Quasi-stationary distribution and metastability for the stochastic Becker-Döring model[*]

Erwan Hingant[†¶]   Romain Yvinec[‡§¶]

Abstract

We study a stochastic version of the classical Becker-Döring model, a well-known kinetic model for cluster formation that predicts the existence of a long-lived metastable state before a thermodynamically unfavorable nucleation occurs, leading to a phase transition phenomena. This continuous-time Markov chain model has received little attention, compared to its deterministic differential equations counterpart. We show that the stochastic formulation leads to a precise and quantitative description of stochastic nucleation events thanks to an exponentially ergodic quasi-stationary distribution for the process conditionally on nucleation has not yet occurred.
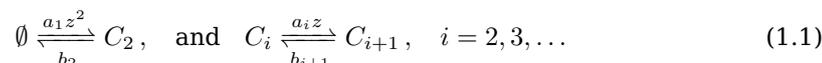
Keywords: quasi-stationary distribution; metastability; exponential ergodicity; Becker-Döring; nucleation; phase transition.
MSC2020 subject classifications: 82C26; 60J27.
Submitted to ECP on August 15, 2020, final version accepted on June 16, 2021.

1 Introduction

The Becker-Döring model is a kinetic model for phase transition phenomenon represented schematically by the reaction network

$$\emptyset \underset{b_2}{\overset{a_1 z^2}{\rightleftharpoons}} C_2\,, \quad \text{and} \quad C_i \underset{b_{i+1}}{\overset{a_i z}{\rightleftharpoons}} C_{i+1}\,, \quad i = 2, 3, \ldots \tag{1.1}$$

We assume an infinite reservoir of monomers, clusters of size 1, represented in (1.1) by $\emptyset$. The parameter z represents the *fixed* concentration of monomers and will play a key role in the sequel. A cluster of size $i \geq 2$, whose population number is represented in (1.1) by C_i, may lengthen to give rise to a cluster of size $i + 1$ at rate $a_i z$, and a cluster of size $i \geq 3$ may shorten to give rise to a cluster of size $i - 1$ at rate b_i. The rate of apparition of a new cluster of size 2 is $a_1 z^2$ (without loss of generality) and a cluster of size 2 may disappear at rate b_2. All parameters are positives.

[†]Departamento de Matemática, Universidad del Bío-Bío, Chile. E-mail: ehingant@ubiobio.cl
[‡]CNRS, IFCE, INRAE, Université de Tours, PRC, F-37380, Nouzilly, France.
  E-mail: romain.yvinec@inrae.fr
[§]Université Paris-Saclay, Inria, Inria Saclay-Île-de-France, 91120, Palaiseau, France.
[¶]Associate member at Cogitamus laboratory.

The Becker-Döring (BD) model goes back to the seminal work *"Kinetic treatment of nucleation in supersaturated vapors"* in [1]. Since then, the model met very different applications ranging from physics to biology. From the mathematical point of view, this model received much more attention in the deterministic literature than the probabilistic one. We refer to our review [4] for historical comments and detailed literature review on theoretical results from the deterministic side. See also [5, 10] for recent results on functional law of large number and central limit theorem.

The model was initially designed to explain critical phase condensation phenomena where macroscopic droplets self-assemble and segregate from an initially supersaturated homogeneous mixture of particles, at a rate that is exponentially small in the excess of particles. This led to important applications in kinetic nucleation theory [9, 5]. Mathematical studies in the 90's showed that (in the deterministic context), departing from certain initial conditions, the size distribution of clusters reaches quickly a metastable configuration composed of "small" clusters, and remains arbitrary close to that state for a very large time, before it converges to the true stationary solution that leads to "infinitely large" clusters (interpreted as droplets) [8, 9].

Our objective in this note is to re-visit the metastability theory in Becker-Döring model in terms of quasi-stationary distribution (QSD) for the associated continuous-time Markov chain. We prove exponential convergence towards a QSD for the BD model conditioned on the event that large clusters have not yet appeared. We prove furthermore that the convergence rate towards the QSD can be faster than the rate of apparition of (sufficiently) large clusters. Quantitative results are obtained thanks to a surprisingly simple analytical formula for the QSD, that provides also an exact rate of apparition of stable large clusters, consistently with the original heuristic development of Becker and Döring.

Outline: Sec. 2: Construction of the BD model. Sec. 3: Collection of few (known) results. Sec. 4: Exponential decay in total variation towards the stationary measure. Sec. 5: Similar result conditionally on no clusters larger than n are formed (QSD). Sec. 6: Estimate on the time for the first cluster larger than n to appear. Sec. 7: Interpretation of the QSD as a long-lived metastable state when n is the critical nucleus size.

Notation: We denote by $\mathbb{N}_i$ the set of non-negatives integers greater or equal to i, $[\![\ ]\!]$ for integers interval. For a set A, $\#A$ its cardinality, $\mathbf{1}_A$ the indicator function on it. $\mathbf{1}$ and $\mathbf{0}$ the constant functions equal to 1 and 0. For two numbers a, b, their minimum is $a \wedge b$. For probability measures μ and ν on a countable state space $\mathcal{S}$, the total variation distance is

$$\|\mu - \nu\| = \frac{1}{2} \sum_{x \in \mathcal{S}} |\mu(x) - \nu(x)| = \inf_{\gamma \in \Gamma} \int_{S \times S} \mathbf{1}_{x \neq y}\, \gamma(dx, dy)\,,$$

where Γ is the set of probability measures on $\mathcal{S} \times \mathcal{S}$ with marginals μ and ν. $\mathbf{E}$ (resp. $\mathbf{E}_\mu$) denotes the expectation with respect to the usual probability measure $\mathbf{P}$ (resp. μ). We set $\mathcal{E} = \ell^1(\mathbb{N}_2, \mathbb{N}_0)$ the space of summable $\mathbb{N}_0$-valued sequences indexed by $\mathbb{N}_2$.

2 The model

The stochastic Becker-Döring (BD) process is a continuous-time Markov chain $\{\mathbf{C}(t),\ t \geq 0\}$ on the countable state space $\mathcal{E}$ with infinitesimal generator $\mathcal{A}$, given for all ψ with finite support on $\mathcal{E}$ and $C \in \mathcal{E}$, by

$$\mathcal{A}\psi(C) = \sum_{i=1}^{+\infty} \Big(a_i z C_i [\psi(C + \Delta_i) - \psi(C)] + b_{i+1} C_{i+1} [\psi(C - \Delta_i) - \psi(C)] \Big)$$

with the convention $C_1 = z$, $\Delta_1 = \mathbf{e}_2$ and $\Delta_i = \mathbf{e}_{i+1} - \mathbf{e}_i$, for each $i \geq 2$, where $\{\mathbf{e}_2, \mathbf{e}_3, \dots\}$ denotes the canonical basis of $\mathcal{E}$ namely, $e_{i,k} = 1$ if $k = i$ and 0 otherwise. In this view, the

BD process is an infinite vector $\mathbf{C}(t) = (C_2(t), C_3(t), \ldots)$ that keeps track of the number of clusters $C_i(t)$ of any size $i \geq 2$ at any time $t \geq 0$.

We shall however use a different approach, modeling explicitly the size of each individual cluster, as a Birth-Death process on $\mathbb{N}_1$ with a birth rate $a_i z$ and a death rate b_i, for $i \geq 2$. Although in the original BD process a cluster can only have a size $i \geq 2$, we augment the size state space with $i = 1$ as an absorbing state of the Birth-Death process, in order to keep track of each cluster of size 2 going back to the infinite reservoir $\emptyset$. Such approach will be called thereafter the *particle description* of the BD process.

We first explain how we can give an equivalent description of the number of clusters of each size in terms of this *particle description* by explicitly labeling each of them. Given $C = (C_2, C_3, \ldots)$ an $\mathcal{E}$-valued random variable, we denote by $N = \sum_{i=2}^{\infty} C_i$ the total number of clusters. By construction $N < \infty$ almost surely (*a.s.*). Then, we define $\{X_k\}$ a denumerable collection of random variables on $\mathbb{N}_2$ such that, *a.s.* for each $i \geq 2$,

$$C_i = \# \{k \in [\![1, N]\!] \mid X_k = i\} . \tag{2.1}$$

Note this construction may be achieved by a bijective *labeling* function[1]. The collection $X_1, \ldots, X_N$ stands for the size of the N individual clusters describing $\mathbf{C}$. For technical purpose, explained below, we let $X_k = 2$ for $k > N$.

We now go through the pathwise construction of the *particle description*. To that, we consider a sufficiently large probability space $(\Omega, \mathcal{F}, \mathbf{P})$. For $\mathbf{C}(0) = (C_2(0), C_3(0), \ldots)$ an $\mathcal{E}$-valued random variable distributed according to an initial probability distribution Π^{in}, we denote $N^{\mathrm{in}} = \sum_{i=2}^{\infty} C_i(0)$ and $X_1(0), X_2(0), \ldots$ the *particle description* defined through (2.1). Then, we introduce:

• $N_1, N_2, \ldots$ a denumerable family of independent Poisson point measures with intensity the Lebesgue measure $ds\,du$ on $\mathbb{R}_+^2$.

• $S_1, S_2, \ldots$ a collection of random times such that the $S_k - S_{k-1}$ are independent exponential random variables of parameter $a_1 z^2$, independent from the above Poisson point measures as well, with $S_0 = 0$.

Finally, we consider the denumerable collection of stochastic processes $X_1, X_2, \ldots$ (denoted by $\{X_k\}$ hereafter), on $\mathbb{N}_1$ solution of the stochastic differential equations, for all $t \geq 0$ and $k \geq 1$,

$$X_k(t) = X_k(0) + \sum_{i=2}^{\infty} \int_0^t \int_{\mathbb{R}^+} \mathbf{1}_{s > S_{k-N^{\mathrm{in}}}} \mathbf{1}_{X_k(s^-)=i} \big(\mathbf{1}_{u \leq a_i z} - \mathbf{1}_{a_i z < u \leq a_i z + b_i} \big) N_k(ds, du) , \tag{2.2}$$

where by convention $S_k = 0$ if $k \leq 0$. We finally define the counting process

$$N(t) = N^{\mathrm{in}} + \sum_{k \geq 1} \mathbf{1}_{t \geq S_k} , \tag{2.3}$$

which gives the total number of clusters at time t, that either were already present initially or that arose at random times S_k from $\emptyset$ trough the reaction in (1.1) of rate $a_1 z^2$. Note that for any $t > 0$ and each $k \leq N(t)$, we have $t > S_{k-N^{\mathrm{in}}}$. For such clusters labelled by $k \leq N(t)$, we see from the propensity in (2.2) that $X_k(t) \in \mathbb{N}_1$ evolves through a Birth-Death process with birth rate $a_i z$ and death rate b_i, for $i \geq 2$, and $i = 1$ is an absorbing state. The clusters labelled with $k > N(t)$ are waiting the random arrival times $S_{k-N^{\mathrm{in}}}$ as an entry from the $\emptyset$. They are fixed to size 2 because they will start at such size. This procedure allows us to deal with a fix collection of clusters, while the BD process has a variable number of clusters.

---

[1] A function (that exists) which associates, to each $C \in \mathcal{E}$ such that $N = \sum_{i=2}^{\infty} C_i < \infty$, a unique sequence $(X_1, \ldots, X_N)$ in $\mathbb{N}_2$ satisfying $C_i = \# \{k \in [\![1, N]\!] \mid X_k = i\}$.

The pathwise construction (2.2-2.3) is the *particle description* of the BD process. The number of clusters of size $i \geq 2$ is recovered through

$$C_i(t) = \# \left\{ k \in [\![1, N(t)]\!] \mid X_k(t) = i \right\} .$$

We now justify the well-posedness of this construction. First, we ensured $X_k(0) < \infty$ *a.s.* because the $C_i(0)$'s are integer-valued random variables and belong to $\mathcal{E}$, the sequence $\mathbf{C}(0)$ is *a.s.* equally 0 from a certain range. Thus, local existence of *càdlàg* processes $t \mapsto X_k(t)$ on $\mathbb{N}_1$ solution to (2.2) can classically be obtained inductively. It is clear from (2.2) that each X_k evolves like a Birth-Death process for $t > S_{k-N^{in}}$ (that will be detailed in the next Sec. 3) and are mutually independent conditionally to their initial value. The Reuter's criterion gives a well-known necessary and sufficient condition so that each process X_k is non-explosive, namely

$$\sum_{n=2}^{\infty} Q_n z^n \left(\sum_{k=n}^{\infty} \frac{1}{a_k Q_k z^{k+1}} \right) = \infty \,, \text{ with } Q_1 = 1 \,, \ Q_i = \frac{a_1 a_2 \cdots a_{i-1}}{b_2 \cdots b_i} \,, \ i \geq 2 \,. \qquad \text{(H0)}$$

Now, noticing that $C_i(t) = \sum_{k=1}^{N(t)} \mathbf{1}_i(X_k(t))$, we can prove from standard stochastic calculus that the process $\mathbf{C}$ given by $\mathbf{C}(t) = (C_2(t), C_3(t), \ldots)$ for all $t \geq 0$ has infinitesimal generator $\mathcal{A}$, and being non-explosive under condition (H0), it is the unique regular jump homogeneous Markov chain on $\mathcal{E}$ with infinitesimal generator $\mathcal{A}$ and initial distribution Π^{in}, say the BD process. The proof is left to the reader and follows from classical theory. In the sequel, $\mathbf{C}$ always denote a BD process, and $\mathbf{P}_{\Pi^{in}} \{ \mathbf{C} \in \cdot \}$ its (unique) finite dimensional probability distribution given that $\mathbf{C}(0)$ is distributed according to Π^{in}. We also set by convention $\mathbf{P}_C = \mathbf{P}_{\delta_C}$ for a deterministic $C \in \mathcal{E}$ and we recover $\mathbf{P}_{\Pi^{in}} \{ \cdot \} = \sum_{C \in \mathcal{E}} \mathbf{P}_C \{ \cdot \} \Pi^{in}(C)$.

We end this section with a small list of important algebraic quantities that will be used repeatedly in the sequel, and are classical in the BD model. All of them depend on z and may be infinite according to its value. We let

$$J := \left(\sum_{k=1}^{\infty} \frac{1}{a_k Q_k z^{k+1}} \right)^{-1} \qquad \text{(2.4)} \qquad J_n := \left(\sum_{k=1}^{n} \frac{1}{a_k Q_k z^{k+1}} \right)^{-1} {}_{(n \geq 2)} \qquad \text{(2.7)}$$

$$f_i := J Q_i z^i \sum_{k=i}^{\infty} \frac{1}{a_k Q_k z^{k+1}} \ {}_{(i \geq 1)} \quad \text{(2.5)} \qquad f_i^n := J_n Q_i z^i \sum_{k=i}^{n} \frac{1}{a_k Q_k z^{k+1}} \ {}_{(n \geq 2, \, 1 \leq i \leq n)} \quad \text{(2.8)}$$

$$K := \left(\sum_{k=2}^{\infty} Q_i z^i \right)^{\frac{1}{2}} \qquad \text{(2.6)} \qquad K_n := \left(\sum_{k=2}^{n} Q_i z^i \right)^{\frac{1}{2}} {}_{(n \geq 2)} \qquad \text{(2.9)}$$

3 Behaviour of one cluster

Let X be the continuous-time Markov chain on $\mathbb{N}_1$ with transition rate matrix $(q_{i,j})_{i,j \geq 1}$ whose nonzero entries are

$$q_{i,i+1} = a_i z \,, \ q_{i,i-1} = b_i \,, \ q_{i,i} = -(a_i z + b_i) \,, \qquad i \geq 2 \,. \qquad \text{(3.1)}$$

Remark that $i = 1$ is absorbing in agreement with (1.1): when a cluster size reaches 1, it "leaves the system". We shall assume standard hypotheses in the BD model [8, 9]:

$$\lim_{i \to \infty} b_i / a_i = z_s > 0 \,, \quad \text{and} \quad \lim_{i \to \infty} b_{i+1} / b_i = 1 \,. \qquad \text{(H1)}$$

Hypothesis (H1) then guarantee (H0) for $z \neq z_s$ for the following reason. The convergence of both series

$$\sum_{k=2}^{\infty} Q_k z^k \quad \text{and} \quad \sum_{k=2}^{\infty} \frac{1}{a_k Q_k z^{k+1}} \, ,$$

depends on the value of z. Indeed, z_s is the radius of convergence of the first series while the second series converges for $z > z_s$ and diverges for $z < z_s$. We have a dichotomy in the long time behavior of X related to this value. The case $z < z_s$ is called the *sub-critical case*, for which absorption at state 1 is certain and the expected time of absorption is finite (also called ergodic absorption). The case $z > z_s$ is called the *super-critical case* and absorption at 1 is not certain (also called transient absorption), and the probability to be absorbed at 1 is, according to [6],

$$\lim_{t \to \infty} p_{i1}(t) = J \sum_{k=i}^{\infty} \frac{1}{a_k Q_k z^{k+1}} = \frac{f_i}{Q_i z^i}$$

where $p_{ij}(t) = \mathbf{P} \{X(t) = j \mid X(0) = i\}$ is the probability transition function of X and f_i is defined in (2.5). The limit case $z = z_s$ is somewhat technical and depends more deeply on the shape of the coefficients. It is not considered in this note.

Following [8], a precise long time estimate on transient states can be obtained, under the hypothesis (H1) and

$$\frac{b_{i+1}}{b_i} - 1 = O(i^{-1}) \, , \quad \frac{a_{i+1}}{a_i} - 1 = O(i^{-1}) \, , \quad a_i = O(i) \quad \text{and} \quad \lim_{i \to +\infty} a_i = +\infty \, . \quad \text{(H2)}$$

In such a case, the infinite matrix $(q_{i,j})_{i,j \geq 2}$ in (3.1) is self-adjoint on the Hilbert space H consisting of the real sequences $\mathbf{x} = (x_2, x_3, \ldots)$ whose norm is $\|\mathbf{x}\|_H^2 = \sum_{i=2}^{\infty} \frac{x_i^2}{Q_i z^i}$. We denote by $\langle \cdot, \cdot \rangle_H$ the associated scalar product. It turns that $(q_{i,j})_{i,j \geq 2}$ has a negative maximum eigenvalue $-\lambda$, and the following estimate holds for any $i \geq 2$,

$$\|(p_{ij}(t))_{j \geq 2}\|_H \leq e^{-\lambda t} \|(p_{ij}(0))_{j \geq 2}\|_H = \frac{e^{-\lambda t}}{\sqrt{Q_i z^i}} \, . \quad (3.2)$$

We will also consider the chain X conditioned to remain below a given state $n + 1 \geq 2$. We define the exit time

$$T_n = \inf \left(t \geq 0 \mid X(t) \notin [\![1, n]\!] \right) = \inf \left(t \geq 0 \mid X(t) \geq n + 1 \right) . \quad (3.3)$$

Let Y be the birth-death process defined by $Y(t) = X(t \wedge T_n)$. Hence, Y is absorbed either in 1 or $n + 1$, and the probability to be absorbed at 1 (without visiting state $n + 1$) is, according to [6, p.387],

$$\lim_{t \to +\infty} p_{i1}^n(t) = J_n \sum_{k=i}^{n} \frac{1}{a_k Q_k z^{k+1}} = \frac{f_i^n}{Q_i z^i} \quad (3.4)$$

where $p_{ij}^n(t) = \mathbf{P} \left(Y(t) = j \mid Y(0) = i \right)$ is the probability transition function of Y and f_i^n is defined in (2.8). Clearly

$$\mathbf{P} \{T_n > t \mid X(0) = i\} \geq \lim_{t \to +\infty} \mathbf{P} \{T_n > t \mid X(0) = i\} = \frac{f_i^n}{Q_i z^i} \, . \quad (3.5)$$

Again, in [8], the author shows that the truncated matrix $(q_{i,j})_{i,j=2,\ldots,n}$ is similar to a symmetric one and then there exists $\gamma_n > 0$ such that for each $i = 2, \ldots, n$,

$$\sqrt{\sum_{j=2}^{n} \frac{p_{ij}^n(t)^2}{Q_j z^j}} \leq \frac{e^{-\gamma_n t}}{\sqrt{Q_i z^i}} \, . \quad (3.6)$$

Note that the probability to be absorbed in 1 before time t, $p_{i1}^n(t)$, is monotonously increasing and $\lim_{t\to+\infty} p_{i1}^n(t) = 1 - \lim_{t\to+\infty} p_{i(n+1)}^n(t)$, thus we deduce that

$$\sum_{j=2}^n \mathbf{P}\{Y(t) = j \mid Y(0) = i,\, T_n > t\} = \frac{\sum_{j=2}^n p_{ij}^n(t)}{1 - p_{i(n+1)}^n(t)}$$

$$\leq \frac{e^{-\gamma_n t}}{\sqrt{Q_i z^i}} \left(\sum_{j=2}^n Q_j z^j\right)^{\frac{1}{2}} \frac{Q_i z^i}{f_i^n} = K_n \frac{\sqrt{Q_i z^i}}{f_i^n} e^{-\gamma_n t}$$

where the inequality is obtained thanks to Cauchy-Schwarz inequality and (3.4), and K_n is defined in (2.9). We end this section, noticing that $X(t) = Y(t)$ on $\{T_n > t\}$, with

$$\mathbf{P}\{X(t) = 1 \mid X(0) = i,\, T_n > t\} \geq 1 - \left(K_n \frac{\sqrt{Q_i z^i}}{f_i^n} e^{-\gamma_n t}\right) \wedge 1. \qquad (3.7)$$

4 Phase transition in long-time behaviour

In this section we are concerned with the long-time behaviour of the BD process. Formally the measure Π^{eq}, given by

$$\Pi^{eq}(C) = \prod_{i=2}^\infty e^{-Q_i z^i} \frac{(Q_i z^i)^{C_i}}{C_i!},$$

for all $C \in \mathcal{E}$, satisfies $\mathbf{E}_{\Pi^{eq}}[\mathcal{A}\psi(C)] = 0$ for any function ψ on $\mathcal{E}$ with finite support[2]. Actually, Π^{eq} satisfies the detailed balance condition $a_i z C_i \Pi^{eq}(C) = b_{i+1}(C_{i+1}+1)\Pi^{eq}(C + \Delta_i)$, for all $i \geq 1$ and all $C \in \mathcal{E}$ (with the convention that $C_1 = z$), as a consequence of the relation $a_i Q_i = b_{i+1} Q_{i+1}$. In the sub-critical case, Π^{eq} is a probability measure on $\mathcal{E}$ (indeed $\Pi^{eq}(\mathbb{N}_0^{\mathbb{N}_2}) = 1$ with support in $\mathcal{E}$ because of $\mathbf{E}_{\Pi^{eq}}\left[\sum_{i=2}^\infty C_i\right] = \sum_{i=2}^\infty Q_i z^i < \infty$) and we prove exponential ergodicity towards Π^{eq}. In the super-critical case, Π^{eq} is not a limiting distribution (and $\sum_{i=2}^\infty Q_i z^i = \infty$) but the measure defined by

$$\Pi^{stat}(C) = \prod_{i=2}^\infty e^{-f_i} \frac{(f_i)^{C_i}}{C_i!},$$

for all $C \in \mathcal{E}$, where f_i is given in (2.5), characterizes the long-time behaviour of any finite-dimensional distributions. Note that $\sum_{i=2}^\infty f_i = \infty$ for $z > z_s$ which suggests that the limiting size distribution has a fat-tail (infinite mean number of clusters above any size), interpreted as a phase-transition, see [5] on a related model. From now on, we note $\mathbf{f} = (f_i)_{i\geq2}$, $\mathbf{Q} = (Q_i z^i)_{i\geq2}$ and $\sqrt{\mathbf{Q}} = (\sqrt{Q_i z^i})_{i\geq2}$.

Theorem 4.1. *Under hypotheses* (H1) *and* (H2). *Let Π^{in} be a probability distribution on $\mathcal{E}$ such that*

$$\mathbf{E}_{\Pi^{in}}\left[\langle C, \sqrt{\mathbf{Q}}\rangle_H\right] < \infty. \qquad (4.1)$$

With $\lambda > 0$ introduced in Sec. 3, see (3.2), *we have:*

- *In the sub-critical case ($z < z_s$), for all $t \geq 0$,*

$$\|\mathbf{P}_{\Pi^{in}}\{\mathbf{C}(t) \in \cdot\} - \Pi^{eq}\| \leq R^{in} e^{-\lambda t},$$

with $R^{in} = K(\mathbf{E}_{\Pi^{in}}\left[\langle C, \sqrt{\mathbf{Q}}\rangle_H\right] + \mathbf{E}_{\Pi^{eq}}\left[\langle C, \sqrt{\mathbf{Q}}\rangle_H\right])$ and K is defined in (2.6);

---

[2] In the sequel C in expectation formula always refers to the free variable of integration.

- *In the super-critical case ($z > z_s$), for all $t \geq 0$, and for all $n \geq 2$,*

$$\|\mathbf{P}_{\Pi^{\mathrm{in}}}\{(C_2(t), \ldots, C_n(t)) \in \cdot\} - \Pi^{stat}(\cdot \times \prod_{k=n+1}^{\infty} \mathbb{N}_0)\| \leq R_n^{\mathrm{in}} e^{-\lambda t},$$

with $R_n^{\mathrm{in}} = K_n \mathbf{E}_{\Pi^{\mathrm{in}}}\left[\langle C, \sqrt{\mathbf{Q}}\rangle_H\right] + \|\mathbf{f}\|_H$ and K_n is defined in (2.9).

Not least, remark that $\mathbf{E}_{\Pi^{eq}}\left[\langle C, \sqrt{\mathbf{Q}}\rangle_H\right] = \sum_{i=2}^{\infty} \sqrt{Q_i z^i} < \infty$ for $z < z_s$ and that $\mathbf{f} \in H$ for $z > z_s$ (see [8]). The proof of Theorem 4.1 is based on a coupling argument between two *particle descriptions* with respective initial distribution $\mathbf{C}(0)$ and $\mathbf{0}$. Thus, controlling the behavior of the initial clusters present in $\mathbf{C}(0)$ is a key point, which is provided in the next Lemma.

Lemma 4.2. *Under the hypothesis of Theorem 4.1. Let the collection of processes $\{X_k\}$ be the* particle description *of the BD process $\mathbf{C}$. We have, for each $n \geq 2$,*

$$\mathbf{P}_{\Pi^{\mathrm{in}}}\left\{\forall k \in [\![1, N^{\mathrm{in}}]\!], \ X_k(t) \notin [\![2, n]\!]\right\} \geq 1 - K_n \mathbf{E}_{\Pi^{\mathrm{in}}}\left[\langle C, \sqrt{\mathbf{Q}}\rangle_H\right] e^{-\lambda t}.$$

In particular, for the sub-critical case,

$$\mathbf{P}_{\Pi^{\mathrm{in}}}\left\{\forall k \in [\![1, N^{\mathrm{in}}]\!], \ X_k(t) = 1\right\} \geq 1 - K \mathbf{E}_{\Pi^{\mathrm{in}}}\left[\langle C, \sqrt{\mathbf{Q}}\rangle_H\right] e^{-\lambda t}.$$

Proof. Fix $n \geq 2$. Let $C \in \mathcal{E}$ be deterministic, define $N = \sum_{i=2}^{\infty} C_i$ and $(i_1, \ldots, i_N) \in \mathbb{N}_2^N$ given by the *labeling* function, *e.g.* $\mathbf{C}(0) = C$ and $X_1(0) = i_1, \ldots, X_N(0) = i_N$ satisfy relation (2.1). Since the processes $X_1, \ldots, X_N$ are independent copies of the chain X given in Sec. 3, conditionally on their initial condition, we have

$$\mathbf{P}_C\left\{\forall k \in [\![1, N]\!], \ X_k(t) \notin [\![2, n]\!]\right\} = \prod_{k=1}^{N} \mathbf{P}\left\{X(t) \notin [\![2, n]\!] \mid X(0) = i_k\right\}, \qquad (4.2)$$

for all $t \geq 0$. Thanks to Cauchy–Schwarz inequality and (3.2),

$$\mathbf{P}\left\{X(t) \in [\![2, n]\!] \mid X(0) = i\right\} = \sum_{j=2}^{n} p_{ij}(t) \leq \left(\frac{1}{\sqrt{Q_i z^i}}\right) K_n e^{-\lambda t} \wedge 1.$$

Hence, with (4.2), we have

$$\mathbf{P}_C\left\{\forall k \in [\![1, N]\!], \ X_k(t) \notin [\![2, n]\!]\right\} \geq 1 - K_n e^{-\lambda t} \sum_{k=1}^{N} \frac{1}{\sqrt{Q_{i_k} z^{i_k}}},$$

remarking that $\prod_{i=1}^{N}(1 - x_i \wedge 1) \geq 1 - \sum_{i=1}^{N} x_i$ for any non-negatives $x_1, \ldots, x_N$. Finally, we conclude that

$$\mathbf{P}_{\Pi^{\mathrm{in}}}\left\{\forall k \in [\![1, N^{\mathrm{in}}]\!], \ X_k(t) \notin [\![2, n]\!]\right\} \geq 1 - K_n e^{-\lambda t} \sum_{C \in \mathcal{E}} \sum_{k=1}^{N} \frac{1}{\sqrt{Q_{i_k} z^{i_k}}} \Pi^{\mathrm{in}}(C)$$

$$= 1 - K_n e^{-\lambda t} \mathbf{E}_{\Pi^{\mathrm{in}}}\left[\sum_{i=2}^{\infty} \frac{\#\left\{k \in [\![1, N^{\mathrm{in}}]\!] \mid X_k(0) = i\right\}}{\sqrt{Q_i z^i}}\right],$$

and the proof ends. $\qquad\qquad\square$

We now compare in total variation distance, by a coupling argument, both processes starting from $\mathbf{0}$ and Π^{in}. In the next Lemma we show that they are exponentially close in large time.

Lemma 4.3. *Under the hypothesis of Theorem 4.1. For all $t \geq 0$, we have*

- *In the subcritical case ($z < z_s$),*

$$\|\mathbf{P}_{\Pi^{in}} \{\mathbf{C}(t) \in \cdot\} - \mathbf{P_0} \{\mathbf{C}(t) \in \cdot\}\| \leq K \mathbf{E}_{\Pi^{in}} \left[\langle C, \sqrt{\mathbf{Q}} \rangle_H \right] e^{-\lambda t};$$

- *In the super-critical case ($z > z_s$), for all $n \geq 2$,*

$$\|\mathbf{P}_{\Pi^{in}} \{(C_2(t), \ldots, C_n(t)) \in \cdot\} - \mathbf{P_0} \{(C_2(t), \ldots, C_n(t)) \in \cdot\}\| \leq K_n \mathbf{E}_{\Pi^{in}} \left[\langle C, \sqrt{\mathbf{Q}} \rangle_H \right] e^{-\lambda t}.$$

Proof. Let the collection of processes $\{X_k\}$ (resp. $\{Y_k\}$) be the *particle description* of the BD process that starts from the initial distribution Π^{in} (resp. from $\delta_{\mathbf{0}}$). We couple the processes $\{X_k\}$ to the processes $\{Y_k\}$ such that all new clusters arise in the system from $\emptyset$ simultaneously and subsequently evolve with the same jumps. Namely, $Y_k(t) = X_{k+N^{in}}(t)$ for all $k \geq 1$ and all $t \geq 0$, where N^{in} is distributed according to Π^{in}.

In the sub-critical case the proof readily follows from Lemma 4.2 and the definition of total variation distance. Indeed

$$\|\mathbf{P}_{\Pi^{in}} \{\mathbf{C}(t) \in \cdot\} - \mathbf{P_0} \{\mathbf{C}(t) \in \cdot\}\| \leq 1 - \mathbf{P}_{\Pi^{in}} \left\{ \forall k \in [\![1, N^{in}]\!], X_k(t) = 1 \right\},$$

because both processes are equal whenever all initial clusters from Π^{in} have reached the absorbing state $i = 1$. A very similar argument holds in the super-critical case. $\square$

Proof of Theorem 4.1. We consider first the sub-critical case. As said, condition (4.1) is satisfied for Π^{eq} since $\mathbf{E}_{\Pi^{eq}} \left[\langle C, \sqrt{\mathbf{Q}} \rangle_H \right] < \infty$, thus Lemma 4.3 applies for Π^{eq} as initial distribution. Because the constructed BD process is regular and Π^{eq} is a stationary distribution (*i.e.* $\mathbf{P}_{\Pi^{eq}} \{\mathbf{C}(t) \in \cdot\} = \Pi^{eq}$), we deduce

$$\|\mathbf{P_0} \{\mathbf{C}(t) \in \cdot\} - \Pi^{eq}\| \leq K \mathbf{E}_{\Pi^{eq}} \left[\langle C, \sqrt{\mathbf{Q}} \rangle_H \right] e^{-\lambda t}.$$

Going back to any Π^{in} satisfying condition (4.1), applying Lemma 4.3 again and the triangular inequality yield the desired result.

Consider now the super-critical case. Π^{stat} is a product of Poisson distribution $\mathcal{P}(f_i)$ on $\mathbb{N}_0$ with mean f_i. According to a classical result on Markov population processes, see e.g. [7, Sec. 4], the law $\mathbf{P_0} \{\mathbf{C}(t) \in \cdot\}$ is also a product of Poisson distribution $\mathcal{P}(c_i(t))$ on $\mathbb{N}_0$ with mean $c_i(t)$. The vector $\mathbf{c}(t) = (c_2(t), c_3(t), \ldots)$ solves the deterministic (linear) Becker-Döring equations. Namely, $\dot{\mathbf{c}}(t) = A\mathbf{c}(t) + a_1 z^2 \mathbf{e}_2$ where $A = (q_{j,i})_{i,j \geq 2}$ the matrix with entries in (3.1), and initial condition $\mathbf{c}(0) = \mathbf{0}$. Thanks to [8, Theorem III], we have

$$\|\mathbf{c}(t) - \mathbf{f}\|_H \leq \|\mathbf{f}\|_H e^{-\lambda t}, \tag{4.3}$$

and we easily obtain that $\|\mathcal{P}(c_i(t)) - \mathcal{P}(f_i)\| \leq |c_i(t) - f_i|$. The latter, with independence of the marginals of Π^{stat} and of $\mathbf{P_0} \{\mathbf{C}(t) \in \cdot\}$, estimate (4.3) and Cauchy-Schwarz inequality, entail

$$\|\mathbf{P_0} \{(C_2(t), \ldots, C_n(t)) \in \cdot\} - \Pi^{stat}(\cdot \times \prod_{k=n+1}^{\infty} \mathbb{N}_0)\| \leq \sum_{i=2}^{n} |c_i(t) - f_i| \leq K_n \|\mathbf{f}\|_H e^{-\lambda t}.$$

We conclude again by Lemma 4.3 and the triangular inequality. $\square$

5 A quasi-stationary distribution for cluster smaller than n

Let $\mathcal{E}_n = \{C \in \mathcal{E} \mid C_i = 0\,, i \geq n+1\}$. We define the first exit time from $\mathcal{E}_n$,

$$\tau_n = \inf\{t \geq 0 \mid \mathbf{C}(t) \notin \mathcal{E}_n\}\,, \tag{5.1}$$

corresponding to the first occurrence of a cluster of size greater than n. In this section, we prove a QSD on $\mathcal{E}_n$ is a quasi-limiting distribution for the BD process conditioned to $\tau_n > t$, with exponential convergence for a wide range of initial distributions. It is remarkable that we have at hand an explicit QSD, given, for all $C \in \mathcal{E}_n$, by

$$\Pi_n^{\mathrm{qsd}}(C) = \prod_{i=2}^n \frac{(f_i^n)^{C_i}}{C_i!} e^{-f_i^n}\,, \tag{5.2}$$

where f_i^n is defined in (2.8), for $i = 2, \ldots, n$. Remark that $\mathbf{P}_{\Pi^{\mathrm{in}}}\{\tau_n > t\} > 0$ for all times $t > 0$ and for any Π^{in} supported on $\mathcal{E}_n$.

Proposition 5.1. *Under assumption* (H0). *The distribution Π_n^{qsd} is a quasi-stationary distribution for the BD process conditioned to stay on $\mathcal{E}_n$ namely,*

$$\mathbf{P}_{\Pi_n^{\mathrm{qsd}}}\{\mathbf{C}(t) \in \cdot \mid \tau_n > t\} = \Pi_n^{\mathrm{qsd}} \quad \textit{and} \quad \mathbf{P}_{\Pi_n^{\mathrm{qsd}}}\{\tau_n > t\} = \exp(-J_n t)\,.$$

Proof. Recall assumption (H0) ensures the BD process is regular. Fix $n \geq 2$. Let the semi-group $P_t^n \psi(C) = \mathbf{E}_C\left[\psi(\mathbf{C}(t))\mathbf{1}_{t < \tau_n}\right]$ for $t \geq 0$ (*i.e.* $\mathbf{C}(0)$ is distributed according to δ_C), whose generator is

$$\mathcal{A}_n \psi(C) = \sum_{i=1}^{n-1} \left(a_i z C_i[\psi(C + \Delta_i) - \psi(c)] + b_{i+1} C_{i+1}[\psi(C - \Delta_i) - \psi(C)] \right) - a_n z C_n \psi(C)\,,$$

for all $C \in \mathcal{E}_n$ (recall $C_1 = z$) and bounded function ψ on $\mathcal{E}_n$. Denote by $\mathcal{A}_n^*$ the dual operator for the generator $\mathcal{A}_n$. Some calculations show that the distribution (5.2) satisfies, for any $C \in \mathcal{E}_n$,

$$\mathcal{A}_n^* \Pi_n^{\mathrm{qsd}}(C) = \Pi_n^{\mathrm{qsd}}(C) \left\{ b_2 f_2^n - a_1 z^2 + \sum_{i=2}^n \frac{C_i}{f_i^n} \left(a_{i-1} z f_{i-1}^n - (a_i z + b_i) f_i^n + \mathbf{1}_{i<n} b_{i+1} f_{i+1}^n \right) \right\}\,,$$

with the convention $f_1^n = z$. Since the f_i^n given by (2.8) verifies $a_i z f_i^n - \mathbf{1}_{i<n} b_{i+1} f_{i+1}^n = J_n$ for all $i \in [\![1, n]\!]$, all terms but the first cancel in the above expression, so that we obtain $\mathcal{A}_n^* \Pi_n^{\mathrm{qsd}} = -J_n \Pi_n^{\mathrm{qsd}}$. The latter ensures, with the fact that $J_n \leq a_1 z^2$, that Π_n^{qsd} is a QSD thanks to the classical spectral criteria [2, Thm 4.4]. $\square$

The next theorem shows the QSD is a quasi-limiting distribution for a wide range of initial distributions supported on $\mathcal{E}_n$, with an exponential rate of convergence and an explicit (non-uniform) pre-factor.

Theorem 5.2. *Under assumption* (H0). *Let Π^{in} be a probability distribution on $\mathcal{E}_n$ such that $\mathbf{E}_{\Pi^{\mathrm{in}}}\left[\sum_{i=2}^{\infty} C_i\right] < \infty$. We have for all $t \geq 0$,*

$$\|\mathbf{P}_{\Pi^{\mathrm{in}}}\{\mathbf{C}(t) \in \cdot \mid \tau_n > t\} - \Pi_n^{\mathrm{qsd}}\| \leq K_n \left(\frac{H_n^{\mathrm{in}}}{\mathbf{P}_{\Pi^{\mathrm{in}}}\{\tau_n > t\}} + e^{J_n t} H_n^{\mathrm{qsd}} \right) e^{-\gamma_n t}\,,$$

where τ_n is defined in (5.1), J_n in (2.7), K_n in (2.9), γ_n in (3.6),

$$H_n^{\mathrm{in}} = \sum_{i=2}^n \sqrt{Q_i} z^i \frac{\mathbf{E}_{\Pi^{\mathrm{in}}}[C_i]}{f_i^n} \quad \textit{and} \quad H_n^{\mathrm{qsd}} = \sum_{i=2}^n \sqrt{Q_i} z^i\,.$$

It is clear that H_n^{in} is finite because $\mathbf{E}_{\Pi^{\text{in}}}\left[\sum_{i=2}^{\infty} C_i\right]$ is. The proof of Theorem 5.2 is similar to the proof of Theorem 4.1 and consists in a coupling argument together with a control of the initial clusters in Π^{in}. We start by the latter, which is the analogous of Lemma 4.2.

Lemma 5.3. *Under the hypothesis of Theorem 5.2. Let the collection of processes $\{X_k\}$ be the* particle description *of the BD process* $\mathbf{C}$. *We have*

$$\mathbf{P}_{\Pi^{\text{in}}}\left\{\forall k \in [\![1, N^{\text{in}}]\!],\ X_k(t) = 1 \mid \tau_n > t\right\} \geq 1 - e^{-\gamma_n t}\frac{K_n H_n^{\text{in}}}{\mathbf{P}_{\Pi^{\text{in}}}\left\{t < \tau_n\right\}}.$$

Proof. We start by observing that the following relation holds true,

$$\tau_n = \min(\tau_n^0, T_n^1, \ldots, T_n^{N^{\text{in}}}),\tag{5.3}$$

where $T_n^k = \inf\left\{t > 0 \mid X_k(t) \geq n + 1\right\}$ for $k = 1, \ldots, N^{\text{in}}$ and

$$\tau_n^0 = \inf\left\{t \geq 0 \mid \exists k > N^{\text{in}},\, X_k(t) \geq n + 1\right\}.\tag{5.4}$$

Let $\mathbf{C}(0) = C \in \mathcal{E}_n$ be deterministic, define $N = \sum_{i=2}^{\infty} C_i$ and $(i_1, \ldots, i_N) \in [\![2, n]\!]^N$ given by the *labeling* function such that $C_i = \#\left\{k \in [\![1, N]\!] \mid i_k = i\right\}$. Conditionally on their initial condition, all clusters $X_k(t)$ (starting at i_k) are independent from each other, thus the event

$$A_t = \{\forall k \in [\![1, N]\!],\ X_k(t) = 1\}\tag{5.5}$$

is independent of $(X_k)_{k>N}$ and thus independent of τ_n^0. Then,

$$\mathbf{P}_C\left\{A_t \mid \tau_n > t\right\} = \mathbf{P}_C\left\{A_t \mid \min(T_n^1, \ldots, T_n^N) > t\right\}.$$

Still by independence of the clusters from each other, we claim that

$$\mathbf{P}_C\left\{A_t \mid \min(T_n^1, \ldots, T_n^N) > t\right\} = \prod_{k=1}^{N} \mathbf{P}\left\{X(t) = 1 \mid X(0) = i_k,\ T_n > t\right\},\tag{5.6}$$

where X is defined in Sec. 3 and T_n in (3.3). This equation is clear for $N = 1$, and is easily proved by induction. We do it only for $N = 2$, for the sake of simplicity. Let i_1, $i_2 \in [\![2, n]\!]$. By definition of A_t, independence of the X_1 and X_2 conditionally on their initial condition, and since they are copy of X in Sec. 3,

$$\mathbf{P}_C\left\{A_t \mid \min(T_n^1, T_n^2) > t\right\} = \prod_{k=1}^{2} \frac{\mathbf{P}\left\{X_k(t) = 1,\ T_n^k > t,\ X_k(0) = i_k\right\}}{\mathbf{P}\left\{T_n^k > t,\ X_k(0) = i_k\right\}}$$

$$= \prod_{k=1}^{2} \mathbf{P}\left\{X(t) = 1 \mid X(0) = i_k,\ T_n > t\right\}.$$

This proves the desired result. Thus, going back to (5.6) and thanks to (3.7) we have

$$\mathbf{P}_C\left\{A_t \mid \tau_n > t\right\} \geq \prod_{k=1}^{N}\left(1 - \left(K_n\frac{\sqrt{Q_{i_k} z^{i_k}}}{f_{i_k}^n}e^{-\gamma_n t}\right) \wedge 1\right) \geq 1 - K_n e^{-\gamma_n t}\sum_{k=1}^{N}\frac{\sqrt{Q_{i_k} z^{i_k}}}{f_{i_k}^n}.$$

Finally, we obtain

$$\mathbf{P}_{\Pi^{\text{in}}}\left\{A_t \mid \tau_n > t\right\} = \sum_{C \in \mathcal{E}_n} \mathbf{P}_C\left\{A_t \mid \tau_n > t\right\}\frac{\mathbf{P}_C\left\{\tau_n > t\right\}}{\mathbf{P}_{\Pi^{\text{in}}}\left\{\tau_n > t\right\}}\Pi^{\text{in}}(C)$$

$$\geq 1 - K_n e^{-\gamma_n t}\sum_{C \in \mathcal{E}_n}\sum_{k=1}^{N}\frac{\sqrt{Q_{i_k} z^{i_k}}}{f_{i_k}^n}\frac{\mathbf{P}_C\left\{\tau_n > t\right\}}{\mathbf{P}_{\Pi^{\text{in}}}\left\{\tau_n > t\right\}}\Pi^{\text{in}}(C),$$

where N and $(i_k)_{k=1..N}$ are given by the *labeling* function for each $C \in \mathcal{E}_n$. Now using that $\mathbf{P}_C \{\tau_n > t\} \leq 1$, we have,

$$\mathbf{P}_{\Pi^{\text{in}}} \{A_t \mid \tau_n > t\} \geq 1 - e^{-\gamma_n t} \frac{K_n}{\mathbf{P}_{\Pi^{\text{in}}} \{\tau_n > t\}} \sum_{C \in \mathcal{E}_n} \sum_{k=1}^{N} \frac{\sqrt{Q_{i_k} z^{i_k}}}{f_{i_k}} \Pi^{\text{in}}(C)$$

$$= 1 - e^{-\gamma_n t} \frac{K_n}{\mathbf{P}_{\Pi^{\text{in}}} \{\tau_n > t\}} \sum_{i=2}^{n} \frac{\sqrt{Q_i z^i}}{f_i} \sum_{C \in \mathcal{E}_n} C_i \Pi^{\text{in}}(C) = 1 - e^{-\gamma_n t} \frac{K_n H_n^{\text{in}}}{\mathbf{P}_{\Pi^{\text{in}}} \{\tau_n > t\}} \,.$$

$\square$

Proof of the Theorem 5.2. Let the collection of processes $\{X_k\}$ (resp. $\{Y_k\}$) be the *particle description* of the BD process that starts from the initial distribution Π^{in} (resp. from $\delta_{\mathbf{0}}$). We couple the processes $\{X_k\}$ to the processes $\{Y_k\}$ as in the proof of Theorem 4.3, namely, $Y_k(t) = X_{k+N^{\text{in}}}(t)$ for all $k \geq 1$ and all $t \geq 0$, where N^{in} is distributed according to Π^{in}.

To avoid confusion, we write τ_n^X and τ_n^Y for the first exit time from $\mathcal{E}_n$ of the collection of processes $\{X_k\}$ and $\{Y_k\}$, respectively. We define $\tau_n^{0,X}$ and $\tau_n^{0,Y}$, respectively to the processes $\{X_k\}$ and $\{Y_k\}$ likewise τ_n^0 in (5.4). Due to the coupling between the $\{X_k\}$ and $\{Y_k\}$, we have

$$\tau_n^{0,Y} = \tau_n^Y = \inf \{t \geq 0 \mid \exists k > 0 , Y_k(t) \geq n+1\}$$
$$= \inf \{t \geq 0 \mid \exists k > N^{\text{in}} , X_k(t) \geq n+1\} = \tau_n^{0,X} \,. \quad (5.7)$$

Remark that each Y_k, for $k \geq 1$, is independent of X_i for $i \leq N^{\text{in}}$, thus independent of $T_n^1, \ldots, T_n^{N^{\text{in}}}$ the exit times arising in (5.3). Hence, by (5.3) and (5.7), the laws of the collection of processes $\{Y_k\}$ conditioned to $\tau_n^Y > t$ equal to the laws of the collection of processes $\{Y_k\}$ conditioned to $\tau_n^X > t$. For any $t > 0$, if the event A_t (given in (5.5)) has occurred, we have, for any $i \geq 2$, $\# \{k \mid X_k(t) = i\} = \# \{k \mid Y_k(t) = i\}$, since all initial particles in $\{X_k\}$ have reached the absorbing state. Finally, we deduce that

$$\|\mathbf{P}_{\Pi^{\text{in}}} \{\mathbf{C}(t) \in \cdot \mid \tau_n > t\} - \mathbf{P}_{\mathbf{0}} \{\mathbf{C}(t) \in \cdot \mid \tau_n > t\}\|$$
$$\leq \mathbf{P} \{\exists i \geq 2, \# \{k \mid X_k(t) = i\} \neq \# \{k \mid Y_k(t) = i\} \mid \tau_n^X > t\} \leq \mathbf{P}_{\Pi^{\text{in}}} \{A_t^c \mid t < \tau_n^X\} \,.$$

The latter, with Lemma 5.3, entails

$$\|\mathbf{P}_{\Pi^{\text{in}}} \{\mathbf{C}(t) \in \cdot \mid \tau_n > t\} - \mathbf{P}_{\mathbf{0}} \{\mathbf{C}(t) \in \cdot \mid \tau_n > t\}\| \leq e^{-\gamma_n t} \frac{K_n H_n^{\text{in}}}{\mathbf{P}_{\Pi^{\text{in}}} \{t < \tau_n\}} \,. \quad (5.8)$$

Then, using the last inequality (5.8) with initial distribution the QSD, Π_n^{qsd}, together with Proposition 5.1, we deduce (recall that $\mathbf{E}_{\Pi_n^{\text{qsd}}}[C_i] = f_i^n$)

$$\|\Pi_n^{\text{qsd}} - \mathbf{P}_{\mathbf{0}} \{\mathbf{C}(t) \in \cdot \mid t < \tau_n\}\| \leq e^{-\gamma_n t} K_n e^{J_n t} H_n^{\text{qsd}} \,.$$

We end the proof by triangular inequality. $\square$

6 Estimates on the time τ_n a cluster larger than n appears

We give in this section a tight lower bound on $\mathbf{P}_{\Pi^{\text{in}}} \{\tau_n > t\}$ in the super-critical case $z > z_s$. The analysis of the τ_n in (5.1) leads off the simple observation

$$\tau_n > t \Leftrightarrow \forall s \leq t, \max_{1 \leq k \leq N(s)} X_k(s) \leq n \,,$$

where $\{X_k\}$ is the *particle description* of the BD process. We prove

Theorem 6.1. *Under assumption* (H0). *Let $z > z_s$ and Π^{in} be a probability distribution on $\mathcal{E}_n$ such that $\mathbf{E}_{\Pi^{\text{in}}}\left[\sum_{i=2}^{\infty} C_i\right] < \infty$. We have*

$$\mathbf{P}_{\Pi^{\text{in}}}\left\{\tau_n > t\right\} \geq G_n^{\text{in}} e^{-J_n t},$$

where

$$G_n^{\text{in}} = \mathbf{E}_{\Pi^{\text{in}}}\left[\prod_{i=1}^{n}\left(\frac{f_i^n}{Q_i z^i}\right)^{C_i}\right].$$

In fact, as in the coupling strategy, (5.3) provides a useful understanding of the statistics of τ_n by decomposing between the initial clusters from the ones that will appear at later times. We start with the statistics of the latter, namely of τ_n^0.

Lemma 6.2. *Under assumption* (H0) *and $z > z_s$. For any probability distribution Π^{in} on $\mathcal{E}_n$ such that for all $i \in [\![2, n]\!]$ and $k \in \mathbb{N}$,*

$$\mathbf{P}_{\Pi^{\text{in}}}\left\{\sum_{i \leq j \leq n} C_j \geq k\right\} \leq \mathbf{P}_{\Pi^{\text{qsd}}}\left\{\sum_{i \leq j \leq n} C_j \geq k\right\}, \tag{6.1}$$

we have

$$\mathbf{P}_{\Pi^{\text{in}}}\left\{\tau_n > t\right\} \geq e^{-J_n t}.$$

Proof of Lemma 6.2. It is classical that condition (6.1) ensures there exists randoms $\mathbf{C}^{\text{in}}$ and $\mathbf{C}^{\text{qsd}}$ distributed according to Π^{in} and Π^{qsd}, respectively, such that for each $i \in [\![2, n]\!]$,

$$\sum_{i \leq j \leq n} C_j^{\text{in}} \leq \sum_{i \leq j \leq n} C_j^{\text{qsd}}, \quad a.s.$$

see *e.g.* [3, Sec. 4.12]. Then, we may construct the collection of processes $\{X_k\}$ (resp. $\{Y_k\}$) as the *particle description* of the BD process associated to $\mathbf{C}^{\text{in}}$ (resp. to $\mathbf{C}^{\text{qsd}}$) such that, *a.s.*, for all $i \geq 1$, $X_i(0) \leq Y_i(0)$. A standard coupling between two copies of the chain X from Sec. 3 consists in having the same jumps in the two copies as soon as they are equal. Such coupling applied to each couple (X_i, Y_i) then ensures that, for all $i \geq 1$ and $t \geq 0$, we have $X_i(t) \leq Y_i(t)$ *a.s.* In particular,

$$\inf\{t > 0 \mid \max_k X_k(t) > n\} \geq \inf\{t > 0 \mid \max_k Y_k(t) > n\} \quad a.s.$$

thus, with Proposition 5.1,

$$\mathbf{P}_{\Pi^{\text{in}}}\left\{\tau_n > t\right\} \geq \mathbf{P}_{\Pi^{\text{qsd}}}\left\{\tau_n > t\right\} = e^{-J_n t}. \qquad \square$$

Proof of Theorem 6.1. Let $n \geq 2$ and $i \in [\![2, n]\!]$. Define $g_{n,i}(t) = \mathbf{P}\left\{T_n > t \mid X(0) = i\right\}$ where T_n and X are given in Sec. 3. Thanks to (3.5), we have

$$g_{n,i}(t) \geq \lim_{t \to +\infty} g_{n,i}(t) = \frac{f_i^n}{Q_i z^i}. \tag{6.2}$$

Let $\psi_n(t, x) = 1$ if $\max_{0 \leq s \leq t} x(s) \leq n$ and 0 otherwise. We have

$$\mathbf{P}_{\Pi^{\text{in}}}\left\{\tau_n > t\right\} = \sum_{C \in \mathcal{E}_n} \mathbf{P}_C\left\{\prod_{i=1}^{N(t)} \psi_n(t, X_i) = 1\right\} \Pi^{\text{in}}(C),$$

where $\{X_k\}$ is the *particle description* of the BD process $\mathbf{C}$ (starting at δ_C). By independence of the particles conditionally to their initial condition,

$$
\mathbf{P}_C \left\{ \prod_{i=1}^{N(t)} \psi_n(t, X_i) = 1 \right\} = \mathbf{P}_C \left\{ \prod_{i=N^{in}+1}^{N(t)} \psi_n(t, X_i) = 1 \right\} \mathbf{P}_C \left\{ \prod_{i=1}^{N^{in}} \psi_n(t, X_i) = 1 \right\}
$$

$$
= \mathbf{P_0} \left\{ \tau_n > t \right\} \prod_{k=1}^{N} g_{n,i_k}(t), \quad (6.3)
$$

where again, N and $(i_k)_{k=1..N}$ are given by the *labeling* function associated to $C \in \mathcal{E}_n$. Combining relations (6.2) and (6.3) with Lemma 6.2 (with the choice $\Pi^{in} = \delta_{\mathbf{0}}$) and summing over all initial conditions ends the proof. $\qquad\square$

7   Metastability for z close to z_s

In this section we interpret the metastability with the results obtained on the QSD. For that purpose, we assume additionally to (H1) and (H2), to fit with [8, 9], that

$$
A' < a_i < A i^\alpha, \quad \frac{b_{i+1}}{a_{i+1}} + \frac{\kappa}{i^\nu} \le \frac{b_i}{a_i} \quad \text{and} \quad z_s e^{G i^{-\gamma}} \le \frac{b_i}{a_i} \le z_s e^{G' i^{-\gamma'}}, \quad (\text{H3})
$$

for all $i \ge 2$, where $\alpha, \gamma \in (0,1)$, $\gamma', \nu > 0$, κ, A', A, G and G' positives. We also use the terminology of [9] namely a quantity $q(z)$ of z is: exponentially small if $q(z)/(z - z_s)^m$ is bounded for all $m > 0$ as $z \searrow z_s$ (z converges to z_s and $z > z_s$); and at most algebraically large if $(z - z_s)^{m_0} q(z)$ is bounded for some $m_0 > 0$ as $z \searrow z_s$.

It is worth mentioning that assumption (H3) ensures the existence of a unique n^* (depending on $z > z_s$) such that $b_{n^*+1}/a_{n^*+1} < z < b_{n^*}/a_{n^*}$, which turns to be the minimizer of $\{a_i Q_i z^i\}$. The size n^* is interpreted as the nucleus size: for a cluster $X(t) \le n^*$, $X(t)$ tends to shorten, while for $X(t) > n^*$, it tends to grow. Moreover, with assumption (H3), further algebraic considerations in [8, 9] allow to obtain quantitative estimates of the key quantities that show up in Theorem 5.2 and Theorem 6.1:

• $n^* \to \infty$ as $z \searrow z_s$ and n^* is at most algebraically large.

• the time scale that quantifies the convergence towards the QSD, $1/\gamma_{n^*}$, is at most algebraically large.

• The rate of formation of cluster of size larger that n^* (nucleation rate), J_{n^*}, is exponentially small.

We gather those estimates with Theorem 5.2 and Theorem 6.1 to quantify the nucleation rate and the convergence towards the QSD, for a wide range of initial distributions, in the next proposition.

Proposition 7.1. *Under hypotheses (H1), (H2) and (H3). Let $j \in \mathbb{N}_1$ and Π^{in} an initial distribution with support in $\mathcal{E}_j$. For all $\varepsilon \in (0, \frac{1}{2})$, there exists $z_0 > z_s$ such that*

$$
\mathbf{P}_{\Pi^{in}} \left\{ \tau_{n^*} > t \right\} \ge (1 - \varepsilon) e^{-J_{n^*} t} \quad (7.1)
$$

for any $z \in (z_s, z_0)$, and

$$
\| \mathbf{P}_{\Pi^{in}} \left\{ \mathbf{C}(t) \in \cdot \mid \tau_{n^*} > t \right\} - \Pi_n^{\text{qsd}} \|
$$
$$
\le \left\{ \frac{2}{1-\varepsilon} \mathbf{E}_{\Pi^{in}} \left[\langle C, \sqrt{\mathbf{Q}} \rangle_H \right] + n^* \sqrt{\frac{a_1 z}{A'}} \right\} \sqrt{\frac{a_1 z}{A'}} n^* e^{J_{n^*} t - \gamma_{n^*} t}. \quad (7.2)
$$

Proof. First, by Theorem 6.1 and the definition of G_n^{in}, it readily follows that

$$\mathbf{P}_{\Pi^{\mathrm{in}}}\left\{\tau_{n^*} > t\right\} \geq \left(1 - \sum_{i=2}^{j} \mathbf{E}_{\Pi^{\mathrm{in}}}\left[C_i\right]\left(1 - \frac{f_i^{n^*}}{Q_i z^i}\right)\right) e^{-J_{n^*}t}\,.$$

We conclude to (7.1) observing that $\lim_{z \searrow z_s} f_i^{n^*}/(Q_i z^i) = 1$ and $f_i^{n^*} \leq Q_i z^i$ for all $i \geq 2$, because $J_{n^*} \to 0$ thanks to [9]. Then, by Theorem 5.2, and inequality (7.1),

$$\|\mathbf{P}_{\Pi^{\mathrm{in}}}\left\{\mathbf{C}(t) \in \cdot \mid \tau_{n^*} > t\right\} - \Pi_n^{\mathrm{qsd}}\| \leq K_n\big(\tfrac{H_{n^*}^{\mathrm{in}}}{1-\varepsilon} + H_{n^*}^{\mathrm{qsd}}\big)e^{J_{n^*}t - \gamma_{n^*}t}\,.$$

By definition of $H_{n^*}^{\mathrm{in}}$, we obtain $H_{n^*}^{\mathrm{in}} \leq 2\mathbf{E}_{\Pi^{\mathrm{in}}}\big[\langle C, \sqrt{\mathbf{Q}}\rangle_H\big]$ when $j \leq n^*$ since $f_i^{n^*}$ goes to $Q_i z^i$. Moreover, since $a_i Q_i z^i$ is decreasing (in i) up to n^*, we have $K_{n^*} \leq \sqrt{\frac{a_1 z}{A'} n^*}$ and $H_{n^*}^{\mathrm{qsd}} \leq \sqrt{\frac{a_1 z}{A'}} n^*$ which yields to (7.2). $\qquad\square$

The interpretation is the following: In the limit $z \searrow z_s$, by (7.1), the occurrence of a first nucleus has probability arbitrarily close to 0 to appear in the range of times $t \ll 1/J_{n^*}$, that is exponentially large. Before nucleation, for times in the range of $1/\gamma_{n^*} \ll t \ll 1/J_{n^*}$, that are at least algebraically large and up to an exponentially large time, the system is "trapped" in the QSD state, since the right-hand side of (7.2) can be made arbitrarily close to 0 for such times. These two considerations fulfills standard notion of metastability [9].

References

[1] R. Becker and W. Döring, *Kinetische behandlung der keimbildung in übersättigten dämpfen*, Annalen der Physik **416** (1935), no. 8, 719–752.

[2] P. Collet, S. Martínez, and J. San Martín, *Quasi-stationary distributions*, Probability and its Applications (New York), Springer, Heidelberg, 2013. MR-2986807

[3] G. R. Grimmett and D. R. Stirzaker, *Probability and random processes*, third ed., Oxford University Press, New York, 2001. MR-2059709

[4] E. Hingant and R. Yvinec, *Deterministic and stochastic Becker-Döring equations: past and recent mathematical developments*, In: Stochastic processes, multiscale modeling, and numerical methods for computational cellular biology, Springer, Cham, 2017, pp. 175–204. MR-3726370

[5] E. Hingant and R. Yvinec, *The Becker-Döring process: pathwise convergence and phase transition phenomena*, J. Stat. Phys. **177** (2019), no. 3, 506–527. MR-4026657

[6] S. Karlin and J. McGregor, *The classification of birth and death processes*, Trans. Amer. Math. Soc. **86** (1957), 366–400. MR-94854

[7] J. F. C. Kingman, *Markov population processes*, J. Appl. Probability **6** (1969), 1–18. MR-254934

[8] M. Kreer, *Classical Becker-Döring cluster equations: rigorous results on metastability and long-time behaviour*, Ann. Physik (8) **2** (1993), no. 4, 398–417. MR-1223849

[9] O. Penrose, *Metastable states for the Becker-Döring cluster equations*, Comm. Math. Phys. **124** (1989), no. 4, 515–541. MR-1014113

[10] W. Sun, *A functional central limit theorem for the Becker-Döring model*, J. Stat. Phys. **171** (2018), no. 1, 145–165. MR-3773855