Analysis and Verification of Robustness Properties in Becker-Döring Model

Lucia Nasti(✉), Roberta Gori, and Paolo Milazzo

Department of Computer Science, University of Pisa, Largo Bruno Pontecorvo, 3, 56127 Pisa, Italy
{lucia.nasti,gori,milazzo}@di.unipi.it

Abstract. Many biochemical processes in living cells involve clusters of particles. Such processes include protein aggregation and the development of intracellular concentration gradients. To study these mechanisms, we can apply coagulation-fragmentation models describing populations of interacting components. In this context, the Becker-Döring equations - theorized in 1935 - provide the simplest kinetic model to describe condensations phenomena. Experimental works on this model reveal that it exhibits robustness, defined as the system's capability to preserve its features despite noise and fluctuations. Here, we verify the robustness of the BD model, applying our notions of initial concentration robustness (α-robustness and β-robustness), which are related to the influence of the perturbation of the initial concentration of one species (i.e., the input) on the concentration of another species (i.e., the output) at the steady state. Then, we conclude that a new definition of robustness, namely the asymptotic robustness, is necessary to describe more accurately the model's behavior.

Keywords: Becker-Döring equations · Robustness · Modeling · Simulation

1 Introduction

Many biochemical processes in living systems are combined with the formation of clusters of particles, such as protein aggregation [23], polymerization [14,15], and formation of intracellular concentration gradients [13]. These biological phenomena have been studied through the application of coagulation-fragmentation models that describe populations of interacting components [12,17].

In this context, one of the most common models is based on the Becker-Döring equations (BD), which were theorized for the first time in 1935, by the two authors who gave the name of the model [6]. They proposed an infinite system of ordinary differential equations as a model for the time evolution of the distribution of cluster sizes for a system [3,8].

In parallel with experimental work [23], it has been observed that coagulation-fragmentation models can show *robustness* with respect to perturbations: changing parameters (such as the particles' concentration) does not change the cluster distribution.

J. Bowles et al. (Eds.): DataMod 2021, LNCS 13268, pp. 27–43, 2022.
https://doi.org/10.1007/978-3-031-16011-0_3

Indeed, at various frequencies and timescales, internal and external fluctuations can alter specific functions or traits of biological systems, causing genetic mutations, loss of structural integrity, diseases, and so on. Nevertheless, many biological networks can maintain their functionalities despite perturbations: this distinct property is known as *robustness* [16]. Robust traits are pervasive in biology: they involve various structural levels, such as gene expression, protein folding, metabolic flux, species persistence. For this reason, the study of robustness is essential for biologists, whose aim is to understand the functioning of a biological system.

In general, investigating biological systems is extremely challenging because they are characterized by non-linearity and non-intuitive behaviors. They can be studied by performing wet-lab (*in vitro*) experiments, or through mathematical or computational (*in silico*) methods on pathway models [4]. Unfortunately, the applicability of these approaches is often hampered by the complexity of the models to be analyzed (often expressed in terms of *ordinary differential equations* (ODEs) or Markov chains). An alternative way is to infer a specific property from the system [5], such as monotonicity [1,10,21] and steady-state reachability [9], by looking only at the structure of the system, without the need of studying or simulating its dynamics. Establishing such properties, indeed, provides information on the Chemical Reaction Network (CRN) dynamics without the need of performing several numerical simulations [19]. Unfortunately, the applicability of these structural approaches is often limited to rather specific classes of CRNs.

We proceed by verifying the robustness of the Becker-Döring model applying different approaches. We first apply the *sufficient condition* proposed by Shinar and Feinberg in [24,25], which allows robustness to be derived directly from a syntactical property of the pathway, without the need of studying or simulating its dynamics. In particular, they investigate the specific notion of the *absolute concentration robustness*, for which a system is robust if there is at least one chemical species that has – at the equilibrium – the same identical concentration even in presence of perturbations. We show that the definition given by Shinar and Feinberg is particularly limiting in the description of the BD model's behavior. Then, we proceed applying our definition of α-robustness [20], which can help us to quantify the influence that the initial concentrations of a species has on the steady state of the system. In this context, a system is α-robust with respect to a given set of initial concentration intervals if the concentration of a chosen output molecule at the steady state varies within an interval of values $[k - \frac{\alpha}{2}, k + \frac{\alpha}{2}]$ for some $k \in \mathbb{R}$. In addition, we can apply the relative notion of β-robustness, which can be obtained easily by dividing α by k, to study which perturbation influences more the distribution of the clusters and in which way.

By using this approach, we find that by increasing the initial concentration of the input in Becker-Döring system, the α value decreases continuously. Thus, in order to describe more accurately the model's behaviour, we introduce the notion of the *asymptotic robustness*. In addition, to support our result obtained by simulations, in this case we are able to study analytically the steady state of

the BD system. Note that, as already mentioned, this analysis is possible only in particular cases because of the complexity of the biological systems.

This work is organized as follows. Before introducing in detail the mathematical properties of Becker-Döring equations in Sect. 2.2, in Sect. 2.1 we present some notions that will be assumed in the rest of the paper about the representation of chemical reactions. In Sect. 3 we verify the robustness of the BD model, we define the new notion of the asymptotic robustness, and we present the analytical study of the steady state. Finally, in Sect. 4 we draw our conclusions and discuss future work.

2 Background

We introduce some notions that will be assumed in the rest of the paper. In the first part, we focus on the representation of chemical reactions, considering one of the main methods that we can use to describe them: the *deterministic approach*. In the second part, we illustrate the characteristics of the BD model, focusing on its mathematical aspects.

2.1 Chemical Reaction

A chemical reaction is a transformation that involves one or more chemical species, in a specific situation of volume and temperature.

We call *reactants* the chemical species that are transformed, while those that are the result of the transformation are called *products*. We can represent a chemical reaction as an equation, showing all the species involved in the process.

A simple example of chemical reaction is the following elementary reaction:

$$a\mathrm{A} + b\mathrm{B} \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} c\mathrm{C} + d\mathrm{D} \tag{1}$$

In this case, A, B, C, D are the species involved in the process: A and B are the reactants, C and D are the products. The parameters a, b, c, d are called *stoichiometric coefficients* and represent the number of reactants and products participating in the reaction. The arrow is used to indicate the direction in which a chemical reaction takes place. In the example, the double arrow means that the reaction is *reversible* (it can take place in both ways). When we have only one arrow, it means that the reaction is *irreversible*, that is it is not possible to have the opposite transformation. The two reaction rates constant k_1 and k_{-1} quantify the rate and direction of a chemical reaction, where the rate is the speed at which a chemical reaction takes place. To describe the dynamical behaviour of the chemical reaction network, we can use the *law of mass action*, which states that the rate of a reaction is proportional to the product of the reactants. Applying the *law of mass action* to the system, we obtain, for each chemical

species, a differential equation describing the production and the consumption of the considered species. Considering the generic chemical Eq. 1, we obtain:

$$\begin{aligned}
\frac{d[A]}{dt} &= \overbrace{-ak_1[A]^a[B]^b}^{\substack{\text{direct reaction}\\ \text{term}}} \overbrace{+ak_{-1}[C]^c[D]^d}^{\substack{\text{inverse reaction}\\ \text{term}}} \\
\frac{d[B]}{dt} &= -bk_1[A]^a[B]^b + bk_{-1}[C]^c[D]^d \\
\frac{d[C]}{dt} &= +ck_1[A]^a[B]^b - ck_{-1}[C]^c[D]^d \\
\frac{d[D]}{dt} &= +dk_1[A]^a[B]^b - dk_{-1}[C]^c[D]^d.
\end{aligned}$$

where, in each equation, we isolated the term describing the direct reaction from the one describing the inverse reaction. With these two terms, we implicitly considered, for each element, the processes of consumption and production.

We can abstract the dynamics of the CRN using the *stoichiometric matrix* $\Gamma = \{n_{ij}\}$, where each row corresponds to a substance S_i and each column to a reaction R_j. The matrix entries n_{ij} are determined considering how the substance is involved in each reaction of the system. If the substance is a reactant or a product of the reaction, the entry is the stoichiometric coefficient of the species, otherwise, the entry is 0. The sign of the stoichiometric coefficient is conventionally assigned as *positive* if the species is a product and *negative* if it is a reactant. The stoichiometric matrix of the CRN (1) is:

$$\Gamma = \begin{array}{c} \\ A \\ B \\ C \\ D \end{array} \begin{array}{c} \begin{array}{cc} \mathcal{R}_1 & \mathcal{R}_{-1} \end{array} \\ \begin{pmatrix} -a & +a \\ -b & +b \\ +c & -c \\ +d & -d \end{pmatrix} \end{array}.$$

2.2 The Becker-Döring equations

The Becker-Döring equations (BD) describe two principal phenomena, namely the *coagulation* and the *fragmentation* of clusters of particles, based on two processes:

1. a *monomer* (or *elementary particle*) is a cluster characterized by a size i equal to 1. Hitting a cluster of size $i \geq 1$, it gives rise to a coagulation phenomenon, producing a cluster of dimension $i + 1$;
2. a cluster of size $i \geq 2$ can be subjected to a spontaneous fragmentation, splitting itself in a cluster of size $i - 1$ and a monomer.

BD equations can be described by a CRN that, for each $i \in \mathbb{N}$, includes the reaction

$$\mathrm{C_1} + \mathrm{C_i} \underset{b_{i+1}}{\overset{a_i}{\rightleftharpoons}} \mathrm{C_{i+1}} \tag{2}$$

where C_i denotes a cluster of i particles, while kinetic coefficients a_i and $b_i + 1$ stand for the rate of aggregation and fragmentation, respectively. Coefficients a_i and $b_i + 1$ may depend on i, but for the moment we assume them to be constant values (denoted a and b).

By applying the *law of mass action* to the BD model we obtain a differential equation for each cluster size i, which can be generalized by the following recursive definition:

$$\begin{cases} \frac{d[C_1]}{dt} = -J_1 - \sum_{k\geq 1}^{C_1(0)} J_k \\ \frac{d[C_i]}{dt} = J_{i-1} - J_i \end{cases} \tag{3}$$

for every $i \geq 2$, where J_i is the flux:

$$J_i = a[C_1][C_i] - b[C_{i+1}]. \tag{4}$$

In Formula (4), it is possible to recognize the generic process of coagulation, as a second order reaction, and the generic process of fragmentation, as a first order linear reaction.

At the basis of this model there are three fundamental assumptions:

- only a monomer coalesces to give rise a cluster;
- a cluster can release spontaneously a (single) monomer;
- at the initial stage of the system, only $C_1(0)$, namely the initial concentration of monomers, is different from 0, hence all the clusters, with size $i \geq 2$, develop successively.

From the last assumption, it follows that the initial concentration of C_1 determines the mass of the system, hence the largest dimension of the cluster that can be formed. Indeed, if we add 5 molecules in an ideal closed pot, the maximum cluster will have size 5, and the only possible clusters will be C_2, C_3, C_4, and C_5. Then, in a model with the initial concentration of monomers $C_1(0) = 5$, we will actually only have 4 enable reversible reactions (so, the size of the CRN is actually finite) as follows:

$$\begin{aligned} \mathrm{C}_1 + \mathrm{C}_1 &\underset{b}{\overset{a}{\rightleftharpoons}} \mathrm{C}_2 \\ \mathrm{C}_1 + \mathrm{C}_2 &\underset{b}{\overset{a}{\rightleftharpoons}} \mathrm{C}_3 \\ \mathrm{C}_1 + \mathrm{C}_3 &\underset{b}{\overset{a}{\rightleftharpoons}} \mathrm{C}_4 \\ \mathrm{C}_1 + \mathrm{C}_4 &\underset{b}{\overset{a}{\rightleftharpoons}} \mathrm{C}_5 \end{aligned} \tag{5}$$

Then, we obtain the following equations:

$$\begin{cases} \frac{d[C_1]}{dt} = -2a[C_1]^2 + 2b[C_2] - a[C_1][C_2] + b[C_3] - a[C_1][C_3] + b[C_4] - a[C_1][C_4] + b[C_5] \\ \frac{d[C_2]}{dt} = +a[C_1]^2 - b[C_2] - a[C_1][C_2] + b[C_3] \\ \frac{d[C_3]}{dt} = +a[C_1][C_2] - b[C_3] - a[C_1][C_3] + b[C_4] \\ \frac{d[C_4]}{dt} = +a[C_1][C_3] - b[C_4] - a[C_1][C_4] + b[C_5] \\ \frac{d[C_5]}{dt} = +a[C_1][C_4] - b[C_5]. \end{cases}$$

We notice that the expression for the biggest cluster is different from the others. In the previous example, indeed, the biggest cluster is C_5, which can only participate in fragmentation, otherwise other interactions it would create cluster bigger than 5, which is the total mass of the system. This is evident, also considering the fluxes.

Considering again a system with $C_1(0) = 5$, there are 4 fluxes, as follows:

$$\begin{aligned} J_1 &= a[C_1]^2 - b[C_2] \\ J_2 &= a[C_1][C_2] - b[C_3] \\ J_3 &= a[C_1][C_3] - b[C_4] \\ J_4 &= a[C_1][C_4] - b[C_5]. \end{aligned} \tag{6}$$

There is not a flux J_5 because the cluster of dimension 5 can be only involved in a spontaneous fragmentation because of the mass conservation.

The mass of the system (ρ) is conserved, hence the considered system has neither sinks nor sources. From the generic Formula (3) of differential equations we can deduce that the mass of system depends on the initial condition of the system and has the following form:

$$\sum_{i \geq 1} iC_i(t) = \rho \tag{7}$$

3 Verification of Robustness Properties in Becker-Döring model

The Becker-Döring equations can be used to study different biological phenomena, as the formation of gradient concentrations, which we can define as the measurement of the variation of quantity of molecules from one area to another in the same system. This phenomenon is experimentally observable in unicellular and multicellular organisms, and it is involved in various processes, such as cellular differentiation, as described in [23,27].

In [23], the authors develop a theoretical model describing cluster aggregation-fragmentation in subcellular systems, based on the Becker-Döring equations, and show that, in particular conditions, the concentration gradient can be robust to relevant biological fluctuations. Therefore, we proceed to verify formally the robustness of the system based on Becker-Döring equations, applying different approaches that we can summarize as follows:

- *Application of Feinberg's Deficiency Theorems*;
- *Application of α-robustness and β-robustness*;
- *Analytical study of the steady state solution.*

3.1 Application of Deficiency Theorems

Biological properties are challenging to study because they require the observation of the system behavior, considering all the possible initial states.

For example, regarding the robustness evaluation of a signaling pathway, all the combinations of initial concentrations of chemical species need to be examined, and this requires many simulations with different hypotheses. For this reason, in literature, numerous papers introduce methods and approaches that assess this property by avoiding simulating the entire system. In particular, the study of how network structure/topology affects its dynamics, known as Chemical Reaction Network Theory (CRNT), has introduced concepts, such as the *deficiency*, and gives conditions for the existence, uniqueness, multiplicity, and stability of equilibrium points on networks endowed with different kinetics (e.g. mass action) [11]. In this context, the work done by Shinar and Feinberg provides a clear example of CRNT's application and lays the foundations to prove how the structure of a CRN characterizes its behavior. In [9], the *Deficiency One Theorem* and *Deficiency Zero Theorem* are presented, and both of them give crucial information about the steady state of the system, using only linear algebra and without any simulation.

In [24,25] the authors identify simple yet subtle structural attributes that impart concentration robustness to a mass action network that owns them. In their framework a biological system shows *absolute concentration robustness* (ACR) for an active molecular species if the concentration of such species is identical in every positive steady state the system admits. To describe their result, we need to introduce some terminology from CRNT. Consider the following mass action toy system, consisting of two species (A, B) and two reactions (R_1, R_2):

$$\begin{aligned} \mathrm{A} + \mathrm{B} &\xrightarrow{k_1} 2\mathrm{B} \\ \mathrm{B} &\xrightarrow{k_2} \mathrm{A}, \end{aligned} \qquad \text{(Example 1)}$$

where (k_1, k_2) are commonly referred to as *kinetic* or *rate constants*.

The authors represent the system above as a directed graph, using the Standard Reaction Diagram (SRD), shown in Fig. 1. Each node of the graph is a *complex* of the network, defined as the group of reactants/products that are linked by arrows. In this case, there are 4 nodes (A+B, 2B, B, A), represented in the solid boxes. A complex is called *terminal*, if it lies always at the tail of reaction arrows, otherwise it is *non-terminal*. In the example, there are two non-terminal nodes (A+B, B), in grey, and two terminal nodes (2B, A), in white. The groups in which the network is divided are its *linkage classes*; here, there are two linkage classes, in the dashed boxes. Finally, we introduce the *deficiency* δ of the network, a non-negative integer index representing the amount of *linear independence* among the reactions of the network and it is calculated as $\delta = n - l - r$, where n and l are the numbers of nodes and linkage classes, respectively, r is the rank of the stoichiometric matrix Γ. The deficiency of the network is equal to 1 ($\delta = 4 - 2 - 1$).

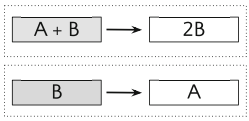

Fig. 1. SRD representation of our toy model

Shinar and Feinberg prove that a mass action system, which admits a positive steady state and is characterized by $\delta = 1$, shows absolute concentration robustness if it has two non-terminal complexes that differ only in one species [24]. In our example, complexes "A+B" and "B" differ in species "A".

On the contrary, if the deficiency is equal to zero, no matter what values the rate constants take, there is no species relative to which the system exhibits absolute concentration robustness, as proved in [25].

Application of the Deficiency One Theorem on BD Model. In order to verify the robustness of the Becker-Döring model, as first step we proceed calculating its deficiency. We consider the set of reactions of Example 5.

As for Example 1, to calculate the deficiency δ, we need to define the number of nodes and the linkage classes, and the rank of the matrix stoichiometric matrix, Γ. In this case, we consider as a species each possible cluster C_i.

- *Nodes* (n) are one or more chemical species involved in forward and backward reaction. We have 8 nodes: $C_1 + C_1$, C_2, $C_1 + C_2$, C_3, $C_1 + C_3$, C_4,$C_1 + C_4$, C_5;
- *linkage classes* (l) are the "groups" of reactions which compose the network, then we have 4 linkage classes;
- Considering the following matrix $r \times N$ (which is the transpose of the stoichiometric matrix Γ), we obtain that the *rank* (r) is 4.

$$\begin{array}{c} \\ R_1 \\ R_{1b} \\ R_2 \\ R_{2b} \\ R_3 \\ R_{3b} \\ R_4 \\ R_{4b} \end{array} \begin{array}{c} \begin{array}{ccccc} C_1 & C_2 & C_3 & C_4 & C_5 \end{array} \\ \left[\begin{array}{ccccc} -2 & +1 & 0 & 0 & 0 \\ +2 & -1 & 0 & 0 & 0 \\ -1 & -1 & +1 & 0 & 0 \\ +1 & +1 & -1 & 0 & 0 \\ -1 & 0 & -1 & +1 & 0 \\ +1 & 0 & +1 & -1 & 0 \\ -1 & 0 & 0 & -1 & +1 \\ +1 & 0 & 0 & +1 & -1 \end{array}\right] \end{array}$$

Then, the deficiency $\delta = n - l - r$ is equal to $\delta = 8 - 4 - 4 = 0$. Therefore, as described in [25] each of the clusters involved in the system cannot be considered robust according to the definition given by Shinar and Feinberg. Therefore, we proceed applying our definition of the *initial concentration robustness*, which extends the definition given in [25].

3.2 Application of α-robustness and β-robustness

As already described in [20], we want to focus on the evaluation of the initial concentration robustness. We want to vary the concentration of at least one chemical species (namely the *input*) and to verify, at the equilibrium, if the concentration of another species (namely the *output*) is included in an interval of possible values. In order to do that, we recall the definition of the initial

concentration robustness, that are formalized by continuous Petri nets. Petri nets has the advantage to provide a graphical support that abstracts away from technical details in the system description. As demonstrated by many biologists, in fact, graphical representations of qualitative trends are often useful to provide intuitions on the network main features [7].

Definition 1 (Continuous Petri net). A *continuous Petri net* N can be defined as a quintuple $\langle P, T, F, W, m_0 \rangle$ where:

- P is the set of continuous *places*, conceptually one for each considered kind of system resource;
- T is the set of continuous *transitions* that consume and produce resources;
- $F \subseteq (P \times T) \bigcup (T \times P) \to \mathbb{R}_{\geq 0}$ represents the set of arcs in terms of a function giving the weight of the arc as result: a weight equal to 0 means that the arc is not present;
- $W : F \to \mathbb{R}_{\geq 0}$ is a function, which associates each transition with a *rate*;
- m_0 is the *initial marking*, that is the initial distribution of *tokens* (representing resource instances) among places. A marking is defined formally as $m : P \to \mathbb{R}_{\geq 0}$. The domain of all markings is M.

Tokens are movable objects, assigned to places, that are consumed by transitions in the input places and produced in the output places. Graphically, a Petri net is drawn as a graph with nodes representing places and transitions. Circles are used for places and rectangles for transitions. Tokens are drawn as black dots inside places. Graph edges represent arcs and are labeled with their weights. To faithfully model biochemical networks, the marking of a place is not an integer (the number of tokens) but a positive real number (called *token value* representing the concentration of a chemical species. Each transition is associated with a kinetic constant, that determines the rate of (continuous) flow of tokens from the input to the output places of the transition.

In order to give the definition, we recall some notions introduced by Nasti et al. in [20]. The initial marking is defined as an assignment of a fixed value to each place p. Now, it is possible to generalize the idea of initial marking by considering a marking as an assignment of a *interval of values* to each place p of the Petri net.

We first recall the definition of the domain of intervals.

Definition 2 (Intervals). The interval domain is defined as

$$\mathcal{I} = \{[min, max] \mid min, max \in \mathbb{R}_{\geq 0} \cup \{+\infty\} \text{ and } min \leq max\}.$$

An interval $[min, max] \in \mathcal{I}$ is *trivial* iff $min = max$. Moreover, $x \in [min, max]$ iff $min \leq x \leq max$.

We now define interval markings.

Definition 3 (Interval marking). Given a set of places P, an *interval marking* is a function $m_{[\,]} : P \to \mathcal{I}$. The domain of all interval markings is $M_{[\,]}$.

An interval marking in which at least one interval is non-trivial represents an infinite set of markings, one for each possible combination of values of the non-trivial intervals. Therefore, given an interval marking, we relate it with the markings as in the original Petri nets formalism in the following way:

$$\text{Given } m \in M \text{ and } m_{[\,]} \in M_{[\,]}, m \in m_{[\,]} \text{ iff } \forall p \in P, m(p) \in m_{[\,]}(p).$$

In a Petri net we assume that there exists *at least one* input place and *exactly one* output place representing input and output species of the modeled biochemical network, respectively. Under this assumption, we can give the formal definition of robustness.

Definition 4 (α-Robustness). A Petri net N with output place O is *α-robust with respect to a given interval marking* $m_{[\,]}$ iff $\exists k \in \mathbb{R}$ such that $\forall m \in m_{[\,]}$, the marking m' corresponding to the concentrations at the steady state reachable from m, is such that

$$m'(O) \in [k - \frac{\alpha}{2}, k + \frac{\alpha}{2}].$$

To compare the α-robustness of different systems or the α-robustness of the same system with different perturbations, we have to introduce another notion: the *relative β-robustness*. First of all, we introduce the concept of normalization of α-robustness defined as:

Definition 5 (Normalized α-robustness). Let N, α ans k be as in Definition 4. The normalized α-robustness of the output O, denoted n_O, is defined as $\frac{\alpha}{k}$.

Definition 6 (Normalized Input). Let N and α be as in Definition 4. Let $[min, max]$ with $min \neq max$ be the interval marking of the input I, defining its initial conditions, and k_I be its midpoint. The normalized input n_I, is defined as $\frac{max-min}{k_I}$.

Therefore, we can state the definition of relative initial concentration robustness as follows:

Definition 7 (Relative β-robustness). Let N be as in Definition 4. The relative initial concentration robustness, denoted as β-robustness, is defined as: $\frac{n_O}{n_I}$, where n_O and n_I are respectively the normalized α-robustness and the normalized input I.

Application α-robustness and β-robustness on BD Model. In order to apply our definition 4, we need to build the Petri net of the BD model, as represented in Fig. 2. We associate to each cluster a place, represented by a circle, and to each reaction a transition, represented by a square. We connect places and transitions by arrows that are defined by the reactions. We identify as input and output of the network the concentration of monomer C_1, in agreement with the assumptions we mentioned in Sect. 2.2. Indeed, C_1 is the only cluster present at the initial state of the system, and it is the cluster involved in every aggregation

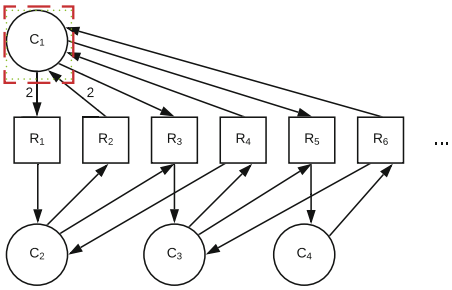

Fig. 2. The Petri net for the BD model. C_1 is both the input and the output of the network. The encountered problems are that the Petri net is potentially infinite and, by changing the initial conditions of the network, we obtain different Petri nets

and fragmentation process. We notice that the Petri net is potentially infinite, because every time we change the initial concentration of the system we change the set of differential equations, describing the system.

We change the mass of the system, in a wide range of values, to study the concentration of monomers at the steady state. As we can notice in Fig. 3, we find out that, assuming as constant the coefficient rates $a = 1$ and $b = 1$, the concentration of C_1 tends to 1 at the steady state, even with very different initial conditions: we change the initial concentration of C_1 in the range $[5, 1500]$.

Then, we proceed applying the definition of β-robustness (Definition 7). Considering as initial conditions the interval $C_1 = [100, 300]$. By looking at the simulation results, we find that at the steady state, the concentration of C_1 is within the interval and $[0.91, 0.95]$.

We calculate the normalized α-robustness (Definition 5), obtaining:

$$n_O = \frac{\alpha}{k} = \frac{0.04}{0.93} = 0.04.$$

The normalized input values is calculated as follows:

$$n_I = \frac{m_I}{k_I} = \frac{[300 - 100]}{\frac{[300+100]}{2}} = 1.$$

Then, the relative β-robustness is trivially calculated as follows:

$$\beta - robustness = \frac{n_O}{n_I} = 0.04.$$

In Table 1, we summarize the initial concentration of C_1, its concentrations at the steady state, the α and the β-robustness. We notice that increasing the initial concentration of the input, we obtain that the α-robustness of the model tends to 0. This result leads us to formalize a new notion of robustness, namely the *asymptotic robustness*.

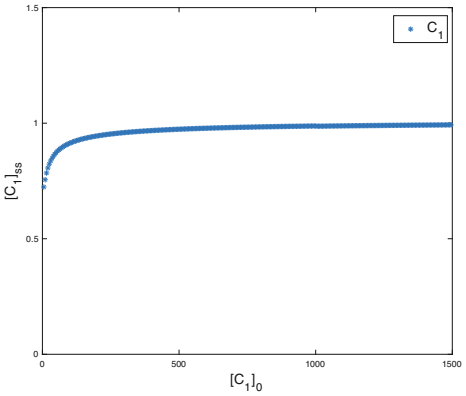

Fig. 3. Simulations result of Becker-Döring model. We plot on the horizontal axis the initial concentration of C_1, in a range $[5, 1500]$, and on the vertical axis the concentration of C_1 at the steady state. We assume the coefficient rates a and b are constant and equal to 1

Definition 8 (Asymptotic Robustness). Let N, α and k be as in Definition 4. A Petri net N with output place O is *asymptotically robust* iff as $k \to \infty$, $\alpha \to 0$.

Table 1. The intervals, the concentrations of monomer reached at the steady state, and the value of the α and β-robustness

Intervals	Steady state concentration	α-robustness	β-robustness
[100, 300]	[0.91, 0.95]	0.03	0.04
[300,500]	[0.95, 0.97]	0.02	0.04
[500,700]	[0.97, 0.98]	0.01	0.03
[700,900]	[0.9806, 0.985]	0.005	0.004

The simulations about the asymptotic robustness persuade us to analytically study the solution of the steady state formula because we want to verify this particular systems behaviour formally. However, this approach cannot be applied indistinctly because of the biological systems' complexity. Indeed, in most cases, it is not possible to derive analytically the steady state formula.

3.3 Analysis of the Steady State

In the simulation result shown in Fig. 3, the concentration of monomer C_1 appears to be robust at the steady state. Indeed, we change, in a wide range of values, the input of the system (C_1) and we notice that the output (C_1) tends

to 1 at the steady state, considering the coefficient rates a and b constant and equal to 1.

Then, in line with the other works presented in literature [3,18,22], our main result in the analytical study of steady state is the following:

Theorem 1 (Monomers Steady State). *Let a and b be the coefficient rates of coagulation and fragmentation process in the Becker-Döring system, ρ the mass of the system and $[C_1]_{ss}$ the concentration of monomers at the steady state. Then, as $\rho \to \infty$, $[C_1]_{ss} \to \frac{b}{a}$.*

Proof As described in Sect. 2.2, the crucial assumption of the Becker-Döring model is mass conservation, which therefore depends on the initial concentration of monomers, hence the mass ρ at the initial state will remain the same at the steady state, that is the sum of the fluxes of the species goes to zero. Then, recalling Formula 7, we can write:

$$\rho(0) = \rho(\infty) = \sum_{i=1}^{k} i \cdot C_i, \tag{8}$$

where k is the maximum number of molecules in the system, hence it is $k = \rho$. Generalizing the fluxes formulas, as in example 6, we deduce the general steady state formula for a cluster of dimension i, as follows:

$$[C_i]_{ss} = \left(\frac{a}{b}\right)^{i-1} [C_1]_{ss}^{i}. \tag{9}$$

Replacing (9) in (8), we obtain:

$$\begin{aligned} \rho &= \sum_{i=1}^{k} i \left(\frac{a}{b}\right)^{i-1} [C_1]_{ss}^{i} \\ &= \frac{b}{a} \sum_{i=1}^{k} i \left(\frac{a}{b}[C_1]_{ss}\right)^{i} \\ &= \frac{b}{a} \frac{k(\frac{a}{b}C_1)^{k+2} - (k+1)(\frac{a}{b}C_1)^{k+1} + (\frac{a}{b}C_1)}{(1-\frac{a}{b}C_1)^2}. \end{aligned} \tag{10}$$

We want to study the asymptotic behaviour of C_1, then we define:

$$\lim_{t\to\infty} C_1(t) := x_k.$$

Equating (10) with its general form, we get:

$$k = x + \frac{a}{b}2x^2 + \ldots + \frac{a^{k-1}}{b^{k-1}}kx^k = \frac{b}{a} \cdot \frac{k \cdot \frac{a}{b}x^{(k+2)} - (k+1) \cdot \frac{a}{b}x^{(k+1)} + \frac{a}{b}x}{(1-\frac{a}{b}x)^2}.$$

We can rearrange the previous expression as follows:

$$1 = \frac{x}{k} + \frac{a}{b}\frac{2}{k}x^2 + ... + \frac{a^{k-1}}{b^{k-1}}x^k = \frac{b}{a} \cdot \frac{k \cdot \frac{a}{b}x^{(k+2)} - (k+1) \cdot \frac{a}{b}x^{(k+1)} + \frac{a}{b}x}{k \cdot (1 - \frac{a}{b}x)^2}, \quad (11)$$

and we define:

$$\begin{aligned} g_k(x) &= \frac{x}{k} + \frac{a}{b}\frac{2}{k}x^2 + ... + \frac{a^{k-1}}{b^{k-1}}x^k, \\ &= \frac{b}{a} \cdot \frac{k \cdot \frac{a}{b}x^{(k+2)} - (k+1) \cdot \frac{a}{b}x^{(k+1)} + \frac{a}{b}x}{k \cdot (1 - \frac{a}{b}x)^2}. \end{aligned}$$

We will need both expressions because they make it simpler to observe different properties.

- $g_k(x)$ is an increasing function of x
- with simple algebraic manipulation, we get:

$$g_k(x) = \frac{1}{k} \cdot \frac{b}{a}\left(\frac{a}{b}x + \frac{a^2}{b^2}2x^2 + ... + k\frac{a^k}{b^k}x^k\right)$$

  from which it is easy to see that:

$$g_k\left(\frac{b}{a}\right) = \frac{1}{k} \cdot \frac{b}{a}(1 + 2 + ... + k) = \frac{b}{a} \cdot \frac{k+1}{2}.$$

- Following the first two items, we notice that there is one and only one solution for the equation $g_k(x) = 1, \forall k$, in the range $\left[0, \frac{b}{a}\right]$.
- Looking at the other expression for the function $g_k(x)$, we find that

$$\lim_{k \to \infty} g_k(x) = 0, \quad \forall x \in \left[0, \frac{b}{a}\right).$$

- Because of the previous limit, if we now take a generic $x^* < \frac{b}{a}$, then, for k large enough we will have $g_k(x^*) < g_k(x_k) \equiv 1$ Since the function is increasing and monotonic with respect to x, then:

$$\frac{g_k(x_k) - g_k(x^*)}{x_k - x^*} > 0 \;\Rightarrow\; x_k > x^*.$$

  This shows that, for k large enough, $x_k > x^*, \forall x^* < \frac{b}{a}$. Therefore $x_k \geq \frac{b}{a}$.

For an intuitive visualization of what just said, in Fig. 4 we show a plot of the $g_k(x)$ in the particular case where $a = b$.

Our result is shown also in Fig. 3, where we consider as rates of the system $a = b$.

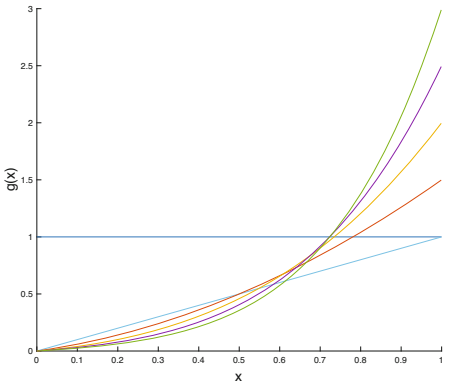

Fig. 4. The plot of the function $g(x)$. Graphically, we see that the curves of function $g(x)$ will be closer to zero, increasing the value of k

4 Conclusion and Future Work

In this paper we focused on the study of robustness of Becker-Döring equations [2], a model that describes condensations phenomena at different pressures. First, we show that the definition given by Shinar and Feinberg is particularly limiting in the description of the BD model's behavior. Then, by applying our definitions of α and β robustness, we show how the BD model is robust with respect to the perturbation on the initial concentration of monomers.

Concerning this, we prove that the concentration of monomer tends to the ratio of coefficients at the steady state and we show this result by simulations. This result leads us to introduce a new notion of robustness, namely the asymptotic robustness. This new notion intends to describe how a system becomes more robust increasing the initial concentration of the input. In order to support our result, in this case, we studied analytically the solution of the steady state formula. Note that, as remarked in the text, it is not always possible to solve the differential equations system.

We have analysed so far the Becker-Döring equations assuming that the coefficient rates of agglomeration and fragmentation are constant real values. We could improve this analysis introducing some physical aspects of the model. As described in [3,22,26], the size and the shape of clusters involved in the reactions influence the dynamical properties of the system. The next step of this preliminary research will be to simulate this system using real rates from specific application domains, useful to describe biological phenomena such as protein aggregation in neurodegenerative diseases.

References

1. Angeli, D., De Leenheer, P., Sontag, E.D.: On the structural monotonicity of chemical reaction networks. In: 2006 45th IEEE Conference on Decision and Control, pp. 7–12. IEEE (2006)

2. Ball, J.M., Carr, J.: Asymptotic behaviour of solutions to the Becker-döring equations for arbitrary initial data. Proc. R. Soc. Edinb. Sect. A Math. **108**(1–2), 109–116 (1988)
3. Ball, J.M., Carr, J., Penrose, O.: The Becker-döring cluster equations: basic properties and asymptotic behaviour of solutions. Commun. Math. Phys. **104**(4), 657–692 (1986)
4. Barbuti, R., Gori, R., Levi, F., Milazzo, P.: Specialized predictor for reaction systems with context properties. In: Suraj, Z., Czaja, L. (eds.) Proceedings of the 24th International Workshop on Concurrency, Specification and Programming, Rzeszow, Poland, September 28–30, 2015. CEUR Workshop Proceedings, vol. 1492, pp. 31–43. CEUR-WS.org (2015)
5. Barbuti, R., Gori, R., Milazzo, P., Nasti, L.: A survey of gene regulatory networks modelling methods: from differential equations, to Boolean and qualitative bioinspired models. J. Membr. Comput. **2**(3), 207–226 (2020)
6. Becker, R., Döring, W.: Kinetische behandlung der keimbildung in übersättigten dämpfen. Ann. Phys. **416**(8), 719–752 (1935)
7. Blanchini, F., Franco, E.: Structurally robust biological networks. BMC Syst. Biol. **5**(1), 74 (2011)
8. Burton, J.: Nucleation theory. In: Berne, B.J. (eds.) Statistical Mechanics. Modern Theoretical Chemistry, vol. 5, pp. 195–234. Springer, Boston, MA (1977). https://doi.org/10.1007/978-1-4684-2553-6_6
9. Feinberg, M.: Chemical reaction network structure and the stability of complex isothermal reactors-i. the deficiency zero and deficiency one theorems. Chem. Eng. Sci. **42**(10), 2229–2268 (1987)
10. Gori, R., Milazzo, P., Nasti, L.: Towards an efficient verification method for monotonicity properties of chemical reaction networks. In: Bioinformatics, pp. 250–257 (2019)
11. Gunawardena, J.: Chemical reaction network theory for in-silico biologists. Notes (2003). http://vcp.med.harvard.edu/papers/crnt.pdf
12. Hingant, E., Yvinec, R.: Deterministic and stochastic Becker–döring equations: past and recent mathematical developments. In: Holcman, D. (ed.) Stochastic Processes, Multiscale Modeling, and Numerical Methods for Computational Cellular Biology, pp. 175–204. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-62627-7_9
13. Hoze, N., Holcman, D.: Coagulation-fragmentation for a finite number of particles and application to telomere clustering in the yeast nucleus. Phys. Lett. A **376**(6–7), 845–849 (2012)
14. Hozé, N., Holcman, D.: Modeling capsid kinetics assembly from the steady state distribution of multi-sizes aggregates. Phys. Lett. A **378**(5–6), 531–534 (2014)
15. Hoze, N., Holcman, D.: Kinetics of aggregation with a finite number of particles and application to viral capsid assembly. J. Math. Biol. **70**(7), 1685–1705 (2014). https://doi.org/10.1007/s00285-014-0819-2
16. Kitano, H.: Biological robustness. Nat. Rev. Genet. **5**(11), 826–837 (2004)
17. Krapivsky, P.L., Redner, S., Ben-Naim, E.: A Kinetic View of Statistical Physics. Cambridge University Press, Cambridge (2010)
18. Kreer, M.: Classical Becker-döring cluster equations: rigorous results on metastability and long-time behaviour. Ann. Phys. **505**(4), 398–417 (1993)
19. Nasti, L.: Verification of Robustness Property in Chemical Reaction Networks. Ph.D. thesis, Universitità di Pisa, Dipartimento di Informatica (2020)

20. Nasti, L., Gori, R., Milazzo, P.: Formalizing a notion of concentration robustness for biochemical networks. In: Mazzara, M., Ober, I., Salaün, G. (eds.) STAF 2018. LNCS, vol. 11176, pp. 81–97. Springer, Cham (2018). https://doi.org/10.1007/978-3-030-04771-9_8
21. Nasti, L., Gori, R., Milazzo, P., Poloni, F.: Efficient analysis of chemical reaction networks dynamics based on input-output monotonicity. arXiv preprint arXiv:2107.00289 (2021)
22. Penrose, O.: Metastable states for the Becker-döring cluster equations. Commun. Math. Phys. **124**(4), 515–541 (1989)
23. Saunders, T.E.: Aggregation-fragmentation model of robust concentration gradient formation. Phys. Rev. E **91**(2), 022704 (2015)
24. Shinar, G., Feinberg, M.: Structural sources of robustness in biochemical reaction networks. Science **327**(5971), 1389–1391 (2010)
25. Shinar, G., Feinberg, M.: Design principles for robust biochemical reaction networks: what works, what cannot work, and what might almost work. Math. Biosci. **231**(1), 39–48 (2011)
26. Slemrod, M.: The becker-döring equations. In:Bellomo, N., Pulvirenti, M. (eds.) Modeling and Simulation in Science, Engineering and Technology, pp. 149–171. Springer, Boston (2000). https://doi.org/10.1007/978-1-4612-0513-5_5
27. Wolpert, L.: Positional information and the spatial pattern of cellular differentiation. J. Theor. Biol. **25**(1), 1–47 (1969)