



Research article

Mean first assembly times in coagulation-fragmentation systems: Theoretical, stochastic, and approximate approaches

Pradeep Kunwar and Katrin Rohlf*

Department of Mathematics, Toronto Metropolitan University, 350 Victoria Street, Toronto, Ontario, M5B 2K3, Canada

* **Correspondence:** Email: krohlf@torontomu.ca; Tel: +1-416-979-5000.

Abstract: Motivated by the application of a stochastic particle-based simulation model to study finite coagulation-fragmentation systems, an approximate coagulation-fragmentation model is introduced for the dynamics. Traditional large and small population examples are explored theoretically, as well as numerically, and approximate models are compared to traditional well-known coagulation-fragmentation frameworks. Our new approximate coagulation-fragmentation models are compared to traditional descriptions both in large population dynamical systems settings, as well as in small population Mean-First assembly time (MFAT) contexts. Stochastic simulations in the form of Gillespie's Stochastic Simulation Algorithm (SSA) for exact and approximate coagulation-fragmentation dynamics, as well as Reactive Multiparticle Collision Dynamics (RMPC) for particle-based versions thereof, are provided, as well as a numerical matrix inversion method for larger systems that is helpful to obtain theoretical MFAT estimates for larger populations. We also consider differences between a chemical system interpretation of the dynamics, Becker-Döring models, and Smoluchowski dynamics. As conventionally explored, formation and break-up rates in the form of constant, additive and multiplicative rates are considered, as well as large initial monomer-only conditions. Finally, a new time-rescaling mechanism is introduced so that MFAT results for RMPC simulations (for both exact and approximate models) can be obtained when the MFAT values are comparable to, or less than, an RMPC time step. For the traditional break-up rates considered in this paper, our findings show that MFATs are fairly comparable for the exact and approximate systems, and that an approximate RMPC model can be used in this context. Since RMPC is a stochastic simulation algorithm that explicitly incorporates positions and velocities of particles in the system, velocity-dependent reaction rates can be easily added to the reaction dynamics, as well as diffusive effects. Since RMPC can be used as a particle-based fluid-flow solver, and since our new approximate self-assembly process can lead to numerical speed-up of the self-assembly process in large population flow-problems, our results thus provide the stepping stone for a simulation paradigm that is capable of furthering our understanding of how self-assembly processes may be affected by diffusion or when subjected to flow conditions. The extent of the performance enhancement and accuracy of the approximate RMPC method applied to the

self-assembly process in flow conditions has yet to be explored and is part of future work.

Keywords: reactive multiparticle collision dynamics; stochastic self-assembly; mean first assembly times; finite systems; coagulation; fragmentation

Mathematics Subject Classification: 65C35, 82C05, 37M05, 92E99

1. Introduction

Understanding the self-assembly process, that is the clustering or aggregation of smaller particles into larger clusters, has important applications in chemistry, biology, polymer science and colloid chemistry [1–5]. Models primarily involve mass-action or mean-field rate laws describing the time-evolution of the concentrations of clusters of different sizes in the form of differential equations. In the case of monomer-only attachment — that is attachment with particles of size 1 only — the Becker-Döring system of differential equations [6] is commonly used. When self-assembly allows for cluster formation between two particles of any size, the Smoluchowski system of differential equations [7, 8] is used, which is incidentally also for an infinite system. Closed-form solutions exist in some special cases for the infinite system, and formation and break-up rates commonly used are constant, additive, or multiplicative in nature [8, 9].

It has recently been shown that self-assembly systems can be sensitive to stochastic effects [4, 5, 10]. First, systems that involve a maximum cluster size constraint, can show significant differences from their mean-field counterparts in that the mean-field equilibria can differ in the stochastic systems compared to the mean-field values. Additionally, the mean-first assembly time (MFAT) in the stochastic system can be highly dependent on the initial monomer concentration relative to the maximal cluster size value, potentially giving rise to trapped states that can lead to large MFAT times when break-up rates are small.

Stochastic self-assembly simulations have recently been carried out using Kinetic Monte-Carlo simulations (KMC), and compared to theoretical expressions derived from the survival probability of the corresponding backward Kolmogorov equation [4, 5, 10]. In this paper, we use the Gillespie algorithm, commonly called the Stochastic Simulation algorithm (SSA) [11], as well as Reactive Multiparticle collision dynamics (RMPC) [12–15]. We present the chemical rate laws representative of corresponding chemical reactive mechanisms for self-assembly processes, as well as an approximate reactive mechanism that would more commonly be used in RMPC dynamics implementation [14–19]. Using the methods outlined in Yvinec et. al. [4], theoretical MFAT expressions are derived here for both approaches, as well as a numerical matrix inversion algorithm. Theoretical MFAT results are compared to exact and approximate stochastic self-assembly reaction dynamics also used in our stochastic simulations. Constant, additive and multiplicative rates are considered. In the case of RMPC dynamics, a time rescaling approach is introduced so that MFATs that are smaller than the RMPC time step, can be obtained from RMPC simulations. Additionally, we provide ODE and MFAT expressions that distinguish between a chemical rate law interpretation of the dynamics, a Becker-Döring (BD) model, as well as dynamics corresponding to the Smoluchowski model.

The benefit of developing an RMPC algorithm for self-assembly processes is that RMPC without reactions, namely MPC (multiparticle collision dynamics), has been successfully used for mesoscopic

simulations for various flow problems [20–23]. Incorporating the self-assembly process as a reaction mechanism into the MPC simulation paradigm using an RMPC approach (which is just MPC with added reactions), will allow for the possibility of understanding how self-assembly processes may be affected by diffusion or when subjected to flow conditions, since velocities of the clusters can be easily incorporated into the self-assembly reaction rates. Adding velocity-dependence into the reaction rates of the self-assembly process in an RMPC flow simulation will allow for the possibility to implement the concept that clusters that have a small relative speed are more likely to stick together to form a larger cluster than those with a larger speed differential. As such, the self-assembly process can be studied subject to flow conditions so as to further our understanding on what shear-rate dependent coalescence and break-up rates should look like for blood flow. Shear-rate dependent rates have been proposed in various works related to blood [24–27], however adequate simulation mechanisms that can provide true insight into the coupling of flow to the self-assembly process do not exist. While diffusive effects could be simulated using more popular stochastic simulation approaches, methods such as the Inhomogeneous Stochastic Simulation Algorithm (ISSA) for example, that can be used to solve the Reaction-diffusion Master equation (RDME) exactly, are not suitable to study fluid flow problems, can be quite slow for the self-assembly process considered here [28], and are not capable to directly tie individual particle-level velocity details to the cluster formation/break-up process. RMPC as proposed here can be used in this context. The reason that an approximate self-assembly process is also introduced here rather than just an exact mechanism, rests on the fact that MPC-flow simulations are computationally demanding, and an approximate self-assembly process as a first assessment can provide a computationally feasible means to begin the exploration of how self-assembly processes are intrinsically coupled to flow conditions.

The main novelty of this work is the introduction of an approximate self-assembly process that can easily be implemented in a particle-based numerical model such as RMPC, and that can lead to computational speed-up in large population applications. We consider exact and approximate chemical rate law interpretations of the self-assembly model, exact and approximate monomer-only Becker-Döring dynamics and related extensions, as well as exact and approximate finite Smoluchowski equations. We compare RMPC simulation results to corresponding exact and approximate theoretical contexts in large-population systems (ODE rate laws), and small-population Mean-first assembly time examples. Our main motivation for the introduction of our approximate RMPC self-assembly model, is that it has the potential to become a foundational building block in a computationally-efficient particle-based RMPC flow model, capable of intrinsically coupling the particle-based self-assembly to the fluid flow.

This paper is organized as follows. In Section 2 we provide the reaction mechanisms for coagulating-fragmenting systems for exact (known) and approximate (novel) reaction dynamics, together with the corresponding rate laws, in the form of ordinary differential equations (ODEs). Section 3 provides a brief review on how a stochastic master equation can be used to obtain theoretical Mean First Assembly times (MFATs). In Section 4, we present theoretical expressions for MFATs when the maximum cluster size is $N = 3$ for constant, additive and multiplicative rate coefficients. The numerical implementation methods in the form of matrix inversion of the transition matrix, as well as stochastic SSA and RMPC simulations are discussed in the next Section 5, including a time rescaling method necessary for RMPC simulations when MFATs are comparable to, or smaller than, the RMPC time step. Section 6 contains our theoretical and numerical results. Finally, a summary and conclusions

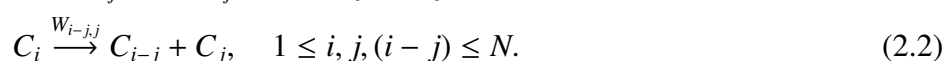
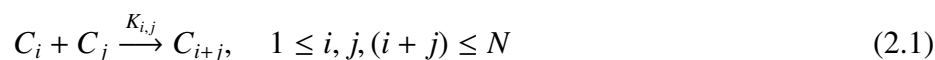
are provided in the last Section 7.

2. Well-stirred coagulation and fragmentation

In this section we provide the various rate laws in the form of ODEs, that have commonly been applied to the self-assembly process in large population systems. As has recently been pointed out in [29,30], use of a stochastic simulation mechanism for a chemical system interpretation of the dynamics, leads to subtle differences in the rate laws, that are worth explicitly pointing out here. We provide the relevant rate laws for dynamics for pairwise interaction possibilities of any two clusters (Sections 2.1), as well as rate laws for monomer-only interactions (Section 2.3), and finally in-between models in which clusters up to a restricted maximum size can take part in formation/break-up interactions (Section 2.4). Our newly proposed approximate dynamics in Section 2.2 can be applied in any of these contexts (limited or unrestricted pair-wise cluster interactions). Since our focus is on the chemical rate law interpretation of the approximate dynamics that can be compared to our RMPC simulations, we explicitly provide the $N = 3$ finite case rate laws for the exact and approximate dynamics in Section 2.6. The corresponding chemical systems from Sections 2.1 and 2.2 will also be used to obtain the theoretical and numerical MFAT results in exact (known) and approximate (novel) dynamics in later parts of this paper, for when we provide smaller-population results and comparisons.

2.1. Chemical rate law for exact dynamics

Coagulation, also called aggregation, can be described by a reactive mechanism in which a cluster of size i joins with a cluster of size j to create a larger cluster of size $(i + j)$ with rate coefficient $K_{i,j}$. Fragmentation, or break-up, occurs when a cluster of size $(i + j)$ breaks into two clusters of size i and j , at rate $W_{i,j}$. In the case of a finite-sized system, there is a maximum cluster size N , which means that two clusters cannot combine if their combined size exceeds this maximum value. The coagulation and fragmentation process can be expressed as a reaction mechanism in the form



where C_i denotes a cluster of size i with constant ($K_{i,j} = p$, $W_{i,j} = q$), additive ($K_{i,j} = p(i + j)$, $W_{i,j} = q(i + j)$) or multiplicative ($K_{i,j} = pij$, $W_{i,j} = qij$) rates.

In a large, well-mixed system, the time evolution of cluster concentrations for this reaction mechanism can be described using a stochastic model with a corresponding Master Equation (ME). The corresponding chemical rate law (chem-exact) takes the form

$$\begin{aligned} \frac{dC_k}{dt} = & \frac{1}{2} \sum_{i=1}^{k-1} (1 + \delta_{i,\frac{k}{2}}) K_{i,k-i} C_i C_{k-i} \\ & - \sum_{i=1}^{N-k} (1 + \delta_{i,k}) K_{i,k} C_i C_k \\ & + \sum_{i=k+1}^N (1 + \delta_{i,2k}) W_{i-k,k} C_i \end{aligned} \quad (2.3)$$

$$-\frac{1}{2} \sum_{i=1}^{k-1} (1 + \delta_{i, \frac{k}{2}}) W_{k-i, i} C_k,$$

for $1 \leq k \leq N$. In (2.3), C_i denotes the concentration of a cluster of size i , and $\delta_{i,j}$ is the Kronecker delta function whose value is 1 whenever $i = j$ and zero otherwise.

It is worth noting that this chemical rate law conserves the number of monomers (or mass) in the system, namely that

$$\sum_{k=1}^{\infty} k C_k(t) = \sum_{k=1}^{\infty} k C_k(0). \quad (2.4)$$

Conservation of the total number of monomers is a mathematical consequence of the equations in the rate law.

2.2. Chemical rate law for approximate dynamics

In large well-mixed systems, it is often convenient and computationally advantageous, to describe the reaction mechanism in an approximate sense. An approximate reaction mechanism for the reactions given in (2.1) and (2.2) is given by



The corresponding chemical rate law (chem-approx) is given by

$$\begin{aligned} \frac{dC_k}{dt} = & \frac{1}{2} \sum_{i=1}^{k-1} (1 + \delta_{i, k-i}) K_{i, k-i} C_i C_{k-i} - \frac{1}{2} \sum_{i=1}^{N-k} (1 + \delta_{i, k}) K_{i, k} C_i C_k \\ & + \frac{1}{2} \sum_{i=k+1}^N (1 + \delta_{i, 2k}) W_{i-k, k} C_i - \frac{1}{2} \sum_{i=1}^{k-1} (1 + \delta_{i, k/2}) W_{k-i, i} C_k, \end{aligned} \quad (2.10)$$

for $1 \leq k \leq N$.

In the implementation of the approximate dynamics, the total number of clusters remains constant in time, rather than the total mass as in the exact dynamics, that is

$$\sum_{k=1}^N C_k(t) = \sum_{k=1}^N C_k(0), \quad (2.11)$$

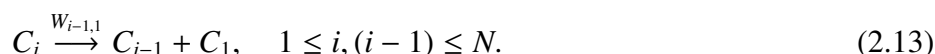
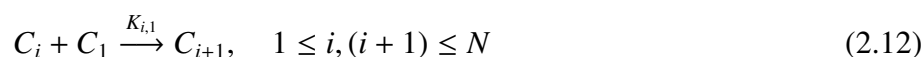
which is true for finite N as well as for $N \rightarrow \infty$. Conservation of the total number of clusters in the approximate dynamics, rather than total number of monomers in the exact dynamics, is a mathematical

consequence of the equations in the approximate rate law. In agreement with this, since the number of reactants is the same as the number of products in the representative chemical reactions for the approximate dynamics, cluster count will be conserved.

The computational efficiency of this approach rests on the particle-based implementation of the approximate reaction mechanism, in which a given particle simply changes species type at a certain rate. In the exact aggregation mechanism, two particles of a specific type, located in the same collision volume, must be found in the system, so that they can both be removed and replaced with the larger-sized cluster. Based on preliminary results (not included), the exact aggregation process can be computationally expensive when there are a large number of particles in the system, namely in the hundreds of thousands. The approximate approach is significantly faster for systems with large numbers of particles, such as what would be necessary for fluid flow applications.

2.3. Becker-Döring (BD) rate laws

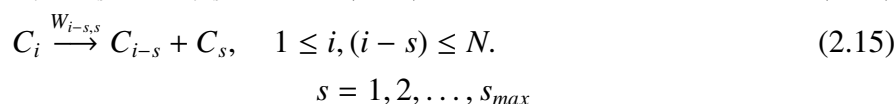
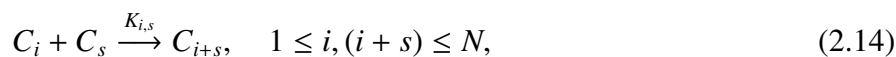
The rate laws corresponding to Becker-Döring (BD) dynamics consider only aggregation events between any-sized clusters with monomers (clusters of size 1), and break-up events that result in the formation of a monomer. The exact reactive BD events are thus



The corresponding BD rate law for these exact reactions (BD-exact) can be obtained from the exact chemical rate law in equation (2.3) by replacing all $(1 + \delta_{i,j})K_{i,j}$ with $K_{i,j}$, and additionally setting appropriate rate coefficients $K_{i,j}$ and $W_{i-j,j}$ to zero whenever $j \neq 1$. It is worth pointing out that the formation rates between monomers in this (well-known) BD rate law, is thus half the formation rate between monomers in the chemical rate law (see also Table 2). Similarly, a BD rate law for approximate reactions (BD-approx) can be obtained by replacing all $(1 + \delta_{i,j})K_{i,j}$ with $K_{i,j}$ in equation (2.10), and additionally setting appropriate rate coefficients $K_{i,j}$ and $W_{i-j,j}$ to zero whenever $j \neq 1$.

2.4. Extended Becker-Döring (EBD) rate laws

The extended Becker-Döring rate laws extend the Becker-Döring aggregation and break-up events between any-sized clusters with monomers, to include events with clusters up to size s_{max} . The reactive events are



The corresponding rate laws (EBD-exact and EBD-approx) can be obtained from the exact and approximate chemical rate laws, i.e. equations (2.3) and (2.10), by setting appropriate rates $K_{i,j}$ and $W_{i-j,j}$ to zero whenever $j > s_{max}$, and halving the formation rates between like-sized clusters by replacing $(1 + \delta_{i,j})K_{i,j}$ with $K_{i,j}$.

2.5. Smoluchowski equation and general rate law comparison

The Smoluchowski rate law, or more commonly called the Smoluchowski equation, has been extensively used to describe the self-assembly process in the form of rate laws. It was first introduced in an infinite context ($N \rightarrow \infty$). Taking the $N \rightarrow \infty$ limit in the chemical rate law for exact dynamics in Section 2.1, the Smoluchowski equation can be obtained from equation (2.3) by replacing $(1 + \delta_{i,j})K_{i,j}$ and $(1 + \delta_{i,j})W_{i,j}$ with $K_{i,j}$ and $W_{i,j}$. Relative to the exact chemical rate law with $N \rightarrow \infty$, formation rates between like-sized clusters is halved, and the break-up rates of a large cluster into two same-sized clusters is also halved. In contrast, relative to the EBD model with $s_{max} \rightarrow \infty$, only the break-up rates of larger clusters into two same-sized clusters are halved. Table 2 provides this relative summary between the rate differences in all models.

While there are slight differences in the rates in the different rate laws, note that for the case $N = 3$, the underlying clustering dynamics of the EBD and BD models are the same as the (full) clustering dynamics for the chemical rate laws. When $N = 4$ or 5 , the reactions for EBD with $s_{max} = 2$ extends the BD dynamics to include the additional $C_2 + C_2 \leftrightarrow C_4$ ($N = 4, 5$), or $C_2 + C_3 \leftrightarrow C_5$ ($N = 5$) reactions, but the EBD has the same full set of reactions as the full set of reactions used in the chemical rate laws. For $N = 6$, the EBD has additional reactions compared to the BD model, but will contain less reactions than the full set for the chemical rate law if $s_{max} = 2$. In general, for $N \geq 6$, the EBD will extend the reactions in BD, but contain less reactions than the full set used for the chemical rate laws, whenever $2 \leq s_{max} \leq \lfloor N/2 \rfloor - 1$. Finally, EBD corresponds to the (finite) Smoluchowski dynamics, whenever $s_{max} = \lfloor N/2 \rfloor$.

All rate laws for exact dynamics (chem-exact, BD-exact, EBD-exact, Smol-exact) satisfy conservation of mass (2.4) mathematically, while all rate laws for approximate dynamics satisfy conservation of the number of clusters (2.11) mathematically.

2.6. The finite case $N = 3$

For the case when $N = 3$, the Chemical Rate Law for the exact dynamics (chem-exact) takes the form

$$\begin{aligned}\frac{dC_1}{dt} &= -2K_{11}C_1^2 - K_{12}C_1C_2 + 2W_{11}C_2 + W_{12}C_3 \\ \frac{dC_2}{dt} &= K_{11}C_1^2 - K_{12}C_1C_2 + W_{12}C_3 - W_{11}C_2 \\ \frac{dC_3}{dt} &= K_{12}C_1C_2 - W_{12}C_3,\end{aligned}$$

and the Chemical Rate Law for the approximate dynamics (chem-approx) is given by

$$\begin{aligned}\frac{dC_1}{dt} &= -K_{11}C_1^2 - \frac{1}{2}K_{12}C_1C_2 + W_{11}C_2 + \frac{1}{2}W_{12}C_3 \\ \frac{dC_2}{dt} &= K_{11}C_1^2 - \frac{1}{2}K_{12}C_1C_2 + \frac{1}{2}W_{12}C_3 - W_{11}C_2 \\ \frac{dC_3}{dt} &= K_{12}C_1C_2 - W_{12}C_3.\end{aligned}$$

For completeness, note that halving K_{11} and W_{11} in these equations leads to the (finite) Smoluchowski dynamics, while only halving K_{11} leads to the BD/EBD model [29] (see also Table 2 and Appendix A).

In the sequel, we will discuss both exact and approximate dynamics in the chemical rate law context, as well as in the (finite) Smoluchowski and BD/EBD contexts, by halving the relevant rates and/or setting some of them to zero, as appropriate.

In Figure 1, the time evolution of the concentration of monomers C_1 , dimers C_2 and trimers C_3 are shown from an initial monomer-only concentration of 20 with constant aggregation and break-up rates. Both exact and approximate rate law solutions are provided, for the chemical rate law, their BD/EBD version and the finite Smoluchowski model. In both the exact (dashed lines) and approximate (solid lines) dynamics, it can be seen that the BD/EBD and (finite) Smoluchowski solution curves are in close agreement initially, while the chemical rate law and finite Smoluchowski solutions are in close agreement at equilibrium. The local maximum in C_2 occurs at approximately the same time for all three models in the exact as well as approximate solution curves, however the maximum occurs earlier in the exact dynamics compared to the approximate model. The local maximum values of dimers of the finite Smoluchowski model is in better agreement with the BD/EBD maximum value, while the equilibrium value is in better agreement with the chemical rate law. While monomer and trimer curves between the three different models (chemical rate law, BD/EBD and finite Smoluchowski) are fairly close to one another, the difference between the approximate and exact equilibrium values increases as the cluster size increases.

It is worth pointing out that one can increase the break-up rates by several orders of magnitude, to reduce the relative error between the exact and approximate models, as shown in Figure 2. It can be seen in the Figure, that the percent relative error in the chemical rate law model, can be reduced significantly for all three chemical species, when the break-up rates are increased to 200 times the formation rates (see blue dashed curves).

Although the rate coefficients are taken to be constant in Figures 1 and 2, for a demonstration of the differences between the chemical rate law, the BD/EBD dynamics and the Smoluchowski model in large population examples, the MFAT results for constant, additive and multiplicative rate coefficients in the Results Section 6 also involve forward rate coefficients that are different in value from the backward rate coefficients. This can also be seen in discussions pertaining to the time rescaling RMPC simulations related to Table 5.

3. Mean first assembly time

In order to validate the approximate RMPC simulation results, together with their corresponding theoretical models that we have discussed, we now consider small-population applications of the approximate dynamics, in the context of Mean First Assembly Times. To briefly review, the Mean First Assembly Time (MFAT) is the average time it takes for the largest cluster to form from an initial concentration of smallest clusters. Using a Master equation approach [5] for probability density functions P , MFATs can be computed from the survival probability $S(\mathbf{m}; t)$, defined as

$$S(\mathbf{m}; t) = \sum_{\{\mathbf{n}, n_N=0\}} P(\mathbf{n}; t|\mathbf{m}; 0).$$

Here $P(\mathbf{n}; t|\mathbf{m}; 0)$ is the probability that the cluster configuration is given by $\mathbf{n} = (n_1, n_2, \dots, n_N)$ at time t , given that it was initially in state $\mathbf{m} = (m_1, m_2, \dots, m_N)$. Each of the bold/vector states has a component for each cluster size, up to a maximum cluster size N , and the N -th entry n_N for example,

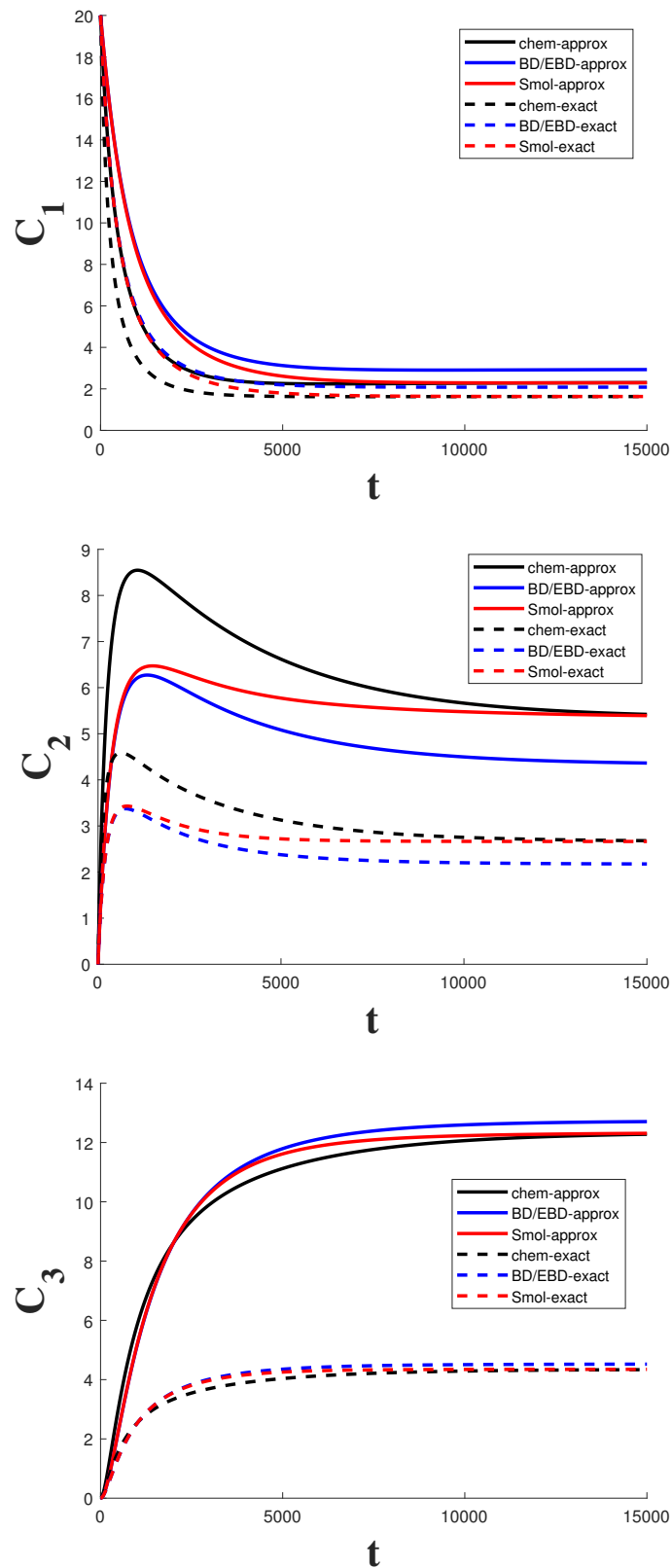


Figure 1. Comparison of the exact and approximate Chemical Rate Law, the BD/EBD rate law and the (finite) Smoluchowski equation ($N = 3$, $K_{ij} = 0.0001$, $W_{ij} = 0.0001$).

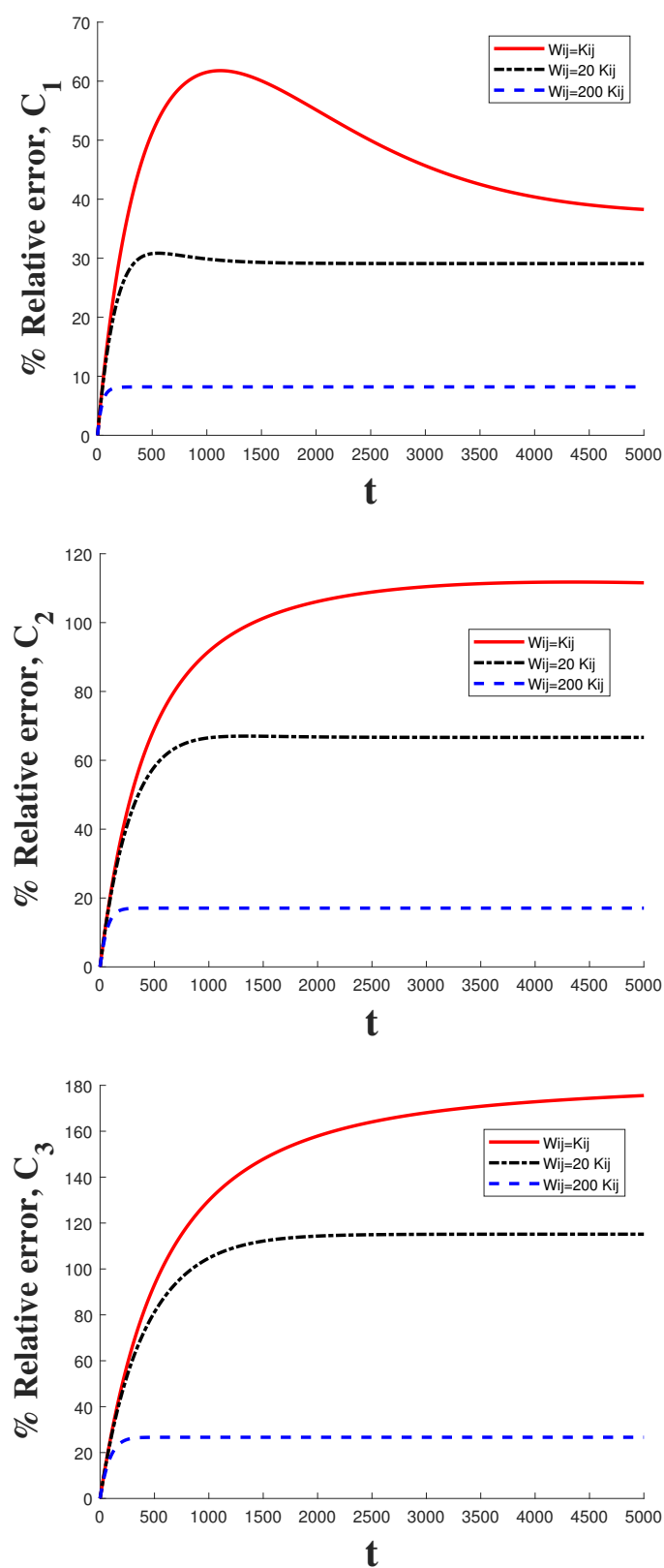


Figure 2. Comparison of the percent relative error of C_1 - C_3 between the Chemical Rate Law for exact (chem-exact) and approximate (chem-approx) dynamics as break-up rates increase ($K_{ij} = 0.0001$).

corresponds to the number of particles of cluster size N . The sum in the survival probability S is thus notably restricted to states for which there is no maximal cluster, so that the first assembly time is not zero.

For a given system that is prescribed to contain M base units, for which the largest cluster size is chosen to be N , there will be a finite number of possible different initial configurations that are free of clusters of maximal size N . One can compute a survival probability $S(\mathbf{m}^i; t)$ for each possible initial state $\mathbf{m}^i$, together with a corresponding mean first assembly time $T(\mathbf{m}^i)$ based on the same initial state. More compactly, each of these survival probabilities with corresponding mean first assembly time can be combined in a survival probability vector $\mathbf{S}(\mathbf{m}; t)$ and mean first assembly time vector $\mathbf{T}(\mathbf{m})$. The time evolution of the survival probability vector satisfies the same time evolution as the probability density of the backward Kolmogorov equation, namely

$$\mathbf{S}'(\mathbf{m}; t) = \mathbf{A}^+ \mathbf{S}(\mathbf{m}; t),$$

where the prime denotes differentiation with respect to time, and $\mathbf{A}^+$ is the transpose of the transition matrix $\mathbf{A}$ of the forward master equation $\mathbf{P}'(\mathbf{n}; t|\mathbf{m}; 0) = \mathbf{A}\mathbf{P}(\mathbf{n}; t|\mathbf{m}; 0)$.

The survival probability satisfies the boundary condition $S(\mathbf{m}|m_N \geq 1; t) = 0$, which corresponds to the time at which a largest cluster is formed, and the initial condition $S(\mathbf{m}|m_N = 0; 0) = 1$ which means that — with certainty — there is no largest cluster in the system initially.

The mean first assembly time vector $\mathbf{T}(\mathbf{m})$ can be constructed using the same $\mathbf{A}^+$ as for the survival probability vector, and is given by [4]

$$\mathbf{T}(\mathbf{m}) = -[(\mathbf{A}^+)^{-1}\mathbf{1}]_{\mathbf{m}}. \quad (3.1)$$

Here $\mathbf{1}$ is a unit column vector of the same dimension as the number of states or configurations. The first component of the vector $\mathbf{T}(\mathbf{m})$ corresponds to the mean first assembly time from the first designated initial configuration, the second component of $\mathbf{T}(\mathbf{m})$ to the mean first assembly time from the second initial configuration, and so on. As such, this expression gives the mean first assembly time from all possible initial configurations of the system.

For additional details on the notation and method to derive theoretical MFAT expressions that were briefly reviewed in this section, the reader is referred to [5], where the method was first introduced.

4. The transition matrix $\mathbf{A}$ for MFAT

In this section, we provide details pertaining to the transition matrix $\mathbf{A}$ that will be inverted to compute the theoretical MFAT expressions as outlined in the previous section. The transition matrix will be derived in a more general setting, for which certain reaction rates can be set to zero to give BD or EBD expressions that may involve less reaction mechanisms than the full coagulation-fragmentation models. The relevant rates, and their reductions to BD and EBD were discussed in Section 2, and can likewise be set to zero in the transition matrix for MFAT calculations. We provide detailed structural properties of the transition matrix, which will provide the foundation for automatic generation for numerical matrix inversion for larger system scenarios. Explicit matrix entries are provided in several cases, and theoretical expression for MFATs through matrix inversion provided. Additional results for closed-form expressions are provided in Appendix C. For larger systems, the properties of the transition

matrix spelled out in this section, form the general numerical matrix inversion process for theoretical MFAT curves in larger systems for future work.

We begin this section by providing the details for small feasible systems so as to spell out the key differences between the exact and approximate reaction approaches. This will allow the ability to explain the fundamental steps implemented in the numerical generation of systems for larger system sizes as well.

For all cases, system states considered include the monomer-only configuration, that would be represented by $\mathbf{m} = (m_1, m_2, \dots, m_N) = (M, 0, \dots, 0)$, where m_i is the number of clusters of size i , N is the largest-cluster size, and M is the total number of monomers (size 1 clusters) in the system available to form larger clusters.

The general procedure constitutes of finding or generating all possible states (or configurations) of the system starting from the monomer-only state, using the prescribed reaction dynamics in the exact case (2.1)-(2.2), or the approximate case (2.5)-(2.9). Once all possible states have been identified for which the maximum cluster is not present, they are aligned in a certain order, and corresponding entries in the transition matrix can be determined. The matrix inversion process will then provide the MFAT from various initial configurations in the state order as specified when writing down the transition matrix. A few examples now follow. In the event that a state is reached that does not include a cluster of maximal size N , that do not lead to any other states in the system through a coagulation process, are termed trapped states. In cases where trapped states exist, the MFAT increases without bound as the break-up rates tend to zero, as being trapped in such a state would mean that a maximal-sized cluster will never be created, or rather, the first assembly time would be infinite.

4.1. $N = 3$, any M , exact dynamics

In the exact dynamics (2.1)-(2.2), reactants are removed from the system, and replaced by reactants in such a way that total mass is conserved. If the maximum cluster size is $N = 3$, then for each M , there are $(\lfloor M/2 \rfloor + 1)$ states that can be ordered as

$$\begin{aligned} \mathbf{m}_1 &= (M, 0, 0) \\ \mathbf{m}_2 &= (M - 2, 1, 0) \\ \mathbf{m}_3 &= (M - 4, 2, 0) \\ &\dots \\ \mathbf{m}_D &= \begin{cases} (0, \frac{M}{2}, 0), & \text{if } M \text{ is even} \\ (1, \frac{M+1}{2}, 0) & \text{if } M \text{ is odd} \end{cases} \end{aligned}$$

where $D = \lfloor \frac{M}{2} \rfloor + 1$. Recall that states that include the maximum cluster ($N = 3$ in this case) are not included in the transition matrix generation, and are thus excluded here.

The transition matrix $\mathbf{A}$ is given in general form [4] by

$$\begin{aligned} [A]_{ii} &= -C_2^{M-2i+2} K_{11} - (i-1)(M-2i+2)K_{12} \\ &\quad - (i-1)W_{11} \\ [A]_{i,i+1} &= iW_{11} \\ [A]_{i,i-1} &= C_2^{M-2i+4} K_{11} \end{aligned}$$

It is worth noting that if M is even, then the last state $(0, M/2, 0)$ is a trapped state if there is no break-up. There are no trapped states if M is odd.

Using (3.1), the Mean First Assembly times for $N = 3$ with various M values can be worked out explicitly. Results for the chemical rate law model, the BD/EBD model, as well as for the Smoluchowski model for $M = 6$ and 7 are presented here, with $M = 8, 9, 10, 50, 51, 100$, and 101 for the BD/EBD model also given in supplementary section 10. For results in this section, we use stochastic rates as spelled out in Table 2 and Appendix A, to distinguish between MFATs for the different models, namely chemical, BD/EBD and Smoluchowski. In addition to examples found in the literature, the algorithm for computing MFAT expressions is further demonstrated in Section 5 where we provide a numerical method for its automation for systems with any N and M .

For $K_{ij} = p$ and $W_{ij} = q$ exact BD/EBD-MFAT expressions are [4]:

$$\begin{aligned} T^{BD/EBD}(6, 0, 0) &= \frac{1}{60} \frac{15p^3 + 95p^2q + 21pq^2 + q^3}{qp^2(5p + q)} \\ T^{BD/EBD}(7, 0, 0) &= \frac{1}{105} \frac{744p^3 + 487qp^2 + 60q^2p + 2q^3}{(27p^2 + 20qp + 2q^2)p^2} \end{aligned}$$

For $K_{ij} = (1 + \delta_{ij})p$ and $W_{ij} = q$ exact chemical rate law (chem-MFAT) expressions are:

$$\begin{aligned} T^{chem}(6, 0, 0) &= \frac{120p^3 + 272p^2q + 36pq^2 + q^3}{120(8p + q)qp^2} \\ T^{chem}(7, 0, 0) &= \frac{1662p^3 + 674p^2q + 51pq^2 + q^3}{6300p^4 + 3360p^3q + 210p^2q^2} \end{aligned}$$

For $K_{ij} = p$ and $W_{ij} = (1 - \frac{1}{2}\delta_{ii})q$ exact Smoluchowski (Smol-MFAT) expressions are

$$\begin{aligned} T^{Smol}(6, 0, 0) &= \frac{120p^3 + 380p^2q + 42pq^2 + q^3}{120(10p + q)qp^2} \\ T^{Smol}(7, 0, 0) &= \frac{2976p^3 + 974p^2q + 60pq^2 + q^3}{11340p^4 + 4200p^3q + 210p^2q^2} \end{aligned}$$

Similarly for additive rates $K_{ij} = p(i + j)$ and $W_{ij} = q(i + j)$, exact BD/EBD-MFATs are

$$\begin{aligned} T^{BD/EBD}(6, 0, 0) &= \frac{1}{180} \frac{126qp^2 + 24q^2p + q^3 + 15p^3}{qp^2(6p + q)} \\ T^{BD/EBD}(7, 0, 0) &= \frac{1}{630} \frac{(1218p^3 + 660qp^2 + 69q^2p + 2q^3)}{(21p^2 + 12qp + q^2)p^2} \end{aligned}$$

Including relevant delta term factors as above for constant rates, gives the additive rate chemical rate law and Smoluchowski MFATs

$$\begin{aligned} T^{chem}(6, 0, 0) &= \frac{120p^3 + 324p^2q + 39pq^2 + q^3}{360(9p + q)qp^2} \\ T^{chem}(7, 0, 0) &= \frac{18015p^3 + 6546p^2q + 444pq^2 + 8q^3}{207900p^4 + 90720p^3q + 5040p^2q^2} \end{aligned}$$

and

$$T^{Smol}(6, 0, 0) = \frac{120p^3 + 504p^2q + 48pq^2 + q^3}{360p^2q(12p + q)}$$

$$T^{Smol}(7, 0, 0) = \frac{4872p^3 + 1320p^2q + 69pq^2 + q^3}{52920p^4 + 15120p^3q + 630p^2q^2}$$

For BD/EBD multiplicative rates $K_{ij} = pij$ and $W_{ij} = qij$, exact MFATs for the three different models are

$$T^{BD/EBD}(6, 0, 0) = \frac{1}{120} \frac{161qp^2 + 27q^2p + q^3 + 15p^3}{qp^2(7p + q)}$$

$$T^{BD/EBD}(7, 0, 0) = \frac{1}{210} \frac{930p^3 + 428qp^2 + 39q^2p + q^3}{(30p^2 + 14qp + q^2)p^2}$$

$$T^{chem}(6, 0, 0) = \frac{120p^3 + 380p^2q + 42pq^2 + q^3}{240(10p + q)qp^2}$$

$$T^{chem}(7, 0, 0) = \frac{2976p^3 + 974p^2q + 60pq^2 + q^3}{22680p^4 + 8400p^3q + 420p^2q^2}$$

$$T^{Smol}(6, 0, 0) = \frac{120p^3 + 644p^2q + 54pq^2 + q^3}{240(14p + q)qp^2}$$

$$T^{Smol}(7, 0, 0) = \frac{7440p^3 + 1712p^2q + 78pq^2 + q^3}{50400p^4 + 11760p^3q + 420p^2q^2}$$

Since for the $N = 3$ case, the full aggregation and Becker-Döring models are equivalent, general expressions for the entries of the transition matrix for any M can also be found in Yvinec et. al. [4].

4.2. $N = 3$, any M , approximate dynamics

In the approximate dynamics (2.5)-(2.9), reactants become products in such a way that the total number of particles, rather than total mass, remains constant in the system. As such, there are many more states in the case of approximate dynamics, and they can be ordered as

$$\begin{aligned} \mathbf{m}_1 &= (M, 0, 0) \\ \mathbf{m}_2 &= (M - 1, 1, 0) \\ \mathbf{m}_3 &= (M - 2, 2, 0) \\ &\dots \\ \mathbf{m}_M &= (1, M - 1, 0) \end{aligned}$$

In this case there are M states with no trapped states existing when there is no break-up.

For the approximate dynamics, the transition matrix $\mathbf{A}$ takes the form

$$[A]_{ii} = -C_2^{M-i+1}K_{11} - (i-1)(M-i+1)K_{12} - (i-1)W_{11}$$

$$[A]_{i,i+1} = iW_{11}$$

$$[A]_{i,i-1} = C_2^{M-i+2}K_{11}$$

Using (3.1), the Mean First Assembly times for $N = 3$ with $M = 6$ and 7 for the approximate dynamics and modifying constant, additive and multiplicative rates with appropriate delta factors are given for the three different models as follows.

For $K_{ij} = p$, $W_{ij} = q$ the three different MFATs for $M = 6$ and $M = 7$ are

$$T_{approx}^{chem}(6, 0, 0) = \frac{32290p^5 + 36625p^4q + 12010p^3q^2 + 1575p^2q^3 + 90pq^4 + 2q^5}{187500p^6 + 213750p^5q + 67650p^4q^2 + 7800p^3q^3 + 300p^2q^4}$$

$$T_{approx}^{BD/EBD}(6, 0, 0) = \frac{12150p^5 + 17517p^4q + 7800p^3q^2 + 1445p^2q^3 + 120pq^4 + 4q^5}{56700p^6 + 80100p^5q + 33675p^4q^2 + 5400p^3q^3 + 300p^2q^4}$$

$$T_{approx}^{Smol}(6, 0, 0) = \frac{97200p^5 + 70068p^4q + 15600p^3q^2 + 1445p^2q^3 + 60pq^4 + q^5}{150p^2(3024p^4 + 2136p^3q + 449p^2q^2 + 36pq^3 + q^4)}$$

and

$$T_{approx}^{chem}(7, 0, 0) = \frac{1256400p^6 + 1599696p^5q + 622998p^4q^2 + 103338p^3q^3 + 8202p^2q^4 + 315pq^5 + 5q^6}{9797760p^7 + 12492144p^6q + 4773384p^5q^2 + 743904p^4q^3 + 50400p^3q^4 + 1260p^2q^5}$$

$$T_{approx}^{BD/EBD}(7, 0, 0) = \frac{1408050p^6 + 2233215p^5q + 1155994p^4q^2 + 265458p^3q^3 + 30008p^2q^4 + 1680pq^5 + 40q^6}{8731800p^7 + 13642650p^6q + 6815592p^5q^2 + 1450008p^4q^3 + 138600p^3q^4 + 5040p^2q^5}$$

$$T_{approx}^{Smol}(7, 0, 0) = \frac{11264400p^6 + 8932860p^5q + 2311988p^4q^2 + 265458p^3q^3 + 15004p^2q^4 + 420pq^5 + 5q^6}{69854400p^7 + 54570600p^6q + 13631184p^5q^2 + 1450008p^4q^3 + 69300p^3q^4 + 1260p^2q^5}.$$

For $K_{ij} = p(i + j)$ and $W_{ij} = q(i + j)$, they are

$$T_{approx}^{chem}(6, 0, 0) = \frac{696060p^5 + 612207p^4q + 161910p^3q^2 + 17520p^2q^3 + 840pq^4 + 16q^5}{10810800p^6 + 9393300p^5q + 2365200p^4q^2 + 223200p^3q^3 + 7200p^2q^4}$$

$$T_{approx}^{BD/EBD}(6, 0, 0) = \frac{344835p^5 + 367767p^4q + 125340p^3q^2 + 18180p^2q^3 + 1200pq^4 + 32q^5}{4054050p^6 + 4199850p^5q + 1336500p^4q^2 + 165600p^3q^3 + 7200p^2q^4}$$

$$T_{approx}^{Smol}(6, 0, 0) = \frac{689670p^5 + 367767p^4q + 62670p^3q^2 + 4545p^2q^3 + 150pq^4 + 2q^5}{8108100p^6 + 4199850p^5q + 668250p^4q^2 + 41400p^3q^3 + 900p^2q^4}$$

and

$$T_{approx}^{chem}(7, 0, 0) = \frac{33881400p^6 + 33017130p^5q + 10236213p^4q^2 + 1384377p^3q^3 + 91074p^2q^4 + 2940pq^5 + 40q^6}{701719200p^7 + 677066040p^6q + 203782068p^5q^2 + 25684344p^4q^3 + 1436400p^3q^4 + 30240p^2q^5}$$

$$T_{approx}^{BD/EBD}(7, 0, 0) = \frac{3258990p^6 + 3794535p^5q + 1488132p^4q^2 + 264705p^3q^3 + 23556p^2q^4 + 1050pq^5 + 20q^6}{50803200p^7 + 57951936p^6q + 21772800p^5q^2 + 3555468p^4q^3 + 264600p^3q^4 + 7560p^2q^5}$$

$$T_{approx}^{Smol}(7, 0, 0)$$

$$= \frac{52143840p^6 + 30356280p^5q + 5952528p^4q^2 + 529410p^3q^3 + 23556p^2q^4 + 525pq^5 + 5q^6}{812851200p^7 + 463615488p^6q + 87091200p^5q^2 + 7110936p^4q^3 + 264600p^3q^4 + 3780p^2q^5}$$

For $K_{ij} = pij$ and $W_{ij} = qij$, they are

$$T_{approx}^{chem}(6, 0, 0) = \frac{97200p^5 + 70068p^4q + 15600p^3q^2 + 1445p^2q^3 + 60pq^4 + q^5}{300(3024p^4 + 2136p^3q + 449p^2q^2 + 36pq^3 + q^4)p^2}$$

$$T_{approx}^{BD/EBD}(6, 0, 0) = \frac{57785p^5 + 49144p^4q + 13640p^3q^2 + 1635p^2q^3 + 90pq^4 + 2q^5}{392700p^6 + 322650p^5q + 82950p^4q^2 + 8400p^3q^3 + 300p^2q^4}$$

$$T_{approx}^{Smol}(6, 0, 0) = \frac{924560p^5 + 393152p^4q + 54560p^3q^2 + 3270p^2q^3 + 90pq^4 + q^5}{300(20944p^4 + 8604p^3q + 1106p^2q^2 + 56pq^3 + q^4)p^2}$$

and

$$T_{approx}^{chem}(7, 0, 0)$$

$$= \frac{11264400p^6 + 8932860p^5q + 2311988p^4q^2 + 265458p^3q^3 + 15004p^2q^4 + 420pq^5 + 5q^6}{139708800p^7 + 109141200p^6q + 27262368p^5q^2 + 2900016p^4q^3 + 138600p^3q^4 + 2520p^2q^5}$$

$$T_{approx}^{BD/EBD}(7, 0, 0)$$

$$= \frac{10698975p^6 + 9892755p^5q + 3140514p^4q^2 + 458622p^3q^3 + 33858p^2q^4 + 1260pq^5 + 20q^6}{96446700p^7 + 87057180p^6q + 26336016p^5q^2 + 3504816p^4q^3 + 214200p^3q^4 + 5040p^2q^5}$$

$$T_{approx}^{Smol}(7, 0, 0)$$

$$= \frac{171183600p^6 + 79142040p^5q + 12562056p^4q^2 + 917244p^3q^3 + 33858p^2q^4 + 630pq^5 + 5q^6}{1543147200p^7 + 696457440p^6q + 105344064p^5q^2 + 7009632p^4q^3 + 214200p^3q^4 + 2520p^2q^5}$$

Approximate reaction mechanisms of this type have often been used for RMPC simulations for different reaction mechanisms [12–19], however theoretical MFAT expressions corresponding to these approximate dynamics with corresponding ODE rate laws, have only been considered in [29] and here.

4.3. MFAT exact theoretical model comparison

In Figure 3, we provide plots of the theoretical expressions from the previous section. For convenience, we take $p = 1$ for all plots in the figure. Figure 3 compares the MFATs for an initial monomer-only state with $M = 6$ (top), namely the $T(6, 0, 0)$ expressions. The vertical asymptote as $q \rightarrow 0$ corresponding to a trapped state is visible in all curves. In all cases (constant, additive and multiplicative rates), the BD/EBD model leads to comparable MFAT values as the other two models (chemical and Smoluchowski), while it leads to much larger MFATs as q increases, compared to the chemical and Smoluchowski models, which are very close to one another. For the $M = 7$ case in the bottom of Figure 3, in the absence of trapped states, all three models (chemical, BD/EBD and Smoluchowski) have comparable $q \rightarrow 0$ MFAT limits in each of the different cases (constant, additive and multiplicative rates), and the BD/EBD model leads to significantly higher MFATs as q increases, with the chemical and Smoluchowski models being in very close agreement for all q values plotted. Worth pointing out is that the BD/EBD multiplicative rate MFAT curve is very similar to the chemical and Smoluchowski constant rate MFAT curve, and the same is true for the BD/EBD additive rate curve and the chemical/Smoluchowski multiplicative rate curves in both plots in the figure.

Comparison of MFATs for the approximate model is done in the Results Section (Figure 6 onwards).

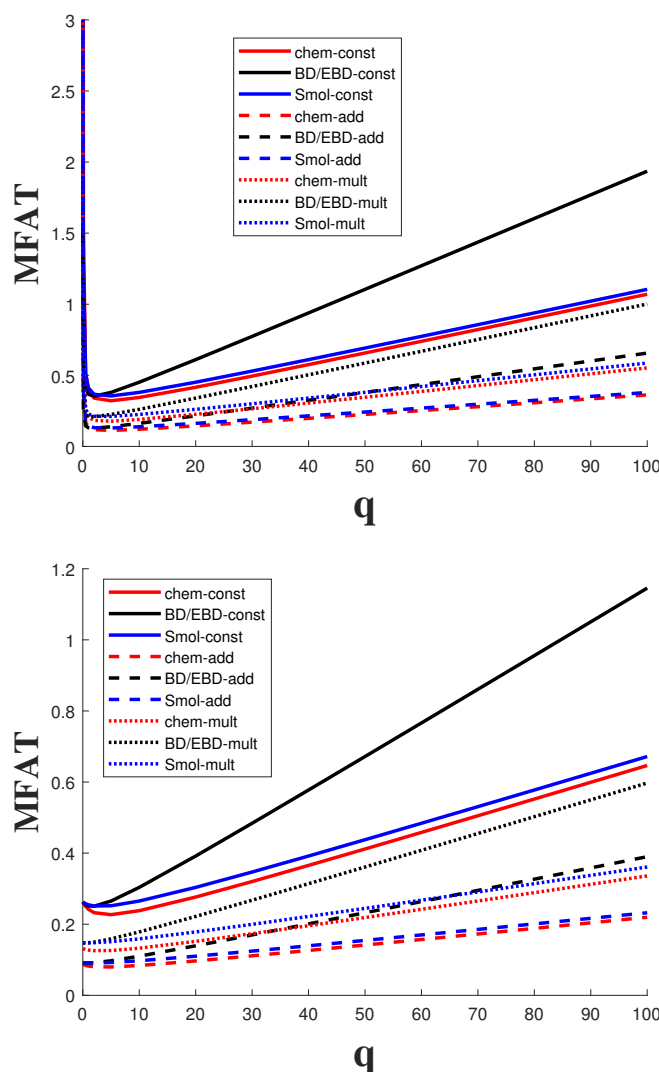


Figure 3. Comparison of the theoretical expressions for $T(6,0,0)$ (top) and $T(7,0,0)$ (bottom) given in Section 4.1 for the exact dynamics ($p = 1$).

5. Numerical methods

5.1. Summary of properties of the transition matrix

In order to create code that automatically generates the states in the given system, as well as the transition matrix, a few key properties can be observed.

First, the $-A_{ii}$ entries along the main diagonal are the sum of all rates from all possible reactions based on state i . For example, in $M = 6$ exact case (see also Appendix A), state 2 is $\mathbf{m}_2 = (4, 1, 0)$. Monomers can aggregate together at a rate $4(3)/2K_{11}$, a monomer can react with a dimer at a rate $4(1)K_{12}$, there aren't enough dimers to aggregate with each other, and dimers can break up at a rate $1w_{11}$. Since we are interested in MFAT, and $N = 3$, trimer break-up is not included in the transition matrix since the trimer is a cluster of maximum size in that case.

Second, the A_{ij} entry (i -th row, j -th column) contains the (positive) rate based on state $\mathbf{m}_j$ that leads

to formation of state $\mathbf{m}_i$. As an explicit example, for the same matrix as per the first property, the $\mathbf{A}_{12}$ entry is the rate based on $\mathbf{m}_2 = (4, 1, 0)$ that leads to state $\mathbf{m}_1 = (6, 0, 0)$. Since that change can only have resulted from a dimer breaking up into two monomers, the matrix entry will be $1w_{11}$.

Third, adding all entries in a column leaves only rates that lead to the formation of the largest cluster, that are negative, and located in the main diagonal. For example, adding all entries in the second column of the same example as in the previous two properties leaves $-4K_{12}$ as the sum. This notably gives rise to a weakly diagonally dominant matrix for $-\mathbf{A}'$ that is to be inverted for MFAT results, that is, $-\mathbf{A}'$ is diagonally dominant, with a strict inequality in at least one row. That row would correspond to the state that leads to the formation of the largest cluster. The MFAT divergence arises from a singularity of the transition matrix for which the main diagonal entry becomes zero in the zero limit of the corresponding break-up rate for that state.

Finally, reordering states leads to a transition matrix with reshuffled entries, but the above three properties remain in place, and thus computations of the matrix inverse for MFAT calculations will be the same for a given starting state, as that from a system whose states were ordered differently.

With that in mind, one can easily write a program to generate the states in the system, and compute the entries in the transition matrix based on the ordering of states created by the program, allowing for theoretical MFAT values for much larger systems than possible to consider by hand, or with limited application through use of approximation techniques used in various papers [4, 5, 10].

These properties are irrespective of the reactions that are part of a given system, they are irrespective of the ordering of the states, and are thus properties of the transition matrix for the exact dynamics, as well as for the approximate dynamics, for any N and M .

5.2. MFAT matrix inversion algorithm

Given that the transition matrix has certain properties as outlined in the previous section, it is easy to construct an algorithm to generate the matrix automatically for any given N and M , and to then generate all entries in the corresponding transition matrix. Subsequent matrix inversion, and summation of the rows, according to (3.1) will then lead to the theoretical value for the MFAT for every initial starting state for that system. The algorithm is outlined in Table 1.

In some cases it is possible and feasible to obtain closed-form expressions for MFAT through manual matrix inversion. We have included the theoretical expressions for $N = 3$ with $M = 6$ and 7 in the previous section for both the exact aggregation mechanism, as well as for the approximate dynamics. Additional theoretical results for $M = 5, 8, 9, 10, 50, 51, 100$ and 101 for the exact reaction mechanism are given in Appendix C.

For the exact mechanism, theoretical expressions are also possible for $N = 4$ and small values of M . However, as M increases even slightly, the dimension of the transition matrix grows quickly, making theoretical expressions difficult to compile and subject to approximation methods used in various papers [4, 5, 10]. Numerical generation of the states, as well as matrix inversion, is possible and necessary for exact results when N or M are large. This is especially true for the approximate dynamics, where the case $N = 4, M = 5$ already contains 20 states.

Table 1. MFAT matrix inversion algorithm.

1. Generate all possible states without a cluster of maximal size, based on the possibilities for a given initial number of monomers M .
2. For each state i , apply each reaction R_k to the state to find the newly create state j' based on the reaction. Then, add to entry $A_{j',i}$ the rate for the reaction R_k based on state i , and subtract this same amount from $A_{i,i}$. In the event that the newly created state has a cluster of maximal size, there is only a subtraction of the rate from the $A_{i,i}$ entry.
3. Compute the matrix inverse of the transpose of the matrix, and multiply by negative 1.
4. Add all entries in a given row i of this inverse matrix to obtain the theoretical MFAT based on starting with state i denoted by $\mathbf{m}_i$.

5.3. Stochastic simulations with SSA and RMPC

We use two stochastic simulation algorithms to generate MFAT values. The first is based on the traditional stochastic simulation algorithm by Gillespie [11], namely the Stochastic Simulation Algorithm (SSA). The second is based on the Reactive Multiparticle Collision (RMPC) dynamics. The details for the the RMPC dynamics used here can be found in Strehl and Rohlf [16]. A key distinction between the two methods is that, in SSA, only total counts of different-sized clusters are kept track of in the entire system volume, while in RMPC, there are individual particles that have individual velocities and positions within the simulation domain, that also propagate through the volume in a diffusive manner as a result of the particle collision and free-streaming steps in the method. In the exact RMPC method, the two particles that join together are removed from the system, and a new particle of their combined size is introduced in the local collision volume in which the aggregation took place. This new particle is placed randomly in the local collision volume, and is assigned a random velocity drawn from the Maxwellian velocity distribution. When a larger cluster breaks up, it is removed from the system, and two new particles of the smaller sizes are placed in the collision volume in which the break-up occurred. In the approximate RMPC method, only the species type of a single particle changes when aggregation or break-up takes place. The positions and velocities remain unchanged in the reaction step in the approximate method, and no particles are added or removed from the system. Finally, in the RMPC results in this paper, unlike in [29], SSA-like propensities are used to decide on a reaction in a local RMPC collision volume. In the case of an approximate aggregation event for $C_i + C_j \rightarrow C_{i+j}$, one of the reactants is chosen with equal probability, to become the larger cluster, explaining why there are additional factors of $1/2$ in the relevant terms in the approximate chemical rate law in Section 2.5. For a break-up event, there is likewise an equal probability for each of the products that the reactant becomes, leading to extra factors of $1/2$ in the break-up terms in the chemical rate law for the approximate dynamics.

In the context of stochastic propensities for the different rate laws (see Table 2), we note that the aggregation event of like clusters for the chemical rate law has twice the propensity than that for the BD/EBD/Smoluchowski models, explaining the difference in factors in the chemical rate laws explained in Section 2. Additionally, the break-up stochastic propensities for the Smoluchowski equation are half that of the BD/EBD/chemical rate law propensities explaining the break-up differences in the ODE/rate laws. This distinction has been pointed out recently in [29, 30].

Table 2. Propensity functions with stochastic rates m_i for aggregation of, or break-up resulting in, two clusters of the same size for the different models.

Reaction	Propensity	Chemical (exact) (K_{ii} or W_{ii})	BD/EBD (K_{ii} or W_{ii})	Smoluchowski (K_{ii} or W_{ii})
$C_i + C_i \rightarrow C_{2i}$	$\frac{N_i(N_i-1)}{2} K_{ii}$	$2p_{ii}/V$	p_{ii}/V	p_{ii}/V
$C_{2i} \rightarrow C_i + C_i$	$N_{2i} W_{ii}$	q_{ii}	q_{ii}	$\frac{1}{2}q_{ii}$

The algorithms in the context of obtaining MFAT results, are outlined in Tables 3 and 4 respectively.

Table 3. Algorithm for generating MFAT using SSA.

1. Initialize the system based on the state from which MFAT is to be found.
2. Perform SSA on the current state of the system, based on all possible reactions in the reaction set.
3. If a maximal cluster is created, record the time, and go back to the starting state, setting time back to zero.
4. Repeat steps 2-3 until a desired number of samples have been generated for FAT (typically 10^3 - 10^6 realizations).
5. Compute MFAT based on averaging the times recorded in the collection of realizations from the previous step 4.

Table 4. Algorithm for generating MFAT using RMPC.

1. Initialize the system based on the state from which MFAT is to be found.
2. Distribute the particles of each species, according to the starting state, uniformly throughout the simulation volume.
3. Update particle species, positions and velocities in each cell based on the reaction mechanisms.
4. If particles free-stream, invoke collisions and free-streaming steps.
5. If a maximal cluster is created, record the time, and go back to the starting state, setting time back to zero.
6. Repeat steps 2-5 until a desired number of samples have been generated for FAT (typically 10^3 - 10^6 realizations).
7. Compute MFAT based on averaging the times recorded in the collection of realizations from the previous step 6.

For RMPC simulations, the dynamics is prescribed at a particle-level and if a random number drawn from a uniform distribution on $[0, 1]$ is less than a comparable SSA-propensity in the reaction volume, the corresponding reaction will take place. If the SSA-time for the next reaction to occur is greater than the RMPC time step of 1, then no reaction happens in the RMPC reaction volume. Otherwise, particles of the corresponding reactant species type are removed, and new particles corresponding to

the products are added. For MFAT results, in which the MFAT time could be smaller than the RMPC time step of 1, a time rescaling method introduced in [29] was used so as to obtain accurate MFAT results (see also [29]). Time rescaling can also be used to obtain access to smaller MFATs than the RMPC time step of 1, as the RMPC time step is fixed at $\Delta t = 1$. Table 5 provides examples of how to choose RMPC values for MFAT simulations in this case. Finally, time rescaling can be useful in lowering the physical RMPC simulation time when MFATs are expected to be very large so long as the rescaled RMPC reaction propensities remain smaller than 1 (compare first and second rows in the Table).

Table 5. Correspondence between desired p and q rate values to probabilistic RMPC values through rescaling time $t^{\text{RMPC}} = \tau t^{\text{ODE}}$.

p	q	MFAT	τ	p^{RMPC} $= \frac{p}{\tau}$	q^{RMPC} $= \frac{q}{\tau}$	MFAT ^{RMPC} $= (\text{MFAT})\tau$
1	100	1	10000	0.0001	0.01	10000
1	0.1	0.1	10000	0.0001	0.00001	1000
1	0.01	1	10000	0.0001	0.000001	10000

6. Results

RMPC concentration curves are compared to the chemical rate law and BD/EBD model, as well the chemical rate law and BD/EBD model for the approximate dynamics in Figures 4. A volume of $1 \times 1 \times 1$ was used, with $k_B T/m = 1$, $\Delta x = 1 = \Delta y = \Delta z$, $\Delta t = 1$, and $K_{ij} = 0.0001$, $W_{ij} = 0.0001$. There are 20 monomers initially, and the results are averaged over 1000 realizations. Both an exact version of RMPC dynamics is considered for which the total number of monomers is conserved in the system (Section 2.1), as well as for the approximate dynamics (Section 2.2). Using constant forward and backward rates, with additional factors as necessary (see Table 2), RMPC results for the exact dynamics follow the corresponding monomer (not shown), dimer (shown) and trimer (not shown) concentration curves fairly closely, capturing both the time and value of the C_2 local maxima, as well as the corresponding equilibria (Figure 4 (top)). In the approximate dynamics, there is a slight discrepancy in the value of the C_2 maximum, and a slight discrepancy in equilibrium values in the C_2 and C_3 curves compared to the corresponding ODE models (Figure 4 (middle)). In the bottom of Figure 4, RMPC results for the chemical rate law model are compared to the corresponding ODE curves, for increasing break-up rates. As the break-up rates increase, RMPC results lead to better agreement in both the exact and approximate models.

Numerical matrix inversion MFAT results were obtained and compared to the theoretical curves where possible (not shown). All matrix inversion results matched with the corresponding theoretical expressions.

Figures 5-8 provide various theoretical MFAT expressions from Sections 4.1 and 4.2, as well as corresponding SSA and/or RMPC results, using the Table 2 rate variations to produce the three different models (Smoluchowski, BD/EBD and chemical). In all figures, $p = 1$, and thus rate rescaling according to Table 5 was necessary for RMPC simulations. Additionally, since for RMPC simulations, we use a time step $\Delta t = 1$, and MFAT values are less than that in many of these Figures, appropriately rescaling

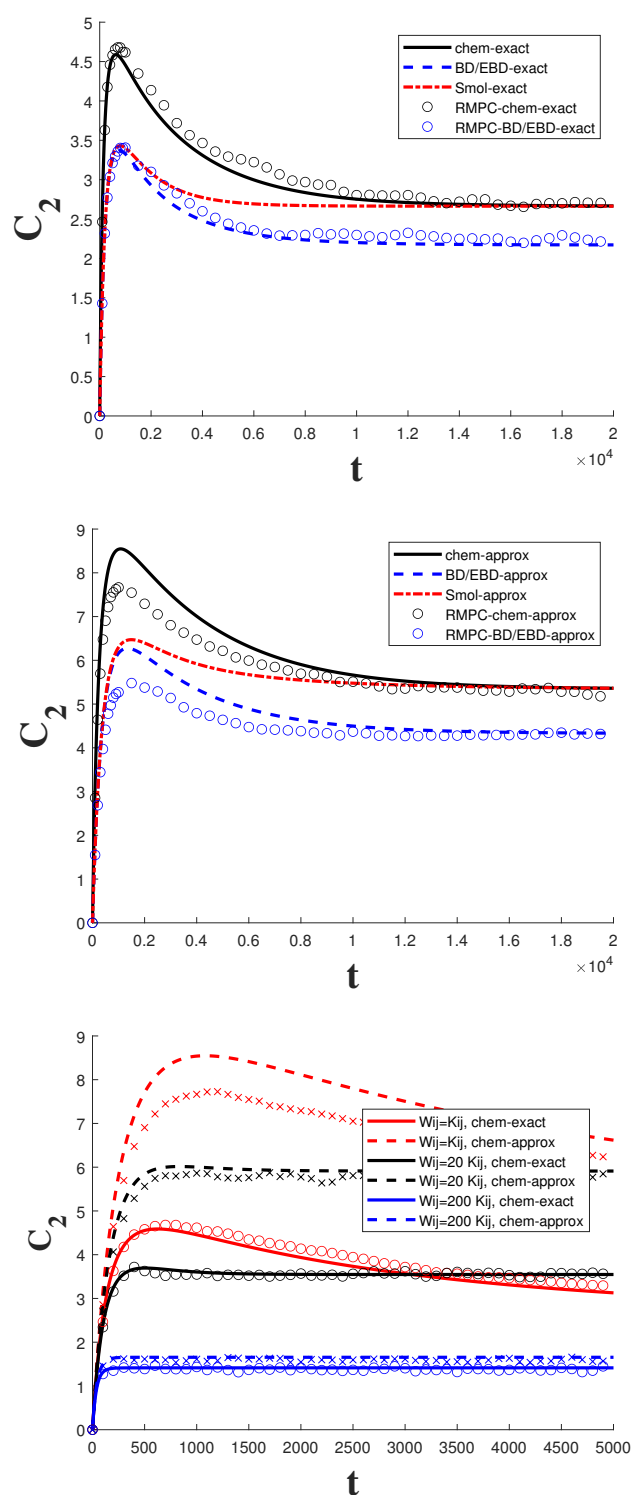


Figure 4. Comparison between the Chemical Rate Law, the BD/EBD and Smoluchowski dynamics, and corresponding exact (top) and approximate (middle) RMPC results for BD/EBD and the chemical rate law ($K_{ij} = 0.0001$, $W_{ij} = 0.0001$) and the effect of increasing the break-up rates in the exact and approximate chemical rate law model (bottom).

the time according to Table 5 discussions was also necessary, and hereby shown to be a feasible means to obtain MFATs that are smaller than the RMPC time step of 1.

In Figure 5, SSA and RMPC data are compared to theoretical MFAT expressions for constant formation and break-up rates. The SSA data (plus symbols), and the RMPC data (circles) agree well with the corresponding theoretical curves in all three models (chemical, BD/EBD and Smoluchowski), both for $M = 6$ (top) and $M = 7$ (bottom). The chemical rate law and Smoluchowski models are in close agreement with one another, while the BD MFATs are much larger in comparison. Note that we provide RMPC data only for the BD/EBD and chemical models.

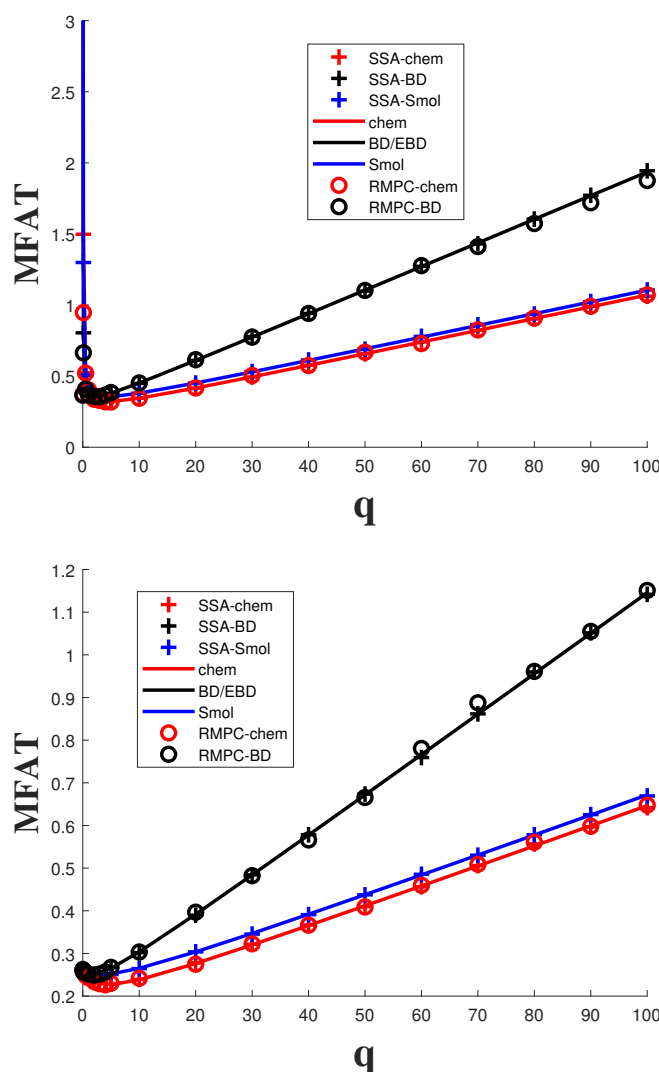


Figure 5. Comparison of SSA and RMPC data to the theoretical expressions for $T(6,0,0)$ (top) and $T(7,0,0)$ (bottom) given in Section 4.1 for the exact dynamics ($p = 1$).

In Figures 6 and 7, we plot theoretical MFAT curves for the exact and approximate theoretical models for the chemical and BD/EBD models, with RMPC data, for constant (top), additive (middle) and multiplicative rates (bottom). The theoretical curves and RMPC data are in good agreement in all cases, both for the exact (circles) and approximate (x 's) model simulations. In all cases, the

approximate model leads to smaller MFATs than the corresponding exact model, the chemical and Smoluchowski models lead to comparable values, and the approximation leads to closer MFAT values in the chemical and Smoluchowski models, compared to the BD/EBD case, with the chemical model always yielding smallest MFAT values. It can also be seen that the BD/EBD approximate model falls between the BD/EBD exact model, and the chem/Smoluchowski exact model curves for q larger than 30. This is true for all cases ($M = 6$ in Figure 6, and $M = 7$ in Figure 7; constant, additive and multiplicative rates). Worth pointing out is that the approximate model does not have a vertical asymptote as $q \rightarrow 0$ in the $M = 6$ case (Figure 6), while the exact model does.

In Figure 8 (top and middle), SSA and RMPC results are compared to theoretical expressions, and the matrix inversion results. While there is little noticeable differences in these graphs between the chemical, BD/EBD and Smoluchowski MFAT values as M increases, RMPC and SSA results lead to excellent agreement with the theoretical expressions in both the exact and approximate models. Figure 8 (bottom) shows the difference in MFAT values as M increases, when comparing the exact and approximate models in the three different cases (chemical, BD/EBD and Smoluchowski). While the actual differences are very small, Figure 9 shows the percent relative error between RMPC and SSA results compared to the theoretical MFAT matrix inversion results. For $M = 4, 5, 6, 7, 8, 9, 10, 15, 20, 25$, time rescaling of $\tau = 10000$ was used for RMPC MFAT results, while for $M = 30, 35, 40, 45, 50, 51, 100, 101$, τ had to be increased to $\tau = 100000$ to decrease the percent relative error from approximately 50 – 60% when $M = 100, 101$ in the chemical model (Fig. 9 (top)), and from approximately 25 – 30% in the BD model (Fig. 9 (bottom)). SSA results were found to provide more accurate MFAT values, while RMPC simulations lead to larger errors as M increases. Increasing τ was found to increase accuracy in the RMPC results, however this comes at a cost of longer simulation time (days compared to hours). Nonetheless, overall, RMPC results agree very well with theoretical predictions.

Overall, it is worth noting in these last Figures, that exact and approximate MFAT matrix inversion results are fairly close to one another, being closer as MFAT values decrease. This then also demonstrates that an approximate self-assembly process as proposed in Section 2.2 can be used as a computationally advantageous RMPC implementation that leads to similar physics as an exact self-assembly process, at least for the rate coefficients considered here. Care should be taken however with this approach in the $q \rightarrow 0$ limit, when the exact system has a trapped state while the approximate dynamics does not (see for example $T(6, 0, 0)$ that has a vertical asymptote as $q \rightarrow 0$, compared to $T^{approx}(6, 0, 0)$ which does not).

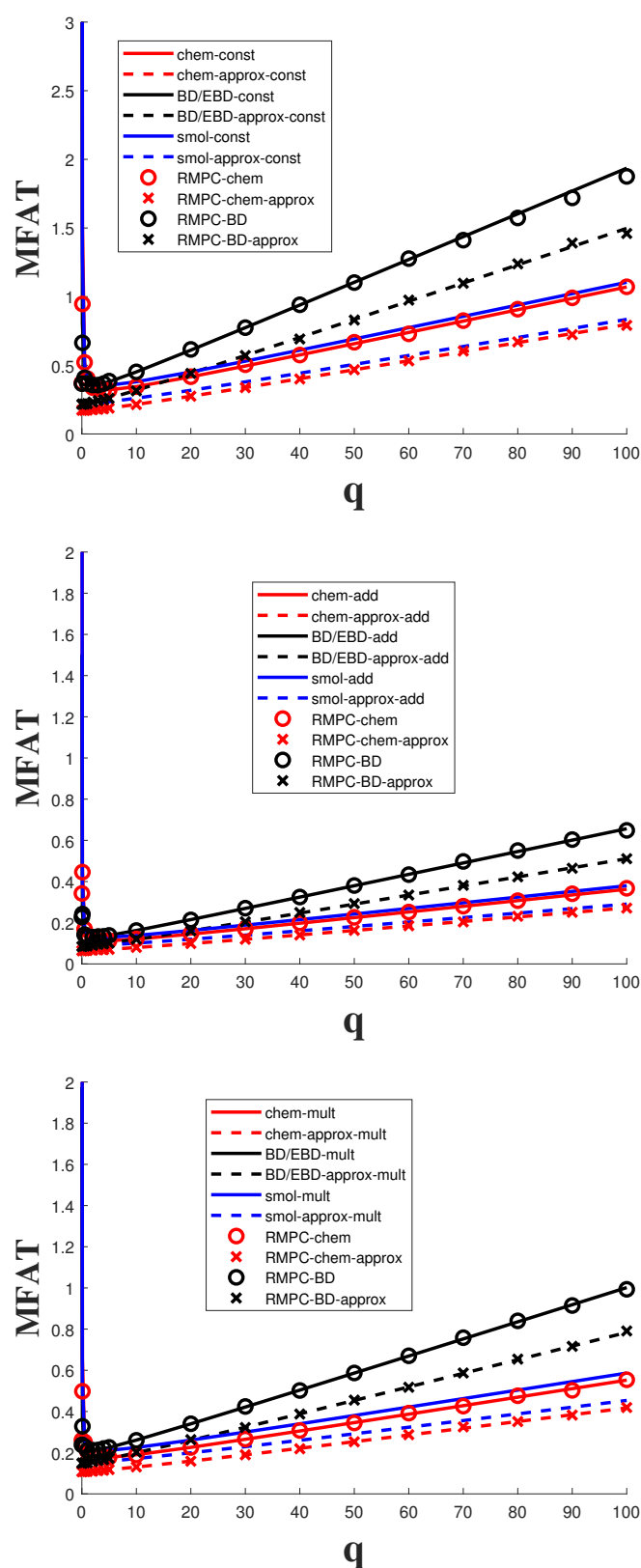


Figure 6. Comparison of RMPC data to the theoretical expressions for $T(6, 0, 0)$ with constant (top), additive (middle) and multiplicative (bottom) rates given in Section 4.1 for the exact, and Section 4.2 for the approximate dynamics ($p = 1$).

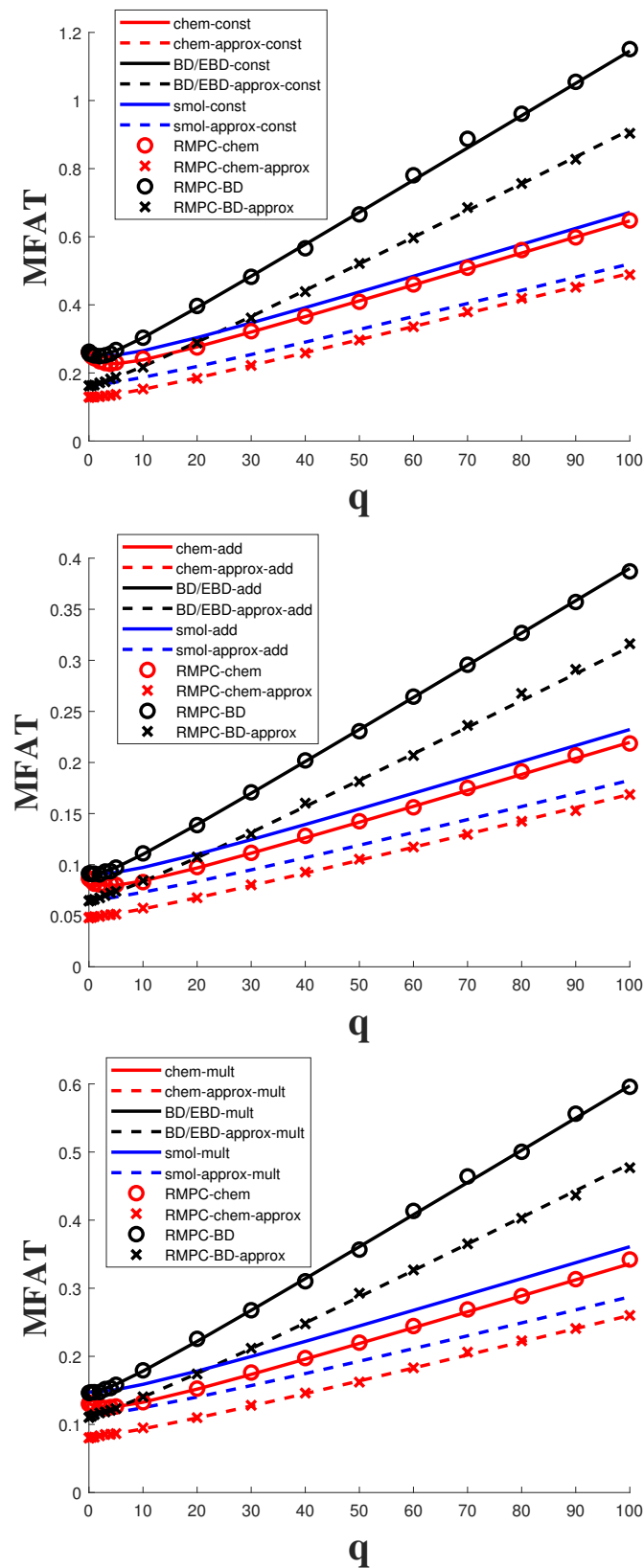


Figure 7. Comparison of RMPC data to the theoretical expressions for $T(7,0,0)$ with constant (top), additive (middle), and multiplicative (bottom) rates given in Section 4.1 for the exact, and Section 4.2 for the approximate dynamics.

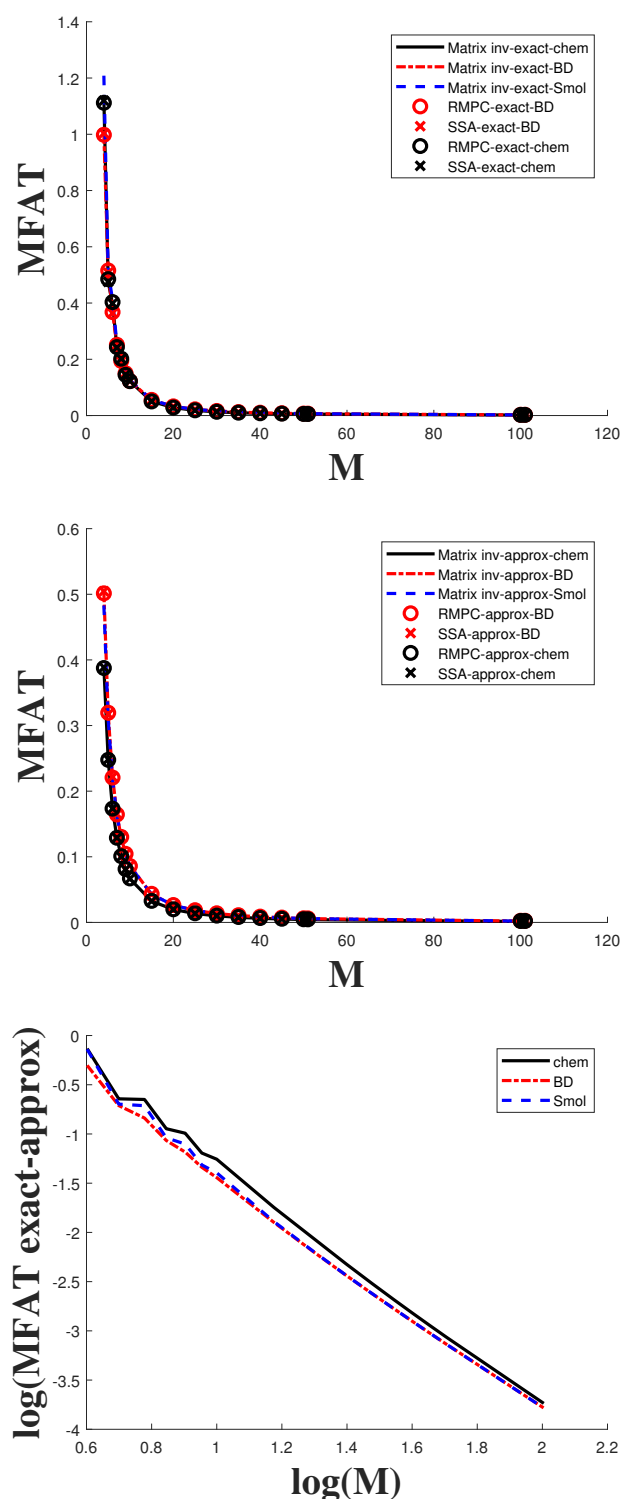


Figure 8. Comparison of MFATs for $N = 3$ and various M from SSA, exact (top) and approximate (middle) RMPC simulations, to theoretical expressions from Section 4 and the MFAT matrix inversion algorithm ($p = 1, q = 1$) with a theoretical difference in MFAT values between the exact and approximate models using matrix inversion data shown in the bottom graph.

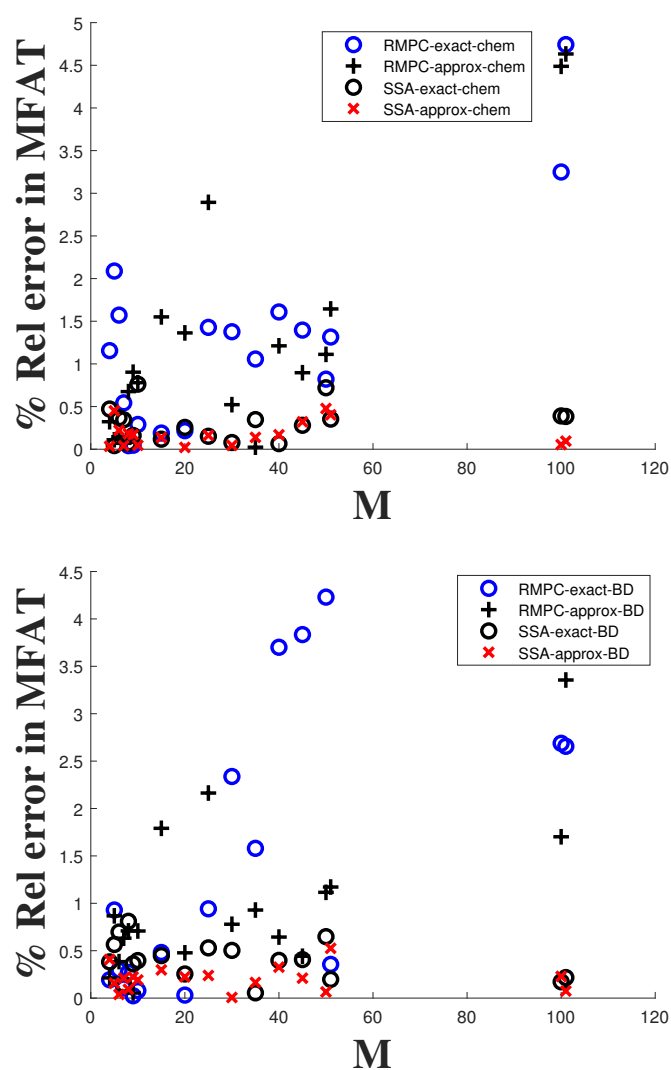


Figure 9. Comparison of the percent relative error in MFAT values between SSA and RMPC compared to the matrix inversion data, for the chemical model (top) and the BD model (bottom). For RMPC results, $\tau = 10000$ is used for time rescaling.

7. Summary and conclusions

In this paper, we have considered the self-assembly process for a finite system, as a set of chemical reactions that were then simulated using Gillespie's Stochastic Simulation algorithm (SSA), as well as with Reactive Multi-particle Collision Dynamics (RMPC). An approximate chemical system was proposed, whose set of chemical reactions more accurately describes a more commonly used approximate RMPC reaction implementation. For a finite system in which the largest cluster size is $N = 3$, it was shown that the approximate RMPC dynamics agrees well with the newly proposed chemical rate law, while the exact RMPC dynamics agrees well with the exact chemical rate law for the largest sized cluster relevant for the corresponding MFAT calculations. By appropriately halving some forward and/or backward rates, we additionally distinguished between a chemical reaction framework, the Smoluchowski model, and Becker-Döring dynamics in both the exact and approximate models, as well as in ODE/rate law and MFAT results.

Using the relevant exact and the proposed approximate reaction mechanisms, theoretical MFAT expressions were obtained for $N = 3$, using a well-established Master equation matrix inversion process. Stochastic MFAT results were compared to the theoretical curves. A time rescaling mechanism was introduced so as to be able to obtain RMPC-MFAT results when the MFAT times are comparable to, or less than, the RMPC time step of 1. Constant, additive and multiplicative rates were considered for MFAT results with monomer-only initial conditions, for a range of initial numbers of monomers, as well as non-constant rates.

For the range of break-up rates considered, constant rates led to largest MFAT values, additive rates led to lower MFATs, and multiplicative rates gave rise to the lowest MFAT values. MFATs for the approximate RMPC reaction mechanism were consistently below MFATs from the corresponding exact reaction mechanism. SSA results agreed well with the theoretical MFATs for exact reaction dynamics, and exact RMPC-MFAT results consistently ended up between the two theoretical MFAT curves. As the initial number of monomers increases, the difference between RMPC, SSA, exact and approximate theoretical MFAT curves decreases. Chemical rate law-based ODEs/MFATs were similar to Smoluchowski-based ODEs/MFATs in both exact and approximate models. Additionally, approximate ODEs become closer to the exact model as the break-up rates increase.

The results in this paper have also shown that RMPC can be used to obtain Mean First Assembly times for self-assembly processes. Since this stochastic simulation algorithm can intrinsically incorporate velocities of particles, velocity-dependent reaction rates can be easily added to the reaction dynamics, as well as diffusive effects. Thus, our results provide the stepping stone for a simulation paradigm that is capable of furthering our understanding of how self-assembly processes may be affected by diffusion or when subjected to flow conditions.

Author contributions

Pradeep Kunwar: Had a lead role in developing the theoretical models and expressions, and created the majority of the numerical results. Had a lead role in the conceptualization, investigation, methodology, formal analysis, validation, software and visualization of the results. Had a supporting role on the writing of the original draft of the manuscript, as well as its subsequent revisions.

Katrin Rohlf: Provided supervision and support in the conceptualization, investigation,

methodology, formal analysis, validation, software and visualization of the results. Had a lead role in writing and editing the original draft of the manuscript, as well as its revisions. Had a small role in software, validation and visualization of the results. Created a small portion of the numerical results.

Use of Generative-AI tools declaration

The authors declare they have not used Artificial Intelligence (AI) tools in the creation of this article.

Acknowledgments

This research has been supported by the Natural Sciences and Engineering Research Council of Canada (NSERC) Discovery Grant RGPIN-2017-04672 and 2007 NSERC Research Tools and Infrastructure (RTI) grant, as well as by the Toronto Metropolitan University (TMU) Faculty of Science Dean's Research Fund and TMU Mathematics Department.

Conflict of interest

The authors do not have any conflicts of interest.

References

1. N. Hoze, D. Holcman, Kinetics of aggregation with a finite number of particles and application to viral capsid assembly, *J. Math. Biol.*, **70** (2015), 1685–1705. <https://doi.org/10.1007/s00285-014-0819-2>
2. S. Chandrasekhar, Stochastic problems in physics and astrophysics, *Rev. Mod. Phys.*, **15** (1943), 1–89. <https://doi.org/10.1103/RevModPhys.15.1>
3. N. Hoze, D. Holcman, Coagulation-fragmentation for a finite number of particles and application to telomere clustering in the yeast nucleus, *Phys. Lett. A*, **376** (2012), 845–849. <https://doi.org/10.1016/j.physleta.2012.01.014>
4. R. Yvinec, M. R. D'Orsogna, T. Chou, First passage times in homogeneous nucleation and self-assembly, *J. Chem. Phys.*, **137** (2012), 244107. <https://doi.org/10.1063/1.4772598>
5. M. R. D'Orsogna, G. Lakatos, T. Chou, Stochastic self-assembly of incommensurate clusters, *J. Chem. Phys.*, **136** (2012), 084110. <https://doi.org/10.1063/1.3688231>
6. R. Becker, W. Döring, Kinetische Behandlung der Keimbildung in übersättigten Dämpfen, *Ann. Phys. (Leipzig)*, **24** (1935), 719–752. <https://doi.org/10.1002/andp.19354160806>
7. M. von Smoluchowski, Drei Vorträge über Diffusion, Brownsche Bewegung und Koagulation von Kolloidteilchen, *Z. Phys.*, **17** (1916), 557–585.
8. P. L. Krapivski, S. Redner, E. Ben-Naim, *A Kinetic view of statistical physics*, Cambridge: Cambridge University Press, 2010. <https://doi.org/10.1017/CBO9780511780516>
9. D. J. Aldous, Deterministic and stochastic models for coalescence (aggregation and coagulation): A review of the mean-field theory for probabilists, *Bernoulli*, **5** (1999), 3–48. <https://doi.org/10.2307/3318611>

10. M. R. D’Orsogna, Q. Lei, T. Chou, First assembly times and equilibration in stochastic coagulation-fragmentation, *J. Chem. Phys.*, **143** (2015), 014112. <https://doi.org/10.1063/1.4923002>
11. D. T. Gillespie, Exact stochastic simulation of coupled chemical reactions, *J. Phys. Chem.*, **81** (1977), 2340–2361. <https://doi.org/10.1021/j100540a008>
12. K. Tucci, R. Kapral, Mesoscopic model for diffusion-influenced reaction dynamics, *J. Chem. Phys.*, **120** (2004), 8262–8270. <https://doi.org/10.1063/1.1690244>
13. K. Tucci, R. Kapral, Mesoscopic multiparticle collision dynamics of reaction-diffusion fronts, *J. Phys. Chem. B*, **109** (2005), 21300–21304. <https://doi.org/10.1021/jp052701u>
14. K. Rohlf, S. Fraser, R. Kapral, Reactive multiparticle collision dynamics, *Comput. Phys. Commun.*, **179** (2008), 132–139. <https://doi.org/10.1016/j.cpc.2008.01.027>
15. K. Rohlf, Stochastic phase-space description for reactions that change particle numbers, *J. Math. Chem.*, **45** (2009), 141–160. <https://doi.org/10.1007/s10910-008-9373-8>
16. R. Strehl, K. Rohlf, Multiparticle collision dynamics for diffusion-influenced signaling pathways, *Phys. Biol.*, **13** (2016), 046004. <https://doi.org/10.1088/1478-3975/13/4/046004>
17. A. Sayyidmousavi, K. Rohlf, Stochastic simulations of the Schnakenberg model with spatial inhomogeneities using reactive multiparticle collision dynamics, *AIMS Math.*, **4** (2019), 1805–1823. <https://doi.org/10.3934/math.2019.6.1805>
18. A. Sayyidmousavi, K. Rohlf, S. Ilie, A hybrid method for micro-mesoscopic stochastic simulation of reaction-diffusion systems, *Math. Biosci.*, **312** (2019), 23–32. <https://doi.org/10.1016/j.mbs.2019.04.001>
19. A. Sayyidmousavi, K. Rohlf, Reactive multi-particle collision dynamics with reactive boundary conditions, *Phys. Biol.*, **15** (2018), 046007. <https://doi.org/10.1088/1478-3975/aabc35>
20. A. Malevanets, R. Kapral, Mesoscopic model for solvent dynamics, *J. Chem. Phys.*, **110** (1999), 8605–8613. <https://doi.org/10.1063/1.478857>
21. S. Bedkihal, J. C. Kumaradas, K. Rohlf, Steady flow through a constricted cylinder by multiparticle collision dynamics, *Biomech. Model. Mechanobiol.*, **12** (2013), 929–939. <https://doi.org/10.1007/s10237-012-0454-z>
22. T. Akhter, K. Rohlf, Quantifying compressibility and slip in multiparticle collision (MPC) flow through a local constriction, *Entropy*, **16** (2014), 418–442. <https://doi.org/10.3390/e16010418>
23. S. Rabba, K. Rohlf, Pressure curves for compressible flows with slip through asymmetric local constrictions, *IJANS*, **3** (2018), 21–40. <https://doi.org/10.1504/IJANS.2018.097324>
24. K. Rohlf, *Theoretical studies on blood flow in small vessels*, Ph.D. Thesis, University of Waterloo, 2002.
25. T. Murata, Theory of non-Newtonian viscosity of blood at low shear rate — Effect of rouleaux, *Biorheol.*, **13** (1976), 287–296. <https://doi.org/10.3233/BIR-1976-13504>
26. T. Murata, T. W. Secomb, Effects of shear rate on rouleau formation in simple shear flow, *Biorheol.*, **25** (1988), 113–122. <https://doi.org/10.3233/BIR-1988-251-218>
27. J. Chen, Z. Huang, Analytical model for effects of shear rate on rouleau size and blood viscosity, *Biophys. Chem.*, **58** (1996), 273–279. [https://doi.org/10.1016/0301-4622\(95\)00105-0](https://doi.org/10.1016/0301-4622(95)00105-0)

28. D. Barrows, *Accurate and efficient stochastic simulation of reaction-diffusion biochemical networks*, Ph.D. Thesis, Toronto Metropolitan University, 2025.
29. P. Kunwar, *Stochastic simulations for extended Becker-Döring models applied to self-assembly processes*, Ph.D. Thesis, Toronto Metropolitan University, 2025.
30. B. Szala-Mendyk, A. Drajkowska, A. Molski, Modified Smoluchowski rate equations for aggregation and fragmentation in finite systems, *J. Phys. Chem. B.*, **127** (2023), 6154–6162. <https://doi.org/10.1021/acs.jpcc.3c02884>

Appendix

A. Number of cluster size ODE comparison

A.1. Exact rate laws

The three different ODEs can alternatively be obtained from

$$\begin{aligned}\frac{dN_1}{dt} &= -K_{11}N_1^2 - K_{12}N_1N_2 + 2W_{11}N_2 + W_{12}N_3 \\ \frac{dN_2}{dt} &= \frac{1}{2}K_{11}N_1^2 - K_{12}N_1N_2 + W_{12}N_3 - W_{11}N_2 \\ \frac{dN_3}{dt} &= K_{12}N_1N_2 - W_{12}N_3,\end{aligned}$$

using rates as follows (see Table 2). Note that the number of clusters of size i , N_i , is related to the concentration by $C_i = \frac{N_i}{V}$ where V is the volume.

Taking $K_{ij} = (1 + \delta_{ii})p_{ij}/V$, and $W_{ij} = q_{ij}$ gives the *chemical rate law* (chem-exact) model

$$\begin{aligned}\frac{dN_1}{dt} &= -2p_{11}N_1^2/V - p_{12}N_1N_2/V + 2q_{11}N_2 + q_{12}N_3 \\ \frac{dN_2}{dt} &= p_{11}N_1^2/V - p_{12}N_1N_2/V + q_{12}N_3 - q_{11}N_2 \\ \frac{dN_3}{dt} &= p_{12}N_1N_2/V - q_{12}N_3.\end{aligned}$$

Taking $K_{ij} = p_{ij}$, and $W_{ij} = q_{ij}$ gives the *BD/EBD* (BD/EBD-exact) model

$$\begin{aligned}\frac{dN_1}{dt} &= -p_{11}N_1^2/V - p_{12}N_1N_2/V + 2q_{11}N_2 + q_{12}N_3 \\ \frac{dN_2}{dt} &= \frac{1}{2}p_{11}N_1^2/V - p_{12}N_1N_2/V + q_{12}N_3 - q_{11}N_2 \\ \frac{dN_3}{dt} &= p_{12}N_1N_2/V - q_{12}N_3.\end{aligned}$$

Taking $K_{ij} = p_{ij}$, and $W_{ij} = (1 - \frac{1}{2}\delta_{ii})q_{ij}$ gives the *Smoluchowski* (Smol-exact) model

$$\frac{dN_1}{dt} = -p_{11}N_1^2/V - p_{12}N_1N_2/V + q_{11}N_2 + q_{12}N_3$$

$$\begin{aligned}\frac{dN_2}{dt} &= \frac{1}{2}p_{11}N_1^2/V - p_{12}N_1N_2/V + q_{12}N_3 - \frac{1}{2}q_{11}N_2 \\ \frac{dN_3}{dt} &= p_{12}N_1N_2/V - q_{12}N_3.\end{aligned}$$

A.2. Approximate rate laws

The three different ODEs for the approximate dynamics can alternatively be obtained from

$$\begin{aligned}\frac{dN_1}{dt} &= -\frac{1}{2}K_{11}N_1^2 - K_{12}N_1N_2 + W_{11}N_2 + W_{12}N_3 \\ \frac{dN_2}{dt} &= \frac{1}{2}K_{11}N_1^2 - K_{12}N_1N_2 + W_{12}N_3 - W_{11}N_2 \\ \frac{dN_3}{dt} &= 2K_{12}N_1N_2 - 2W_{12}N_3,\end{aligned}$$

using rates as follows (see Table 2). Note that the number of clusters of size i , N_i , is related to the concentration by $C_i = \frac{N_i}{V}$ where V is the volume.

Taking $K_{ij} = (1 + \delta_{ii})p_{ij}/V$, and $W_{ij} = q_{ij}$ gives the *approximate chemical rate law* (chem-approx) model

$$\begin{aligned}\frac{dN_1}{dt} &= -p_{11}N_1^2 - \frac{1}{2}p_{12}N_1N_2 + q_{11}N_2 + \frac{1}{2}q_{12}N_3 \\ \frac{dN_2}{dt} &= p_{11}N_1^2 - \frac{1}{2}p_{12}N_1N_2 + \frac{1}{2}q_{12}N_3 - q_{11}N_2 \\ \frac{dN_3}{dt} &= p_{12}N_1N_2 - q_{12}N_3.\end{aligned}$$

Taking $K_{ij} = p_{ij}$, and $W_{ij} = q_{ij}$ gives the *approximate BD/EBD* (BD/EBD-approx) model

$$\begin{aligned}\frac{dN_1}{dt} &= -\frac{1}{2}p_{11}N_1^2 - \frac{1}{2}p_{12}N_1N_2 + q_{11}N_2 + \frac{1}{2}q_{12}N_3 \\ \frac{dN_2}{dt} &= \frac{1}{2}p_{11}N_1^2 - \frac{1}{2}p_{12}N_1N_2 + \frac{1}{2}q_{12}N_3 - q_{11}N_2 \\ \frac{dN_3}{dt} &= p_{12}N_1N_2 - q_{12}N_3.\end{aligned}$$

Taking $K_{ij} = p_{ij}$, and $W_{ij} = (1 - \frac{1}{2}\delta_{ii})q_{ij}$ gives the *approximate Smoluchowski* (Smol-approx) model

$$\begin{aligned}\frac{dN_1}{dt} &= -\frac{1}{2}p_{11}N_1^2 - \frac{1}{2}p_{12}N_1N_2 + \frac{1}{2}q_{11}N_2 + \frac{1}{2}q_{12}N_3 \\ \frac{dN_2}{dt} &= \frac{1}{2}p_{11}N_1^2 - \frac{1}{2}p_{12}N_1N_2 + \frac{1}{2}q_{12}N_3 - \frac{1}{2}q_{11}N_2 \\ \frac{dN_3}{dt} &= p_{12}N_1N_2 - q_{12}N_3.\end{aligned}$$

B. $N = 3$, any M

B.1. $N = 3$, any M , exact dynamics

As a demonstration, for $M = 6$, the matrix $\mathbf{A}$ with the ordering of the four states as indicated above each column of the matrix, is

$$\begin{array}{c} \begin{array}{cccc} (6, 0, 0) & (4, 1, 0) & (2, 2, 0) & (0, 3, 0) \end{array} \\ \left(\begin{array}{c|c|c|c} -\frac{6(5)}{2}K_{11} & W_{11} & 0 & 0 \\ \hline \frac{6(5)}{2}K_{11} & -\frac{4(3)}{2}K_{11} - 4K_{12} - W_{11} & 2W_{11} & 0 \\ \hline 0 & 6K_{11} & -4K_{12} - K_{11} - 2W_{11} & 3W_{11} \\ \hline 0 & 0 & K_{11} & -3W_{11} \end{array} \right) \end{array}$$

with the trapped state $(0, 3, 0)$ in case of no break-up. For $M = 7$ it is

$$\begin{array}{c} \begin{array}{cccc} (7, 0, 0) & (5, 1, 0) & (3, 2, 0) & (1, 3, 0) \end{array} \\ \left(\begin{array}{c|c|c|c} -\frac{7(6)}{2}K_{11} & W_{11} & 0 & 0 \\ \hline \frac{7(6)}{2}K_{11} & -\frac{5(4)}{2}K_{11} - 5K_{12} - W_{11} & 2W_{11} & 0 \\ \hline 0 & 10K_{11} & -6K_{12} - 3K_{11} - 2W_{11} & 3W_{11} \\ \hline 0 & 0 & 3K_{11} & -3K_{12} - 3W_{11} \end{array} \right) \end{array}$$

with no trapped states.

B.2. $N = 3$, any M , approximate dynamics

For comparison purposes the $M = 6$ transition matrix takes the form

$$\begin{array}{c} \begin{array}{cccccc} (6, 0, 0) & (5, 1, 0) & (4, 2, 0) & (3, 3, 0) & (2, 4, 0) & (1, 5, 0) \end{array} \\ \left(\begin{array}{c|c|c|c|c|c} -\frac{6(5)}{2}K_{11} & W_{11} & 0 & 0 & 0 & 0 \\ \hline \frac{6(5)}{2}K_{11} & -\frac{5(4)}{2}K_{11} - 5K_{12} - W_{11} & 2W_{11} & 0 & 0 & 0 \\ \hline 0 & \frac{5(4)}{2}K_{11} & -\frac{4(3)}{2}K_{11} - 8K_{12} - 2W_{11} & 3W_{11} & 0 & 0 \\ \hline 0 & 0 & \frac{4(3)}{2}K_{11} - \frac{3(2)}{2}K_{11} - 9K_{12} - 3W_{11} & 4W_{11} & 0 & 0 \\ \hline 0 & 0 & 0 & \frac{3(2)}{2}K_{11} - \frac{2(1)}{2}K_{11} - 8K_{12} - 4W_{11} & 5W_{11} & 0 \\ \hline 0 & 0 & 0 & 0 & \frac{2(1)}{2}K_{11} - 5K_{12} - 5W_{11} & 0 \end{array} \right) \end{array}$$

The $M = 7$ transition matrix is given by

$(7, 0, 0)$	$(6, 1, 0)$	$(5, 2, 0)$	$(4, 3, 0)$	$(3, 4, 0)$	$(2, 5, 0)$	$(1, 6, 0)$
$-\frac{7(6)}{2}K_{11}$	W_{11}	0	0	0	0	0
$\frac{7(6)}{2}K_{11}$	$-\frac{6(5)}{2}K_{11} - 6K_{12} - W_{11}$	$2W_{11}$	0	0	0	0
0	$\frac{6(5)}{2}K_{11}$	$-\frac{5(4)}{2}K_{11} - 10K_{12} - 2W_{11}$	$3W_{11}$	0	0	0
0	0	$\frac{5(4)}{2}K_{11}$	$-\frac{4(3)}{2}K_{11} - 12K_{12} - 3W_{11}$	$4W_{11}$	0	0
0	0	0	$\frac{4(3)}{2}K_{11}$	$-\frac{3(2)}{2}K_{11} - 12K_{12} - 4W_{11}$	$5W_{11}$	0
0	0	0	0	$\frac{3(2)}{2}K_{11}$	$-K_{11} - 10K_{12} - 5W_{11}$	$6W_{11}$
0	0	0	0	0	K_{11}	$-6K_{12} - 6W_{11}$

C. Theoretical MFATs

C.1. $N = 3$, $K_{ij} = p$, $W_{ij} = q$ (BD/EBD model)

$$\begin{aligned}
 T(5, 0, 0) &= \frac{31p^2 + 14pq + q^2}{30p^2(2p + q)} \\
 T(6, 0, 0) &= \frac{15p^3 + 95p^2q + 21pq^2 + q^3}{60qp^2(5p + q)} \\
 T(7, 0, 0) &= \frac{744p^3 + 487qp^2 + 60q^2p + 2q^3}{105(27p^2 + 20qp + 2q^2)p^2} \\
 T(8, 0, 0) &= \frac{105p^4 + 1526qp^3 + 488q^2p^2 + 40q^3p + q^4}{8qp^2(1029p^2 + 336pq + 21q^2)} \\
 T(9, 0, 0) &= \frac{6277p^4 + 5224p^3q + 909p^2q^2 + 52pq^3 + q^4}{4p^2(10080p^3 + 9072qp^2 + 1512q^2p + 63q^3)} \\
 T(10, 0, 0) &= \frac{945p^5 + 31707qp^4 + 13074q^3p^3 + 1512q^3p^2 + 65q^4p + q^5}{360p^2q(729p^3 + 309qp^2 + 33q^2p + q^3)} \\
 T(50, 0, 0) &= \frac{n_{50}}{d_{50}} \\
 n_{50} &= (105172353q^{25} + 191939544225q^{24} + 174068483506731q^{23} + 102389240246134842q^{22} \\
 &\quad + 43196576261540468501q^{21} + 13752438268249448129152q^{20} \\
 &\quad + 3403389045062481396810595q^{19}
 \end{aligned}$$

$$\begin{aligned}
& +666993729030176880117240224q^{18} + 104763921933641397388630126458q^{17} \\
& +13287650936577566423248112717559q^{16} \\
& +1366667257242769568047907052426438q^{15} \\
& +114163166052512019474890814839521260q^{14} \\
& +7739073452794465598359628957080025586q^{13} \\
& +424458910321134109479354920941594080272q^{12} \\
& +18735391087469385571620644912840791288846q^{11} \\
& +660371296878930350965095555037404737195032q^{10} \\
& +18389704241602708337569990292822391129971085q^9 \\
& +398875420016046265200492091453722539526852599q^8 \\
& +6612568502036372523176355254154091236335130863q^7 \\
& +81690457234006104234471444269924763734731233194q^6 \\
& +726277140989770090209660984623033201955165033593q^5 \\
& +4419590822515325390359455570453103880833602456320q^4 \\
& +17032410741248272709240114227904915908833393306031q^3 \\
& +36252659451722343588210233759117388001512518752632q^2 \\
& +31033638603168321275080667375864313849209540152056q \\
& +6145834944405196915793054022393272390625) \\
d_{50} = & 58800q(105172353q^{23} + 171746452449q^{22} + 135848534909472q^{21} \\
& +68562010335501837q^{20} \\
& +24570150432582919745q^{19} + 6603633827709903307057q^{18} \\
& +1373994714251335765934366q^{17} \\
& +225678437451180757071559853q^{16} \\
& +29618932073266799879197571458q^{15} \\
& +3128544358686403002596644075010q^{14} \\
& +266893962146468558354140534961128q^{13} \\
& +18399333878116859853092654241762554q^{12} \\
& +1022906408653600565678819240286974602q^{11} \\
& +45650042345449788115243436722664153002q^{10} \\
& +1623533409955835562054890865187839918260q^9 \\
& +45543751427275194575418620361186258405922q^8 \\
& +993776945392339585463083094096084206722781q^7 \\
& +16555299591771415144714881259137095234337533q^6 \\
& +205329025174488343937330533282130276019369928q^5 \\
& +1831273554097904923175567439533386072657510745q^4 \\
& +11171569278156813960667591593054389738578591573q^3
\end{aligned}$$

$$\begin{aligned}
&+43135833700071166856106057940021788314207820869q^2 \\
&+91941296067832480190705285009933296968560864862q \\
&+78779175478331397259823357325010988107318886097)
\end{aligned}$$

$$T(51, 0, 0) = \frac{n_{51}}{d_{51}}$$

$$\begin{aligned}
n_{51} = & (25948518875136q^{25} + 49302185862758400q^{24} + 46586076213236465664q^{23} \\
& + 28589565096754583961600q^{22} + 12606002592279147743150080q^{21} \\
& + 4203139628106439224301977600q^{20} + 1091888059553064973578122956800q^{19} \\
& + 225205854819313789771214785945600q^{18} \\
& + 37333690158001782659660001756134400q^{17} \\
& + 5013550940491173512546256690934886400q^{16} \\
& + 547904798441056884281836496861759635200q^{15} \\
& + 48824947482511605351159348499298021952000q^{14} \\
& + 3546817322356271212531208260700644325409600q^{13} \\
& + 209539832116725597983935103070534564954560000q^{12} \\
& + 10022619094110477969845814386722448174706360000q^{11} \\
& + 385532394806775070905748929899624567538598320000q^{10} \\
& + 11816043657655539938876088898678928372786647450000q^9 \\
& + 285000206652288268596802056686535068378829773880000q^8 \\
& + 5322514969736114669750349761518061785498126078050000q^7 \\
& + 75328842268479960812806515530410552751219921624700000q^6 \\
& + 784971750264287164255919337779801581407823191399937500q^5 \\
& + 5786880794904923847896266408605256029031829384970000000q^4 \\
& + 28475749802939548172657082772022646638300613293766812500q^3 \\
& + 85344731802783495623789849508743157903928065826115625000q^2 \\
& + 132141403692064706539569597403160584721974734440572734375q \\
& + 71766925586528307103299342207989697895256946859581187500) \\
d_{51} = & 62475(25948518875136q^{24} + 44216276163231744q^{23} + 36540262821840814080q^{22} \\
& + 19297685260045225820160q^{21} + 7250153004219621806571520q^{20} \\
& + 2047275487616944122193510400q^{19} \\
& + 448634791919987781855616307200q^{18} \\
& + 77822514717481262768647503667200q^{17} \\
& + 10820094590162082937591384572902400q^{16} \\
& + 1214947155006491149792904862369792000q^{15} \\
& + 110614321354447006845774142518896582400q^{14}
\end{aligned}$$

$$\begin{aligned}
&+8174724311660159745752033889927981657600q^{13} \\
&+489706710050739380953448524031452757640000q^{12} \\
&+23689905025423917856333541697846153956000000q^{11} \\
&+919730370089633235001974210844489046875000000q^{10} \\
&+28403104660274627031988414008824301406250000000q^9 \\
&+689351910421901887619677225236412365722656250000q^8 \\
&+12939720553711001881116806806766354980468750000000q^7 \\
&+183895783883000770853344210229928779602050781250000q^6 \\
&+1922740452864650111014844289754243850708007812500000q^5 \\
&+14212525876178072317995833596941083669662475585937500q^4 \\
&+70081820948957647292700362438336014747619628906250000q^3 \\
&+210369787065023202794628923584241420030593872070312500q^2 \\
&+326071576371675451496366804349236190319061279296875000q \\
&+177194539196445440065730281276046298444271087646484375)
\end{aligned}$$

$$\begin{aligned}
T(100, 0, 0) &= \frac{n_{100}}{d_{100}} \\
n_{100} &= (2.0614 \times 10^{-6} (2.7980 \times 10^{102} p^{50} + 6.6488 \times 10^{122} p^{49} q + 1.0005 \times 10^{123} p^{48} q^2 \\
&\quad + 6.4644 \times 10^{122} p^{47} q^3 \\
&\quad + 2.4425 \times 10^{122} p^{46} q^4 + 6.1596 \times 10^{121} p^{45} q^5 + 1.1175 \times 10^{121} p^{44} q^6 + 1.5313 \times 10^{120} p^{43} q^7 \\
&\quad + 1.6404 \times 10^{119} p^{42} q^8 + 1.4093 \times 10^{118} p^{41} q^9 + 9.9043 \times 10^{116} p^{40} q^{10} + 5.7845 \times 10^{115} p^{39} q^{11} \\
&\quad + 2.8438 \times 10^{114} p^{38} q^{12} + 1.1894 \times 10^{113} p^{37} q^{13} + 4.2702 \times 10^{111} p^{36} q^{14} + 1.3259 \times 10^{110} p^{35} q^{15} \\
&\quad + 3.5841 \times 10^{108} p^{34} q^{16} + 8.4808 \times 10^{106} p^{33} q^{17} + 1.7652 \times 10^{105} p^{32} q^{18} + 3.2452 \times 10^{103} p^{31} q^{19} \\
&\quad + 5.2892 \times 10^{101} p^{30} q^{20} + 7.6662 \times 10^{99} p^{29} q^{21} + 9.9084 \times 10^{97} p^{28} q^{22} + 1.1446 \times 10^{96} p^{27} q^{23} \\
&\quad + 1.1841 \times 10^{94} p^{26} q^{24} + 1.0989 \times 10^{92} p^{25} q^{25} + 9.1598 \times 10^{89} p^{24} q^{26} + 6.8650 \times 10^{87} p^{23} q^{27} \\
&\quad + 4.6296 \times 10^{85} p^{22} q^{28} + 2.8105 \times 10^{83} p^{21} q^{29} + 1.5362 \times 10^{81} p^{20} q^{30} + 7.5592 \times 10^{78} p^{19} q^{31} \\
&\quad + 3.3470 \times 10^{76} p^{18} q^{32} + 1.3325 \times 10^{74} p^{17} q^{33} + 4.7643 \times 10^{71} p^{16} q^{34} + 1.5275 \times 10^{69} p^{15} q^{35} \\
&\quad + 4.3830 \times 10^{66} p^{14} q^{36} + 1.1226 \times 10^{64} p^{13} q^{37} + 2.5587 \times 10^{61} p^{12} q^{38} + 5.1695 \times 10^{58} p^{11} q^{39} \\
&\quad + 9.2131 \times 10^{55} p^{10} q^{40} + 1.4398 \times 10^{53} p^9 q^{41} + 1.9584 \times 10^{50} p^8 q^{42} + 2.2963 \times 10^{47} p^7 q^{43} \\
&\quad + 2.2923 \times 10^{44} p^6 q^{44} + 1.9158 \times 10^{41} p^5 q^{45} + 1.3092 \times 10^{38} p^4 q^{46} + 7.0612 \times 10^{34} p^3 q^{47} \\
&\quad + 2.8361 \times 10^{31} p^2 q^{48} + 7.5972 \times 10^{27} p q^{49} + 1.0267 \times 10^{24} q^{50}) \\
d_{100} &= p^2 q (6.3374 \times 10^{119} p^{48} + 9.5337 \times 10^{119} p^{47} q + 6.1579 \times 10^{119} p^{46} q^2 + 2.3257 \times 10^{119} p^{45} q^3 \\
&\quad + 5.8618 \times 10^{118} p^{44} q^4 + 1.0628 \times 10^{118} p^{43} q^5 + 1.4553 \times 10^{117} p^{42} q^6 + 1.5576 \times 10^{116} p^{41} q^7 \\
&\quad + 1.3370 \times 10^{115} p^{40} q^8 + 9.3861 \times 10^{113} p^{39} q^9 + 5.4754 \times 10^{112} p^{38} q^{10} + 2.6883 \times 10^{111} p^{37} q^{11} \\
&\quad + 1.1227 \times 10^{110} p^{36} q^{12} + 4.0242 \times 10^{108} p^{35} q^{13} + 1.2473 \times 10^{107} p^{34} q^{14} + 3.3649 \times 10^{105} p^{33} q^{15} \\
&\quad + 7.9447 \times 10^{103} p^{32} q^{16} + 1.6496 \times 10^{102} p^{31} q^{17} + 3.0247 \times 10^{100} p^{30} q^{18} + 4.9154 \times 10^{98} p^{29} q^{19}
\end{aligned}$$

$$\begin{aligned}
& +7.1015 \times 10^{96} p^{28} q^{20} + 9.1458 \times 10^{94} p^{27} q^{21} + 1.0524 \times 10^{93} p^{26} q^{22} + 1.0840 \times 10^{91} p^{25} q^{23} \\
& + 1.0011 \times 10^{89} p^{24} q^{24} + 8.3006 \times 10^{86} p^{23} q^{25} + 6.1844 \times 10^{84} p^{22} q^{26} + 4.1432 \times 10^{82} p^{21} q^{27} \\
& + 2.4968 \times 10^{80} p^{20} q^{28} + 1.3536 \times 10^{78} p^{19} q^{29} + 6.5989 \times 10^{75} p^{18} q^{30} + 2.8914 \times 10^{73} p^{17} q^{31} \\
& + 1.1376 \times 10^{71} p^{16} q^{32} + 4.0127 \times 10^{68} p^{15} q^{33} + 1.2667 \times 10^{66} p^{14} q^{34} + 3.5702 \times 10^{63} p^{13} q^{35} \\
& + 8.9557 \times 10^{60} p^{12} q^{36} + 1.9919 \times 10^{58} p^{11} q^{37} + 3.9091 \times 10^{55} p^{10} q^{38} + 6.7287 \times 10^{52} p^9 q^{39} \\
& + 1.0081 \times 10^{50} p^8 q^{40} + 1.3015 \times 10^{47} p^7 q^{41} + 1.4294 \times 10^{44} p^6 q^{42} + 1.3121 \times 10^{41} p^5 q^{43} \\
& + 9.8187 \times 10^{37} p^4 q^{44} + 5.7714 \times 10^{34} p^3 q^{45} + 2.5081 \times 10^{31} p^2 q^{46} + 7.1948 \times 10^{27} p q^{47} \\
& + 1.0267 \times 10^{24} q^{48}) \\
T(101, 0, 0) &= \frac{n_{101}}{d_{101}} \\
n_{101} &= 2.0002 \times 10^{-6} (2.0738 \times 10^{134} p^{50} + 4.5373 \times 10^{134} p^{49} q + 3.8206 \times 10^{134} p^{48} q^2 \\
& + 1.7767 \times 10^{134} p^{47} q^3 \\
& + 5.3248 \times 10^{133} p^{46} q^4 + 1.1213 \times 10^{133} p^{45} q^5 + 1.7539 \times 10^{132} p^{44} q^6 + 2.1177 \times 10^{131} p^{43} q^7 \\
& + 2.0309 \times 10^{130} p^{42} q^8 \\
& + 1.5810 \times 10^{129} p^{41} q^9 + 1.0165 \times 10^{128} p^{40} q^{10} + 5.4736 \times 10^{126} p^{39} q^{11} + 2.4969 \times 10^{125} p^{38} q^{12} \\
& + 9.7433 \times 10^{123} p^{37} q^{13} + 3.2787 \times 10^{122} p^{36} q^{14} + 9.5811 \times 10^{120} p^{35} q^{15} + 2.4459 \times 10^{119} p^{34} q^{16} \\
& + 5.4832 \times 10^{117} p^{33} q^{17} + 1.0842 \times 10^{116} p^{32} q^{18} + 1.8985 \times 10^{114} p^{31} q^{19} + 2.9538 \times 10^{112} p^{30} q^{20} \\
& + 4.0953 \times 10^{110} p^{29} q^{21} + 5.0726 \times 10^{108} p^{28} q^{22} + 5.6256 \times 10^{106} p^{27} q^{23} + 5.5961 \times 10^{104} p^{26} q^{24} \\
& + 5.0008 \times 10^{102} p^{25} q^{25} + 4.0194 \times 10^{100} p^{24} q^{26} + 2.9085 \times 10^{98} p^{23} q^{27} + 1.8960 \times 10^{96} p^{22} q^{28} \\
& + 1.1139 \times 10^{94} p^{21} q^{29} + 5.8978 \times 10^{91} p^{20} q^{30} + 2.8140 \times 10^{89} p^{19} q^{31} + 1.2093 \times 10^{87} p^{18} q^{32} \\
& + 4.6768 \times 10^{84} p^{17} q^{33} + 1.6257 \times 10^{82} p^{16} q^{34} + 5.0715 \times 10^{79} p^{15} q^{35} + 1.4169 \times 10^{77} p^{14} q^{36} \\
& + 3.5361 \times 10^{74} p^{13} q^{37} + 7.8585 \times 10^{71} p^{12} q^{38} + 1.5490 \times 10^{69} p^{11} q^{39} + 2.6951 \times 10^{66} p^{10} q^{40} \\
& + 4.1143 \times 10^{63} p^9 q^{41} + 5.4694 \times 10^{60} p^8 q^{42} + 6.2712 \times 10^{57} p^7 q^{43} + 6.1249 \times 10^{54} p^6 q^{44} \\
& + 5.0105 \times 10^{51} p^5 q^{45} \\
& + 3.3529 \times 10^{48} p^4 q^{46} + 1.7714 \times 10^{45} p^3 q^{47} + 6.9716 \times 10^{41} p^2 q^{48} + 1.8303 \times 10^{38} p q^{49} \\
& + 2.4242 \times 10^{34} q^{50}) \\
d_{101} &= p^2 (1.9489 \times 10^{131} p^{49} + 4.2632 \times 10^{131} p^{48} q + 3.5888 \times 10^{131} p^{47} q^2 + 1.6682 \times 10^{131} p^{46} q^3 \\
& + 4.9973 \times 10^{130} p^{45} q^4 \\
& + 1.0518 \times 10^{130} p^{44} q^5 + 1.6441 \times 10^{129} p^{43} q^6 + 1.9836 \times 10^{128} p^{42} q^7 + 1.9006 \times 10^{127} p^{41} q^8 \\
& + 1.4782 \times 10^{126} p^{40} q^9 \\
& + 9.4935 \times 10^{124} p^{39} q^{10} + 5.1057 \times 10^{123} p^{38} q^{11} + 2.3260 \times 10^{122} p^{37} q^{12} + 9.0626 \times 10^{120} p^{36} q^{13} \\
& + 3.0446 \times 10^{119} p^{35} q^{14} + 8.8807 \times 10^{117} p^{34} q^{15} + 2.2625 \times 10^{116} p^{33} q^{16} + 5.0608 \times 10^{114} p^{32} q^{17} \\
& + 9.9826 \times 10^{112} p^{31} q^{18} + 1.7433 \times 10^{111} p^{30} q^{19} + 2.7042 \times 10^{109} p^{29} q^{20} + 3.7369 \times 10^{107} p^{28} q^{21} \\
& + 4.6121 \times 10^{105} p^{27} q^{22} + 5.0945 \times 10^{103} p^{26} q^{23} + 5.0456 \times 10^{101} p^{25} q^{24} + 4.4870 \times 10^{99} p^{24} q^{25} \\
& + 3.5871 \times 10^{97} p^{23} q^{26} + 2.5802 \times 10^{95} p^{22} q^{27} + 1.6708 \times 10^{93} p^{21} q^{28} + 9.7431 \times 10^{90} p^{20} q^{29})
\end{aligned}$$

$$\begin{aligned}
&+5.1162 \times 10^{88} p^{19} q^{30} + 2.4184 \times 10^{86} p^{18} q^{31} + 1.0284 \times 10^{84} p^{17} q^{32} + 3.9299 \times 10^{81} p^{16} q^{33} \\
&+1.3476 \times 10^{79} p^{15} q^{34} + 4.1387 \times 10^{76} p^{14} q^{35} + 1.1356 \times 10^{74} p^{13} q^{36} + 2.7753 \times 10^{71} p^{12} q^{37} \\
&+6.0177 \times 10^{68} p^{11} q^{38} + 1.1521 \times 10^{66} p^{10} q^{39} + 1.9356 \times 10^{63} p^9 q^{40} + 2.8319 \times 10^{60} p^8 q^{41} \\
&+3.5726 \times 10^{57} p^7 q^{42} + 3.8358 \times 10^{54} p^6 q^{43} + 3.4437 \times 10^{51} p^5 q^{44} + 2.5215 \times 10^{48} p^4 q^{45} \\
&+1.4507 \times 10^{45} p^3 q^{46} + 6.1730 \times 10^{41} p^2 q^{47} + 1.7343 \times 10^{38} p q^{48} + 2.4242 \times 10^{34} q^{49})
\end{aligned}$$



AIMS Press

© 2026 the Author(s), licensee AIMS Press. This is an open access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0>)