



HAL
open science

Inverse populations balance problems: heuristics for aggregation/breakage of solid clusters in liquid metals

Jean-Marc Auger, Sylvain Martin, Frédéric Gruy

► To cite this version:

Jean-Marc Auger, Sylvain Martin, Frédéric Gruy. Inverse populations balance problems: heuristics for aggregation/breakage of solid clusters in liquid metals. *Industrial and engineering chemistry research*, 2020, 59 (22), pp.10363-10378. 10.1021/acs.iecr.0c00859 . hal-02572811

HAL Id: hal-02572811

<https://hal.science/hal-02572811>

Submitted on 16 Jul 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

**Inverse populations balance problems: heuristics for aggregation/breakage of
solid clusters in liquid metals**

Jean-Marc Auger a, Sylvain Martin b, Frédéric Gruy b

a j-mauger@laposte.net

b Mines Saint-Étienne, Univ Lyon, CNRS, UMR 5307 LGF, Centre SPIN, 42023 Saint-Étienne, France

ABSTRACT

Populations balance models have been applied to the simultaneous aggregation and breakage of solid clusters in agitated liquid metal. Challenges for these systems lie in kinetics constants determination (kernels) from experimental clusters concentrations measurements. However, due to difficulties inherent to molten metals experimentation (severely limiting available data points and replication), and lack of literature information concerning the physical phenomena involved, more advanced inverse methods were needed. Taking into account physical characteristics, theoretical approaches using reactions networks proved the existence of at least one stable positive equilibrium state. This result allowed the construction of two distinct fitting algorithms, aimed at solving the corresponding inverse problem (kernels determination from experimental data). The first of these heuristic methods accurately identifies kernels parameters from perfect steady state data, while the second, based on transitory states, reliably leads to correct results with as few data points as 2, and measurements errors as high as 5%.

KEYWORDS

populations balance, aggregation, breakage, reaction networks, inverse methods, kernels, heuristics

1 Introduction

Populations balances are a common class of models with a wide range of applications, from biology and population ecology, to chemical engineering ¹. This paper will address specifically the use of populations balance models in the context of particulate clusters aggregation and breakage in agitated liquid media.

Generally speaking, populations balance problems are characterized by three main elements:

- a set of boundary conditions (initial state, eventual entrance or exit fluxes for clusters)
- a reaction network (describing every possible interaction between clusters)
- a set of kernels (linking reactions kinetics to physical data and parameters) – in our case, the aggregation kernel will describe the mechanisms by which clusters colliding through liquid drag merge into larger clusters, while the breakage kernel will address clusters separation under liquid turbulence

While literature shows numerous examples of kernels models built upon theoretical considerations, experimental validations are notoriously less frequent and much more complicated to implement. Determination of experimental kernels characteristics (kernel model type, and the corresponding physical parameters) is thus a difficult but critical task, needed to advance comprehension in this domain. Usual methods involve optical in-situ measurements of clusters concentrations during aggregation experiments, with the resulting concentrations evolutions with time being matched to those provided by populations balance simulations. While this approach is suitable for ordinary transparent media (such as water and most common solvents), the opacity of other fluids, such as liquid metals, makes it highly impractical, when at all possible. This obstacle is compounded by the uneasy transposition of results obtained in simple cases to more difficult ones, owing to:

- considerable quantitative changes in fluid properties (viscosity – ranging from 2×10^{-4} for Rb to 5.5×10^{-3} Pa.s for Co-, volumetric mass...)
- qualitative modifications of involved physics: wettability (and thus surface tension), for instance, has differing meaning and mechanisms in water and liquid steel (due to the different bonds – hydrogen or metallic – present in these two media)

General investigations of the inverse problem linked to population balances (e.g. the determination of aggregation or breakage kernels from clusters concentrations data) revealed it as ill-posed, adding theoretical difficulties to the experimental ones. Following works from Narsimhan et al.², Sathyagal et al.³, and Wright et al.⁴, many studies managed to devise successful protocols for kernels determination under the hypothesis of scaling behavior. Other regularization approaches involved approximation of kernel expressions with parametric functions such as Hermite splines⁵, or by using the method of moments^{6,7}. All these procedures typically require access to a large amount of experimental data, either to validate their hypotheses, or to carry out their calculations. Sensitivity to experimental noise has only been investigated in some rare studies, such as⁸ or⁹ (limited to elementary kernels forms: sum, product, and constant). The corresponding observations seem to imply that most solving methodologies for inverse population balances additionally suffer from ill-conditioning.

In this work, we will provide tools and algorithms to determine kernels characteristics (general kernels models, then orders of magnitude for their parameters), *in the context of realistic laboratory model experiments*, and under the most commonly encountered conditions in liquid metals:

- non repeatability of initial conditions; while some experimental conditions such as temperature or media hydrodynamics can easily be imposed, *controlled* introduction of particles in liquid metals has been shown by decades of research to be a currently impossible task (^{10,11, 12, 13, 14, 15, 16} providing examples of such attempts). The initial clusters concentrations distribution can thus be measured, but not set.

- *limited number of experimental points*; this restriction follows from the use of samples for measurements, typically through metallographic analysis. As sampling disturbs the media, it must thus be kept to a minimum (up to a maximum 3 or 4 samples per experiment).

- concentrations measurements errors: metallographic analyses are carried out on 2-dimensional cross-sections from sampled ingots, while clusters are 3-dimensional structures, leading to unavoidable errors. Furthermore, the method is destructive, forbidding measurements duplication to lessen uncertainty through statistical methods. X-ray tomography could provide similar data but seems more appropriate to study individual cluster 3D characteristics (as shown in¹⁷) than global concentration distributions.

The second part of this paper will detail the formalism and hypotheses used for populations balance problems under the aforementioned conditions, then discuss the existence and nature of potential steady states. The third part will concern itself with new heuristic algorithms to determine kernels from realistic experimental data, and their limitations.

2 Theoretical analysis of the population balance problem

Chemical kinetics formalism will be used for both aggregation and breakage reactions, even though these phenomena are physical in nature for the cases considered.

2.1 Model hypotheses:

H1 Quantization of clusters size

Existence of a minimal cluster size is postulated, and all clusters are considered aggregates of these elementary particles. This hypothesis is both realistic (as monocrystalline oxide can be considered discrete, and ultimately is at atomic scale) and not (aggregation phenomena and properties for atoms differ widely from those of macroscopic clusters, leading to transitions in models long before reaching atomic scale). However, experimental use of non-reactive (thus unbreakable) elementary particles of calibrated size coincides well with this modeling.

H2 Continuity of clusters concentrations

While classically used, this approach is highly unrealistic, as single clusters can only react in full, and concentrations values should thus be discrete. However, literature generally postulates that long term stochastic variations are similar to the corresponding deterministic steady state, even if formal proofs have only been given under very restrictive conditions¹⁸.

H3 Closed system / elementary particles conservation

While highly unrealistic for most practical cases, this hypothesis can be justified in the context of model experiments where unitary particles are introduced once in the medium and not removed. Results from these simplified experiments could be transposed to more complex situations. One example for such adaptation would be to take clusters entrance into account via additional nucleation mechanisms. From this hypothesis, with V the total volume considered, and n the total number of elementary particles, conservation Eq. 1 can be deduced.

$$\sum_{i=1}^{i=\infty} i \cdot [i] = \sum_{i=1}^{i=N \leq n} i \cdot [i] = \frac{n}{V} \quad (1)$$

Following from hypotheses H1, H3, and the necessary finite volume of experimental devices, one can deduce the existence of a maximum cluster size $N=n$ (corresponding to the totality of elementary particles aggregated as a single cluster). An interesting consequence is that simulations involving finite size kernels are sufficient to completely describe the aggregation/breakage problem as formulated in Eq. 2.

$$\forall (i,j) \in \mathbb{N}^2, \quad i+j > N, \quad \left\{ \begin{array}{l} k_{ij} = 0 \\ k_{ij}^* = 0 \end{array} \right\} \quad (2)$$

Using the notations from symbols table, the most general expression for quantized aggregation/breakage population balance problems is thus given by Eq. 3.

$$\dot{[i]} = \sum_{j=1}^{\lfloor \frac{i}{2} \rfloor} \phi_{j,i-j} - \sum_{j=1}^{j=N-i} (1 + \delta_{ij}) \cdot \phi_{ij} \quad (3)$$

H4 Aggregation and breakage phenomena as binary reversible reactions (see Eq. 4).

Reversibility accounts for some unusual aspects of this work, as pure aggregation or breakage have been studied separately in several instances (¹⁹, ²⁰, ²¹, or ²², for example), but much more scarcely together. This hypothesis implies the existence of at least one reversible aggregation reaction in the system, but does not assume reversibility for each.

$$R_{i,j}: \quad i + j \Leftrightarrow (i+j), \quad \forall (i,j) \in \mathbb{N}^2 \quad (4)$$

H5 First order kinetics for both aggregation and breakage reactions (e.g. use of mass action kinetics).

This hypothesis is commonly used for diluted suspensions, and expressed as Eq. 5, with k_{ij} and k_{ij}^* the kinetics coefficients corresponding to aggregation and breakage kernels respectively. H5 also implies that kinetics coefficients are independent from clusters concentrations.

$$\phi_{ij} = k_{ij} \cdot [i] \cdot [j] - k_{ij}^* \cdot [i+j], \quad \forall (i,j), \quad i+j \leq N \quad (5)$$

2.2 Existence of equilibrium states

Theorems developed for chemical reactions networks can be used to evaluate the existence and uniqueness of equilibrium states for coupled chemical reactions systems. The peculiarity of this approach resides in its independence from kinetics coefficients and initial conditions, being based on the structure of reactions networks alone. While no general method was discovered to determine networks equilibrium properties, various theorems have been devised to conclude under more restrictive conditions (summarized for instance in ²³ or ²⁴). Examples include deficiency zero and deficiency one theorems ²⁵, higher deficiency algorithm ²⁶, or injectivity ²⁷ and concordance ²⁸ tests.

In the context of chemical reactions networks, *species* refer to pure chemical species (in our case, a species includes all clusters of a given size), while *complexes* correspond to sets of species appearing on at least one side of a reaction (they are thus every single clusters sizes and every couples of 2 clusters sizes).

Reaction graphs (see ²⁹ for other possible representations) corresponding to the population

balance problem with the aforementioned hypotheses are summarized in Figure 1. General properties and concepts pertaining to these networks will be detailed in the following sections:

2.2.1 Reversibility

Weak reversibility defines the existence, for each linked complexes couple (C1, C2), of oriented paths linking C1 to C2, and C2 to C1. As population balance reaction graphs present no cycle, the network is either fully reversible (when each effective reaction is reversible), or not reversible at all (when at least one reaction is not reversible).

Weak reversibility is equivalent to complete reversibility for the population balance networks studied.

2.2.2 Number of linkage classes l

A linkage class is defined as a set of complexes that are linked together through reactions. In the case of population balance problems, there are $(N-1)$ linkage classes (as expressed in Eq. 6), each one corresponding to all reversible reactions leading to the formation of each clusters size ranging from 1 to $(N-1)$.

$$l(N) = N - 1 \quad (6)$$

2.2.3 Number of complexes m

As visible in Figure 1, the reaction network contains $(N-1)$ single species complexes, and k binary complexes per linkage class linked to k -size single clusters complex, summed in Eq. 7.

$$m(N) = (N - 1) + \sum_{k=2}^{k=N} \left\lfloor \frac{k}{2} \right\rfloor = (N - 1) + \left\lfloor \frac{N}{2} \right\rfloor \cdot \left(N - \left\lfloor \frac{N}{2} \right\rfloor \right) \quad (7)$$

2.2.4 Stoichiometric sub-space dimension s

The stoichiometric sub-space is defined as the vector space spanned by all chemical reactions belonging to the network. In the case of population balance problems, the easiest way to determine its dimension is to exhibit a (mathematical) basis for this space.

It can be noticed that each reaction of the network defined in Eq. 4 can be decomposed in a sum of reactions involving the aggregation of single elementary particle to existing clusters, as expressed in Eq. 8.

$$R_{i,j} = \sum_{l=j}^{k=i+j-1} R_{1,l} - \sum_{k=1}^{k=i} R_{1,k} \quad (8)$$

These $(N-1)$ reactions are independent, and thus form a basis for the stoichiometric sub-space, leading to the conclusion (Eq. 9) that its dimension equals $(N-1)$:

$$s(N) = N - 1 \quad (9)$$

2.2.5 Stoichiometric compatibility class

A stoichiometric compatibility class is defined as the set of clusters concentrations distributions that can be obtained from a given initial state (initial concentrations distribution) through any combination of the network reactions. Due to the H3 hypothesis of total elementary particles conservation, the stoichiometric compatibility class associated to a starting concentrations distribution with a total concentration of unitary particles n/V , is a subset of concentrations distributions with (obviously) the same concentration of elementary particles of n/V .

From the following observations:

- the set of concentrations distributions with n/V elementary particles concentration is of dimension $(N-1)$ (N clusters sizes concentrations which are linked by Eq. 1)
- any stoichiometric compatibility class of the population balance problem shows the same dimension as the stoichiometric sub-space, which was determined in the preceding paragraph as $(N-1)$

It can be concluded about the *identity between the stoichiometric compatibility class associated with any clusters concentrations distribution totaling n/V elementary particles per volume unit, and clusters concentrations distributions with n/V total elementary particles concentration.*

The stoichiometric compatibility class of n/V elementary particles concentration can be represented, in the vector space of clusters concentrations distributions, by a simplex bounded by N vertices (each corresponding to a concentrations distribution with a unique non-zero clusters size concentration).

2.2.6 Deficiency

Deficiency is an important function of reaction networks, as its values can help determine the existence and uniqueness of equilibrium states. Its expression (using definitions and notations detailed previously) is given as Eq. 10.

$$\delta = m - l - s \quad (10)$$

This allows the determination of deficiency as a function of maximum cluster size N (Eq. 11).

$$\delta(N) = \left\lfloor \frac{N}{2} \right\rfloor \cdot \left(N - \left\lfloor \frac{N}{2} \right\rfloor \right) - N + 1 \quad (11)$$

Deficiency will be evaluated for populations balance networks as a function of maximum cluster size N, and its consequences in terms of equilibrium states will be detailed, using deficiency zero and one theorems (as reminded below, from ²⁵) for simple cases, and properties of embedded networks when N value increases:

Deficiency Zero Theorem

If a mass action system is applied to a weakly reversible and deficiency zero reaction network, then, independently from kinetics constants, the system has exactly one equilibrium concentration in each positive stoichiometric compatibility class.

Deficiency One Theorem

If a weakly reversible reaction network shows the three following properties:

- deficiency for each of its linkage classes is less than 2
- deficiency of the whole network equals the sum of its linkage classes deficiencies
- each of its linkage class contains only one terminal strongly linked component

Then every mass action system permitted by the network admits a positive equilibrium.

2.2.6.1 N=2 case

Deficiency equals zero. If the corresponding network is weakly reversible (e.g. the only reaction in the network is reversible – *all kinetics constants are strictly positive*), according to zero deficiency theorem, the reaction network for these population balance problems has then exactly one stable positive equilibrium state.

2.2.6.2 N=3 case

Deficiency equals zero. If the corresponding network is weakly reversible (e.g. the 2 reactions in the network are reversible – *all kinetics constants are strictly positive*), according to zero deficiency theorem, the reaction network for these population balance problems has then exactly one stable positive equilibrium state.

2.2.6.3 N=4 case

Deficiency equals one. According to deficiency one theorem²⁵, deficiency of each linkage class in the network should be evaluated in order to prove existence and/or uniqueness of equilibrium states. From the data summarized in Table 1, each linkage class has deficiency equal to zero, while the sum of all deficiencies is one, which precludes the use of deficiency one theorem. However, the deficiency one algorithm³⁰, used through the CRNToolbox implementation (developed by Martin Feinberg Group for Chemical Reaction Network Theory at The Ohio State University), can successfully determine that such networks admit either exactly one stable positive equilibrium state, or none, depending on the reversibility of individual reactions, as summarized in Table 2.

2.2.6.4 N>4 case

For N greater than 4, the corresponding deficiency is greater than 1 and increasing with N. Simple theorems and classical approaches (such as^{31, 32, 33, 34, 35, 36, 23, 37, 38}, or²⁹) cannot be applied anymore. According to³⁹, it is however possible to conclude on the existence of steady states using the following property: a reaction network admits at least the same number of stable positive equilibrium states as one of its embedded networks with which it shares the exact same stoichiometric sub-space. In the context of reactions networks, an embedded network is constructed from an initial network by the removal of any number of reactions and/or chemical species.

An example of such an embedded network applied to population balance is presented in Figure 2, and obtained through the removal of every reaction not corresponding to the aggregation of single elementary particles to clusters. As presented before, dependency of characteristic functions for the embedded network can be expressed in Eq. 12, Eq. 13, and Eq. 14 with N as a variable.

$$m(N) = 2 \cdot (N - 1) \quad (12)$$

$$l(N) = N - 1 \quad (13)$$

$$s(N) = N - 1 \quad (14)$$

Deficiency can then be expressed in Eq. 15 as a function of the maximum cluster size.

$$\delta(N) = m(N) - l(N) - s(N) = 0 \quad (15)$$

By construction, the embedded network stoichiometric sub-space is included into the original network stoichiometric sub-space. Furthermore, these sub-spaces share the same dimension (Eq. 9 and Eq. 14), and are subsequently one and the same. Consequently, *under the condition of embedded network weak reversibility* (equivalent to the reversibility of the aggregation reactions of single elementary particles to clusters, with this choice of embedded network), zero deficiency theorem applies to the embedded network, which then admits exactly one stable positive equilibrium state. From ³³, the reaction network for population balance problems has then at least one stable positive equilibrium state.

It should be noted that such reasoning can be applied to any embedded network comprised of exactly one (effectively) reversible reaction per linkage class of the original network.

As a conclusion for $N > 4$ population balance problems, the existence of at least one stable positive equilibrium state is guaranteed, provided (sufficient condition) there is at least one effectively reversible aggregation/breakage reaction leading to the formation of each clusters size ranging from 2 to N .

2.3 Multiplicity of equilibrium states

An example will show the potential for multiple equilibrium states in a given population balance problem. Kernel data are summarized in Table 3, and the corresponding boundary conditions are following: maximum cluster size of $N=5$, total elementary particles concentration of 1 particle per volume unit. In this case, two equilibrium states shown in Table 4 are allowed, depending on initial conditions (e.g. starting clusters concentrations distribution) of the problem.

In order to assess the dependency of equilibrium states multiplicity with boundary conditions (and especially total elementary particles concentration), population balance simulations were carried out (until steady state) using this

same kernel and multiple starting points inside various stoichiometric compatibility classes. As demonstrated before, in the case of population balance problems verifying H3 hypothesis, a compatibility class corresponds to all clusters concentrations distributions with a given elementary particles total concentration.

As mentioned in 1.2.5, each stoichiometric compatibility class can be seen as a simplex in the clusters concentrations distributions vector space. As a way to maximize the regularity of simulation starting points, they were chosen as evenly spaced points on each edge of the simplex.

Edges are defined as linking two distinct vertices of the simplex, which in turn correspond to clusters concentrations distributions with only one non-zero concentration. Concentrations distributions belonging to an edge can thus be expressed as Eq. 16.

$$(i,j) \in (1:N=5)^2, i < j, \left\{ \forall l \in [0,1], \left([1]=0, \dots, [i]=\frac{n}{V \cdot i} \cdot l, \dots, [j]=\frac{n}{V \cdot j} \cdot (1-l), \dots, [N]=0 \right) \right\} \quad (16)$$

Five evenly spaced concentrations distributions on each edge for every compatibility class have been used as starting points. For N=5, the simplex has 10 edges, for a total of 35 starting concentrations distributions, defined by Eq. 17.

$$(i,j) \in (1:N=5)^2, i < j, \left\{ \forall l \in \left\{ 0, \frac{1}{4}, \frac{1}{2}, \frac{3}{4}, 1 \right\}, \left([1]=0, \dots, [i]=\frac{n}{V \cdot i} \cdot l, \dots, [j]=\frac{n}{V \cdot j} \cdot (1-l), \dots, [N]=0 \right) \right\} \quad (17)$$

From these initial distributions, steady states were computed through population balance simulations, using tools detailed in part 3: “Kernels determination through experimental data fitting”. Three consecutive time step sizes ($2 \cdot 10^{-3}$ s; $2 \cdot 10^{-6}$ s; $2 \cdot 10^{-9}$ s), have been applied 10^4 times each. Final instant concentrations variations were verified to be less than 10^{-10} particles per volume unit and second, confirming that steady states were actually reached.

Comparison between equilibrium states derived from starting concentrations distributions *corresponding to the same stoichiometric class* has been carried out through Bhattacharyya distance calculations. The upper bound of computed distances for each class (equivalent to a given total elementary particles concentration value) is represented in Figure 3. As visible on the figure, low and high values of particles concentration are characterized by a single equilibrium state, while intermediary values allow the existence of multiple (in fact two) balance states. Since the kernel used is

the same for all concentrations, this last parameter alone cannot be used to determine the multiplicity of solutions for population balance problems with clusters sizes higher than 4. *As a conclusion, while the reactions network structure allows the possibility of multiple equilibrium states for N greater than 4, boundary conditions (total particles number, maximum clusters size) impose the actual number of steady states.*

2.4 Kernel models

Apart from the 3 classical elementary kernel forms (namely, constant, sum, and product), which have the advantage of providing analytical solutions ⁴⁰, but do not correspond to any realistic experimental setup, several kernels formalisms have been developed for various applications (coagulation of solid suspensions in liquids, polymerization kinetics, liquid drops coagulation in gas...). Generally speaking, aggregation kernels were far more investigated than breakage ones, hence explaining their higher mathematical complexity and number of variants.

In the limited scope of solid clusters coagulation and breakage in liquids, 6 general kernel classes (3 for aggregation, and 3 for breakage) can be derived from literature reviews (⁴¹, for example). They are detailed in Table 5. Each of these mathematical expressions corresponds to several physical models (for example, Kruis-Kuster and Saffman-Turner models are both variants of the turbulence type 1 general model), the physical meanings of coefficients varying wildly with the hypotheses used in their construction.

As easily visible from the expressions in Table 5, aggregation (respectively breakage) kernel coefficients k_{ij} (respectively k_{ij}^*) can locally (e. g. for an (i,j) clusters couple) take a value of zero, only if the corresponding kernel is a constant and null. From this observation, it can be deduced that either each aggregation reaction is reversible, or none is. In combination with H4 hypothesis, it can be concluded that classical kernel expressions lead to reversible reactions exclusively, which in turn (using demonstrations from the preceding paragraphs) ensures the existence at least one stable positive equilibrium state, independently from maximum clusters size.

2.5 Nature of equilibrium states

The most general form of equilibrium is the complex balanced state, which is characterized by the lack of evolution of complexes (and consequently species) concentrations with time. A more restrictive case is the so-called detailed

balanced state, in which each elementary aggregation/breakage reaction is balanced. As such, detailed balanced states are much simpler to investigate theoretically, since they can be studied through the classical tools of single reaction kinetics. We will identify some necessary conditions for the existence of detailed balanced state in the context of population balance problems.

If an equilibrium state is detailed balanced, using the chemical kinetics formalism, and *under the hypothesis of reversibility for each reaction*, a general expression for equilibrium constants is given by the Eq. 18 system.

$$\left\{ K_{ij} = \frac{[(i+j)]_{eq}}{[i]_{eq} \cdot [j]_{eq}} = \frac{k_{ij}}{k_{ij}'} \right\}_{(i=1 \dots N-1, j=1 \dots i, i+j \leq N)} \quad (18)$$

Easily verified recursive reasoning using Eq. 18 gives relationships between generic equilibrium constants, and constants corresponding to the reactions involving the aggregation of elementary particles. The resulting expressions in Eq. 19 constitute a necessary condition for detailed balanced states.

$$\left\{ K_{ij} = \frac{\prod_{l=\max(i,j)}^{i+j-1} K_{1l}}{\prod_{k=1}^{\min(i,j)-1} K_{1k}} \right\}_{(i,j) \in |\mathbb{R} \setminus \{1\}|^2} \quad (19)$$

A simpler incremental formulation, as shown in Eq. 20, can be used as a criterion to rule out the possibility of detailed balanced states. If this criterion is different from 1 for at least one (i,j) couple in a given kernel, balanced states can only be of the complex variety for this kernel.

$$\left\{ \frac{K_{1,i} \cdot K_{i+1,j}}{K_{1,i+j} \cdot K_{i,j}} = 1 \right\}_{(i,j), i+j < N} \quad (20)$$

Formal expressions for this criterion have been calculated for each of the nine combinations of physical kernel models detailed in the preceding paragraphs. As easily visible in Table 6, criterion equality with 1 can only be satisfied when the maximum clusters size N equals 2. This configuration is, by definition, that of the equilibrium of a single reversible reaction, which can be considered as both detailed and complex balanced. For i or j strictly higher than 1, and independently from aggregation and breakage kernels models parameters, criterion values differ from 1.

It can thus be concluded that *realistic kernels models restrict equilibrium states for population balance problems, under the aforementioned hypotheses, to complex balanced states.*

As a general conclusion for this part, chemical reactions networks considerations proved the existence of at least one positive equilibrium state for each population balance problem, independently from kinetics models (kernels) involved. Balance states can thus be used as a basis for inverse methods (e.g. determination of kernels models and parameters from experimental data) in all cases, which is simpler than using transitory states. Furthermore, general kernels models types constructed from physical consideration have been identified in literature, reducing the scope of kernels identification to 9 aggregation/breakage models combinations. Finally, a criterion has been devised to identify the nature of population balances equilibrium states. Its application demonstrated that physically realistic kernels can only attain complex balanced states.

From these results, inverse methods applied to populations balance problems convey two main challenges. The first is the determination of best fitting constants for each of the 9 kernels types combinations based on a set of experimental data. The other resides in the ranking of these models against each other, in order to identify which one is the best suited to describe the aggregation/breakage phenomena studied. These problems will be addressed in the following part.

3 Kernels determination through experimental data fitting

Data fitting can be carried out along two distinct strategies, involving either the use of equilibrium state concentrations distributions (which existence was demonstrated in part 2), or the analysis of multiple transitory concentrations distributions. Each approach shows marked and distinctive advantages:

- Equilibrium data is fundamentally easier to obtain, as each populations balance experiment will finally reach this state. Time and duration measurements are also unnecessary since equilibrium is by definition a steady state.
- Dynamic concentrations distributions are more difficult to assess (in-situ experimental measurements are needed), and more sensitive to errors (for both concentrations and time). However, the potential to enforce any experimental time frame (duration between measurements) can be especially useful in the analysis of complex experiments involving multiple aggregation/breakage mechanisms that occur at different time scales. By choosing the right experimental time intervals, some phenomena can be made negligible, thus simplifying results interpretation.

The following paragraphs will first provide some information about simulation tools and their use to generate realistic pseudo-experimental data. Algorithms constructions corresponding to each of the aforementioned strategies will then be detailed and validated using simulated data.

3.1 Simulation tools

Between the lack of generalized theoretical tools, and minimal number of population balance problems allowing analytical solutions, numerical simulations currently constitute the only reliable method when studying both transitory and steady states. To this end, the C++ framework for population balances (PopulationBalance: ⁴²) has been devised to provide simulation and analysis tools, with performance in terms of CPU time and memory usage as the main objective.

Tools in this framework were devised with the hypotheses mentioned in the first part of this paper in mind and are thus optimized for similar cases. Adaptation to other situations is nonetheless possible by using clever boundary conditions (Neumann and Dirichlet types), at the expense of CPU time. Hypotheses H1, H4 and H5, however, are too ingrained in the internal computing algorithms to be easily bypassed.

This framework allows the manipulation of two different types of objects: *kernels* and *clusters concentrations distributions*.

Three different kernel types can be generated: simple mathematical expressions (such as sum or product), general physical models as detailed in the previous part, and specific models based on physical parameters (Saffman-Turner, or Levitch kernels for example) were implemented.

Basic operations between kernels (corresponding to coefficient to coefficient operations) allow the construction of hybrid or even discontinuous kernels models.

Manipulation of clusters concentrations distributions follows the rules classically applied to probability distributions, with one exception: *distances computations*. Probability distributions distances can only be calculated under the assumption that the two distributions are normalized, a condition that is incompatible with clusters conservation. Definitions for Total Variation Distance, Hellinger, Bhattacharyya, and Manhattan distances are thus modified by first normalizing the concerned clusters distributions, before applying the standard formulas reproduced in Eq. 21.

$$\begin{aligned}
 d_{TVD} \left(\left([i]_1 \right)_{i \leq N}, \left([i]_2 \right)_{i \leq N} \right) &= \max_{i \leq N} \left(|[i]_1 - [i]_2| \right) \\
 d_{Manhattan} \left(\left([i]_1 \right)_{i \leq N}, \left([i]_2 \right)_{i \leq N} \right) &= \sum_{i \leq N} |[i]_1 - [i]_2| \\
 d_{Hellinger} \left(\left([i]_1 \right)_{i \leq N}, \left([i]_2 \right)_{i \leq N} \right) &= \sqrt{1 - \sum_{i \leq N} \sqrt{[i]_1 \cdot [i]_2}} \\
 d_{Bhattacharyya} \left(\left([i]_1 \right)_{i \leq N}, \left([i]_2 \right)_{i \leq N} \right) &= -\ln \left(\sum_{i \leq N} \sqrt{[i]_1 \cdot [i]_2} \right)
 \end{aligned} \tag{21}$$

Interactions between clusters concentrations distributions and kernels correspond to population balance advancement. Such time steps are applied using either an explicit Euler method, with the objective to minimize CPU time, or a second order Runge-Kutta method. The latter case is classically more stable, but configurations involving at least one concentration reaching zero can lead to unavoidable negative concentrations values on the following time step, and must thus be carefully monitored. Typical durations for complete populations balance simulations with clusters size reaching several thousand elementary particles can be estimated (depending on precise simulation conditions) as a few minutes to a few hours. A time step adaptation mechanism was also added to the Euler algorithm, in order to forbid clusters concentrations from reaching negative values.

Sources for this framework, documentation, and examples of use are provided at <https://gitlab.com/AugerJM/populationbalance>.

3.2 Pseudo-experimental data generation

As a means to validate inverse methods algorithms, data obtained through controlled experimental parameters (kernel types, kernel constants, initial state, experiment duration, measurement errors...) would be needed. Due to the practical impossibility to carry out such experiments (experimental kernels are by definition unknown), clusters concentrations distributions were generated using populations balance simulations.

Time steps used for simulations of experiments (“experimental simulations”) are deliberately very small (10^{-8} s) in comparison to those (10^{-6} s) of simulations used in the various fitting algorithms (“analysis simulations”), to better reflect that experimental results correspond to infinitely small time increments.

Steady states are generated from a starting clusters concentrations distribution (unless noted otherwise, this initial state correspond to a maximal number of clusters, e.g. all clusters are of elementary size). Time steps (10^{-8} s) are then successively applied until the Bhattacharyya distance between clusters distributions at two consecutive steps is less than a limit value. This limit is set to 10^{-18} , and by analogy to Cauchy criterion, ensures that simulations are stopped close enough to the equilibrium state.

Transitory states are similarly obtained from a starting clusters concentrations distribution corresponding to a maximal number of clusters. The difference is that time steps of 10^{-8} s are only applied until a fixed total duration has elapsed, to obtain an out of balance clusters concentrations distribution. Such distributions are typically generated in pairs, the first corresponding to a simulated duration from the initial state, and the second to a simulated time interval following the first state. This time interval is considered a part of transitory states data, while duration from initial distribution is not (duration between experimental measurements can easily be chosen, but precise control of the initial state is far more difficult).

Concentrations distributions thus generated simulate perfect experimental data, which is highly unrealistic, as measurements are always subject to noise. Therefore, a good fitting algorithm must not only identify the right kernel models (kernels expressions and constants values both) from perfect data, but also demonstrate robustness when

applied to noisy information. Considering that concentrations are measured independently for each cluster size, three hypotheses have been used to accurately model the noise added to simulated concentrations distributions:

- noise is applied independently for each clusters size concentration
- noise follows a centered normal distribution
- noise standard deviation is proportional to the concentration it is applied to, with the same factor for all clusters concentrations in a given distribution

From now on, a 4% error (or 4% noise) on concentrations distribution will describe a distribution for which the standard deviation of the error applied to each cluster concentration equals 4% of the undisturbed concentration value.

Considering that proper kernels models are currently unidentified for liquid metals, computation parameters (durations, constants, clusters concentrations...) used thereafter in this paper are voluntarily unrealistic, as the aim of this part is to provide generalized examples and validate methods.

3.3 Equilibrium states fitting

Using data corresponding to steady states presents several advantages, one being that equilibrium is practically easily reached (prolonged experiments always lead to equilibrium by definition), and another laying in the limitation to only one necessary measurement point. Most methods aimed at fitting theoretical models to experimental results rely on the use of some objective function, whose parameters can be divided between the experimental data set to fit and model parameters (constrained multivariate optimization). In the case of inverse populations balance problems, the data set is a measured equilibrium clusters concentrations distribution $([i])_{i \leq N}$, and the model parameters consist of aggregation and breakage kernels K and K^* (kernels mathematical expressions and their corresponding constants).

From Eq. 3 and 5, a possible definition for an objective function C would be given by Eq. 22.

$$C\left(\left([i]\right)_i, K, K^*\right) = \sum_{i=1}^{i=N} \left| \sum_{j=1}^{j=i-1} \left(k_{j,i-j} \cdot [j] \cdot [i-j] - k_{j,j-i}^* \cdot [i] \right) - \sum_{j=1}^{j=N-i} \left(1 + \delta_{i,j} \right) \cdot \left(k_{i,j} \cdot [i] \cdot [j] - k_{i,j}^* \cdot [i+j] \right) \right| \quad (22)$$

By construction, criterion C has the following properties:

1- positive values

2- zero is reached only when the clusters concentrations distribution $([i])_{i \leq N}$ corresponds to an equilibrium state for kernels K and K* (distribution and kernels will be called compatible)

3- if kernels K, K* and distribution $([i])_{i \leq N}$ are compatible, then for any positive real number a, kernels a.K, a.K* (a.K defined as the kernel whose coefficients equals those of K multiplied by a) and distribution $([i])_{i \leq N}$ are also compatible (direct consequence of complex balanced state, which was demonstrated as the only equilibrium type possible in Part 2)

For a chosen equilibrium distribution $([i]_{eq})_{i \leq N}$, criterion C can thus be used as an objective function whose minimization with regards to K and K* allows the identification of compatible kernels. Due to the general multiplicative coefficients common to all kernel models (α coefficients in Table 5), property 3 implies that the set of parameters allowing compatibility for an aggregation/breakage kernel models couple is either infinite or empty. Consequently, minimizing algorithms cannot be applied directly to criterion C with the chosen kernel models. In order to avoid this effect, the *multiplicative coefficient of aggregation kernel expressions can be locked to a single value* (1 will be used), reducing the number of parameters to determine.

The algorithm used for C criterion minimization (a non-linear problem) is the Nelder-Mead heuristic method (detailed in ⁴³), which is based on simplexes properties. Implementation simplicity and efficiency when applied to few parameters (2 or 3 depending on the kernels considered) make this method particularly apt to the fitting of kernels models to experimental data.

3.3.1 Examples of application for kernels determination: perfect experimental data

Pseudo-experimental data were generated using the following basis:

Kernels:

- Turbulence type 1 aggregation, $\alpha=1$
- Exponential law 2 breakage, $n=0.5$, $\alpha=0.25$, $\beta=3$

Step size: 10^{-8} s

Total elementary particles: 5/volume unit

Maximum clusters size: $N=200$

Initial state: $[1]=5$ clusters/volume unit

Noise: 0% (no noise)

Using the previously described fitting method, optimal parameters for each of the 9 kernel models combinations have been determined, and the results summarized in Table 7. Two interesting constataions can be made:

- first, *kernel constants determined by fitting the right models* (Turbulence type 1 and Exponential law 2) *match perfectly* (less than 0.2% error) the values used for pseudo-experimental data generation.
- second, the correct kernel models couple is easily identified, as it corresponds to the lowest value for C . Furthermore, the C criterion value reached at the end of the minimization procedure shows a *marked discriminating effect*, being six orders of magnitude smaller for the right models than for non-matching kernels.

Generality of these results was easily confirmed, by applying the same fitting method to pseudo-experimental data generated using various kernels models and parameters.

3.3.2 Examples of application for kernels determination: noisy experimental data

In order to assess the effect of noise on the fitting procedure, a slightly disturbed pseudo-experimental data set was generated from:

Kernels:

- Turbulence type 1 aggregation, $\alpha=1$
- Exponential law 2 breakage, $n=0.5$, $\alpha=0.25$, $\beta=3$

Step size: 10^{-8} s

Total elementary particles: 5/volume unit

Maximum clusters size: $N=200$

Initial state: [1]= 5 clusters/volume unit

Noise: 1%

As shown in Table 8, the turbulence type 1 and exponential law 1 kernels combination is falsely identified as best fitting the pseudo-experimental data. Furthermore, even using the right theoretical models leads to highly incorrect constants (the minimal error is around two orders of magnitude). This effect could be linked to two different phenomena: an important sensitivity of the algorithm to entered data (ill-conditioning), or a systematic error following the use of noisy data (bias). In order to discriminate between these causes, 200 pseudo-experimental data sets were generated using the same kernels and parameters as previously mentioned, with the fitting results corresponding to the correct kernels models being summarized in table Table 9 for 25 of them. Errors on kernels parameters determined by the algorithm do not follow a systematic pattern, and seem dispersed between several orders of magnitude. This effect hints at a *low condition number (ratio between relative error of computed parameters and relative error of entered data) for this approach*.

As a general conclusion, the inverse method devised for kernels determination from equilibrium states is highly accurate for both kernel models discrimination and constants determination, *provided that equilibrium data is noiseless*. However, *even small magnitude noise lead to incorrect kernels identification, and huge errors in constants determination*. This sensitivity to data accuracy is a telling symptom of an ill-conditioned problem-solving algorithm.

3.4 Transitory states fitting

As mentioned before, experimental or pseudo-experimental data used to define transitory states consist in two consecutive clusters concentrations distributions, together with the time interval separating these 2 measurements. The issue is thus to identify an aggregation/breakage kernels couple, such that there exist a population balance “trajectory” linking the two aforementioned distributions with the correct time interval. This kind of problem is generally the province of least squares methods. However, a population balance trajectory is defined not only by two kernels models and the corresponding parameters (3 or 4 in total), but also by a starting concentrations distribution (100 or more additional parameters). The considerable number of parameters involved, together with the non-linear

nature of the problem, renders classical multivariate fitting methods inapplicable. The following paragraphs detail a variant version of the least squares method devised to circumvent this limitation.

Using the first measured clusters concentrations distribution $([i]_{initial})_i$ as an initial state, a populations balance simulation (PBS) is carried out with kernels candidates K and K^* for a time interval Δt corresponding to the duration between experimental (or pseudo-experimental) measurements (formally expressed as Eq. 23). The resulting simulated distribution $([i]_{simul})_i$ is then compared to the second (pseudo-)experimental distribution $([i]_{final})_i$ through their Bhattacharyya distance, which constitute the objective function T (Eq. 24) to minimize with respect to K and K^* .

$$PBS(K, K^*, \Delta t): ([i]_{initial})_{i \leq N} \rightarrow ([i]_{simul})_{i \leq N} \quad (23)$$

$$T(K, K^*, \Delta t, ([i]_{initial})_{i \leq N}, ([i]_{final})_{i \leq N}) = d_{Bhattacharyya}([i]_{final})_{i \leq N}, ([i]_{simul})_{i \leq N} \quad (24)$$

Similarly to equilibrium states fitting, the method used for T criterion minimization (also a non-linear problem) is the Nelder-Mead heuristic approach. Due to the more complex definition of objective function (which includes a complete population balance simulation), this algorithm displays longer CPU times.

This fitting method provides results like those obtained using equilibrium states (some reproducibility examples are given in Table 10 for lightly disturbed transitory distributions):

- accurate kernel models determination and parameters identification are possible, when based on exact (pseudo-)experimental data

- an inability to identify the correct kernel models using even slightly noisy (1% error) (pseudo-)experimental data.

Determination of kernel parameters in these conditions suffers from a reproducibility problem, with instability of computed values ranging several orders of magnitude.

In order to understand the reasons for this demonstrated sensitivity of the fitting algorithm to imprecise data and determine potential workarounds, statistics of errors in clusters concentrations distributions had to be further investigated.

3.5 Measurements errors analysis

Models for errors in clusters concentrations distributions were selected for their agreement with experimental measurement procedures. Each clusters size concentration is measured separately, leading to independent errors that are often considered the result of multiple random phenomena. From central limit theorem, individual concentration errors can be realistically thought-out as following a normal distribution, centered on an undisturbed concentration value. Hypotheses pertaining to pseudo-experimental data generation can thus be summarized as Eq. 25.

$$\forall i \in (1 ; N), [\hat{i}] = [i] + X_i \text{ with } X_i = N(0, k \cdot [i]) \quad (25)$$

Variable k is the measurement error as a proportionality factor to clusters concentration.

In order to check an eventual bias linked to concentrations errors, knowledge of the probability distribution of noisy concentrations distributions is needed. As literature does not provide any example of theoretical study for this problem, it will be qualitatively evaluated through numerical simulations. Starting from a base concentrations distribution (maximum clusters size N=100), 100000 disturbed distributions were generated by adding 5% noise to individual concentrations. Bhattacharyya distances between base and disturbed distributions were calculated, and the corresponding frequencies determined using ten evenly spaced classes, as presented in Figure 4. It should be noted that the use of different distance definitions, such as Hellinger, Manhattan, Total Variation Distance or Mahalanobis, leads to the same characteristics, whose consequences will be discussed in detail:

- the probability distribution does not follow a normal law (no symmetry around the maximum probability distance). Statistical properties of errors in clusters concentrations distributions can be likened to those of multivariate normal distributions.
- highest probability is not reached for a distance of zero. This observation reveals the existence of a *statistical bias* in disturbed experimental data when considering clusters distributions distances. The actual clusters distribution is most probably situated on a sphere at some distance around the measured distribution.
- probability around a distance of zero is very low. This effect is of extreme import to the accuracy of fitting algorithms. It is highly improbable that any measured distribution closely matches the real distribution. Since the transitory states fitting method is based on distance minimization between experimental (thus noisy) points and

simulated points, the algorithm as presented will ensure (statistically) a high error in kernels parameters determination.

This kind of bias is classically dealt with through the multiplication of experiments under similar conditions. Provided enough data points, the corrected clusters concentrations distribution will be at the center of the noisy points cloud. In the case of clusters concentrations, such correction is obtained by a separate averaging of the concentration for each cluster size. This approach, however, is limited by experimental repeatability. Sufficient data accumulation is impossible in the context of liquid metal tests, due to uncontrollable initial conditions and destructive measurement methods (metallography). As a replacement, another corrective method based on statistical maximum likelihood can be used instead, as described in the following paragraph.

3.6 Maximum likelihood statistical correction

The objective of this part will be to construct a function (the likelihood function) of an experimental (noisy) data point (e.g. clusters concentrations distribution), which is statistically minimal when simulated points approach the corresponding undisturbed concentrations distribution.

Under the hypotheses already detailed in the previous part in Eq. 25, normalizing and centering of the normal law leads to the expression of Eq. 26.

$$\frac{\hat{[i]} - [i]}{k \cdot \hat{[i]}} \simeq \frac{\hat{[i]} - [i]}{k \cdot [i]} = N(0,1) \quad (26)$$

By definition of X² law, summation of Eq. 26 for all values of i gives Eq. 27.

$$\sum_{i=1}^{i=N} \left(\frac{\hat{[i]} - [i]}{k \cdot \hat{[i]}} \right)^2 \simeq X^2(N) \quad (27)$$

If *N* is sufficiently large (*N* > 100), X² law can be accurately approximated by a normal law leading to Eq. 28.

$$\sum_{i=1}^{i=N} \left(\frac{\hat{[i]} - [i]}{k \cdot \hat{[i]}} \right)^2 \approx N(N, 2 \cdot N) \quad (28)$$

Centering this normal law (Eq. 29), the probability of the expression becomes maximal at zero.

$$\left(\sum_{i=1}^{i=N} \left(\frac{[\hat{i}] - [i]}{k \cdot [\hat{i}]} \right)^2 \right) \cdot \frac{1}{N} - 1 \approx N(0,2) \quad (29)$$

Finally leading to the maximum likelihood function expression f (Eq. 30), which follows the half-normal distribution.

$$f\left(\left([\hat{i}]\right)_i, \left([i]\right)_i, k, N\right) = \left| \left(\sum_{i=1}^{i=N} \left(\frac{[\hat{i}] - [i]}{k \cdot [\hat{i}]} \right)^2 \right) \cdot \frac{1}{2 \cdot N} - 0.5 \right| \approx |N(0,1)| \quad (30)$$

f has the required properties to be used in the transitory states algorithm in place of Bhattacharyya distance (confirmed by 100000 simulations, $N=100$, $k=5\%$ - see Figure 5):

- f is positive, and consequently f minimal value is zero.
- f follows the half-normal law, ensuring a maximal probability when f is minimal.

Conclusion:

For a given disturbed concentrations distribution (maximum clusters size N , $k\%$ error) $\left([\hat{i}]\right)_i$, the clusters concentrations distribution $\left([i]\right)_i$ such as the function f defined by Eq. 31 is minimal, offers the most probable correspondence with the undisturbed concentrations distribution.

$$f\left(\left([\hat{i}]\right)_i, \left([i]\right)_i, k, N\right) = \left| \left(\sum_{i=1}^{i=N} \left(\frac{[\hat{i}] - [i]}{k \cdot [\hat{i}]} \right)^2 \right) \cdot \frac{1}{2 \cdot N} - 0.5 \right| \quad (31)$$

3.7 Transitory states fitting with maximum likelihood correction

Maximum likelihood correction can be used to improve kernels fitting methods without relying on measurements multiplication. The algorithm previously described for transitory states fitting is altered by substituting the objective function T with another, TC , constructed from the maximum likelihood function f (as defined in the previous paragraph). This modified algorithm is described by Eq. 32 and TC function in Eq. 33.

$$PBS(K, K^*, \Delta t): ([i]_{initial})_{i \leq N} \rightarrow ([i]_{simul})_{i \leq N} \quad (32)$$

$$TC(K, K^*, \Delta t, N, k, ([i]_{initial})_{i \leq N}, ([i]_{final})_{i \leq N}) = f\left(\left([i]_{final}\right)_{i \leq N}, \left([i]_{simul}\right)_{i \leq N}, k, N\right) \\ = \left| \left(\sum_{i=1}^{i=N} \left(\frac{[i]_{final} - [i]_{simul}}{k \cdot [i]_{final}} \right)^2 \right) \cdot \frac{1}{2 \cdot N} - 0.5 \right| \quad (33)$$

Minimization of TC objective function with respect to K and K* is similarly carried out using Nelder-Mead heuristic approach.

In order to assess the effect of noise on the maximum likelihood corrected fitting procedure, a disturbed pseudo-experimental data set was generated from:

Kernels:

- Turbulence type 1 aggregation, $\alpha=4$
- Exponential law 1 breakage, $n=0.5, \alpha=1, \beta=3$

Step size: 10^{-8} s

Simulated time from initial state: 0.01 s

Simulated duration between transitory states: 0.025 s

Total elementary particles: 5/volume unit

Maximum clusters size: $N=150$

Initial state: $[1]=5$ clusters/volume unit

Noise: 5%

As illustrated in Table 11, the *kernels are accurately identified by the minimum of criterion TC*. Discrimination of kernels models is however less clear-cut than for undisturbed base data.

Reproducibility tests for kernels parameters determination have been carried out for a large pseudo-experimental data set (100 clusters distributions couples) generated using:

Kernels:

- Turbulence type 2 aggregation, $\alpha=1$
- Exponential law 2 breakage, $n=0.5$, $\alpha=1$, $\beta=3$

Step size: 10^{-8} s

Simulated time from initial state: 0.01 s

Simulated duration between transitory states: 0.025 s

Total elementary particles: 5/volume unit

Maximum clusters size: $N=150$

Initial state: $[1]=5$ clusters/volume unit

Noise: 2%

Fitting results samples using correct kernels models are detailed by tests N°1 to 5 in Table 12. While still much higher (around one order of magnitude) than the original data errors (2% only), the variability of computed kernels parameters was dramatically diminished using maximum likelihood corrected algorithm. It should be noted that anomalous parameters values are still computed in some cases (α value for breakage kernel model in N°4 test, for example), but the frequency of these happenings is decreased compared to that shown by using T criterion (see Table 10). Generally speaking, multiplicative factor α tends to be more precise than exponential factor β , which in turn is more precise than exponent n .

As a conclusion, *maximum likelihood correction greatly improves the accuracy of kernels parameters determination*. However, as a statistics-based approach, this method is still dependent on actual measured data (two clusters concentrations distributions), that can occasionally take values too far removed from the most probable ones to be effectively corrected. In order to establish the limits of this improved algorithm (and thus determine the best experimental setup to optimize its use), sensitivity to two easily controlled experimental conditions has been investigated: the time interval between transitory states measurements, and how far removed they are from equilibrium state.

3.7.1 Sensitivity to time duration between (pseudo-)experimental points:

Cases N°11 to 15 in Table 12 were generated using the following characteristics (corresponding to a duration decrease between the pseudo-experimental transitory states in contrast to cases N°1 to 5):

Kernels:

- Turbulence type 2 aggregation, $\alpha=1$
- Exponential law 2 breakage, $n=0.5$, $\alpha=1$, $\beta=3$

Step size: 10^{-8} s

Simulated time from initial state: 0.01 s

Simulated duration between transitory states: 0.005 s

Total elementary particles: 5/volume unit

Maximum clusters size: $N=150$

Initial state: $[1]=5$ clusters/volume unit

Noise: 2%

Comparison of results N°1 to 5 and N°11 to 15 shows of marked decrease in fitting accuracy when the time interval between measurements for transitory states diminishes. This effect is like the increasing slope error observed if base points are closer from each other, when performing linear regression algorithms.

3.7.2 Sensitivity to distance from equilibrium state:

Cases N°6 to 10 in Table 12 were generated using the following characteristics (corresponding to an increase in duration from initial simulation state compared to cases N°1 to 5):

Kernels:

- Turbulence type 2 aggregation, $\alpha=1$
- Exponential law 2 breakage, $n=0.5$, $\alpha=1$, $\beta=3$

Step size: 10^{-8} s

Simulated time from initial state: 0.05 s

Simulated duration between transitory states: 0.025 s

Total elementary particles: 5/volume unit

Maximum clusters size: $N=150$

Initial state: $[1]=5$ clusters/volume unit

Noise: 2%

Accuracy of kernels parameters determined by the algorithm is visibly degraded with data points nearer to the equilibrium state (comparison of cases N°1-5 to 6-10), in terms of both aberrant values frequency and amplitude (for example, computed β coefficient values range between 24 orders of magnitude in only 5 tests).

4 Conclusions

Part 2 investigated the specifics of population balance problems, notably using reaction networks theory, with a special interest toward eventual equilibrium states and their attributes.

- A set of hypotheses that closely matches *model experimental setups* (clusters quantization, particles conservation, concentrations continuity, binary reversible aggregations, and mass action kinetics) allowed the use of embedded reaction networks for population balance problems with maximum cluster size N . The existence of at least one stable positive equilibrium state was demonstrated, provided (sufficient condition) there is at least one effectively reversible aggregation/breakage reaction leading to the formation of each clusters size ranging from 2 to N .

- Boundary conditions (total elementary particles concentration) and kernels mathematical formulations have been shown to determine the multiplicity of equilibrium states. However, no simulation involving physical kernels expressions has been observed to generate multiple equilibrium states during this study.

- Possibility for detailed balanced equilibrium state can be assessed through a simple criterion based on kernels formulae. This criterion was constructed under the hypothesis of reversible elementary aggregation reactions.

- Using the 6 general kernels expressions found in literature for aggregation and breakage of solid clusters in liquid media, *equilibrium states for realistic population balance problems were thus demonstrated to be of the complex balanced variety.*

Part 3 concerned itself with the population balances inverse problem (e.g. determination of aggregation and breakage kernels from experimental data). Making use of the specifically developed PopulationBalance C++ framework for implementation and testing, two different fitting heuristics have been devised to solve this non-linear problem:

- The first algorithm, based on equilibrium states, is a simple and highly efficient method to both identify kernels types and compute their parameters, when applied to undisturbed (perfect) experimental data. However, extreme sensitivity to measurement errors makes it unusable in any realistic experimental context and exposes the inverse problem as an ill-conditioned one.

- Similar observations were made on a second algorithm constructed from least squares methods applied to transitory states. Sensitivity to measurements uncertainties evidenced a systematic bias that could be traced back to statistical errors spread in clusters concentrations measurements.

A statistically unbiased objective function (TC) was devised as a substitute to distributions distances in the algorithm. This improvement reduced error ranges for computed parameters to a single order of magnitude, and allowed accurate kernels identification using minimal experimental data (only two points) with 5% or less errors.

Study of the corrected algorithm sensitivity to experimental measurement parameters led to the following (and opposite) recommendations, in order to optimize kernels determination:

- measurement points must be separated by the greatest time interval possible
- measurement points must be as far removed from equilibrium as possible

ACKNOWLEDGMENTS

This research work was carried out in the context of French/German FLOTINC project (ref. ANR-15-CE08-0040), funded by ANR (Agence Nationale de la Recherche) and DFG (Deutsche Forschungsgemeinschaft).

FIGURES

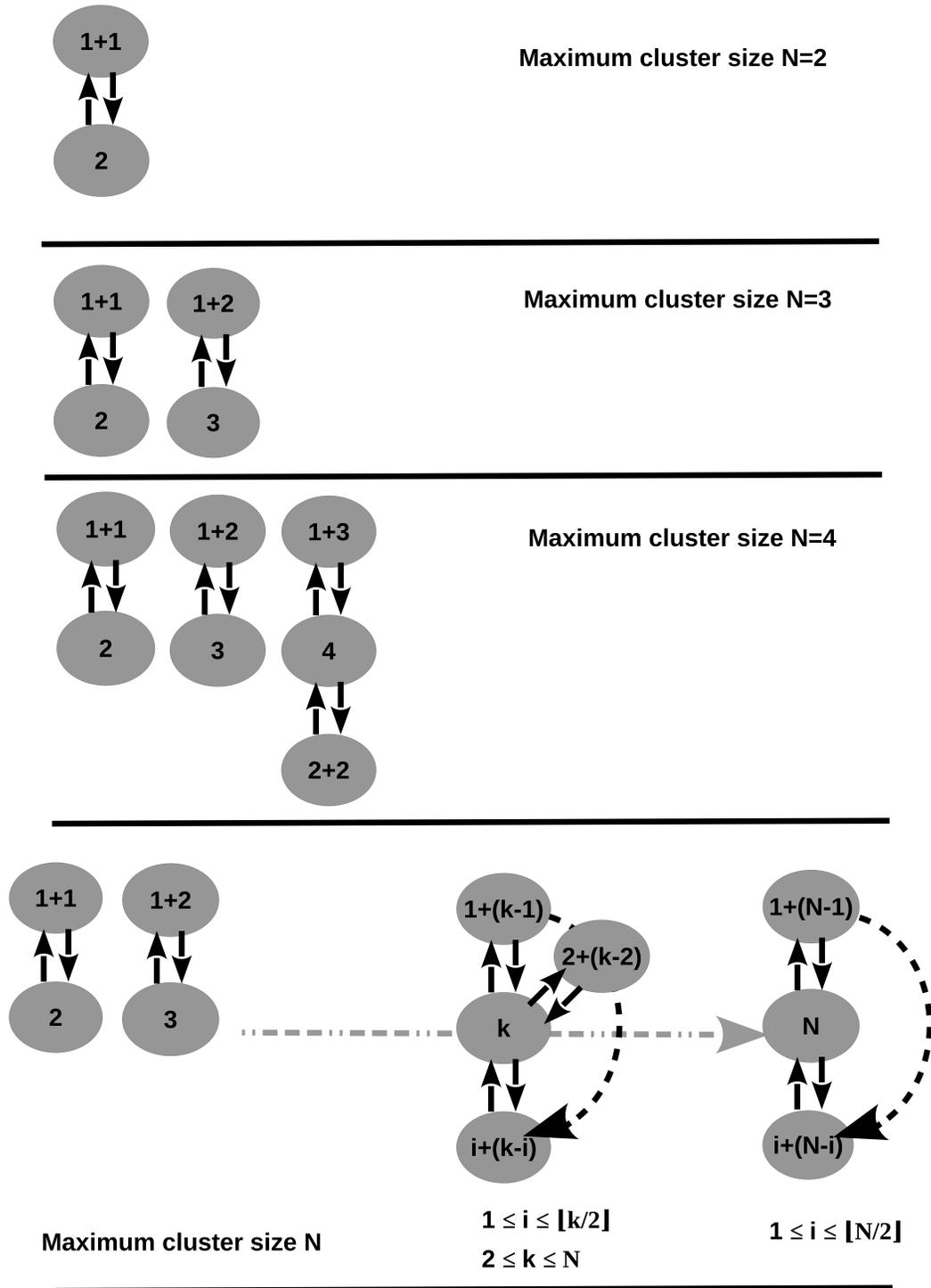


Figure 1. Reaction graphs for population balance problems ($N=2$, $N=3$, $N=4$ and general graph) – Boxes correspond to various complexes – Arrows are aggregation or breakage reactions

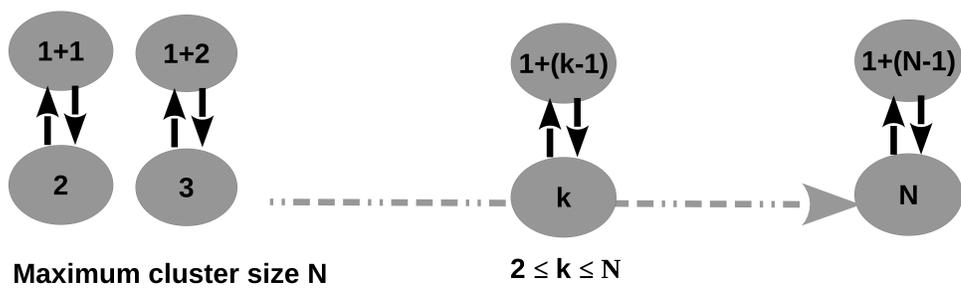


Figure 2. Example of deficiency zero embedded network extracted from the population balance problem reaction network

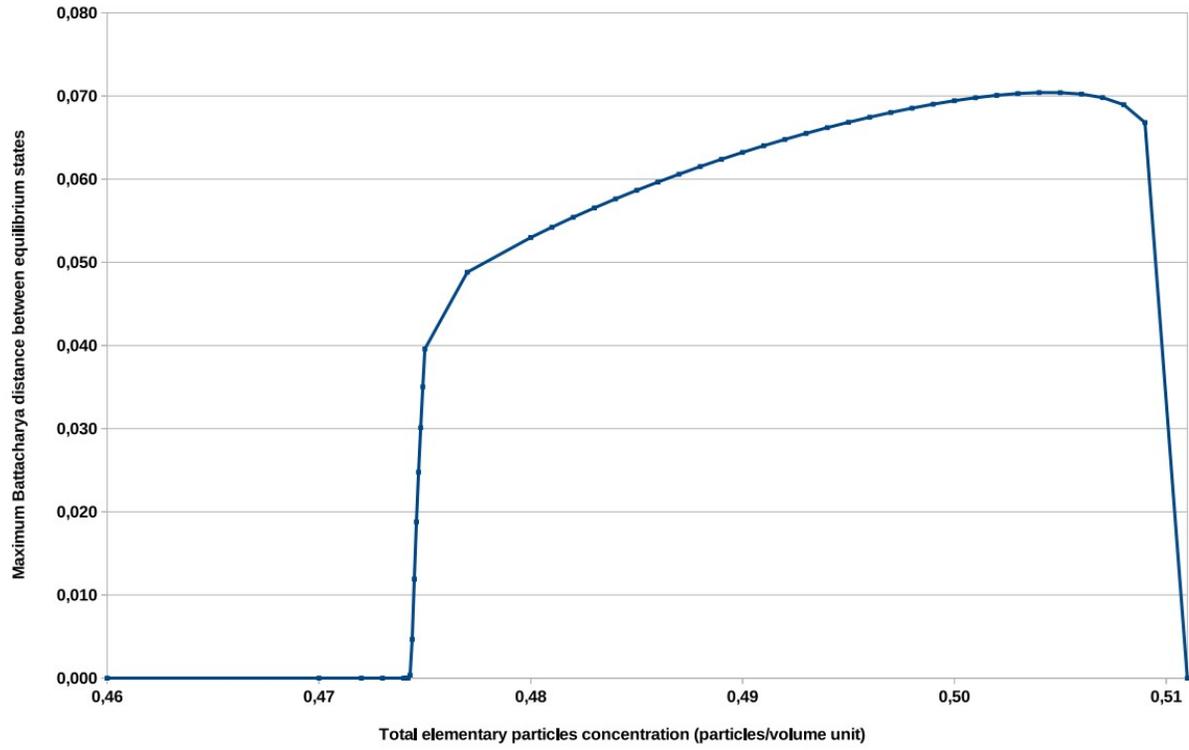


Figure 3. Maximum distance between equilibrium states as a function of total elementary particles concentration

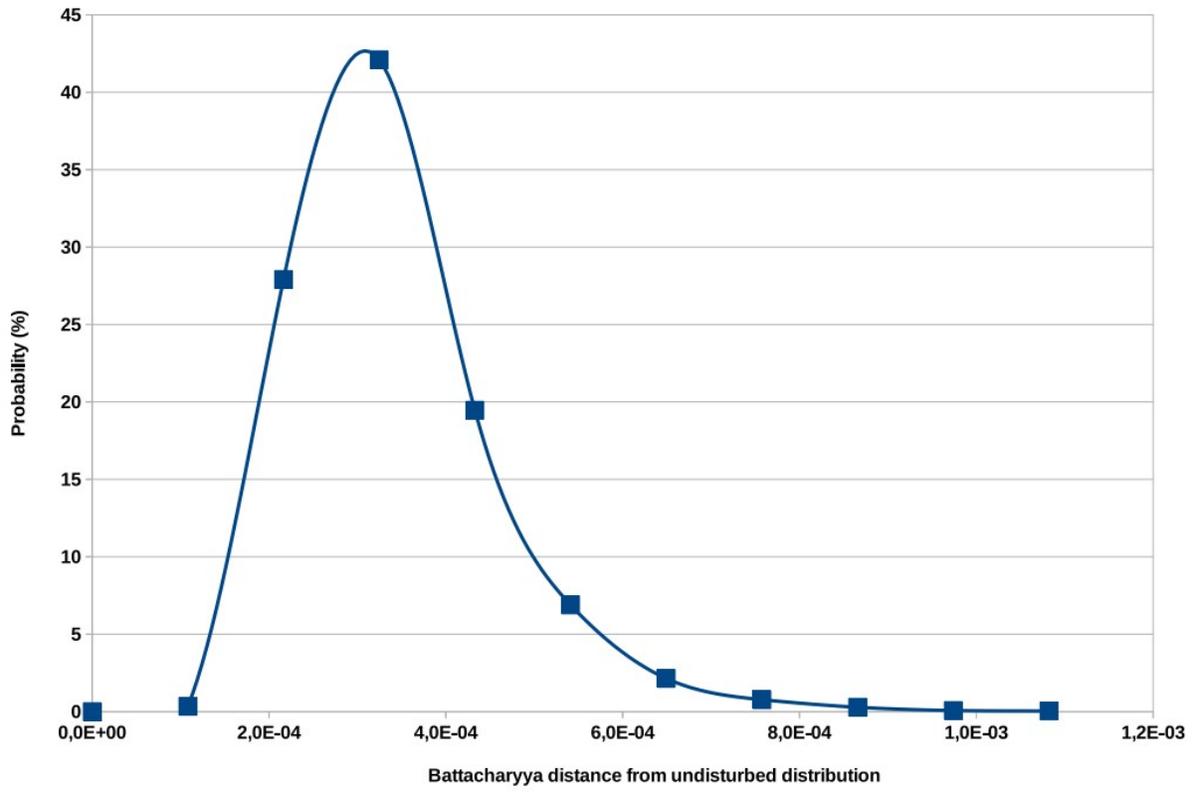


Figure 4. Simulated probability distribution of Bhattacharyya distances between disturbed and undisturbed clusters distributions (5% noise)

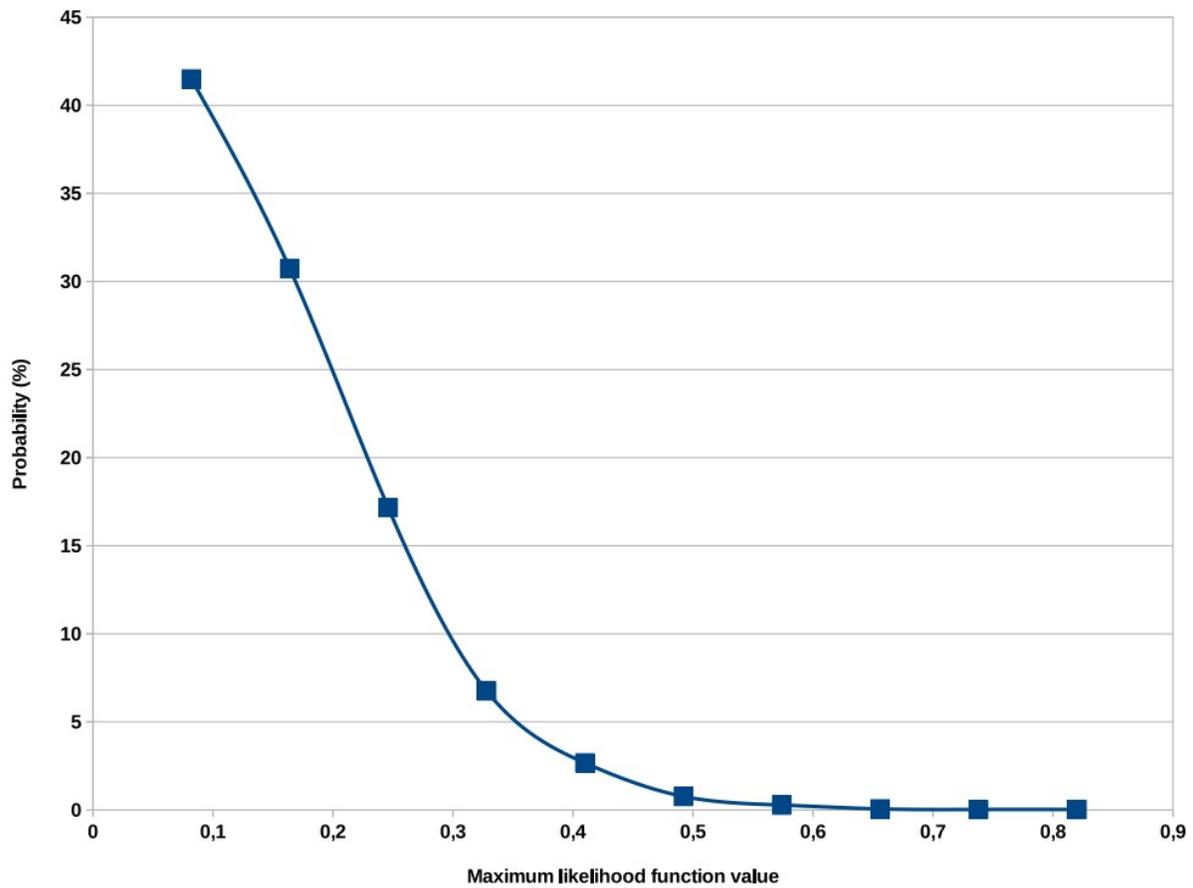


Figure 5. Simulated probability distribution of the maximum likelihood function (N=100, k=5% noise)

ABBREVIATIONS AND NOTATIONS

| | |
|----------------------------|--|
| (i) | Size i cluster (i elementary particles aggregated) |
| $(i+j)$ | Size $(i+j)$ cluster |
| $(i)_{i \in \mathbb{N}}$ | Cluster sizes set |
| $[i]$ | Volumetric concentration for size i clusters |
| $[i]_{\text{eq}}$ | Volumetric concentration for size i clusters at equilibrium |
| $\dot{[i]}$ | Evolution rate for volumetric concentration for size i clusters |
| $\hat{[i]}$ | Volumetric concentration for size i clusters measurement (random variable) |
| $([i])_{i \in \mathbb{N}}$ | Volumetric concentration distribution of clusters by sizes |
| $(\dot{[i]})_{i \leq N}$ | Evolution rate for volumetric concentration distribution of clusters by sizes |
| k_{ij} | Kinetic constant for i and j clusters aggregation into $(i+j)$ clusters |
| k_{ij}^* | Kinetic constant for $(i+j)$ clusters breakage into i and j clusters |
| K_{ij} | Equilibrium constant for i and j clusters aggregation into $(i+j)$ clusters reaction |
| Φ_{ij} | Reaction rate for aggregation/breakage of i and j clusters |
| N | Maximum clusters size (as a number of elementary particles) |
| n | Total number of identical elementary (unbreakable) particles |
| V | Total volume considered |
| δ_{ij} | Kronecker delta |
| δ | Deficiency of population balance network |
| l | Number of linkage classes of population balance network |
| m | Number of complexes in population balance network |
| s | Dimension of stoichiometric sub-space of population balance network |
| K | Aggregation kernel |

| | |
|-------|-----------------|
| K^* | Breakage kernel |
| [] | Floor function |

TABLES

Table 1. Deficiency data for each linkage class in N=4 reaction networks

| Network decomposition linkage classes | Number of linkage classes | Complexes number | Stoichiometric subspace dimension | Deficiency |
|---------------------------------------|---------------------------|------------------|-----------------------------------|------------|
| 1st linkage class | 1 | 2 | 1 | 0 |
| 2nd linkage class | 1 | 2 | 1 | 0 |
| 3rd linkage class | 1 | 3 | 2 | 0 |
| Full network | 3 | 7 | 3 | 1 |

Table 2. Equilibrium states conditions in N=4 reaction networks

| Clusters reactions reversibility | | | | Stable positive equilibrium states |
|----------------------------------|-------|-------|-------|------------------------------------|
| 1+1→2 | 1+2→3 | 1+3→4 | 2+2→4 | |
| Yes | Yes | Yes | Yes | Unique |
| No | Yes | Yes | Yes | Unique |
| Yes | Yes | Yes | No | Unique |
| Yes | No | Yes | Yes | Unique |
| Yes | Yes | No | Yes | Unique |
| No | Yes | Yes | No | Unique |
| No | No | Yes | Yes | None |
| No | Yes | No | Yes | None |
| Yes | Yes | No | No | None |
| Yes | No | No | Yes | Unique |
| Yes | No | Yes | No | None |
| Yes | No | No | No | None |
| No | Yes | No | No | None |
| No | No | Yes | No | None |
| No | No | No | Yes | None |
| No | No | No | No | None |

Table 3. Example of potentially multi-stable aggregation/breakage kernel

| | |
|------------------------------|-----------------------|
| Maximum cluster size $N = 5$ | |
| Kinetic constants | |
| $k_{1,1}=1$ | $k_{1,1}^*=592.48963$ |
| $k_{1,2}=95.437698$ | $k_{1,2}^*=201.38279$ |
| $k_{1,3}=419.0933$ | $k_{1,3}^*=2074.8396$ |
| $k_{1,4}=1421.4501$ | $k_{1,4}^*=368.28533$ |
| $k_{2,2}=34076.512$ | $k_{2,2}^*=753.42078$ |
| $k_{2,3}=507535.96$ | $k_{2,3}^*=11190.471$ |

Table 4. Equilibrium states corresponding to the same population balance problem (kernels described in Table 3, total elementary clusters concentration of 1 per volume unit)

| Clusters size | Equilibrium state 1 (clusters/volume unit) | Equilibrium state 2 (clusters/volume unit) |
|---------------|---|---|
| [1] | 0.93783354 | 0.84858688 |
| [2] | 9.8266×10^{-3} | 1.6201×10^{-2} |
| [3] | 4.3674×10^{-3} | 1.0742×10^{-2} |
| [4] | 4.3674×10^{-3} | 1.0742×10^{-2} |
| [5] | 2.3881×10^{-3} | 8.7629×10^{-3} |

Table 5. Models types for aggregation and breakage kernels (clusters are considered spherical)

| Aggregation | | | |
|-------------------|---|----------------------------------|---|
| Kernel type | Expression as functions of clusters sizes | Constants | Specific examples |
| Brownian | $k_{i,j} = \alpha \cdot \frac{\left(\frac{1}{i^3} + \frac{1}{j^3}\right)^2}{\frac{1}{i^3} \cdot \frac{1}{j^3}}$ | $\alpha \geq 0$ | Smoluchowski ⁴⁴ |
| Turbulence type 1 | $k_{i,j} = \alpha \cdot \left(\frac{1}{i^3} + \frac{1}{j^3}\right)^2$ | $\alpha \geq 0$ | Kruis-Kuster ⁴⁵ , Saffman-Turner ⁴⁶ |
| Turbulence type 2 | $k_{i,j} = \alpha \cdot \left(\frac{1}{i^3} + \frac{1}{j^3}\right)^3$ | $\alpha \geq 0$ | Camp-Stein ⁴⁷ , Levich ⁴⁸ |
| Breakage | | | |
| Kernel type | Expression as functions of clusters sizes | Constants | Specific examples |
| Power law | $k_{i,j}^* = \alpha \cdot (i+j)^{(n-1)}$ | $n \geq 0, \alpha \geq 0$ | Valentas ⁴⁹ Ramkrishna ⁵⁰ |
| Exponential law 1 | $k_{i,j}^* = \alpha \cdot e^{-\beta \cdot \left(\frac{1}{i+j}\right)^n}$ | $n \geq 0, \alpha \geq 0, \beta$ | Delichatsios-Probstein ⁵¹ |
| Exponential law 2 | $k_{i,j}^* = \frac{\alpha}{(i+j)^{(2/9)}} \cdot e^{-\beta \cdot \left(\frac{1}{i+j}\right)^n}$ | $n \geq 0, \alpha \geq 0, \beta$ | Marchisio ⁵² Selomulya ⁵³ |

Table 6. Detailed balance criterion expressions for each physical kernels models combination

| Aggregation kernel type | Breakage kernel type | Criterion expression |
|-------------------------|----------------------|--|
| Brownian | Power law | $\frac{\left(\frac{l}{i^{\frac{1}{3}}+l}\right)^2 \cdot (i+l)^{\left(\frac{2}{3}-n\right)} \cdot \left(j^{\frac{1}{3}}+(i+1)^{\frac{1}{3}}\right)^2 \cdot (i+j)^{\left(n-\frac{2}{3}\right)}}{\left(\frac{1}{i^{\frac{1}{3}}+j^{\frac{1}{3}}}\right)^2 \cdot \left((i+j)^{\frac{1}{3}}+1\right)^2}$ |
| Brownian | Exponential law 1 | $\frac{\left(\frac{l}{i^{\frac{1}{3}}+l}\right)^2 \cdot \left(j^{\frac{1}{3}}+(i+1)^{\frac{1}{3}}\right)^2 \cdot (i+j)^{\left(\frac{1}{3}\right)} \cdot e^{\frac{\beta}{(1+i)^n} - \frac{\beta}{(i+j)^n}}}{(i+1)^{\frac{1}{3}} \cdot \left(\frac{1}{i^{\frac{1}{3}}+j^{\frac{1}{3}}}\right)^2 \cdot \left((i+j)^{\frac{1}{3}}+1\right)^2}$ |
| Brownian | Exponential law 2 | $\frac{\left(\frac{l}{i^{\frac{1}{3}}+l}\right)^2 \cdot \left(j^{\frac{1}{3}}+(i+1)^{\frac{1}{3}}\right)^2 \cdot (i+j)^{\left(\frac{1}{9}\right)} \cdot e^{\frac{\beta}{(1+i)^n} - \frac{\beta}{(i+j)^n}}}{(i+1)^{\frac{1}{9}} \cdot \left(\frac{1}{i^{\frac{1}{3}}+j^{\frac{1}{3}}}\right)^2 \cdot \left((i+j)^{\frac{1}{3}}+1\right)^2}$ |
| Turbulence type 1 | Power law | $\frac{\left(\frac{l}{i^{\frac{1}{3}}+l}\right)^2 \cdot (i+l)^{ l-n } \cdot \left(j^{\frac{1}{3}}+(i+1)^{\frac{1}{3}}\right)^2 \cdot (i+j)^{ n-1 }}{\left(\frac{1}{i^{\frac{1}{3}}+j^{\frac{1}{3}}}\right)^2 \cdot \left((i+j)^{\frac{1}{3}}+1\right)^2}$ |
| Turbulence type 1 | Exponential law 1 | $\frac{\left(\frac{l}{i^{\frac{1}{3}}+l}\right)^2 \cdot \left(j^{\frac{1}{3}}+(i+1)^{\frac{1}{3}}\right)^2 \cdot e^{\frac{\beta}{(1+i)^n} - \frac{\beta}{(i+j)^n}}}{\left(\frac{1}{i^{\frac{1}{3}}+j^{\frac{1}{3}}}\right)^2 \cdot \left((i+j)^{\frac{1}{3}}+1\right)^2}$ |
| Turbulence type 1 | Exponential law 2 | $\frac{\left(\frac{l}{i^{\frac{1}{3}}+l}\right)^2 \cdot (i+l)^{\frac{2}{9}} \cdot \left(j^{\frac{1}{3}}+(i+1)^{\frac{1}{3}}\right)^2 \cdot e^{\frac{\beta}{(1+i)^n} - \frac{\beta}{(i+j)^n}}}{\left(\frac{1}{i^{\frac{1}{3}}+j^{\frac{1}{3}}}\right)^2 \cdot (i+j)^{\frac{2}{9}} \cdot \left((i+j)^{\frac{1}{3}}+1\right)^2}$ |
| Turbulence type 2 | Power law | $\frac{\left(\frac{l}{i^{\frac{1}{3}}+l}\right)^3 \cdot (i+l)^{ l-n } \cdot \left(j^{\frac{1}{3}}+(i+1)^{\frac{1}{3}}\right)^3 \cdot (i+j)^{ n-1 }}{\left(\frac{1}{i^{\frac{1}{3}}+j^{\frac{1}{3}}}\right)^3 \cdot \left((i+j)^{\frac{1}{3}}+1\right)^3}$ |

| | | |
|-------------------|-------------------|---|
| Turbulence type 2 | Exponential law 1 | $\frac{\left(\frac{1}{i^{\frac{1}{3}}+1}\right)^3 \cdot \left(\frac{1}{j^{\frac{1}{3}}+(i+1)^{\frac{1}{3}}}\right)^3 \cdot e^{\frac{\beta}{(1+i)^n} - \frac{\beta}{(i+j)^n}}}{\left(\frac{1}{i^{\frac{1}{3}}+j^{\frac{1}{3}}}\right)^3 \cdot \left((i+j)^{\frac{1}{3}}+1\right)^3}$ |
| Turbulence type 2 | Exponential law 2 | $\frac{\left(\frac{1}{i^{\frac{1}{3}}+1}\right)^3 \cdot (i+1)^{\frac{2}{9}} \cdot \left(\frac{1}{j^{\frac{1}{3}}+(i+1)^{\frac{1}{3}}}\right)^3 \cdot e^{\frac{\beta}{(1+i)^n} - \frac{\beta}{(i+j)^n}}}{\left(\frac{1}{i^{\frac{1}{3}}+j^{\frac{1}{3}}}\right)^3 \cdot (i+j)^{\frac{2}{9}} \cdot \left((i+j)^{\frac{1}{3}}+1\right)^3}$ |

Table 7. Example of kernel parameters fitting based on equilibrium state data (0% error)

| Kernel models | | Best fitting aggregation kernel constants | Best fitting breakage kernel constants | | | Fitting criterion |
|---------------|---------------|---|--|---------|--------|------------------------|
| Aggregation | Breakage | α | α | β | n | C |
| Turbulence 1 | Exponential 2 | 1 | 0.25 | 3.00 | 0.501 | 1.62×10^{-12} |
| Turbulence 1 | Power law | 1 | 0.0421 | - | 0.146 | 5.98×10^{-7} |
| Brownian | Power law | 1 | 0.246 | - | -0.915 | 1.01×10^{-5} |
| Turbulence 2 | Exponential 1 | 1 | 0.313 | 785 | 2.98 | 1.07×10^{-4} |
| Turbulence 2 | Exponential 2 | 1 | 0.703 | 183 | 2.21 | 1.08×10^{-4} |
| Turbulence 2 | Power law | 1 | 0.0573 | - | 0.518 | 5.04×10^{-4} |
| Brownian | Exponential 1 | 1 | 0.00231 | 4.31 | 0.604 | 1.05×10^{-3} |
| Brownian | Exponential 2 | 1 | 0.00231 | 4.31 | 0.604 | 1.13×10^{-3} |
| Turbulence 1 | Exponential 1 | 1 | 0.00231 | 4.31 | 0.604 | 1.95×10^{-2} |

Table 8. Example of kernel parameters fitting based on disturbed equilibrium state data (1% error)

| Kernel models | | Best fitting aggregation kernel constants | Best fitting breakage kernel constants | | | Fitting criterion |
|---------------|---------------|---|--|---------|-----------------------|-----------------------|
| Aggregation | Breakage | α | α | β | n | C |
| Turbulence 1 | Exponential 1 | 1 | 2.31×10^{-3} | 4.31 | 0.604 | 5.98×10^{-7} |
| Turbulence 1 | Exponential 2 | 1 | 3.98×10^{67} | 159 | 2.32×10^{-3} | 5.57×10^{-6} |
| Turbulence 1 | Power law | 1 | 0.0422 | - | 0.143 | 1.01×10^{-5} |
| Turbulence 2 | Exponential 2 | 1 | 0.742 | 77.8 | 1.84 | 1.07×10^{-4} |
| Turbulence 2 | Exponential 1 | 1 | 0.742 | 77.8 | 1.84 | 1.08×10^{-4} |
| Turbulence 2 | Power law | 1 | 0.056 | - | 0.523 | 5.04×10^{-4} |
| Brownian | Exponential 2 | 1 | 0.00231 | 4.31 | 0.604 | 1.05×10^{-3} |
| Brownian | Exponential 1 | 1 | 0.00231 | 4.31 | 0.604 | 1.13×10^{-3} |
| Brownian | Power law | 1 | 0.26 | - | -0.935 | 1.95×10^{-2} |

Table 9. Examples of kernel parameters fitting based on disturbed equilibrium state data (1% error) – reproducibility tests using correct theoretical kernel models

| | Best fitting aggregation kernel constants | Best fitting breakage kernel constants | | | Fitting criterion |
|--------------------|---|--|--------------------------|---------|-------------------|
| | α | α | β | n | C |
| Theoretical values | 1 | 0.25 | 3 | 0.5 | 0 |
| N°1 test | 1 | 0.1546 | 2400 | 3.23446 | 0.00051 |
| N°2 test | 1 | 0.135 | -1.6×10^{-14} | 46.68 | 0.00299 |
| N°3 test | 1 | 0.1278 | 1.34×10^{89} | 114.103 | 0.0027 |
| N°4 test | 1 | 0.134876 | 13.34 | 2.18957 | 0.00593 |
| N°5 test | 1 | 0.134956 | 87733 | 5.16 | 0.0065858 |
| N°6 test | 1 | 0.14008 | -2.2369×10^{15} | 50 | 0.014257 |
| N°7 test | 1 | 0.166198 | 86.2013 | 2.06 | 0.021 |
| N°8 test | 1 | 1.42×10^{160} | 372 | 0.00169 | 0.016 |
| N°9 test | 1 | 0.13748 | 7×10^{235} | 211 | 0.0284 |
| N°10 test | 1 | 0.1436 | -286 | 6.63 | 0.0579 |
| N°11 test | 1 | 0.132886 | -100 | 5.49 | 0.03299 |
| N°12 test | 1 | 0.113926 | -5.80392 | 1.418 | 0.0565 |
| N°13 test | 1 | 0.123744 | -5.56×10^{11} | 37 | 0.0893 |
| N°14 test | 1 | 2.93 | 10.17 | 0.3929 | 0.0615 |
| N°15 test | 1 | 0.1236 | 9.4×10^{307} | 284 | 0.0357 |
| N°16 test | 1 | 0.131805 | 2.4×10^{149} | 156 | 0.147 |
| N°17 test | 1 | 4.49×10^{80} | 189 | 0.0028 | 0.181 |

| | | | | | |
|-----------|---|------------------------|-----------------------|---------|--------|
| N°18 test | 1 | 0.1297 | 1.61×10^{46} | 93.89 | 0.285 |
| N°19 test | 1 | 0.136 | -629.6 | 8.37 | 0.174 |
| N°20 test | 1 | 1.33×10^{202} | 471 | 0.00284 | 0.0921 |
| N°21 test | 1 | 0.134 | 1.37×10^{12} | 10.9 | 0.17 |
| N°22 test | 1 | 0.144 | 1.2×10^6 | 5.64 | 0.2977 |
| N°23 test | 1 | 49.28 | 7.13 | 0.0552 | 0.1458 |
| N°24 test | 1 | 0.1897 | 30.48 | 1.477 | 0.153 |
| N°25 test | 1 | 1.8×10^{234} | 547 | 0.00298 | 0.466 |

Table 10. Examples of kernel parameters fitting based on disturbed equilibrium state data (2% error) – reproducibility tests using correct theoretical kernel models – T criterion

| | Best fitting aggregation kernel constants | Best fitting breakage kernel constants | | |
|--------------------|---|--|---------------------------|--------------------------|
| | α | α | β | n |
| Theoretical values | 1 | 1 | 3 | 1.5 |
| N°1 test | 1099.2 | 6.47607×10^{-27} | 28.8895 | 250.435 |
| N°2 test | 1094.63 | 5.54897×10^{-28} | 7.5477 | 470.113 |
| N°3 test | 1098.46 | 2.52139×10^{-27} | 28.8702 | 250.264 |
| N°4 test | 49.9474 | 4.40021×10^{-5} | 6.73752×10^{-13} | 0.000330121 |
| N°5 test | 30.8188 | 1.98245×10^{-13} | 1.79745×10^{-13} | 1.22582×10^{-5} |

Table 11. Example of kernel parameters fitting based on disturbed transitory states data (5% error)

| Kernel models | | Best fitting aggregation kernel constants | Best fitting breakage kernel constants | | | Fitting criterion |
|---------------|---------------|---|--|--------------------------|----------|-------------------|
| Aggregation | Breakage | α | α | β | n | T |
| Turbulence 1 | Exponential 1 | 4.05888 | 0.842453 | 1.45318 | 0.815055 | 3.91454 |
| Turbulence 1 | Exponential 2 | 3.99292 | 7.77817 | 3.4312 | 0.214202 | 3.92658 |
| Turbulence 1 | Power law | 4.07621 | 0.68399 | 0.0399478 | - | 3.94365 |
| Turbulence 2 | Exponential 1 | 1.32957 | 3.10842 | 17.3232 | 1.67806 | 4.48134 |
| Turbulence 2 | Power law | 1.572811 | 3.76433 | 1.318×10^{-14} | - | 5.31992 |
| Turbulence 2 | Exponential 2 | 4.14059 | 28.1234 | 12.1649 | 38.7169 | 6.7654 |
| Brownian | Exponential 2 | 27.2112 | 0.299807 | 3.9293×10^{-12} | 1.37684 | 102.421 |
| Brownian | Exponential 1 | 25.7857 | 3.1725×10^{-14} | 5.7908 | 0.185404 | 102.438 |
| Brownian | Power law | 25.7857 | 8.27797×10^{-15} | 0.00251735 | - | 102.438 |

Table 12. Examples of kernel parameters fitting based on disturbed transitory states data (2% error- reproducibility tests using correct theoretical kernel models - TC criterion minimization) : effects of proximity to equilibrium state and time interval between measurements

| | Best fitting aggregation kernel constants | Best fitting breakage kernel constants | | |
|--|---|--|---------------------------|----------|
| | α | α | β | n |
| Theoretical values | 1 | 1 | 3 | 1.5 |
| Time from initial state 0.01s - Time interval 0.025s | | | | |
| N°1 test | 0.997673 | 1.18519 | 2.4185 | 3.19936 |
| N°2 test | 0.961397 | 0.854704 | 1.86165 | 4.0272 |
| N°3 test | 0.982993 | 1.24039 | 3.3828 | 2.21797 |
| N°4 test | 0.94833 | 0.071937 | 5.11014 | 3.54243 |
| N°5 test | 1.03542 | 1.15425 | 3.97135 | 2.38127 |
| Time from initial state 0.05s - Time interval 0.025s | | | | |
| N°6 test | 0.95 | 0.857545 | 6.44983 | 0.981753 |
| N°7 test | 0.983844 | 0.934813 | 1.75802×10^{12} | 26.2046 |
| N°8 test | 1.01492 | 1.12928 | 25.8302 | 2.14533 |
| N°9 test | 0.970671 | 0.914566 | 5.63656×10^{-12} | 1.67278 |
| N°10 test | 0.981762 | 0.922952 | 1.95157 | 3.49263 |
| Time from initial state 0.01s - Time interval 0.005s | | | | |
| N°11 test | 1.0203 | 3.54121 | 14.1176 | 0.703108 |
| N°12 test | 0.978134 | 1.49193×10^{-13} | 2.05187 | 4.75196 |
| N°13 test | 1.06494 | 9.41688 | 477.607 | 3.29802 |
| N°14 test | 1.00915 | 2.24547 | 12.0581 | 0.968625 |

| | | | | |
|-----------|---------|---------------------------|---------|---------|
| N°15 test | 0.98008 | 3.33347×10^{-14} | 3.57407 | 1.08501 |
|-----------|---------|---------------------------|---------|---------|

REFERENCES

- (1) Ramkrishna, D. Population Balances: Theory and Applications to Particulate Systems in Engineering. Elsevier, 2000.
- (2) Narsimhan, G.; Ramkrishna, D.; Gupta, J. P. Analysis of drop size distributions in lean liquid-liquid dispersions. *AIChE J.* **1980**, *26* (6), 991–1000.
- (3) Sathyagal, A. N.; Ramkrishna, D.; Narsimhan, G. Solution of Inverse Problems in Population Balances-II. Particle Break-Up. *Comput. Chem. Eng.* **1995**, *19* (4), 437–451.
- (4) Wright, H.; Ramkrishna, D. Solutions of Inverse Problems in Population Balances—I. Aggregation Kinetics. *Comput. Chem. Eng.* **1992**, *16* (12), 1019–1038.
- (5) Dürr, R. Parameter Estimation and Method of Moments for Multi Dimensional Population Balance Equations with Application to Vaccine Production Processes. Ph.D. Dissertation, Otto-von-Guericke-Universität, Magdeburg, 2016.
- (6) Nguyen, T. T.; Laurent, F.; Fox, R.; Massot, M. Solution of Population Balance Equations in Applications with Fine Particles: Mathematical Modeling and Numerical Schemes. *J. Comput. Phys.* **2016**, *325*, 129–156.
- (7) Singh, M. Forward and Inverse Problems in Population Balances. Ph.D. Dissertation, IIT, Kharagpur, 2015.
- (8) Vikhansky, A.; Kraft, M.; Simon, M.; Schmidt, S.; Bart, H.-J. Droplets Population Balance in a Rotating Disc Contactor: An Inverse Problem Approach. *AIChE J.* **2006**, *52* (4), 1441–1450.
- (9) Chakraborty, J.; Kumar, J.; Singh, M.; Mahoney, A.; Ramkrishna, D. Inverse Problems in Population Balances. Determination of Aggregation Kernel by Weighted Residuals. *Ind. Eng. Chem. Res.* **2015**, *54* (42), 10530–10538.
- (10) Tahamtan, S.; Halvae, A.; Emamy, M.; Zabihi, M. S. Fabrication of Al/A206–Al2O3 Nano/Micro Composite by Combining Ball Milling and Stir Casting Technology. *Mater. Des.* **2013**, *49*, 347–359.

- (11) Agarwala, V.; Dixit, D. Fabrication of Aluminium Base Composite by Foundry Technique. *Trans. Jpn. Inst. Met.* **1981**, *22* (8), 521–526.
- (12) Pai, B. C.; Ray, S.; Prabhakar, K. V.; Rohatgi, P. K. Fabrication of Aluminium-Alumina (Magnesia) Particulate Composites in Foundries Using Magnesium Additions to the Melts. *Mater. Sci. Eng.* **1976**, *24* (1), 31–44.
- (13) Auger, J. M.; Martin, S.; Gruy, F. Wettability-Enabling Coating on Oxide Particles through Controlled Milling. *Powder Technol.* **2019a**, *344*, 302–306.
- (14) Pai, B. C.; Ramani, G.; Pillai, R. M.; Satyanarayana, K. G. Role of Magnesium in Cast Aluminium Alloy Matrix Composites. *J. Mater. Sci.* **1995**, *30* (8), 1903–1911.
- (15) Auger, J.-M.; Martin, S.; Gruy, F. Preparation of Alumina Particle Suspension in Liquid Tin Using a Pre-Coating Process. *Metall. Res. Technol.* **2019b**, *116* (5), 510.
- (16) Schultz, B. F.; Ferguson, J. B.; Rohatgi, P. K. Microstructure and Hardness of Al₂O₃ Nanoparticle Reinforced Al–Mg Composites Fabricated by Reactive Wetting and Stir Mixing. *Mater. Sci. Eng. A* **2011**, *530*, 87–97.
- (17) Li, T.; Shimasaki, S.; Taniguchi, S.; Uesugi, K.; Narita, S. Morphology of Nonmetallic-Inclusion Clusters Observed in Molten Metal by X-Ray Micro-Computed Tomography (CT). *ISIJ Int.* **2013**, *53* (11), 1943–1952.
- (18) Anderson, D. F.; Craciun, G.; Kurtz, T. G. Product-Form Stationary Distributions for Deficiency Zero Chemical Reaction Networks. *Bull. Math. Biol.* **2010**, *72* (8), 1947–1970.
- (19) Lachin, K. De La Stabilité à La Coagulation de Latex Acrylique : Étude Des Mécanismes et Mise En Oeuvre En Milliréacteur. Ph.D. Dissertation, INPT, Toulouse, 2016.
- (20) Cheng, Z.; Redner, S. Kinetics of Fragmentation. *J. Phys. Math. Gen.* **1990**, *23* (7), 1233–1258.
- (21) Lachin, K.; Le Sauze, N.; Di Miceli Raimondi, N.; Aubin, J.; Gourdon, C.; Cabassud, M. Aggregation and Breakup of Acrylic Latex Particles inside Millimetric Scale Reactors. *Chem. Eng. Process. Process Intensif.* **2017**, *113*, 65–73.

- (22) Celada Murillo, A. T.; Salcido, A. A Mean Field Model for Brownian and Turbulent Coagulation of Polydispersed Aerosols. *Rev. Mex. Física* **2005**, *51*, 379.
- (23) Ji, H. Uniqueness of Equilibria for Complex Chemical Reaction Networks, Ph.D. Dissertation, The Ohio State University, 2011.
- (24) Joshi, B.; Shiu, A. A Survey of Methods for Deciding Whether a Reaction Network Is Multistationary. *Math. Model. Nat. Phenom.* **2015**, *10* (5), 47–67.
- (25) Feinberg, M. Chemical Reaction Network Structure and the Stability of Complex Isothermal Reactors—I. The Deficiency Zero and Deficiency One Theorems. *Chem. Eng. Sci.* **1987**, *42* (10), 2229–2268.
- (26) Ellison, P. R. The Advanced Deficiency Algorithm and Its Applications to Mechanism Discrimination. PhD Thesis, The University of Rochester, Eastman School of Music, 1998.
- (27) Craciun, G.; Feinberg, M. Multiple Equilibria in Complex Chemical Reaction Networks: I. The Injectivity Property. *SIAM J. Appl. Math.* **2005**, *65* (5), 1526–1546.
- (28) Shinar, G.; Feinberg, M. Concordant Chemical Reaction Networks. *Mathematical Biosciences* **2012**, *240* (2), 92–113.
- (29) Craciun, G.; Feinberg, M. Multiple Equilibria in Complex Chemical Reaction Networks: II. The Species-Reaction Graph. *SIAM J. Appl. Math.* **2006**, *66*, 1321–1338.
- (30) Feinberg, M. Chemical Reaction Network Structure and the Stability of Complex Isothermal Reactors—II. Multiple Steady States for Networks of Deficiency One. *Chem. Eng. Sci.* **1988**, *43* (1), 1–25.
- (31) Gopalkrishnan, M.; Miller, E.; Shiu, A. A Geometric Approach to the Global Attractor Conjecture. *SIAM J. Appl. Dyn. Syst.* **2014**, *13* (2), 758–797.
- (32) Angeli, D.; Sontag, E. A Tutorial on Monotone Systems- With an Application to Chemical Reaction Networks. In *Proceedings of the 16th International Symposium on Mathematical Theory of Networks and Systems*; Katholieke Universiteit Leuven, 2004; pp 1–14.

- (33) Anderson, D. F. Boundedness of Trajectories for Weakly Reversible, Single Linkage Class Reaction Systems. *J. Math. Chem.* **2011**, *49* (10), 2275.
- (34) Shinar, G.; Feinberg, M. Concordant Chemical Reaction Networks and the Species-Reaction Graph. *Math. Biosci.* **2013**, *241* (1), 1–23.
- (35) Craciun, G. Toric Differential Inclusions and a Proof of the Global Attractor Conjecture. arXiv preprint, arXiv:1501.02860, 2016, 1–41.
- (36) Boros, B. Notes on the Deficiency-One Theorem: Multiple Linkage Classes. *Math. Biosci.* **2012**, *235* (1), 110–122.
- (37) Akle, S.; Dalal, O.; Fleming, R. M. T.; Saunders, M.; Taheri, N.; Ye, Y. Existence of Positive Steady States for Mass Conserving and Mass-Action Chemical Reaction Networks with a Single Terminal-Linkage Class. arXiv preprint, arXiv:1105.2359, 2011, 1–17.
- (38) Otero-Muras, I.; Szederkényi, G.; Alonso, A. A.; Hangos, K. M. Local Dissipative Hamiltonian Description of Reversible Reaction Networks. *Syst. Control Lett.* **2008**, *57* (7), 554–560.
- (39) Joshi, B.; Shiu, A. Atoms of Multistationarity in Chemical Reaction Networks. *J. Math. Chem.* **2013**, *51* (1), 153–178.
- (40) Dubovskii, P. Mathematical Theory of Coagulation, Volume 23 of Lecture Notes. Global Analysis Research Center, Seoul National University 1994.
- (41) Collet, J.-F. Some Modelling Issues in the Theory of Fragmentation-Coagulation Systems. *Commun. Math. Sci.* **2004**, *2* (Supplemental Issue), 35–54.
- (42) Auger, J.-M. PopulationBalance C++ Library for Population Balance Problems Simulations, Gitlab, <https://gitlab.com/AugerJM/PopulationBalance>, 2020.
- (43) Nelder, J. A.; Mead, R. A Simplex Method for Function Minimization. *Comput. J.* **1965**, *7* (4), 308–313.

- (44) Smoluchowski, M. V. Drei Vortrage Uber Diffusion, Brownsche Bewegung Und Koagulation von Kolloidteilchen. *Z. Phys.* **1916**, *17*, 557–585.
- (45) Kruis, F. E.; Kusters, K. A. The Collision Rate of Particles in Turbulent Flow. *Chem. Eng. Commun.* **1997**, *158* (1), 201–230.
- (46) Saffman, P. G.; Turner, J. S. On the Collision of Drops in Turbulent Clouds. *J. Fluid Mech.* **1956**, *1* (1), 16–30.
- (47) Camp, T. R.; Stein, P. C. Velocity Gradients and Internal Work in Fluid Motion. *J. Boston Soc. Civ. Eng.* **1943**, *85*, 219–237.
- (48) Levich, V. G. *Physicochemical Hydrodynamics.*; Prentice-Hall: Englewood Cliffs, N.J., 1962.
- (49) Valentas, K. J.; Bilous, O.; Amundson, N. R. Analysis of Breakage in Dispersed Phase Systems. *Ind. Eng. Chem. Fundam.* **1966**, *5* (2), 271–279.
- (50) Ramkrishna, D. Drop-Breakage in Agitated Liquid—Liquid Dispersions. *Chem. Eng. Sci.* **1974**, *29* (4), 987–992.
- (51) Delichatsios, M. A.; Probstein, R. F. Coagulation in Turbulent Flow: Theory and Experiment. *J. Colloid Interface Sci.* **1975**, *51* (3), 394–405.
- (52) Marchisio, D. L.; Soos, M.; Sefcik, J.; Morbidelli, M. Role of Turbulent Shear Rate Distribution in Aggregation and Breakage Processes. *AIChE J.* **2006**, *52* (1), 158–173.
- (53) Selomulya, C.; Bushell, G.; Amal, R.; Waite, T. Understanding the Role of Restructuring in Flocculation: The Application of a Population Balance Model. *Chem. Eng. Sci.* **2003**, *58*, 327–338.