi_P \int_0^t l_0(u)du, \\
l_1(t) &= 0, \\
l_2(t) &= l_2(0) + \Phi_P \int_0^t l_0(u)du - \lambda_F^+ \pi_M(1) \int_0^{t_0} l_2(u)du.
\end{aligned}
$$

As l_0 and l_2 are two continuous functions, the functions $t \mapsto \int_0^t l_0(u)du$ and $t \mapsto \int_0^t l_2(u)du$ are differentiable and for any $t \leq t_0$, we have

$$
\begin{aligned}
l_0'(t) =&\ \lambda^- - \pi_M(0)\lambda^+ l_0(t) - \Phi_P\, l_0(t), \\
l_1'(t) =&\ \pi_M(0)\lambda^+ l_0(t) + \pi_M(1)\lambda_F^+ l_2(t) - \lambda^-, \\
l_2'(t) =&\ \Phi_P l_0(t) - \pi_M(1)\lambda_F^+ l_2(t),
\end{aligned}
\tag{28}
$$

while for any $t > t_0$,

$$
l_0'(t) = -\Phi_P l_0(t), \qquad l_1'(t) =\ 0, \qquad l_2'(t) =\ \Phi_P l_0(t).
$$

Solving these systems of equations, we obtain that the limiting process is necessarily the one given in Theorem 1.1 and therefore this limit is unique. Since $(\overline{L^N})_{N\geq 1}$ is tight by Proposition 2.2, we can conclude that the full sequence converges in distribution and Theorem 1.1 is proved. □

It remains to prove Lemmas 2.3 and 2.4. As the proofs are similar, we only give the proof of Lemma 2.3.

Proof of Lemma 2.3. Let $t > 0$. We split the interval $[0, t]$ into intervals of length $\varepsilon > 0$. This gives us

$$
\begin{aligned}
&\left| \int_0^t (\mathbb{1}_{\{M(Ns)=0\}} - \pi_M(0))\overline{L_0^N}(s)ds \right| \\
&\leq \left| \sum_{k=1}^{\lfloor \frac{t}{\varepsilon} \rfloor} \int_{(k-1)\varepsilon}^{k\varepsilon} (\mathbb{1}_{\{M(Ns)=0\}} - \pi_M(0))\overline{L_0^N}(s)ds \right| + \left| \int_{\lfloor \frac{t}{\varepsilon} \rfloor \varepsilon}^t (\mathbb{1}_{\{M(Ns)=0\}} - \pi_M(0))\overline{L_0^N}(s)ds \right| \\
&\leq \sum_{k=1}^{\lfloor \frac{t}{\varepsilon} \rfloor} \left| \int_{(k-1)\varepsilon}^{k\varepsilon} (\mathbb{1}_{\{M(Ns)=0\}} - \pi_M(0))\overline{L_0^N}((k-1)\varepsilon)ds \right| \\
&\qquad\qquad + \sum_{k=1}^{\lfloor \frac{t}{\varepsilon} \rfloor} \left| \int_{(k-1)\varepsilon}^{k\varepsilon} (\mathbb{1}_{\{M(Ns)=0\}} - \pi_M(0)) \Big(\overline{L_0^N}(s) - \overline{L_0^N}((k-1)\varepsilon) \Big) ds \right| \\
&\quad + \left| \int_{\lfloor \frac{t}{\varepsilon} \rfloor \varepsilon}^t (\mathbb{1}_{\{M(Ns)=0\}} - \pi_M(0))\overline{L_0^N}(\lfloor t/\varepsilon \rfloor \varepsilon)ds \right| \\
&\qquad\qquad + \left| \int_{\lfloor \frac{t}{\varepsilon} \rfloor \varepsilon}^t (\mathbb{1}_{\{M(Ns)=0\}} - \pi_M(0)) \Big(\overline{L_0^N}(s) - \overline{L_0^N}(\lfloor t/\varepsilon \rfloor \varepsilon) \Big) ds \right|.
\end{aligned}
\tag{29}
$$

Now since $L_0^N((k-1)\varepsilon) \in [0, 1]$ for every k, we have

$$
\left| \int_{(k-1)\varepsilon}^{k\varepsilon} (\mathbb{1}_{\{M(Ns)=0\}} - \pi_M(0))\overline{L_0^N}((k-1)\varepsilon)ds \right| \leq \left| \int_{(k-1)\varepsilon}^{k\varepsilon} (\mathbb{1}_{\{M(Ns)=0\}} - \pi_M(0))ds \right|. \tag{30}
$$

and likewise

$$\left| \int_{(k-1)\varepsilon}^{k\varepsilon} (\mathbb{1}_{\{M(Ns)=0\}} - \pi_M(0))(\overline{L_0^N}(s) - \overline{L_0^N}((k-1)\varepsilon) ds \right| \leq \int_{(k-1)\varepsilon}^{k\varepsilon} \left| \overline{L_0^N}(s) - \overline{L_0^N}((k-1)\varepsilon) \right| ds. \quad (31)$$

We now control each of these terms.

First, since M is an irreducible Markov jump process on a finite state space, the ergodic theorem applies and gives us that for every $0 < s_1 < s_2$, we have the almost sure convergence

$$\int_{s_1}^{s_2} \mathbb{1}_{\{M(Ns)=0\}} ds = \frac{1}{N} \int_{Ns_1}^{Ns_2} \mathbb{1}_{\{M(s)=0\}} ds \xrightarrow[N\to+\infty]{a.s.} \pi_M(0)(s_2 - s_1).$$

Applying this result with $s_1 = (k-1)\varepsilon$ and $s_2 = k\varepsilon$ (or $s_2 = t$ when $k = \lfloor t/\varepsilon \rfloor + 1$) for each of the finitely many values of k of interest, we obtain that

$$\sum_{k=1}^{\lfloor t/\varepsilon \rfloor} \left| \int_{(k-1)\varepsilon}^{k\varepsilon} (\mathbb{1}_{\{M(Ns)=0\}} - \pi_M(0)) ds \right| + \left| \int_{\lfloor \frac{t}{\varepsilon} \rfloor \varepsilon}^{t} (\mathbb{1}_{\{M(Ns)=0\}} - \pi_M(0)) ds \right| \xrightarrow[N\to+\infty]{a.s.} 0. \quad (32)$$

Second, recall from (20a) that for every $0 \leq s' \leq s$, we have

$$\left| \overline{L_0^N}(s) - \overline{L_0^N}(s') \right| \leq (\lambda^- + \lambda^+ + \Phi_P)(s - s') + \left| \overline{X_0^N}(s) - \overline{X_0^N}(s') \right|.$$

Applying this result with $s' = (k-1)\varepsilon$ and $s \in [(k-1)\varepsilon, k\varepsilon]$ and combining them with (24) (with $T = t$), we obtain that for $\mathcal{C} := \lambda^- + \lambda^+ + \Phi_P + 2/9$,

$$\lim_{N\to\infty} \mathbb{P}\left(\int_{(k-1)\varepsilon}^{k\varepsilon} \left| \overline{L_0^N}(s) - \overline{L_0^N}((k-1)\varepsilon) \right| ds \geq \mathcal{C}\varepsilon^2 \right) = 0.$$

Consequently,

$$\begin{aligned}
&\lim_{N\to\infty} \mathbb{P}\left(\sum_{k=1}^{\lfloor \frac{t}{\varepsilon} \rfloor} \int_{(k-1)\varepsilon}^{k\varepsilon} \left| \overline{L_0^N}(s) - \overline{L_0^N}((k-1)\varepsilon) \right| ds + \int_{\lfloor \frac{t}{\varepsilon} \rfloor \varepsilon}^{t} \left| \overline{L_0^N}(s) - \overline{L_0^N}(\lfloor t/\varepsilon \rfloor \varepsilon) \right| ds \geq \mathcal{C}\varepsilon^2 \left(\left\lfloor \frac{t}{\varepsilon} \right\rfloor + 1 \right) \right) \\
&\quad \leq \lim_{N\to\infty} \left\{ \sum_{k=1}^{\lfloor \frac{t}{\varepsilon} \rfloor} \mathbb{P}\left(\int_{(k-1)\varepsilon}^{k\varepsilon} \left| \overline{L_0^N}(s) - \overline{L_0^N}((k-1)\varepsilon) \right| ds \geq \mathcal{C}\varepsilon^2 \right) \right. \\
&\qquad\qquad \left. + \mathbb{P}\left(\int_{\lfloor \frac{t}{\varepsilon} \rfloor \varepsilon}^{t} \left| \overline{L_0^N}(s) - \overline{L_0^N}(\lfloor t/\varepsilon \rfloor \varepsilon) \right| ds \geq \mathcal{C}\varepsilon^2 \right) \right\} \\
&\quad = 0
\end{aligned} \quad (33)$$

Coming back to (29), combining (30), (31), (32) and (33), and observing that $\mathcal{C}\varepsilon^2(\lfloor t/\varepsilon \rfloor + 1) = \mathcal{O}(\mathcal{C}\varepsilon t)$ as ε tends to 0, we obtain that

$$\lim_{N\to\infty} \mathbb{P}\left(\left| \int_0^t (\mathbb{1}_{\{M(Ns)=0\}} - \pi_M(0)) \overline{L_0^N}(s) ds \right| \geq (\mathcal{C}t + 1)\varepsilon \right) = 0.$$

Since ε was arbitrary, this proves Lemma 2.3. □

3 Large population asymptotics

In this section, we prove Theorem 1.2. Recall that for every $N \in \mathbb{N}$, the Markov process $\overline{\mathcal{Z}^N}$ takes its values in the space $\mathcal{M}_{\mathcal{I}}^* \subset \mathcal{M}_{\mathcal{I}}^{(1)}$, respectively defined in (15) and (16). Recall also our notation

$$\langle \overline{\mathcal{Z}_t^N}, f \rangle = \int_{\mathcal{I}} f(i) \overline{\mathcal{Z}_t^N}(di),$$

whenever this integral makes sense. Let $\mathcal{D}$ be the set of all functions on $\mathcal{M}_{\mathcal{I}}^{(1)}$ of the form

$$\varphi_f : Z \mapsto \varphi\big(\langle Z, f \rangle\big), \tag{34}$$

with f a bounded measurable function on $\mathcal{I}$ and φ a function of class $\mathcal{C}^1$ on $\mathbb{R}$.

Since the rescaled process $\left(\overline{\mathcal{Z}_t^N}\right)_{t \in \mathbb{R}_+}$ is a Markov jump process with bounded rates for every $N \in \mathbb{N}$, its infinitesimal generator $\mathcal{G}^N$ applied to any $\varphi_f \in \mathcal{D}$ is given by: for every $Z \in \mathcal{M}_{\mathcal{I}}^{(1)}$,

$$\begin{aligned}
\mathcal{G}^N \varphi_f(Z) =& \lambda^+ Z(M,1) \sum_{l=1}^{N-1} NZ(S,l) \left(\varphi_f \left(Z + \frac{\delta_{(S,l+1)}}{N} - \frac{\delta_{(S,l)}}{N} - \frac{\delta_{(M,1)}}{N} \right) - \varphi_f(Z) \right) \\
&+ \lambda_F^+ Z(P,1) \sum_{l=1}^{N-1} NZ(F,l) \left(\varphi_f \left(Z + \frac{\delta_{(F,l+1)}}{N} - \frac{\delta_{(F,l)}}{N} - \frac{\delta_{(P,1)}}{N} \right) - \varphi_f(Z) \right) \\
&+ \lambda^- \sum_{T \in \{S,F\}} \sum_{l=1}^{N} NZ(T,l) \left(\varphi_f \left(Z + \frac{\delta_{(T,l-1)}}{N} \mathbb{I}_{\{l \geq 2\}} - \frac{\delta_{(T,l)}}{N} + \frac{\delta_{(M,1)}}{N} \right) - \varphi_f(Z) \right) \\
&+ \Phi_P NZ(M,1) \left(\varphi_f \left(Z + \frac{\delta_{(P,1)}}{N} - \frac{\delta_{(M,1)}}{N} \right) - \varphi_f(Z) \right) \\
&+ \Phi_F^+ \sum_{l=1}^{N} NZ(S,l) \left(\varphi_f \left(Z + \frac{\delta_{(F,l)}}{N} - \frac{\delta_{(S,l)}}{N} \right) - \varphi_f(Z) \right) \\
&+ \Phi_F^- \sum_{l=1}^{N} NZ(F,l) \left(\varphi_f \left(Z + \frac{\delta_{(S,l)}}{N} - \frac{\delta_{(F,l)}}{N} \right) - \varphi_f(Z) \right),
\end{aligned} \tag{35}$$

where for simplicity we have written $Z(T,l)$ for $Z(\{(T,l)\})$. Furthermore, $\left(\overline{\mathcal{Z}_t^N}\right)_{t \in \mathbb{R}_+}$ satisfies the property that for every $\varphi_f \in \mathcal{D}$, the process $\left(\overline{\mathcal{Y}_t^N}(\varphi_f)\right)_{t \in \mathbb{R}_+}$ defined by:

$$\left(\overline{\mathcal{Y}_t^N}(\varphi_f)\right)_{t \in \mathbb{R}_+} \stackrel{def}{=} \left(\varphi_f\left(\overline{\mathcal{Z}_t^N}\right) - \varphi_f\left(\overline{\mathcal{Z}_0^N}\right) - \int_0^t \mathcal{G}^N \varphi_f\left(\overline{\mathcal{Z}_s^N}\right) ds \right)_{t \in \mathbb{R}_+} \tag{36}$$

is a mean-zero martingale with predictable quadratic variation given for all $t \geq 0$ by

$$\langle \overline{\mathcal{Y}^N}(\varphi_f) \rangle_t = \int_0^t y_s^N(\varphi_f) ds, \tag{37}$$

where, writing this time $\overline{K_t^N}(T,l)$ for $\overline{\mathcal{Z}_t^N}(\{(T,l)\})$,

$$
\begin{aligned}
y_t^N(\varphi_f) \stackrel{def}{=} & \lambda^+\overline{K_t^N}(M,1)\sum_{l=1}^{N-1} N\overline{K_t^N}(S,l)\left(\varphi_f\left(\overline{\mathcal{Z}_t^N}+\frac{\delta_{(S,l+1)}}{N}-\frac{\delta_{(S,l)}}{N}-\frac{\delta_{(M,1)}}{N}\right)-\varphi_f\left(\overline{\mathcal{Z}_t^N}\right)\right)^2 \\
& + \lambda_F^+\overline{K_t^N}(P,1)\sum_{l=1}^{N-1} N\overline{K_t^N}(F,l)\left(\varphi_f\left(\overline{\mathcal{Z}_t^N}+\frac{\delta_{(F,l+1)}}{N}-\frac{\delta_{(F,l)}}{N}-\frac{\delta_{(P,1)}}{N}\right)-\varphi_f\left(\overline{\mathcal{Z}_t^N}\right)\right)^2 \\
& + \lambda^-\sum_{T\in\{S,F\}}\sum_{l=1}^{N} N\overline{K_t^N}(T,l)\left(\varphi_f\left(\overline{\mathcal{Z}_t^N}+\frac{\delta_{(T,l-1)}}{N}\mathbb{1}_{\{l\geq 2\}}-\frac{\delta_{(T,l)}}{N}+\frac{\delta_{(M,1)}}{N}\right)-\varphi_f\left(\overline{\mathcal{Z}_t^N}\right)\right)^2 \\
& + \Phi_P N\overline{K_t^N}(M,1)\left(\varphi_f\left(\overline{\mathcal{Z}_t^N}+\frac{\delta_{(P,1)}}{N}-\frac{\delta_{(M,1)}}{N}\right)-\varphi_f\left(\overline{\mathcal{Z}_t^N}\right)\right)^2 \\
& + \Phi_F^+\sum_{l=1}^{N} N\overline{K_t^N}(S,l)\left(\varphi_f\left(\overline{\mathcal{Z}_t^N}+\frac{\delta_{(F,l)}}{N}-\frac{\delta_{(S,l)}}{N}\right)-\varphi_f\left(\overline{\mathcal{Z}_t^N}\right)\right)^2 \\
& + \Phi_F^-\sum_{l=1}^{N} N\overline{K_t^N}(F,l)\left(\varphi_f\left(\overline{\mathcal{Z}_t^N}+\frac{\delta_{(S,l)}}{N}-\frac{\delta_{(F,l)}}{N}\right)-\varphi_f\left(\overline{\mathcal{Z}_t^N}\right)\right)^2.
\end{aligned}
$$

Hence, the semimartingale decomposition of $\varphi_f(\overline{\mathcal{Z}^N})$ reads

$$
\varphi_f\left(\overline{\mathcal{Z}_t^N}\right)=\varphi_f\left(\overline{\mathcal{Z}_0^N}\right)+\overline{\mathcal{V}_t^N}(\varphi_f)+\overline{\mathcal{Y}_t^N}(\varphi_f),\qquad t\geq 0 \tag{38}
$$

with

$$
\overline{\mathcal{V}_t^N}(\varphi_f)\stackrel{def}{=}\int_0^t \mathcal{G}^N\varphi_f\left(\overline{\mathcal{Z}_s^N}\right)ds.
$$

3.1 Tightness

As a first step in the proof of Theorem 1.2, we show the tightness of $(\overline{\mathcal{Z}^N})_{N\in\mathbb{N}}$.

Let $\mathcal{I}^\Delta$ be the one-point compactification of $\mathcal{I}$. In this subsection, we use the classical trick of seeing our sequence of processes as taking values in the space $\mathcal{M}_{\mathcal{I}^\Delta}$ of all finite measures on $\mathcal{I}^\Delta$. In the next subsection, we shall show that any limit point actually takes its values in $\mathcal{M}_{\mathcal{I}}^{(1)}$.

Proposition 3.1 (Tightness). *The sequence $\left(\overline{\mathcal{Z}^N}\right)_{N\in\mathbb{N}}$ is tight in $D_{\mathcal{M}_{\mathcal{I}^\Delta}}[0,\infty)$.*

Proof of Proposition 3.1. Since the total mass of each $\overline{\mathcal{Z}_t^N}$ is bounded by 1 and since the set of all measures on $\mathcal{I}^\Delta$ with total mass bounded by 1 is compact (hence the fact that we consider the one-point compactification of $\mathcal{I}$), the compact containment condition holds for $(\overline{\mathcal{Z}^N})_{N\geq 1}$ seen as a sequence in $D_{\mathcal{M}_{\mathcal{I}^\Delta}}[0,\infty)$. By Theorem 3.9.1 in [11], tightness of $(\overline{\mathcal{Z}^N})_{N\geq 1}$ will follow from the tightness of $(\varphi_f(\overline{\mathcal{Z}^N}))_{N\geq 1}$ for every $\varphi_f\in\mathcal{D}$.

Let thus f be a bounded measurable function on $\mathcal{I}^\Delta$ and $\varphi\in\mathcal{C}^1(\mathbb{R})$. We use the standard Aldous-Rebolledo criterion [1, 20]. First, the tightness of $(\varphi_f(\overline{\mathcal{Z}_t^N}))_{N\geq 1}$ for every $t\geq 0$ is obvious from the boundedness of f and the fact that the total mass of $\overline{\mathcal{Z}_t^N}$ is bounded by 1.

Second, we have to prove that for any $T>0$, any sequence of stopping times $(T_N)_{N\geq 1}$ bounded by T and for every $\varepsilon>0$, there exists $\delta>0$ and $N_0\in\mathbb{N}$ such that for all $\forall N\geq N_0$ and $\theta\in[0,\delta]$,

$$
\mathbb{P}\left(\left|\int_{T_N}^{T_N+\theta}\mathcal{G}^N\varphi_f\left(\overline{\mathcal{Z}_t^N}\right)dt\right|>\varepsilon\right)\leq\varepsilon, \tag{39}
$$

and

$$\mathbb{P}\left(\left|\int_{T_N}^{T_N+\theta} y_t^N(\varphi_f)dt\right| > \varepsilon\right) \leq \varepsilon. \tag{40}$$

Plugging the definition of $\varphi_f(Z) = \varphi(\langle Z, f\rangle)$ in the expression for $\mathcal{G}^N\varphi_f$, Taylor expanding φ and writing $\|\varphi'\|_f$ for the supremum of φ' over $[-\|f\|_\infty, \|f\|_\infty]$, we obtain that

$$\begin{aligned}\left|\mathcal{G}^N\varphi_f\left(\overline{\mathcal{Z}_t^N}\right)\right| \leq & \|\varphi'\|_f\|f\|_\infty\Bigg(3\lambda^+\overline{K_t^N}(M,1)\sum_{l=1}^{N-1}\overline{K_t^N}(S,l) + 3\lambda_F^+\overline{K_t^N}(P,1)\sum_{l=1}^{N-1}\overline{K_t^N}(F,l) \\ & + 3\lambda^-\sum_{T\in\{S,F\}}\sum_{l=1}^{N}\overline{K_t^N}(T,l) + 2\Phi_P\overline{K_t^N}(M,1) + 2\Phi_F^+\sum_{l=1}^{N}\overline{K_t^N}(S,l) \\ & + 2\Phi_F^-\sum_{l=1}^{N}\overline{K_t^N}(F,l)\Bigg) + o(1),\end{aligned} \tag{41}$$

where the remainder term tends to 0 uniformly in N and t. By construction (since the total mass of each $\overline{\mathcal{Z}_t^N}$ is bounded by 1), for every type $T \in \{S, F\}$ we have

$$\overline{K_t^N}(M,1) \leq 1, \qquad \overline{K_t^N}(P,1) \leq 1, \qquad \sum_{l=1}^{N}\overline{K_t^N}(T,l) \leq 1,$$

and hence

$$\left|\mathcal{G}^N\varphi_f\left(\overline{\mathcal{Z}_s^N}\right)\right| \leq \|\varphi'\|_f\|f\|_\infty\left(3\lambda^+ + 3\lambda_F^+ + 6\lambda^- + 2\Phi_P + 2\Phi_F^+ + 2\Phi_F^-\right) + o(1) \stackrel{def}{=} C_1 + o(1). \tag{42}$$

Therefore, taking $\delta = \varepsilon/(2C_1)$, we see that there exists $N_0 \in \mathbb{N}$ such that for every $N \geq N_0$ and $\theta \in [0, \delta]$,

$$\mathbb{P}\left(\left|\int_{T_N}^{T_N+\theta}\mathcal{G}^N\varphi_f\left(\overline{\mathcal{Z}_t^N}\right)dt\right| > \varepsilon\right) = 0$$

and (39) is satisfied.

We prove (40) in exactly the same way, noticing that the squared increments in the expression for $y_t^N(\varphi_f)$ give rise to a bound on $|y_t^N(\varphi_f)|$ of the form $C_2/N + o(1/N)$, where again the remainder term is uniform in t. Proceeding this way, we obtain a stronger result than (40), namely that the predictable quadratic variation of $\varphi_f(\overline{\mathcal{Z}^N})$ vanishes as $N \to \infty$, uniformly over compact time intervals: by the Markov inequality and then Doob's maximal inequality,

$$\mathbb{P}\left(\sup_{t\in[0,T]}\left|\overline{\mathcal{Y}_t^N}(\varphi_f)\right| > \varepsilon\right) \leq \frac{1}{\varepsilon^2}\mathbb{E}\left[\sup_{t\in[0,T]}\left|\overline{\mathcal{Y}_t^N}(\varphi_f)\right|^2\right] \leq \frac{2}{\varepsilon^2}\mathbb{E}\left[\langle\overline{\mathcal{Y}^N}(\varphi_f)\rangle_T\right] \leq \frac{2C_2T}{N\varepsilon^2} + o\left(\frac{1}{N}\right), \tag{43}$$

which tends to 0 as $N \to \infty$. As a corollary to this result, we obtain that any potential limit for $(\overline{\mathcal{Z}^N})_{N\geq 1}$ must be deterministic. Since (43) implies (40), all the points in the Aldous-Rebolledo criterion have now been checked and we can conclude that $(\varphi_f(\overline{\mathcal{Z}^N}))_{N\geq 1}$ is tight in $D_{\mathbb{R}}[0,\infty)$ for every $\varphi_f \in \mathcal{D}$, and finally that the sequence $(\overline{\mathcal{Z}^N})_{N\geq 1}$ is tight in $D_{\mathcal{M}_{\mathcal{I}^\Delta}}[0,\infty)$. □

3.2 Large population limit

We now complete the proof of Theorem 1.2 by checking that all the assumptions of Theorem 4.8.10 in [11] are met. Tightness in $D_{\mathcal{M}_{\mathcal{I}^\Delta}}[0,\infty)$ was obtained in the previous subsection and we now want to prove that, if we define the operator $\mathcal{G}^\infty$ on $\mathcal{D}$ (the set of all functions of the form $\varphi(\langle\cdot, f\rangle)$ with $\varphi \in \mathcal{C}^1(\mathbb{R})$ and $f : \mathcal{I} \to \mathbb{R}$ bounded measurable) by: for every $Z \in \mathcal{M}_{\mathcal{I}}$,

$$\begin{aligned}\mathcal{G}^\infty\varphi_f(Z) = \varphi'(\langle Z, f\rangle)\Big\{&\lambda^+ Z(M,1)\sum_{l=1}^{+\infty} Z(S,l)\left(f(S,l+1) - f(S,l) - f(M,1)\right)\\ &+ \lambda_F^+ Z(P,1)\sum_{l=1}^{+\infty} Z(F,l)\left(f(F,l+1) - f(F,l) - f(P,1)\right)\\ &+ \lambda^- \sum_{T\in\{S,F\}}\sum_{l=1}^{+\infty} Z(T,l)\left(f(T,l-1)\mathbb{I}_{\{l\geq 2\}} - f(T,l) + f(M,1)\right)\\ &+ \Phi_P Z(M,1)\left(f(P,1) - f(M,1)\right)\\ &+ \Phi_F^+\sum_{l=1}^{+\infty} Z(S,l)\left(f(F,l) - f(S,l)\right) + \Phi_F^-\sum_{l=1}^{+\infty} Z(F,l)\left(f(S,l) - f(F,l)\right)\Big\}, \qquad (44)\end{aligned}$$

then for every $\varphi_f \in \mathcal{D}$ and $t, s > 0$, every $k \in \mathbb{N}$, $0 \leq t_1 < \cdots < t_k \leq t < t+s$ and $\beta_i \in \mathcal{C}_b(\mathcal{M}_{\mathcal{I}})$, we have

$$\lim_{N\to\infty} \mathbb{E}\left[\left(\varphi_f\big(\overline{\mathcal{Z}^N_{t+s}}\big) - \varphi_f\big(\overline{\mathcal{Z}^N_t}\big) - \int_t^{t+s} \mathcal{G}^\infty\varphi_f\big(\overline{\mathcal{Z}^N_u}\big)du\right)\left(\prod_{i=1}^k \beta_i\big(\overline{\mathcal{Z}^N_{t_i}}\big)\right)\right] = 0. \qquad (45)$$

Once this result is shown, we shall prove that any potential limiting process takes its value in $\mathcal{M}_{\mathcal{I}}^{(1)}$, on which every $\mathcal{G}^\infty\varphi_f$ is a continuous and bounded function, which will also prove that tightness of $(\overline{\mathcal{Z}^N})_{N\geq 1}$ holds in $D_{\mathcal{M}_{\mathcal{I}}^{(1)}}[0,\infty)$ (by Corollary 3.9.3 in [11]). Finally, we shall prove that there is at most one possible limit for $(\overline{\mathcal{Z}^N})_{N\geq 1}$, which is the deterministic solution to the $D_{\mathcal{M}_{\mathcal{I}}^{(1)}}[0,\infty)$-martingale problem associated to $(\mathcal{G}^\infty, \delta_{\mathcal{Z}_0})$ or, equivalently, the $\mathcal{M}_{\mathcal{I}}^{(1)}$-valued solution to the system of equations stated in Theorem 1.2. Combining these arguments, we shall be able to use Theorem 4.8.10 in [11] to conclude that the limiting process exists and $(\overline{\mathcal{Z}^N})_{N\geq 1}$ indeed converges to it in $D_{\mathcal{M}_{\mathcal{I}}^{(1)}}[0,\infty)$.

Let us thus start by showing (45). Using the decomposition (38), we have for every $N \in \mathbb{N}$,

$$\mathbb{E}\left[\left(\varphi_f\big(\overline{\mathcal{Z}^N_{t+s}}\big) - \varphi_f\big(\overline{\mathcal{Z}^N_t}\big) - \int_t^{t+s} \mathcal{G}^N\varphi_f\big(\overline{\mathcal{Z}^N_u}\big)du\right)\left(\prod_{i=1}^k \beta_i\big(\overline{\mathcal{Z}^N_{t_i}}\big)\right)\right] = 0. \qquad (46)$$

It thus remains to prove that

$$\lim_{N\to\infty} \mathbb{E}\left[\int_t^{t+s} \left|\mathcal{G}^N\varphi_f\big(\overline{\mathcal{Z}^N_u}\big) - \mathcal{G}^\infty\varphi_f\big(\overline{\mathcal{Z}^N_u}\big)\right|du\right] = 0, \qquad (47)$$

since then

$$\begin{aligned}
&\left|\mathbb{E}\Big[\Big(\varphi_f(\overline{\mathcal{Z}^N_{t+s}}) - \varphi_f(\overline{\mathcal{Z}^N_t}) - \int_t^{t+s} \mathcal{G}^\infty\varphi_f(\overline{\mathcal{Z}^N_u})du\Big)\Big(\prod_{i=1}^k \beta_i(\overline{\mathcal{Z}^N_{t_i}})\Big)\Big]\right| \\
&\qquad = \left|\mathbb{E}\Big[\Big(\int_t^{t+s} (\mathcal{G}^N\varphi_f(\overline{\mathcal{Z}^N_u}) - \mathcal{G}^\infty\varphi_f(\overline{\mathcal{Z}^N_u}))du\Big)\Big(\prod_{i=1}^k \beta_i(\overline{\mathcal{Z}^N_{t_i}})\Big)\Big]\right| \\
&\qquad \le \Big(\prod_{i=1}^k \|\beta_i\|_\infty\Big)\mathbb{E}\Big[\int_t^{t+s} \big|\mathcal{G}^N\varphi_f(\overline{\mathcal{Z}^N_u}) - \mathcal{G}^\infty\varphi_f(\overline{\mathcal{Z}^N_u})\big|du\Big] \to 0
\end{aligned}$$

as $N \to \infty$, and we obtain (45).

Coming back to the definition of $\mathcal{G}^N\varphi_f$ and performing a Taylor expansion of φ, it is straightforward to obtain that there exists a sequence $(\epsilon_N)_{N\ge1}$ tending to 0 such for any $Z \in \mathcal{M}_{\mathcal{I}}^{(1)}$,

$$\big|\mathcal{G}^N\varphi_f(Z) - \mathcal{G}^\infty\varphi_f(Z)\big| \le \epsilon_N.$$

Using this result, the fact that for every $N \in \mathbb{N}$ and $t \ge 0$ we have $\langle\overline{\mathcal{Z}^N_t}, 1\rangle \le 1$, the bound (42) and the dominated convergence theorem, we can conclude that (47) holds true and the proof of (45) is complete.

Second, let us show that any potential limit point $(z_t)_{t\ge0}$ for $(\overline{\mathcal{Z}^N})_{N\ge1}$ satisfies that $z_t \in \mathcal{M}_{\mathcal{I}}^{(1)}$ for every $t \ge 0$. We already know from the subsection on tightness that each z_t is deterministic and that its total mass is bounded by 1. It only remains to show that z_t gives no mass to the "infinity" point of $\mathcal{I}^\Delta$ (*i.e.*, to polymers of infinite length). Let thus $t \ge 0$ and N large enough for (17) to hold. Since the elongation rate of each polymer is bounded by $\lambda^* = \max\{\lambda^+, \lambda_F^+\}$, the probability that at time t a given polymer is of length at least $K_0 + k$ is bounded by the probability that the sum of k i.i.d exponential random variables with parameter λ^* is less than t, that is:

$$\mathbb{P}\Big(\sum_{i=1}^k E_i < t\Big) = \mathbb{P}\Big(\exp\Big(-\sum_{i=1}^k E_i\Big) > e^{-t}\Big) \le e^t\Big(\frac{\lambda^*}{\lambda^*+1}\Big)^k,$$

where the last inequality uses the Markov inequality. Since by construction $\overline{\mathcal{Z}^N_t}$ has at most N atoms (or "polymers"), another use of the Markov inequality gives us that for every $k \in \mathbb{N}$, $\varepsilon \in (0,1)$ and N large enough (independently of k and ε)

$$\mathbb{P}\Big(\langle\overline{\mathcal{Z}^N_t}, \mathbb{I}_{\{l\ge K_0+k\}}\rangle > \varepsilon\Big) \le \frac{1}{\varepsilon}\,\mathbb{E}\Big[\langle\overline{\mathcal{Z}^N_t}, \mathbb{I}_{\{l\ge K_0+k\}}\rangle\Big] \le \frac{1}{\varepsilon N}\,Ne^t\Big(\frac{\lambda^*}{\lambda^*+1}\Big)^k.$$

Since we suppose that $\overline{\mathcal{Z}^N_t}$ converges in distribution to z_t, we can take the limit as $N \to \infty$ in the above to obtain

$$\mathbb{P}\Big(\langle z_t, \mathbb{I}_{\{l\ge K_0+k\}}\rangle > \varepsilon\Big) \le \frac{1}{\varepsilon}\,e^t\Big(\frac{\lambda^*}{\lambda^*+1}\Big)^k. \tag{48}$$

Together with the fact that z_t is not random, we obtain that for every $\varepsilon \in (0,1)$, there exists $k(\varepsilon) \in \mathbb{N}$ such that $\varepsilon^{-1}e^t(\lambda^*/(\lambda^*+1))^{k(\varepsilon)} < 1$ and so

$$\langle z_t, \mathbb{I}_{\{l\ge K_0+k(\varepsilon)\}}\rangle \le \varepsilon.$$

Taking ε to 0, we can conclude that z_t puts no mass on infinitely long polymers, *i.e.*, $z_t \in \mathcal{M}_{\mathcal{I}}^{(1)}$.

Thirdly, let us check that any potential limit $\mathcal{Z}$ for $(\overline{\mathcal{Z}^N})_{N\geq 1}$ satisfies the system of equations stated in Theorem 1.2. As recalled earlier, we know from (43) that $\mathcal{Z}$ is deterministic (since we assumed that the limit of the initial value $\overline{\mathcal{Z}_0^N}$ was deterministic). Using (45) along the subsequence of $(\overline{\mathcal{Z}^N})_{N\geq 1}$ converging to $\mathcal{Z}$, we can conclude that for every bounded measurable function f on $\mathcal{I}$ (taking $\varphi = \mathrm{Id}$),

$$\left(\langle \mathcal{Z}_t, f\rangle - \langle \mathcal{Z}_0, f\rangle - \int_0^t \mathcal{G}^\infty \mathrm{Id}_f(\mathcal{Z}_s)ds\right)_{t\geq 0}$$

is a martingale, which is thus constant equal to 0. Taking now $f = \mathbb{1}_{\{(T,l)\}}$ for $(T,l)\in\mathcal{I}$, we obtain the integrated version of the differential equation for $k_t(T,l) = \mathcal{Z}_t(\{(T,L)\})$ given in Theorem 1.2, namely

$$k_t(T,L) = k_0(T,l) + \int_0^t \mathcal{G}^\infty \mathrm{Id}_{\mathbb{1}_{(T,l)}}(\mathcal{Z}_s)ds. \tag{49}$$

Since the integrand in the above expression is bounded, the function $t\mapsto k_t(T,l)$ is differentiable and we can therefore differentiate it to obtain the corresponding line in the system of Theorem 1.2.

Finally, let us prove that this system has at most one solution with trajectories in $\mathcal{M}_{\mathcal{I}}^{(1)}$. Let thus $(z_t)_{t\geq 0}$ and $(\tilde{z}_t)_{t\geq 0}$ be two solutions to the limiting system of equations in Theorem 1.2 such that $z_0 = \tilde{z}_0$ and $z_t, \tilde{z}_t \in \mathcal{M}_{\mathcal{I}}^{(1)}$ for all $t\geq 0$. Let f be a bounded measurable function on $\mathcal{I}$ such that $\|f\|_\infty \leq 1$. From the above, for every $t\geq 0$, we have

$$\big|\langle z_t, f\rangle - \langle \tilde{z}_t, f\rangle\big| = \left|\int_0^t \big(\mathcal{G}^\infty \mathrm{Id}_f(z_s) - \mathcal{G}^\infty \mathrm{Id}_f(\tilde{z}_s)\big)ds\right|. \tag{50}$$

Let us consider each term in $\mathcal{G}^\infty \mathrm{Id}_f(z_s) - \mathcal{G}^\infty \mathrm{Id}_f(\tilde{z}_s)$ separately to obtain a bound on the integral on the r.h.s. of (50). We have

$$\begin{aligned}
&\left|\lambda^+ \int_0^t \left(\sum_{l=1}^\infty \big\{z_s(M,1)z_s(S,l) - \tilde{z}_s(M,1)\tilde{z}_s(S,l)\big\} f(S,l+1)\right)ds\right| \\
&= \lambda^+ \left|\int_0^t \left(\sum_{l=1}^\infty \big\{z_s(M,1)(z_s(S,l) - \tilde{z}_s(S,l)) + (z_s(M,1) - \tilde{z}_s(M,1))\tilde{z}_s(S,l)\big\} f(S,l+1)\right)ds\right| \\
&= \lambda^+ \left|\int_0^t \left(z_s(M,1)\langle z_s - \tilde{z}_s, \mathbb{1}_{(S,\cdot)} f(S,\cdot+1)\rangle + \langle z_s - \tilde{z}_s, \mathbb{1}_{(M,1)}\rangle \sum_{l=1}^\infty \tilde{z}_s(S,l) f(S,l+1)\right)ds\right|.
\end{aligned} \tag{51}$$

By assumption, the total mass of z_s and $\tilde{z}_s$ is bounded by 1, and so

$$0\leq z_s(M,1)\leq 1 \qquad \text{and} \qquad \left|\sum_{l=1}^\infty \tilde{z}_s(S,l) f(S,l+1)\right| \leq \|f\|_\infty \langle \tilde{z}_s, 1\rangle \leq 1.$$

As a consequence, the r.h.s. of (51) is bounded by

$$\lambda^+ \int_0^t \Big\{\big|\langle z_s - \tilde{z}_s, \mathbb{1}_{(S,\cdot)} f(S,\cdot+1)\rangle\big| + \big|\langle z_s - \tilde{z}_s, \mathbb{1}_{(M,1)}\rangle\big|\Big\}ds \leq 2\lambda^+ \int_0^t \sup_{\|\phi\|_\infty\leq 1} \big|\langle z_s,\phi\rangle - \langle \tilde{z}_s,\phi\rangle\big|ds.$$

Proceeding in the same way for all the other terms in the generator, we obtain that

$$\big|\langle z_t, f\rangle - \langle \tilde{z}_t, f\rangle\big| \leq \big(6\lambda^+ + 6\lambda_F^+ + 6\lambda^- + 2\Phi_P + 2\Phi_F^+ + 2\Phi_F^-\big) \int_0^t \sup_{\|\phi\|_\infty\leq 1} \big|\langle z_s,\phi\rangle - \langle \tilde{z}_s,\phi\rangle\big|ds,$$

and since this inequality holds for every bounded measurable f such that $\|f\|_\infty \leq 1$, we can write that for every $t \geq 0$,

$$\sup_{\|\phi\|_\infty \leq 1} \left|\langle z_t, \phi\rangle - \langle \tilde{z}_t, \phi\rangle\right| \leq \left(6\lambda^+ + 6\lambda_F^+ + 6\lambda^- + 2\Phi_P + 2\Phi_F^+ + 2\Phi_F^-\right) \int_0^t \sup_{\|\phi\|_\infty \leq 1} \left|\langle z_s, \phi\rangle - \langle \tilde{z}_s, \phi\rangle\right| ds.$$

By Gronwall's lemma,

$$\sup_{\|\phi\|_\infty \leq 1} \left|\langle z_t, \phi\rangle - \langle \tilde{z}_t, \phi\rangle\right| \leq \left(\sup_{\|\phi\|_\infty \leq 1} \left|\langle z_0, \phi\rangle - \langle \tilde{z}_0, \phi\rangle\right|\right) e^{(6\lambda^+ + 6\lambda_F^+ + 6\lambda^- + 2\Phi_P + 2\Phi_F^+ + 2\Phi_F^-)t} = 0.$$

Since every bounded measurable function can be written as a multiple of a function ϕ such that $\|\phi\|_\infty \leq 1$, we obtain that $\langle z_t, f\rangle = \langle \tilde{z}_t, f\rangle$ for every bounded measurable f, and hence $z_t = \tilde{z}_t$ for all $t \geq 0$ and the system of Theorem 1.2 has at most one solution with trajectories in $\mathcal{M}_{\mathcal{I}}^{(1)}$.

We have now checked that all the conditions of Theorem 4.8.10 in [11] were satisfied and we can therefore conclude that, as N tends to infinity, $(\overline{\mathcal{Z}^N})_{N\geq 1}$ converges in distribution in $D_{\mathcal{M}_{\mathcal{I}}^{(1)}}[0,\infty)$ to the unique solution with values in $\mathcal{M}_{\mathcal{I}}^{(1)}$ of the system stated in Theorem 1.2. □

We end this section with some heuristics that are confirmed by the simulations discussed in the next section. In order to compare the large population limit of Theorem 1.2 with the fluid limit of Theorem 1.1, recall that informally, for a given N, L_0^N in the single polymer model corresponds to $K^N(M,1)$, L_2^N corresponds to $K^N(P,1)$ and L_1^N corresponds to M_{inF}^N (defined in (13)). The same correspondence can be made between the limiting objects, if we define

$$m_{inF}(t) \stackrel{def}{=} \sum_{l=1}^{+\infty} l(k_t(S,l) + k_t(F,l)).$$

(This quantity may be infinite.) Moreover, writing for any $t \geq 0$,

$$\overline{F^\infty}(t) = \lim_{N\to+\infty} \overline{F^N}(t) = \sum_{l=1}^{+\infty} k_t(S,l) + \sum_{l=1}^{+\infty} k_t(F,l)$$

for the amount of "real" polymers at time t in the limiting population process, we see that the quantity $\sum_{l=1}^{+\infty} k_t(S,l)/\overline{F^\infty}(t)$ corresponds to the proportion of those polymers which are simple and we expect it to be close to $\pi_M(0)$ when t is large (recall that $\pi_M(0)$ is the proportion of time that a single polymer spends not bound with a formin at stationarity). Although we have no proof of this fact, Figures 6 and 7 show two examples where this convergence happens.

Finally, summing the corresponding equations in the system of ODE stated in Theorem 1.2, we obtain the following system which can be more easily compared with the result of Theorem 1.1.

Corollary 3.2. *In the limit as $N \to +\infty$, the distribution of monomers among the different sup-populations (free monomers, inserted in a polymer or in a G-actin/profilin complex) is described by the following system of ODE:*

$$\begin{cases} \frac{d}{dt}k_t(M,1) &= \lambda^-\overline{F^\infty}(t) - \lambda^+\overline{F^\infty}(t)k_t(M,1)\sum_{l=1}^{+\infty} k_t(S,l) - \Phi_P k_t(M,1) \\ \frac{d}{dt}m_{inF}(t) &= \lambda^+\overline{F^\infty}(t)k_t(M,1)\sum_{l=1}^{+\infty} k_t(S,l) + \lambda_F^+\overline{F^\infty}(t)k_t(P,1)\sum_{l=1}^{+\infty} k_t(F,l) - \lambda^-\overline{F^\infty}(t) \\ \frac{d}{dt}k_t(P,1) &= \Phi_P k_t(M,1) - \lambda_F^+\overline{F^\infty}(t)k_t(P,1)\sum_{l=1}^{+\infty} k_t(F,l). \end{cases}$$

4 Simulations

4.1 Single actin polymer

In this section, we exhibit an interesting example of behaviour for a single long polymer, that highlights the role of the different proteins. The parameters chosen in these simulations are fairly arbitrary and do not come from biological measurements. Investigating more biologically-driven examples and the effect of sudden parameter changes on the behaviour of the system will be the object of future work.

Figure 3 shows the evolution of the system with $N = 1000$, $\lambda^+ = 10$, $\lambda^- = 2$, $\lambda_F^+ = 100$, $\Phi_P = 10$, $\Phi_F^+ = 1$ and $\Phi_F^- = 1$, starting with a unique actin polymer of length 334, 333 free monomers and 333 G-actin/profilin complexes. That is, initially a third of the monomers in the system are free, another third are associated with a profilin and the rest sits in the polymer.

Figure 3 is divided into three graphs. On the top left, we show the evolution in time of the number of free monomers, $L_0^N(t)$. The graph on the top right presents the evolution of the length of the polymer, $L_1^N(t)$, and the graph at the bottom corresponds to the evolution of the number of G-actin/profilin complexes, $L_2^N(t)$. Each graph displays four plots. The oscillating solid line shows a single realisation of the stochastic system of Section 1.2.1 with $N = 1000$ and the parameters expounded above. The smooth solid line represents N times the fluid limit described in Theorem 1.1 (so that we compare numbers instead of proportions of monomers). The dashed line represents (N times) the fluid limit that we would obtain if we assumed that the polymer was constantly in normal mode. Finally, the alternated dashed line represents (N times) the fluid limit that we would obtain if we assumed that the polymer was always in fast mode.

In this particular example, we observe that if the system stays in normal mode, the length of the polymer converges quickly to 0 because too many free monomers are captured by the polymer or by profilins and the rate at which depolymerisation re-creates free monomers is not fast enough to compensate. In contrast, if the polymer is constantly in fast mode or if it switches modes, the length of the polymer converges to a nonzero equilibrium. Noticeably, the polymer length at equilibrium is longer in the case of mode switching than in the case when the polymer remains in the state with the highest rate of polymerisation. As can be remarked on the two other graphs, this phenomenon is due to the fact that the switching between modes enables the polymer to use the two types of "resources" (free monomers and G-actin/profilin complexes) more efficiently, allowing the pool of free monomers not to be depleted (so that they can be consumed during the elongation of the polymer in normal mode, or be used to replenish the pool of complexes).

4.2 Population of actin polymers

Figure 4 displays the evolution of a population of actin polymers with the same set of parameters as in Figure 3, namely $\lambda^+ = 10$, $\lambda^- = 2$, $\lambda_F^+ = 100$, $\Phi_P = 10$, $\Phi_F^+ = 1$ and $\Phi_F^- = 1$. In order to compare the limiting models for a single polymer and for a population of competing polymers, we start with an analogous initial distribution of monomers into free monomers, and monomers sequestered in polymers or complexes: $33,3\%$ of free monomers, $33,4\%$ of monomers included in a polymer and $33,3\%$ of monomers in G-actin/ profilin complexes. Moreover, to match the conditions of Theorem 1.2 we start with 50% of the polymers being simple and 50% being associated with a formin, all of them being of length 30. Finally, we use the Euler explicit method with a time pitch $dt = 0.001$. Note that here we only compare the limits obtained in

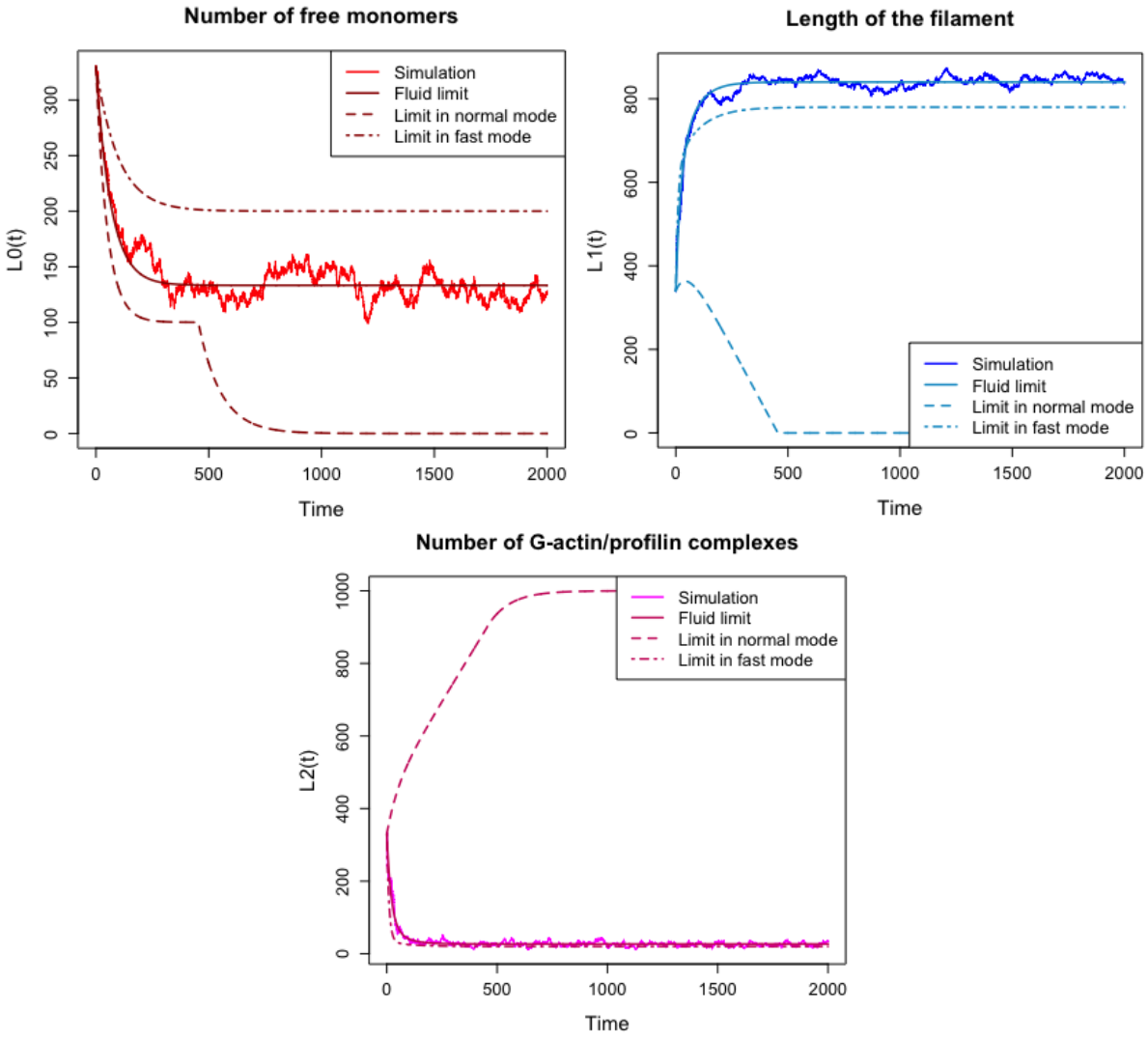


Figure 3: Evolution of the system described by the model for a single actin polymer with $N = 1000$, $\lambda^+ = 10$, $\lambda^- = 2$, $\lambda_F^+ = 100$, $\Phi_P = 10$, $\Phi_F^+ = 1$ and $\Phi_F^- = 1$. Initially, there are a (unique) polymer of length 334, 333 free monomers and 333 G-actin/profilin complexes. Left: evolution of the number of free monomers. Middle: evolution of the length of the polymer. Right: evolution of the number of G-actin/profilin complexes. On each graph, we plot a single realisation of the stochastic system for $N = 1000$, the fluid limit given in Theorem 1.1 (*i.e.*, N times $l_i(t)$), the fluid limit one would obtain by assuming that the system stays in normal mode and the fluid limit one would obtain by assuming that the system remains in fast mode all the time.

Theorems 1.1 and 1.2, and therefore we directly consider $N = \infty$.

In Figure 4, the same quantities are shown as in Figure 3, but for a population of actin polymers. Figure 4 is divided into three graphs. On the top left, we show the evolution in time of the proportion of free monomers in the population, $k_t(M, 1)$. The graph on the top right

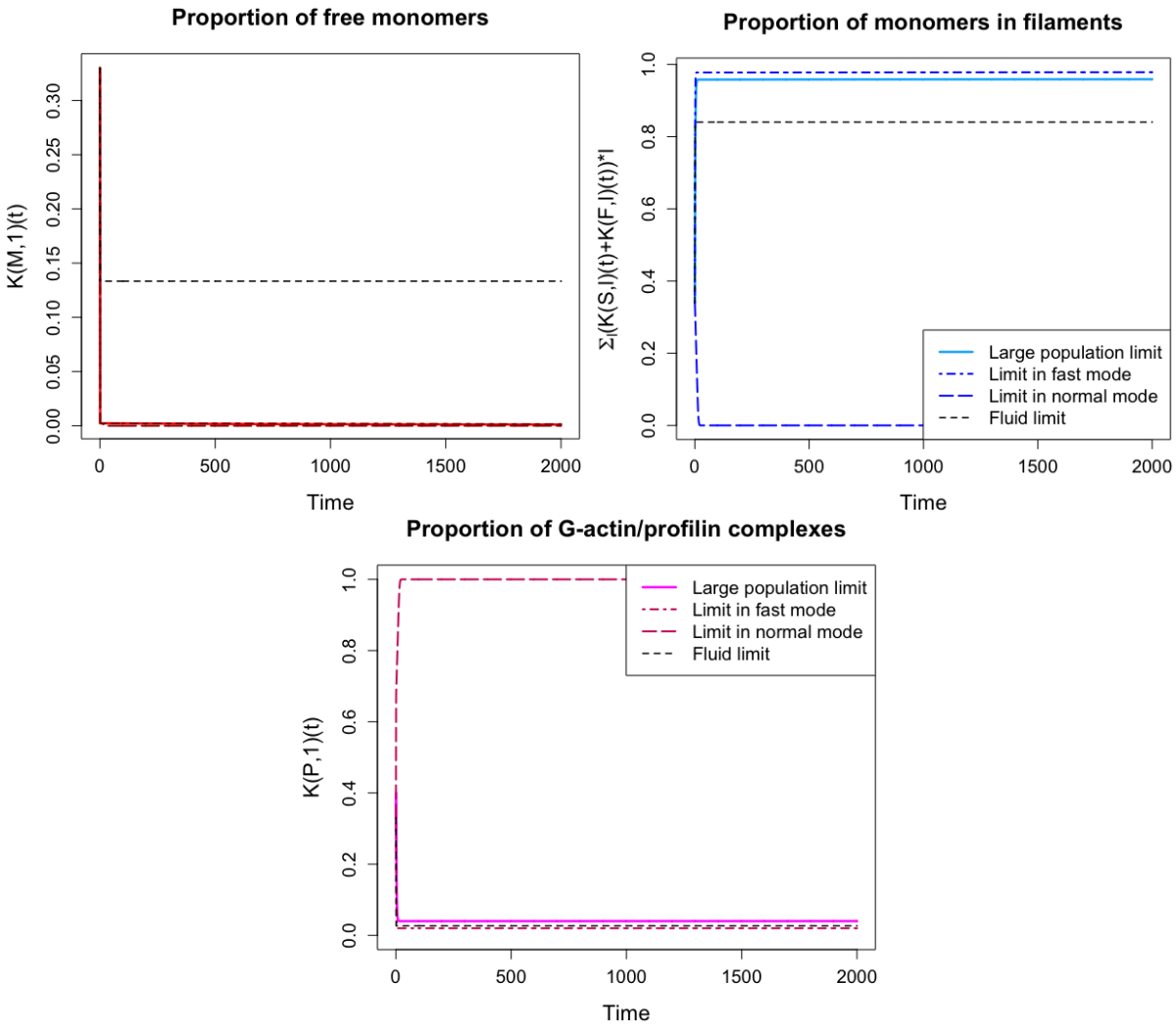
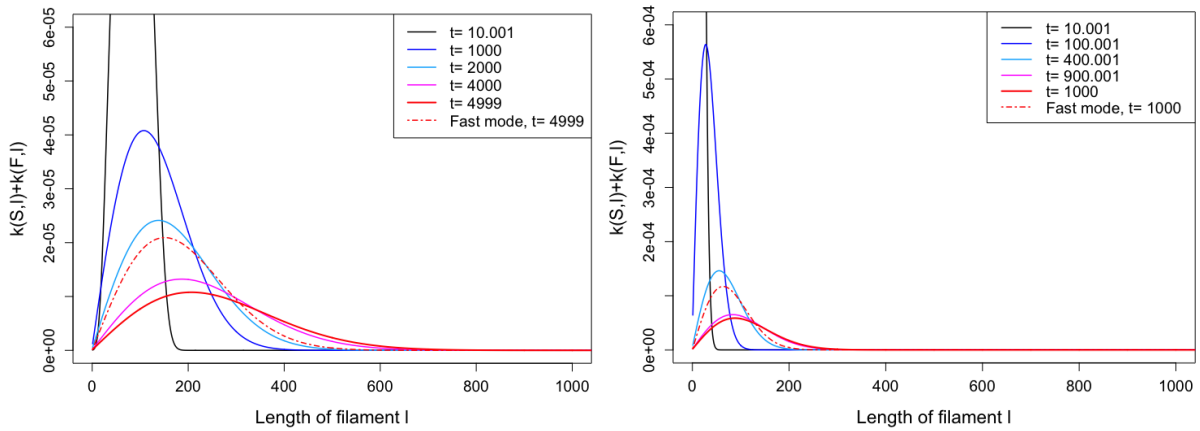


Figure 4: Evolution of the system described by the limiting model for a population of actin polymers with $dt = 0.001$, $\lambda^+ = 10$, $\lambda^- = 2$, $\lambda_F^+ = 100$, $\Phi_P = 10$, $\Phi_F^+ = 1$ and $\Phi_F^- = 1$, starting with 33% of free monomers, 33% of G-actin/profilin complexes and 34% of monomers included within polymers of length 30. Left: evolution of the proportion of free monomers. Middle: evolution of the proportion of monomers included in a polymer. Right: evolution of the proportion of G-actin/profilin complexes. On each graph, we plot the fluid limit given in Theorem 1.1, the large population limit given in Theorem 1.2, the large population limit obtained by assuming that the system stays in normal mode and the large population limit obtained by assuming that the system stays in fast mode.

presents the evolution of the proportion of monomers included in a polymer,

$$m_{inF}(t) = \sum_{l=1}^{+\infty} l(k_t(S,l) + k_t(F,l)),$$

and the graph at the bottom gives the evolution in time of the proportion G-actin/profilin complexes available in the population, $k_t(P,1)$. On each graph, we display four plots. The



(a) Initial polymers of length 30

(b) Initial polymers of length 3

Figure 5: Evolution of the distribution of the polymer lengths with $dt = 0.001$, $\lambda^+ = 10$, $\lambda^- = 2$, $\lambda_F^+ = 100$, $\Phi_P = 10$, $\Phi_F^+ = 1$ and $\Phi_F^- = 1$, starting with 33% of free monomers, 33% of G-actin/profilin complexes and 34% of monomers included in polymers. In each graph, we plot the distribution at different times. The red dashed line represents the distribution after stabilisation when assuming that all polymers are blocked in fast mode.

solid line represents the large population limit described in Theorem 1.2. The coloured dashed line (*resp.* coloured alternated dashed line) represents the large population limit that we would obtain if we assumed that all polymers were stuck in normal mode (*resp.* fast mode - in these cases, initially all the polymers are simple, *resp.*, associated with a formin). Finally, the thin black dashed line represents the equilibrium state in the fluid limit obtained in the single polymer model and shown in Figure 3.

In this particular example, we can observe that the qualitative behaviour of a population of polymers bears some similarity with the fluid limit for a single long polymer. Indeed, if all polymers are blocked in normal mode, the proportion of monomers inserted in a polymer converges quickly to 0 because a large fraction of free monomers is quickly captured by profilins and cannot be converted again into free monomers (since there is no accelerated polymerisation), so that they can no longer be used in polymer elongation. In contrast, when the polymers can switch modes or when they all remain in fast mode, the proportion of monomers inserted in a polymer stabilises to a non-zero value. However, this time the "equilibrium" is lower in the case with switching than in the constantly fast mode. This can be explained by the fact that the large number of polymers in the system compete for the same order of magnitude of free monomers (themselves also captured by profilins) and so the pool of free monomers is constantly kept at a low level, which impedes the elongation of the simple polymers. In this case, mode switching thus does not allow a better use of the different pools of monomers.

Let us now consider the polymer lengths in our large population limit to see what is the impact of mode switching on their distribution. Figure 5 presents the evolution of the distribution of the polymer lengths with the same set of parameters as in Figure 4. We start with the same proportion as previously, $33,3\%$ of free monomers, $33,3\%$ of G-actin/profilin complexes and $33,4\%$ of monomers included in a polymer. In Figure 5a we start with polymers of length 30, whereas on Figure 5b we start with polymers of length 3. In each graph, we plot the distribution of the polymer lengths obtained in Theorem 1.2 at different times and in dashed line we plot the

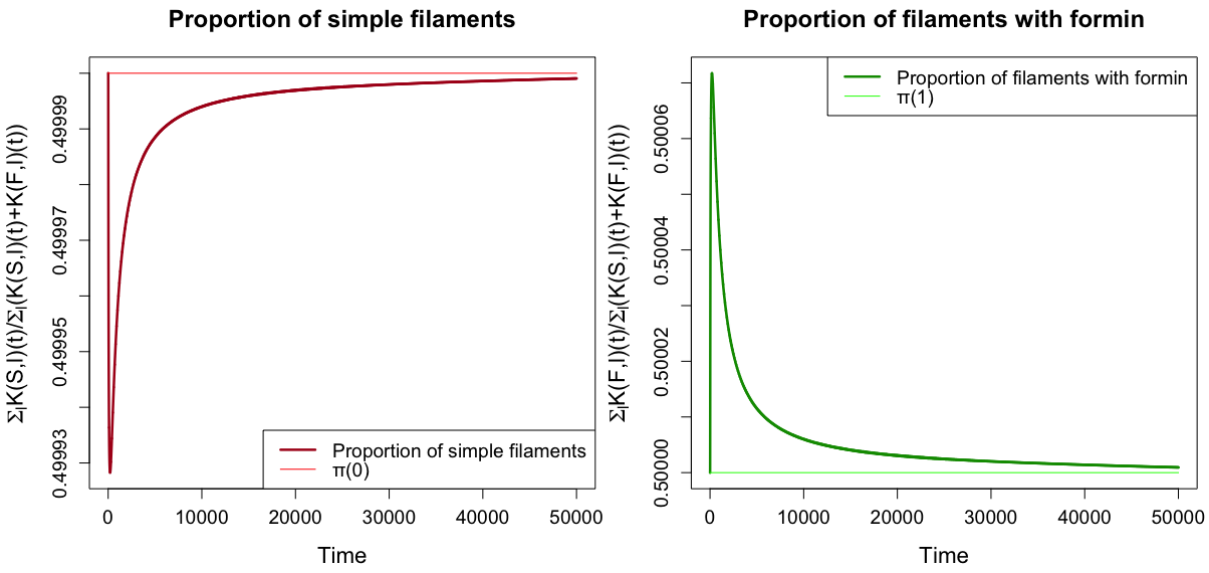


Figure 6: Convergence of the proportion of simple polymers (*resp.*, polymers bound with a formin) towards $\pi_M(0)$ (*resp.*, $\pi_M(1)$) with $dt = 0.001$, $\lambda^+ = 10$, $\lambda^- = 2$, $\lambda_F^+ = 100$, $\Phi_P = 10$, $\Phi_F^+ = 1$ and $\Phi_F^- = 1$, starting with 33% of free monomers, 33% of G-actin/profilin complexes and 34% of monomers included in polymers. Initially, we start with 50% of simple polymers and 50% of polymers bound with a formin, all of length 30.

distribution at which the system seems to stabilise (in the simulations) when we assume that all polymers are blocked in fast mode.

First, comparing Figure 5a and Figure 5b, we see that the distribution in length depends on the initial condition. Indeed, if the polymers are initially shorter, at later times they are globally shorter but more numerous. In contrast, if we start with longer polymers, at later times they are longer, the variance in polymer length is larger but there are fewer polymers.

Second, let us compare the distributions of the polymer lengths at different times, the distribution after stabilisation (in red) and the distribution assuming that all polymers remain in fast mode (red dashed line) in Figure 5a. As in Figure 4, we see that the proportion of polymers is larger when the polymers are blocked in fast mode. But we can also observe that the polymer lengths are globally shorter than when polymers switch modes. We can therefore conclude that mode switching allows to have longer polymers but limits the proportion of monomers included in polymers, and thus it limits the number of polymers.

Although we are not able to prove analytically that, in the limiting process obtained in Theorem 1.2, the proportion of polymers bound with a formin approaches the stationary probability $\pi_M(1)$ that a single polymer is in "fast" mode (in the terminology of the single polymer model of Section 1.2), we checked this intuition by simulation. In Figure 6, we plot the evolution of the quantities

$$\frac{\sum_{l=1}^{+\infty} k_t(S,l)}{\overline{F^\infty}(t)} \qquad \text{and} \qquad \frac{\sum_{l=1}^{+\infty} k_t(F,l)}{\overline{F^\infty}(t)},$$

where as earlier

$$\overline{F^\infty}(t) = \sum_{l=1}^{+\infty} k_t(S,l) + \sum_{l=1}^{+\infty} k_t(F,l)$$

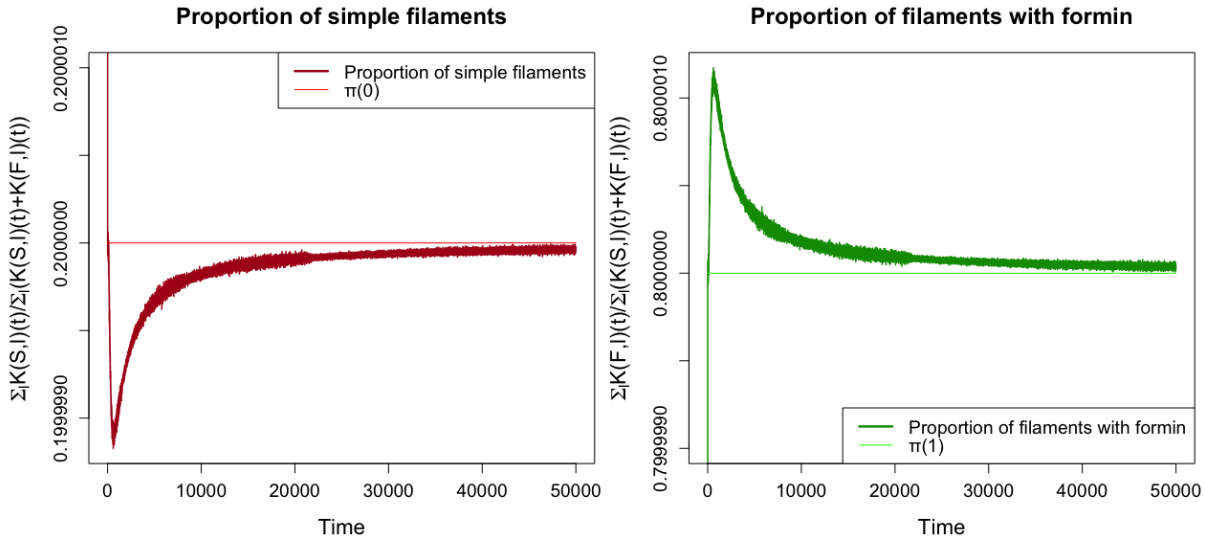


Figure 7: Convergence of the proportion of simple polymers (*resp.*, polymers bound with a formin) towards $\pi_M(0)$ (*resp.*, $\pi_M(1)$) with $dt = 0.001$, $\lambda^+ = 10$, $\lambda^- = 2$, $\lambda_F^+ = 100$, $\Phi_P = 10$, $\Phi_F^+ = 8$ and $\Phi_F^- = 2$, starting with 33% of free monomers, 33% of G-actin/profilin complexes and 34% of monomers included in polymers. Initially, we start with 50% of simple polymers and 50% of polymers with formin, all of length 30.

is the total amount of "real" polymers (thus excluding monomers and complexes) in the system at time t. The parameters and initial state are the same as before. Since $\Phi_F^+ = \Phi_F^-$, in this case we have $\pi_M(0) = 0.5 = \pi_M(1)$ and the initial proportion of polymers bound with a formin is also 0.5. As we see in Figure 6, after an initial very small shift (whose origin is not clear), the proportion of polymers converges again to the value 0.5, which confirms our intuition.

Using this approximation, the system of ODE proposed in Corollary 3.2 can be rewritten for large times as

$$\begin{cases} \frac{d}{dt}k_t(M,1) &= \lambda^-\overline{F^\infty}(t) - \lambda^+\pi_M(0)\overline{F^\infty}(t)^2 k_t(M,1) - \Phi_P k_t(M,1) \\ \frac{d}{dt}m_{inF}(t) &= \lambda^+\pi_M(0)\overline{F^\infty}(t)^2 k_t(M,1) + \lambda_F^+\pi_M(1)\overline{F^\infty}(t)^2 k_t(P,1) - \lambda^-\overline{F^\infty}(t) \\ \frac{d}{dt}k_t(P,1) &= \Phi_P k_t(M,1) - \lambda_F^+\pi_M(1)\overline{F^\infty}(t)^2 k_t(P,1). \end{cases}$$

If we compare this system to the limiting evolution obtained in (28) for a single polymer, we observe that in the above, the elongation rates (spontaneous or with a formin) are multiplied by the *square* of the amount of polymers $\overline{F^\infty}$. Therefore, it appears that elongation is more slowed down by the competition between polymers than the other transitions.

Since in Figure 6 the initial proportion of polymers bound with a formin was set to be equal to its (supposed) equilibrium value, we explored what happens in other cases by choosing a different set of parameters. Figure 7 displays the same quantities as Figure 6 but with the following set of parameters: $\lambda^+ = 10$, $\lambda^- = 2$, $\lambda_F^+ = 100$, $\Phi_P = 10$, $\Phi_F^+ = 8$ and $\Phi_F^- = 2$. We start again with the same initial state (33% of free monomers, 34% of monomers included in a polymer, 33% of monomers in G-actin/profilin complexes, 50% of polymers being simple, 50% being associated with a formin and all polymers being of length 30). By (3), this time we have $(\pi_M(0), \pi_M(1)) = (0.2, 0.8)$ and Figure 7 confirms again the suspected convergence (with another initial overshoot by a very small amount that we cannot explain).

Acknowledgements

A.V.G. and A.V. were supported by the chaire program Modélisation Mathématique et Biodiversité of Veolia Environnement–École Polytechnique–Museum National d'Histoire Naturelle–FX.

References

[1] Aldous, D. Stopping times and tightness. *The Annals of Probability 6*, 2 (1978), 335–340.

[2] Bansaye, V., and Vatutin, V. On the survival probability for a class of subcritical branching processes in random environment. *Bernoulli 23*, 1 (2017), 58–88.

[3] Benaïm, M., and Lobry, C. Lotka–volterra with randomly fluctuating environments or "How switching between beneficial environments can make survival harder". *The Annals of Applied Probability 26*, 6 (2016), 3754–3785.

[4] Billingsley, P. *Convergence of probability measures*, second ed. Wiley Series in Probability and Statistics, John Wiley and Sons Inc., 1999.

[5] Bindschadler, M. Modeling actin dynamics. *Wiley interdisciplinary reviews. Systems biology and medicine 2*, 4 (July 2010), 481–488.

[6] Bindschadler, M., Osborn, E. A., Dewey, C. F., and McGrath, J. L. A mechanistic model of the actin cycle. *Biophysj 86*, 5 (May 2004), 2720–2739.

[7] Burke, T. A., Christensen, J. R., Barone, E., Suarez, C., Sirotkin, V., and Kovar, D. R. Homeostatic actin cytoskeleton networks are regulated by assembly factor competition for monomers. *Current biology : CB 24*, 5 (Mar. 2014), 579–585.

[8] Civelekoglu, G., and Edelstein-Keshet, L. Modelling the dynamics of F-actin in the cell. *Bulletin of Mathematical Biology 56*, 4 (1994), 587–616.

[9] Edelstein-Keshet, L., and Ermentrout, G. B. Models for the length distributions of actin filaments: I. Simple polymerization and fragmentation. *Bulletin of Mathematical Biology 60*, 3 (May 1998), 449–475.

[10] Ermentrout, G. B., and Edelstein-Keshet, L. Models for the length distributions of actin filaments: II. Polymerization and fragmentation by gelsolin acting together. *Bulletin of Mathematical Biology 60*, 3 (May 1998), 477–503.

[11] Ethier, S. N., and Kurtz, T. G. Markov Processes: Characterization and Convergence. *John Wiley & Sons Inc.* (1986).

[12] Hu, J., and Othmer, H. G. A theoretical analysis of filament length fluctuations in actin and other polymers. *Journal of Mathematical Biology 63*, 6 (2011), 1001–1049.

[13] Kovar, D. R. Molecular details of formin-mediated actin assembly. *Current Opinion in Cell Biology 18*, 1 (Feb. 2006), 11–17.

[14] Kovar, D. R., Harris, E. S., Mahaffy, R., Higgs, H. N., and Pollard, T. D. Control of the Assembly of ATP- and ADP-Actin by Formins and Profilin. *CELL 124*, 2 (Jan. 2006), 423–435.

[15] Lecuit, T., and Lenne, P.-F. Cell surface mechanics and the control of cell shape, tissue patterns and morphogenesis. *Nature reviews Molecular cell biology 8*, 8 (2007), 633.

[16] Mogilner, A., and Edelstein-Keshet, L. Regulation of actin dynamics in rapidly moving cells: a quantitative analysis. *Biophysj 83*, 3 (Sept. 2002), 1237–1258.

[17] Mohapatra, L., Lagny, T. J., Harbage, D., Jelenkovic, P. R., and Kondev, J. The limiting-pool mechanism fails to control the size of multiple organelles. *Cell systems 4*, 5 (2017), 559–567.

[18] Pfaffelhuber, P., and Popovic, L. How spatial heterogeneity shapes multiscale biochemical reaction network dynamics. *Journal of the Royal Society Interface 12*, 104 (2015), 20141106.

[19] Pollard, T. D., and Berro, J. Mathematical models and simulations of cellular processes based on actin filaments. *Journal of Biological Chemistry 284*, 9 (2009), 5433–5437.

[20] Rebolledo, R. Sur l'existence de solution à certains problèmes de semimartingales. *CR Acad Sci Paris 290* (1980).

[21] Robert, P., and Sun, W. On the asymptotic distribution of nucleation times of polymerization processes. *arXiv preprint arXiv:1712.08347* (2017).

[22] Rotty, J. D., Wu, C., Haynes, E. M., Suarez, C., Winkelman, J. D., Johnson, H. E., Haugh, J. M., Kovar, D. R., and Bear, J. E. Profilin-1 Serves as a Gatekeeper for Actin Assembly by Arp2/3-Dependent and -Independent Pathways. *Developmental Cell 32*, 1 (Jan. 2015), 54–67.

[23] Suarez, C., Carroll, R. T., Burke, T. A., Christensen, J. R., Bestul, A. J., Sees, J. A., James, M. L., Sirotkin, V., and Kovar, D. R. Profilin Regulates F-Actin Network Homeostasis by Favoring Formin over Arp2/3 Complex. *Developmental Cell 32*, 1 (Jan. 2015), 43–53.

[24] Suarez, C., and Kovar, D. R. Internetwork competition for monomers governs actin cytoskeleton organization. *Nature Reviews Molecular Cell Biology 17*, 12 (2016), 799.

[25] Yarmola, E. G., and Bubb, M. R. How depolymerization can promote polymerization: the case of actin and profilin. *BioEssays 31*, 11 (Nov. 2009), 1150–1160.

[26] Yarmola, E. G., Dranishnikov, D. A., and Bubb, M. R. Effect of Profilin on Actin Critical Concentration: A Theoretical Analysis. *Biophysj 95*, 12 (Dec. 2008), 5544–5573.