On The Asymptotic Distribution of Nucleation Times of Polymerization Processes

SUN Wen

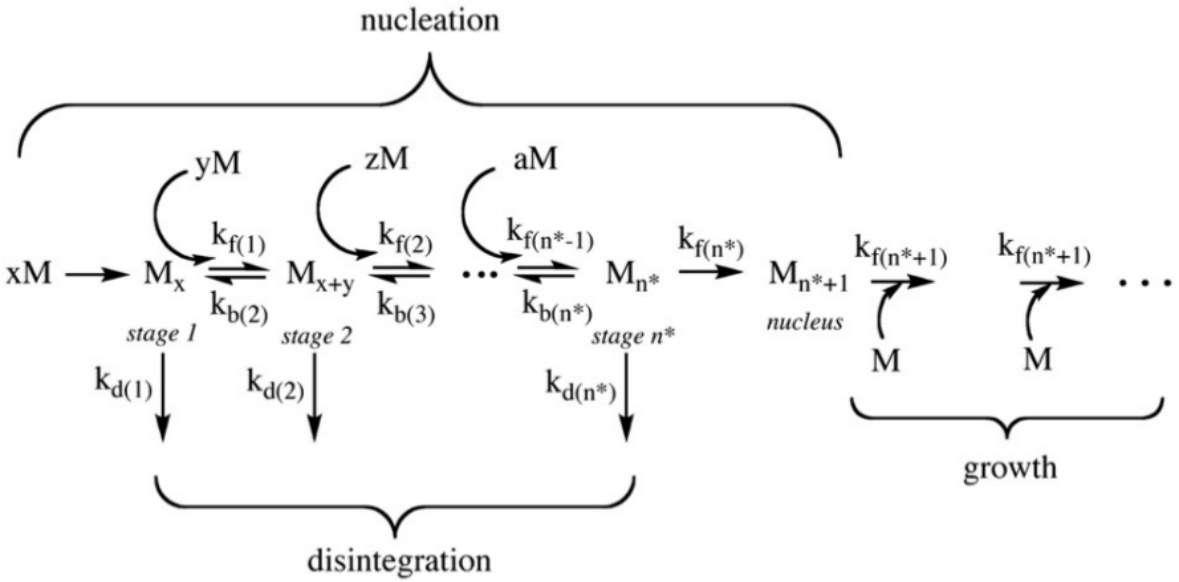
Joint work with Philippe Robert

Les probabilités de demain, Paris, May 2018

Overview

Polymerization & Nucleation

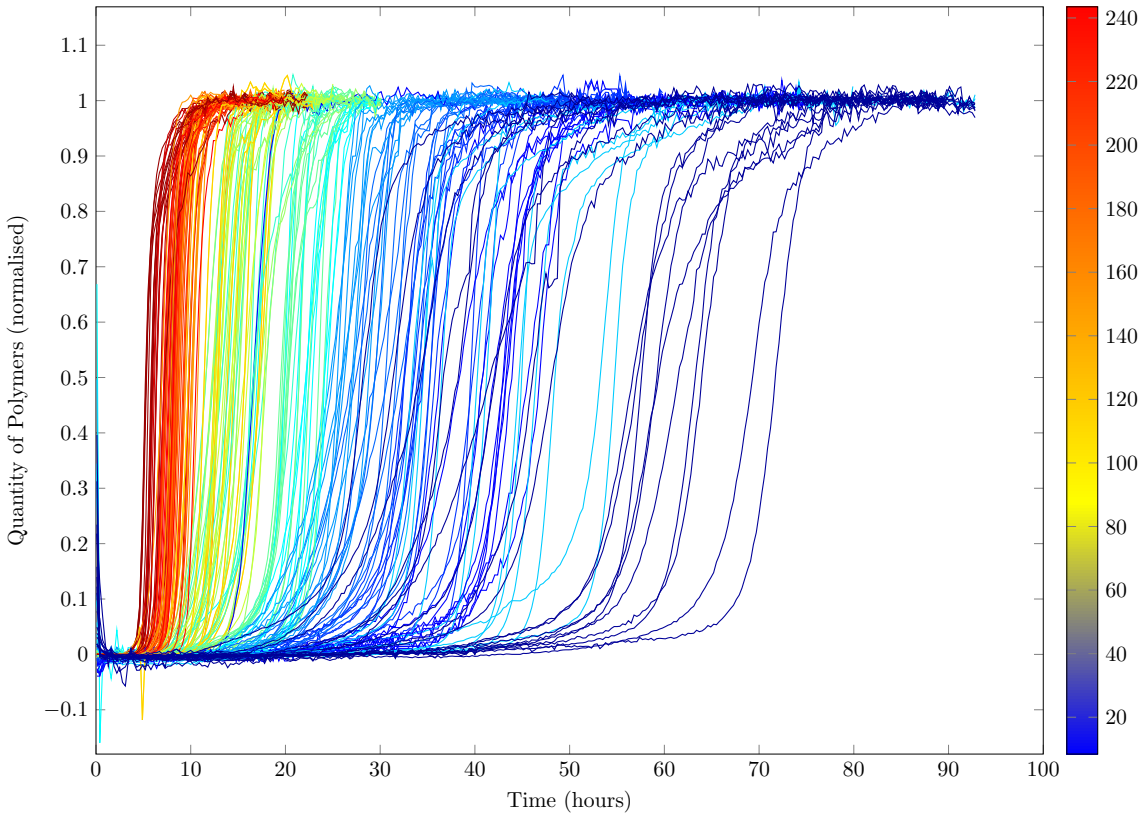
$$\text{small particles} \underset{\textit{decomposition}}{\overset{\textit{synthesis}}{\rightleftharpoons}} \text{Big (stable) clusters.}$$



- Physics: Aerosols...
- Chemistry: Polymers/monomers
- Biology: Protein/Peptide v.s. amino acid monomers

Figure: Flyvbjerg, Jobs, and Leibler's model (96' PNAS) for the self-assembly of microtubules, retrieved from Morris et al. (09' Biochimica et Biophysica Acta)

Experiments: large variability in nucleation



Observations:

- sharp curve;
- huge variance in time.

Figure: Experiments for the evolution of polymerized mass. From data published in Xue et al.(08' PNAS).

Goals of our study

- Explain **sharp phase transition** in nucleation;
- Explain **high variance** of the transition moment.

Literature: coagulation and fragmentation models

- Particles are identified by their sizes.
- Reactions: for $m = \sum_{i=1}^{n} m_i$,

$$\begin{cases} (m_1)+(m_2)+\ldots+(m_n) \to (m) & (\text{coagulation}), \\ (m) \to (m_1)+(m_2)+\ldots+(m_n) & (\text{fragmentation}), \end{cases}$$

- Binary reaction: **Smoluchowski Model**

$$(i) + (j) \underset{F(i,j)}{\overset{K(i,j)}{\rightleftharpoons}} (i+j).$$

where $(F(i,j))$, $(K(i,j))$ are reaction rates.

Literature: coagulation and fragmentation models

- Deterministic studies:
  Oosawa et al. (75), Ball et al. (86'), Penrose (89',08'),
  Jabin et al. (03'), Niethammer (04') ...
- Stochastic studies:
  Jeon (98'), Durrett et al. (99'), Norris (99'), Ranjbar et al. (10'),
  Bertoin (06',17'), Calvez et al. (12'), Sun (18') ...
- Survey:
  Aldous (99'), Hingant & Yvinec (16')

Can not explain the high variance observed in the experiments! (CLT is not enough!)

Model with the nucleus

- Reaction:

$$\begin{cases} (1)+(k) & \xrightarrow{\kappa_+^k} (k+1), \\ (k) & \xrightarrow{\kappa_-^{k,a}} (a_1)+(a_2)+\cdots+(a_p),\ \forall p\geq 2,\ a_1+\cdots+a_p=k, \end{cases}$$

- Critical Nucleus size: n_c
- Polymers larger than the nucleus are more stable than the smaller polymers: $\forall s < n_c < \ell$,

$$\frac{\kappa_-^s}{\kappa_+^s} \gg \frac{\kappa_-^\ell}{\kappa_+^\ell}.$$

where $\kappa_-^k = \sum_a \kappa_-^{k,a}$.

Assumptions & Markovian description

- Only monomers at $t = 0$ with total mass N;
- **Scaling assumption:** for two positive sequences (λ_k), (μ_k) and $\mu > 0$,

$$\kappa_+^k = \lambda_k, \qquad \kappa_-^k = \begin{cases} \mathbf{N}\ \mu_{\mathbf{k}}, & \text{if } \mathrm{k}<\mathrm{n_c} \\ \mu, & \text{if } \mathrm{k}\geq\mathrm{n_c} \end{cases}$$

- $U_k^N(t) :=$ number of polymers of size k at time t;
- Markov process $(U_k^N(t), k \in \mathbb{N})$ with generator

$$\Omega_N(f)(u) = \sum_{k=1}^{+\infty} \lambda_k u_k \frac{u_1}{N} \left[f(u+e_{k+1}-e_k-e_1)-f(u)\right]$$

$$+\sum_{k=2}^{+\infty} \left(N\mu_k \mathbb{1}_{\{k<n_c\}} + \mu \mathbb{1}_{\{k\geq n_c\}}\right) u_k \int_{\mathcal{S}_k} \left[f(u+y-e_k)-f(u)\right] \nu_k(\mathrm{d}y)$$

where (ν_k) are fragmentation measures and $(\mathcal{S}_k)$ are the set of all possible fragmentations.

Mathematical Interpretation

- **Lag time:** for any fraction $\delta \in (0,1)$,

$$L_\delta^N := \inf\{t \geq 0 : \sum_{k \geq n_c} k U_k^N(t) \geq \delta N\}.$$

- Observations in terms of lag time:

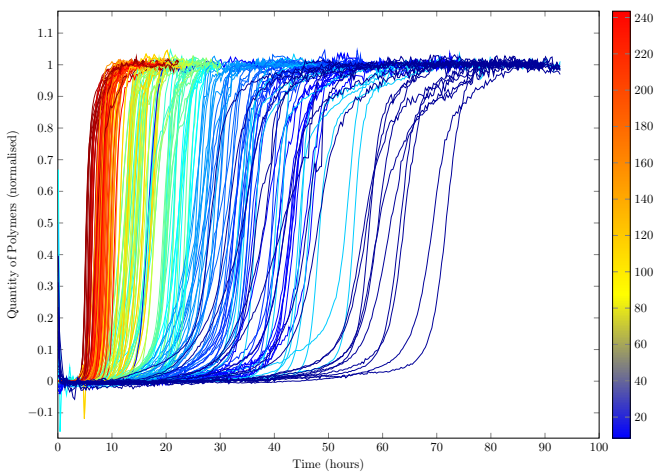


Figure: Xue et al.(08' PNAS).

- sharp phase transition: for any $\delta_1, \delta_2 \in (0,1)$,

$$L_{\delta_1}^N \sim L_{\delta_2}^N$$

- high variance:

$$\mathcal{O}\left(\sqrt{\mathrm{Var}(L_\delta^N)}\right) \sim \mathcal{O}\left(\mathbb{E}(L_\delta^N)\right)$$

Main Results (for $n_c \geq 3$)

- The moment of the first nucleus:

$$T^N := \inf\{t \geq 0 : U_{n_c}^N(t) = 1\}.$$

With high probability, for any $\delta \in (0, \delta_0)$,

$$\boxed{\mathbf{L}_\delta^{\mathbf{N}} \sim \mathcal{O}\left(\mathbf{T}^{\mathbf{N}} + \mathbf{log}(\mathbf{N})\right).}$$

- For the convergence in probability

$$\lim_{N\to\infty} \left(\frac{T^N}{N^{n_c-3}}\right) = E_\rho,$$

where E_ρ is an **exponential random variable** with parameter ρ only depends on $(\lambda_k, \mu_k, k \leq n_c - 1)$.

Main Results (for $n_c \geq 3$)

- The moment of the first nucleus:

$$T^N := \inf\{t \geq 0 : U_{n_c}^N(t) = 1\}.$$

With high probability, for any $\delta \in (0, \delta_0)$,

Therefore, $\boxed{\mathbf{L}_\delta^{\mathbf{N}} \sim \mathcal{O}\left(\mathbf{E}_\rho \mathbf{N}^{\mathbf{n_c}-3} + \mathbf{log(N)}\right).}$

- For the convergence in probability

$$\lim_{N\to\infty} \left(\frac{T^N}{N^{n_c-3}} \right) = E_\rho,$$

where E_ρ is an **exponential random variable** with parameter ρ only depends on $(\lambda_k, \mu_k, k \leq n_c - 1)$.

Main Results (for $n_c \geq 3$)

- The moment of the first nucleus:

$$T^N := \inf\{t \geq 0 : U_{n_c}^N(t) = 1\}.$$

With high probability, for any $\delta \in (0, \delta_0)$,

If $n_c > 3$, $\boxed{\dfrac{\mathbf{L}_\delta^{\mathbf{N}}}{\mathbf{N}^{\mathbf{n_c}-\mathbf{3}}} \sim \mathcal{O}(\mathbf{E}_\rho)}$ $\longleftarrow$ **Not depends on δ & Large var**

- For the convergence in probability

$$\lim_{N\to\infty} \left(\frac{T^N}{N^{n_c-3}} \right) = E_\rho,$$

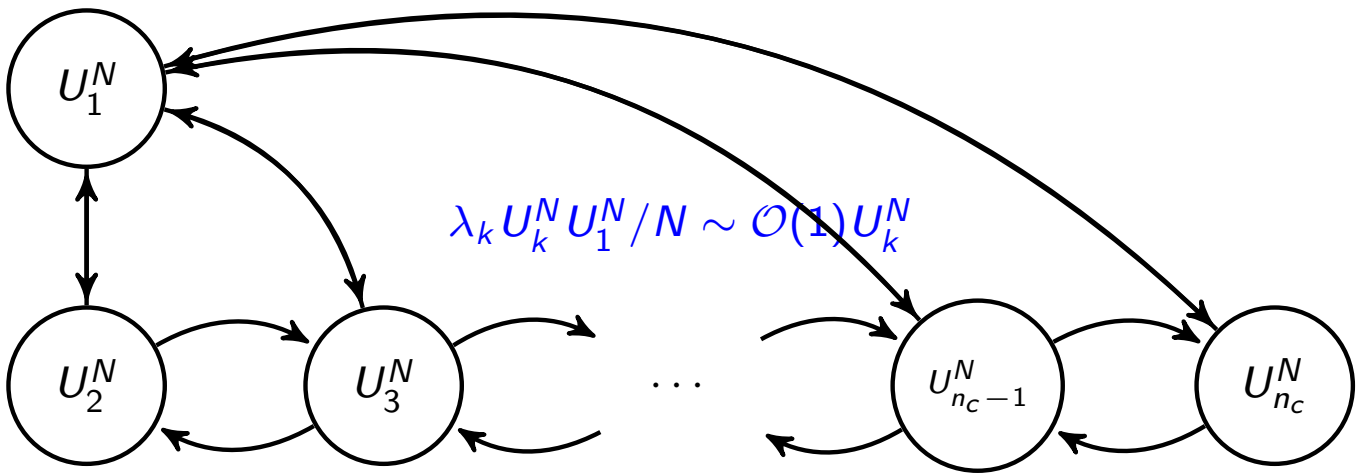
where E_ρ is an **exponential random variable** with parameter ρ only depends on $(\lambda_k, \mu_k, k \leq n_c - 1)$.

Sketch of proofs (Step I)

Before T^N, $U_k^N(t) \equiv 0$, for all $k > n_c + 1$.
A simple example, Becker-Döring reactions:

$$(1) + (k) \underset{N\mu_{k+1}}{\overset{\lambda_k}{\rightleftharpoons}} (k+1).$$

fast process



slow process

Sketch of proofs (Step I)

- Distribution of T^N only depends on the fast-slow system $(U_1^N(t), \ldots, U_{n_c}^N(t))$.
- Study the dynamic on the very large time interval $[0, N^{n_c-3}t]$ by using marked Poisson point processes;
- Main Difficulties:
  - very large fluctuations
    (Time scale N^{n_c-3} *v.s.* Space scale N).
  - multi-dimensional stochastic averaging system:
    hard to identify the limit of occupation measures
- Techniques: coupling, flow balance equations...
- Proofs work for general fragmentation measures under reasonable conditions.

Sketch of proofs (Step II)

After time T^N, by coupling, number of stable polymers

$$(U^N_{n_c}(t), U^N_{n_c+1}(t), \dots)$$

could be lower bounded by a supercritial branching process.

- The lag time of the branching process is less than $K \log N$ with probability $p_0 > 0$.
- Therefore, stochastically

$$T^N \le L^N_\delta \le \sum_{i=1}^{G_{p_0}} (T^N_i + K \log N)$$

where G_{p_0} is a geometric random variable.

Future work

- Experiments: fragmentation rates are more likely sublinear for smaller polymers, *i.e.*, for k small,

$$\kappa_-^k = \mathcal{O}\left(N^\alpha\right), \text{ for an } 0 < \alpha < 1.$$

  See the biology review Morris et al. (09').

- In the general case $\kappa_-/\kappa_+ \sim \phi(N)$, the nucleation time should be $\mathcal{O}(\phi(N)^{n_c-2}/N)$.

- Nucleation in a multi-type polymers environment.

References

- (Eugène, Xue, Robert & Doumic, 16') Insights into the variability of nucleated amyloid polymerization by a minimalistic model of stochastic protein assembly. (Journal of Chemical Physics)
- (Doumic, Eugène & Robert, 16') Asymptotics of stochastic protein assembly models. (SIAM Journal on Applied Mathematics)
- (Robert & Sun, 17') On the Asymptotic Distribution of Nucleation Times of Polymerization Processes. (arXiv:1712.08347)

Thank you!