Fast Reactive Brownian Dynamics

Aleksandar Donev
Chiao-Yu Yang (now at UC Berkeley)

Courant Institute, *New York University*

Spatially Distributed Stochastic Dynamical Systems in Biology
Newton Institute, Cambridge University
June 21st 2016

Chemical Reactions in Solution

- Many chemical reactions occur in a viscous solvent and are **affected by diffusion**, making a mean-field or "well-mixed" deterministic reaction-diffusion PDE approximation inappropriate.
- Classical examples where the **Law of Mass Action** (LMA) reaction-diffusion equations **fails spectacularly** is annihilation $A + B \to 0$
  But even in $A + B \rightleftarrows C$ there are power-law tail signatures in the dynamics even at chemical equilibrium.
- **Spatial fluctuations play a key role** and spatial diffusion must be accounted for; this is different from (in addition to) fluctuations coming from there being **very few reactants** of certain species.
- Primarily interested in the case when **fluctuations are weak**, i.e., lots of molecules are involved, but fluctuations still make a difference.

Grid-Based Methods

- The traditional approach to simulation of reaction-diffusion problems is to solve the **Reaction-Diffusion Master Equation** (RDME), which has the following issues:
    - Diffusion events/hops dominate when cells are "well-mixed" (they must be!) making solving the RDME exactly **expensive** (but one can do **multinomial diffusion** and/or **tau leaping** to speed things up).
    - Diffusion is modeled by a jump process instead of the more physical continuous random walk leading to a **fluctuating Fick's law** (fluctuating hydrodynamics [1]).
    - LMA is postulated instead of following from the model; this requires **effective macroscopic rates** instead of microscopic ones.
      LMA is missing a length scale (**reactive distance**).
    - The results depend strongly on the cell size and are thus **not grid-independent**: binary reactions lost as cell size shrinks [2]
      See Sam Isaacson's **Convergent RDME** (CRDME) [3] for one *grid-based* fix (reactions between neighboring cells).

Particle-Based Methods

- Particles are modeled as species-labeled spheres that diffuse as **independent Brownian walkers** (but note importance of **hydrodynamic interactions**) and react based on a microscopic reaction rule.
- Particle methods are **grid-free** and **closure-free** (they take reactive length scale as input).
- Key problem is **lack of efficiency**. This is what I address in this talk.
- To handle reactions two models are commonly used:
  - In the **surface-reactivity / Smoluchowski model** particles react **upon touching**.
    This automatically includes **steric repulsion** but lacks a mechanism to control reaction rate (but one can introduce unbinding).
  - In the **volume-reactivity / Doi model** particles react with a certain Poisson reaction rate **while they overlap**.
    This allows to separately and independently control the reactive distance and the effective reaction rate [2].

Particle Algorithms

- The Smoluchowski model can be simulated **exactly** by the event-driven **First Passage Kinetic Monte Carlo** method (FPKMC) [4], called **eGFRD** in the biochemical community.
- FPKMC becomes **inefficient** at larger densities, it is **hard to generalize**, and it is **quite complicated** to implement.
- Approximate **reactive Brownian Dynamics** exist (e.g., Smoldyn) but they make **uncontrolled approximations** in diffusion and reactions.
- The Doi model is much simpler and more flexible so we use it here. Consider $A + B \to \ldots$

  Particles are spheres of a given reactive radius R_A and R_B. They diffuse as independent Brownian walkers, and **while two particles overlap** ($r_{AB} \leq R_A + R_B$) they react as a Poisson process with a given rate λ. Only binary reactions are allowed.
- Use our **Split Reactive Brownian Dynamics (SRBD)** to simulate Doi model efficiently with controlled accuracy!

Time Splitting

SRBD is a combination of ideas taken from the **Isotropic Direct Simulation Monte Carlo** (IDSMC) algorithm (used for simulating binary collisions in low-density gases) and the **next subvolume method** for solving the RDME:

- Strang **time splitting** is used to separate diffusion from reaction (**D/2+R+D/2**). This is the only error introduced so **error is controlled** by reducing the time step size Δt.
- Simulating the diffusion exactly without reactions is trivial and inexpensive for independent Brownian walkers:

  $$\mathbf{q}_k(t+\Delta t) = \mathbf{q}_k(t) + \sqrt{2D_k \Delta t}\,\mathcal{N}(0,1)$$

  One can include hydrodynamic interactions (expensive!).
- The difficult part is to **simulate reactions exactly** while particles are stationary (fixed). This is our key contribution.
  It is possible to make some approximations and speed this up greatly but we want to control the error by a single parameter Δt.

Selecting Reactions

- An obvious but very slow method is to first make a list of all overlapping pairs of particles, and then use a Gillespie-like / SSA / KMC algorithm to select pairs to react in sequence.
- The key idea is to accomplish the same (in law) without making a list of all overlapping pairs, by using an **event-driven algorithm** not on particle pairs but on grid cells!
- Introduce back a **computational grid** which is not part of model (think of neighbor search in MD) with spacing larger than all potential reactive distances:

  Particles can only overlap/react with a particle in their own cell or neighboring cells.
- For each cell i, schedule the next potential binary reaction between a particle in cell i and a particle in the neighborhood of cell i (9 cells in 2D or 27 cells in 3D).
- Make an **event queue** (heap) of all cells, and then process reactions by choosing the next cell (next subvolume) in which to try a reaction.

Basic Time Stepping Algorithm

1. Diffuse for half a time step

$$\mathbf{q}_k^{n+\frac{1}{2}} = \mathbf{q}_k^n + \sqrt{D_k \Delta t}\,\mathcal{N}(0,1)$$

2. Prepare: Build **linked-list cells** (LLCs) and schedule next reaction for each cell (if before time Δt) and **build** an **event queue**.
3. Event Loop: Until the event queue is empty, do:
    1. Select cell i on top of queue with time stamp $t^n \leq t \leq t^n + \Delta t$.
    2. Select next reaction to happen in cell i using usual KMC/SSA method.
    3. Process the reaction (if particles overlap for binary), creating/destroying/updating particles+LLCs as necessary.
    4. For each cell i (potentially) affected by reaction, compute the total reaction rate α, sample an exponentially-distributed δt with mean α^{-1}. If $t_i = t + \delta t < t + \Delta t$ schedule next event at time t_i and update event queue, otherwise delete cell i from queue.
4. Diffuse remaining/new particles for half a time step

$$\mathbf{q}_k^{n+1} = \mathbf{q}_k^{n+\frac{1}{2}} + \sqrt{D_k \Delta t}\,\mathcal{N}(0,1)$$

Scheduling Reactions

- For reactions with different reactants $A + B \to ...$ with rate λ, we schedule separately $A + B \to ...$ and $B + A \to ...$
  - For binary reaction r of form $A + B \to ...$ (order matters!), the propensity function (rate) for cell i is
    $$\alpha_r = \frac{\lambda}{2} N_A N'_B$$
    where N_A is the number of A particles in cell i, and
    N'_B is the total number of B particles in the neighborhood of i.
  - For $A + A \to 0$, the rate is $\alpha_r = \frac{\lambda}{2} N_A N'_A$ since pairs are counted twice (we reject self-reactions later).
- We add all the rates in each cell, $\alpha = \sum_{r=1}^{N_r} \alpha_r$ (as in ordinary SSA).
- Note that this over-estimates the actual rate since it does not account for whether the particles actually overlap; we correct for this using rejection: If a pair is selected to react does not overlap (or the same particle A is selected twice) we reject the pair.

Processing Reactions

Once we select the next cell i to *potentially* have a reaction using the event queue, we need to:

- Select a particle at random from cell i of the first reactant species, and another particle of the second reactant species from the neighborhood of i (can be the same particle twice!).
- Test if the two particles are within their reactive distance, and if not, do nothing.
- Otherwise, process the reaction by deleting and adding particles (see next slide) depending on the reaction.
- While doing this, keep track of whether any event is processed that changes the population of cell i (number of particles of each species), and also whether the population of a neighboring cell j changes.
- Recompute reaction rates and schedule a new event for cell i.
- If population of i changed, update the event prediction for all neighbor cells of i, and, if population of j changed, update the event prediction for all cell neighbors of j that are not neighbors of i.

Reversible Reaction Rules

We have made a set of rules that obey **microscopic reversibility** (detailed balance) that dictate how reactions are processed, e.g.,

1. **Annihilation**: $A+B \to \emptyset$ if within distance $R_{AB} = R_A + R_B$.
2. **Birth**: $\emptyset \to A+B$. The A is born uniformly in the system, and B is born with position uniformly chosen within a reactive sphere of radius R_{AB} around the A.
3. **Merge**: $A+A \to B$ or $A+B \to C$. One of the reactants changes species and the others disappears.
4. **Replication**: $A \to B+C$, where B/C can be equal to A: A becomes a B or a C or remains as is, and another particle is born uniformly in a sphere centered at the A with radius R_{BC}.
5. **Transform**: $A+B \to C+D$ or $A+B \to A+C$ (catalysis): No particle changes position or new particles are created, only species are changed (e.g., B becomes a C or D).

From microscopic to macroscopic rates

- For a single reaction $A + B \to \dots$, Erban and Chapman [2] derive that **in 3D** the **macroscopic reaction rate** k (units m^3/s) is related to the **microscopic rate** λ (units s^{-1}),

$$k = 4\pi D_{AB} R_{AB} \left[1 - \sqrt{\frac{D_{AB}}{\lambda R_{AB}^2}} \tanh\left(\sqrt{\frac{\lambda R_{AB}^2}{D_{AB}}} \right) \right], \qquad (1)$$

  where $D_{AB} = D_A + D_B$ is the mutual diffusion coefficient and $R_{AB} = R_A + R_B$ is the reactive radius.

- For $A + A \to$ one just divides the rate by two, and using $D_{AA} = 2D_A$ gives

$$k_{AA} = 4\pi D_A R_{AA} \left[1 - \sqrt{\frac{2D_A}{\lambda R_{AA}^2}} \tanh\left(\sqrt{\frac{\lambda R_{AA}^2}{2D_A}} \right) \right], \qquad (2)$$

Diffusion-Limited vs Reaction-Limited?

- This shows that the important parameter is the dimensionless number

$$r = \frac{\lambda R_{AB}^2}{D_{AB}} \quad \text{and} \quad k < 4\pi D_{AB} R_{AB}.$$

- If $r \gg 1$ (**diffusion-limited**), then $k \to 4\pi D_{AB} R_{AB}$, which is the **Smoluchowski rate**, i.e., particles react upon first touching.
- For $r \ll 1$ (**reaction-limited**), then we get the result expected if the particle positions are uncorrelated, i.e., the system is "uniformly mixed" at microscopic scales:

$$k \approx \frac{4\pi}{3} R_{AB}^3 \lambda \quad \text{for A+B}$$
$$k \approx \frac{2\pi}{3} R_{AB}^3 \lambda \quad \text{for A+A}$$

Diffusion-Limited RDME

- Similarly, **for RDME** in 3D for $A+B \to \dots$, the effective macroscopic rate is related to input (microscopic) rate k_{RDME} via [2, 5]
$$\frac{1}{k} = \frac{1}{k_{RDME}} + \frac{\beta = 0.25273}{hD},$$
where $D = D_{AB}$ for $A+B$ and $D = D_A$ for $A+A$.
- This explains the **loss of bimolecular reactions** as $h \to 0$ (more precisely, in 3D, when $h \ll k/D_{AB}$).
- **Renormalization** theory suggests that for $A+A \to A$ the **law of mass action at finite densities is non-analytic** [5]
$$\frac{k}{k_0} = 1 + \alpha \left(\frac{k_0}{D_A}\right)^{\frac{3}{2}} n_A^{\frac{1}{2}} = 1 + \beta f^{3/2} \phi^{\frac{1}{2}},$$
where $k_0 = \lim_{\phi \to 0} k$, $f = k_0 / (4\pi D_A R_{AA})$, $\phi = n_A \cdot \left(4\pi R_{AA}^3/3\right)$ for Doi or $\phi = n_A h^3$ for RDME, and for RDME $\alpha = 1/2\pi\sqrt{2}$ [5].

Many-Body Effects (3D): $A + A \to A$

Coagulation $A + A \to A, \quad 0 \to A.$

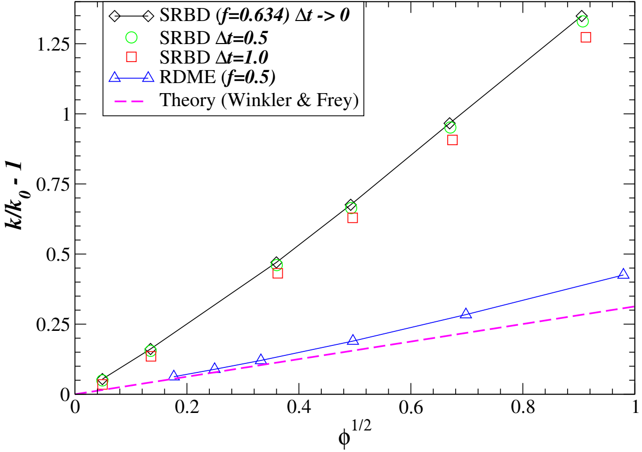


Figure: Here $k_0 = 0.634 \cdot (2\pi D_{AA} R_{AA})$ is computed from (2), and $\phi = n_A \cdot \left(4\pi R_{AA}^3/3\right)$ is the packing density.

Many-Body Effects (3D): $A + B \to B$

We find that (1) holds **only for very dilute many-body systems** with $A + B \to B, \quad 0 \to A$ (B is conserved).

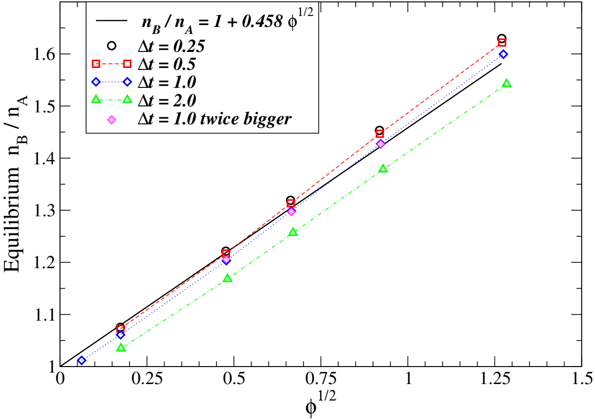


Figure: If (1) is correct, then $k = (1/2) \cdot 4\pi D_{AB} R_{AB}$ and $\langle n_A \rangle = n_B$ at equilibrium. Here $R_A = R_B$, $D_A = D_B$, and $\phi = (n_A + n_B) \cdot \left(4\pi R_{AB}^3 / 3\right)$.

Long-Time Tail: $A + B \leftrightarrow C$

Obeys detailed balance: $k = \frac{4\pi}{3} R_{AB}^3 \lambda$ **independent of kinetics**!

For diffusion-limited there are long-time tails in the ACF [6].

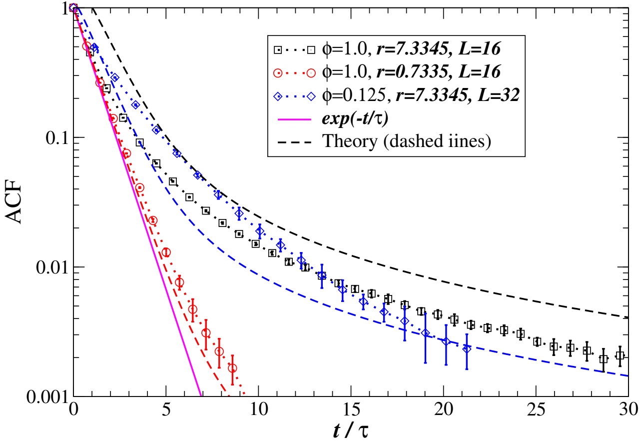


Figure: Here $n_A^{eq} = n_B^{eq} = n_C^{eq}$, $D_{A/B/C} = D$, $r = \lambda R^2/D$, and $\phi = n \cdot \left(4\pi R_{AB}^3/3\right)$.

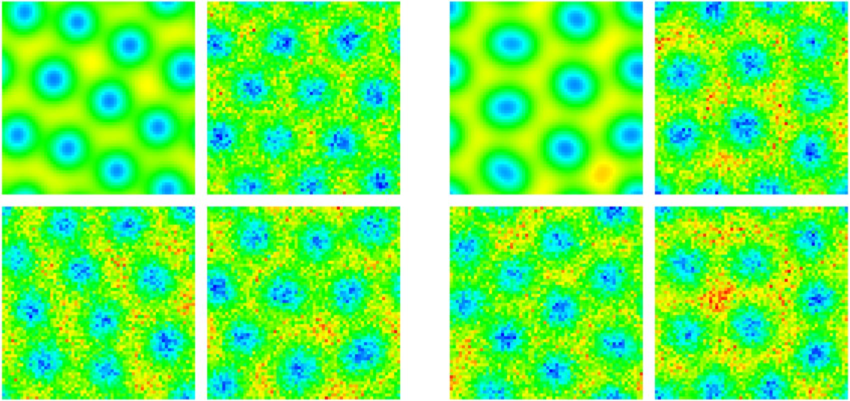
BPM Model

- We have studied the Baras-Pearson-Mansour (BPM) reaction network (3 species, 7 reactions):

$$
\begin{aligned}
U + W &\to V + W \\
V + V &\leftrightarrows W \\
V &\leftrightarrows 0 \\
U &\leftrightarrows 0.
\end{aligned}
$$

- This system has only binary reactions but can exhibit bimodal states (**bistability**) and also **limit cycles**.
- We use parameters giving a limit cycle together with $D_V = D_W = D_U/10$ to get a **Turing-like pattern**.
- We want to understand the role fluctuations play in the formation of the Turing pattern.
- We do simulations in 2D to make them more feasible and to simplify visualization, but the code works in 3D as well.

Turing-like Patterns



Top (grid only): (Left) Deterministic reaction-diffusion.
(Right) RDME using multinomial diffusion [7] + SSA.
Bottom (particle+grid): (Left) **S-BD-RME** [8]. (Right) **SRBD**
Periodic 256 × 256 grid in all cases coarsened to 64 × 64
$R_U = R_V = R_W = h/2$ for SRBD

Diffusion-Limited or Reaction-Limited?

- Define the **chemical penetration depth**, related to the typical distance a molecule travels between successive reactions,

$$L=\sqrt{\frac{D_{AB}}{kn_{AB}}},\quad \text{where} \quad n_{AB}=n_A+n_B.$$

- Define also the **packing fraction**

$$\phi=\frac{4\pi}{3}n_{AB}R_{AB}^3\gtrsim 0.1.$$

If $\phi\ll 1$ use FPKMC-style algorithms and *not* SRBD!

- Let the number of molecules in a penetration volume be $N_L=n_{AB}L^3\gg 1$ (if this were not true then fluctuations would be too large to see a Turing pattern).

Order-of-Magnitude Estimates

- Then for SRDB we have

$$\frac{k}{R_{AB}D_{AB}} = \frac{1}{N_L}\left(\frac{L}{R_{AB}}\right) = \left(\frac{4\pi}{3\phi N_L^2}\right)^{\frac{1}{3}} \ll 1 \quad \Rightarrow$$

$$r = \frac{3}{4\pi}\frac{k}{R_{AB}D_{AB}} \ll 1$$

- And for RDME we would have $k \approx k_{RDME}$ since

$$\frac{k}{hD_{AB}} \ll \frac{k}{R_{AB}D_{AB}} \ll 1 \text{ for RDME since } h > R_{AB}.$$

From (macroscopic) RDME to (microscopic) SRBD rates

- The relation

$$k \approx \frac{4\pi}{3} R_{AB}^3 \lambda \quad \text{for A+B}$$
$$k \approx \frac{2\pi}{3} R_{AB}^3 \lambda \quad \text{for A+A}$$

  has an obvious generalization in 1D and 2D as well and we will use it to determine λ given k in order to compare different methods.
- Note that for our Turing test to see any corrections from diffusion-limited effects we would need grids larger than 1024×1024, so assuming $r \ll 1$ is well justified.
- We have studied grids 64^2, 128^2 and 256^2, keeping the physical parameters (diffusion, macroscopic reaction rates, domain size) fixed.

Comparison of methods

- We compare to a **deterministic/fluctuating hydrodynamics solver** developed with the group of Bell/Garcia at LBNL [1].
- We implement **RDME** using splitting with **multinomial diffusion** (Tartakovsky [7]) + SSA for reactions – much simpler and much more efficient than next-reaction method.
- We also implement a particle algorithm that we call **Split-Brownian Dynamics-Reaction Master Equation (S-BD-RME)** where diffusion is done using a continuum Brownian walk as in SRBD but reactions are done on a grid using SSA (see Tartakovsky [8]).
- Before carrying out SSA in S-BD-RME, we randomly shift the reaction grid to improve Galilean invariance.
- In SRBD we use $R_U = R_V = R_W = h/2$ where h is the grid spacing used in the RME-based methods.
- Time step in *all* methods is **limited by fast diffusion**; we set Courant number $D_U \Delta t / h^2 \approx 0.3$.

Changing Resolution

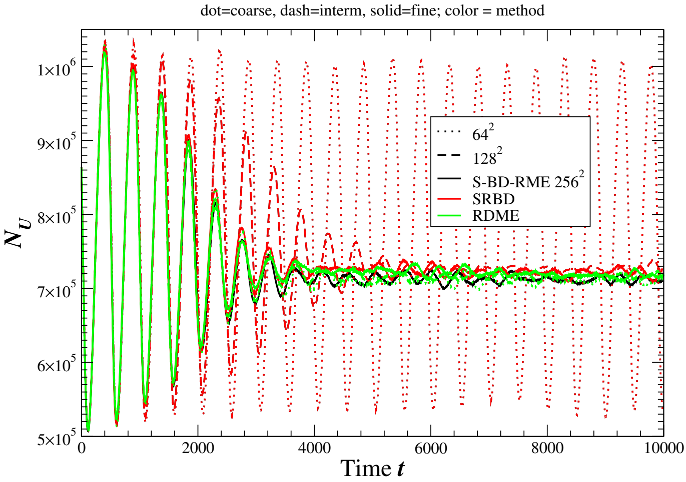


Figure: For 256^2 resolution, total running time is 4.5h for RDME, 12h for S-BD-RME, and 19h for SRBD. For 512^2 resolution, 54h for S-BD-RME and 48h for SRBD (but 43h for 256^2 reaction grid).

Walk versus Hops (coarse grid)

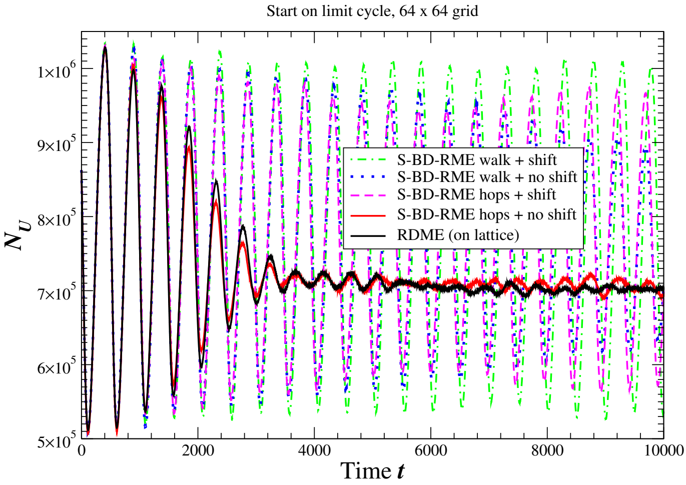


Figure: Importance of microscopic details of diffusion (**hops versus continuum walk**) and reactions (reaction grid randomly shifted or not) for **under-resolved simulations** (cells not uniformly mixed).

Walk versus Hops (finer grid)

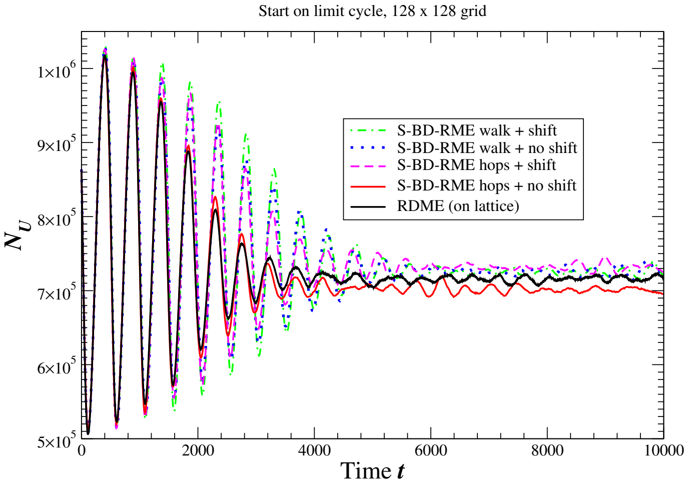


Figure: Importance of microscopic details of diffusion (**hops versus continuum walk**) and reactions (reaction grid randomly shifted or not) for **marginally-resolved simulations** (cells not uniformly mixed).

Fine Resolutions (Resolved)

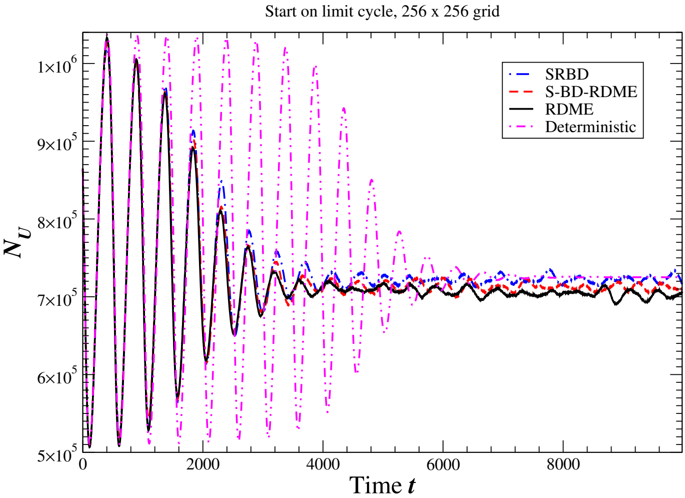


Figure: If the resolution is sufficiently high (resolved) we get matching between the fluctuating methods (including fluctuating hydrodynamics, not shown).

Quantitative Comparison

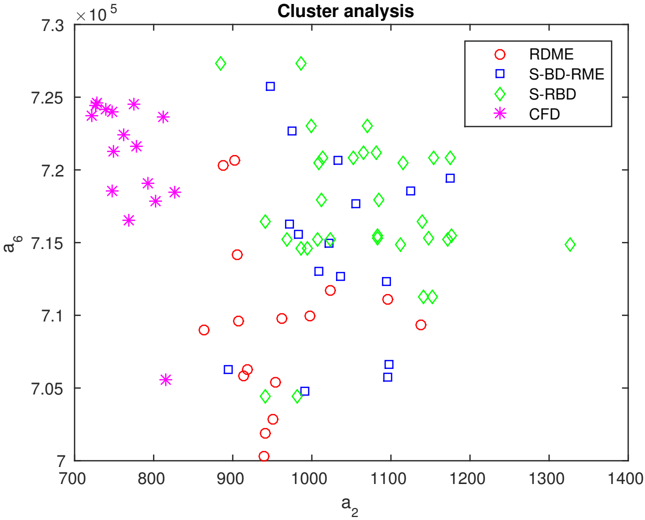


Figure: Fit $N_U(t) = (1 - \tanh((t - a_0)/a_2))\,(a_1\ \sin(a_3 x + a_4) + a_5) + a_6.$

Conclusions

- Our **Split Reactive Brownian Dynamics** (SRBD) introduces a **time splitting error** but handles diffusion and reaction exactly.
- Efficiency was gained by never computing neighbor lists for particles, instead, **scheduling potential reactions** and only looking for neighbors when a reaction actually happens.
- The algorithm is best at higher densities and slow reactions (otherwise use FPKMC).
- How one handles diffusion (hops vs walk) and reactions (fixed grids, shifting grids, grid-free) microscopically affects the macroscopic behavior strongly if the cells are not sufficiently well-mixed (i.e., much smaller than chemical penetration depth).
- For resolved simulations we were able to match the macroscopic behavior between RDME, S-BD-RME, SRBD and FHD.
- Under what conditions can one use coarse-grained models like RDME/FHD and avoid tracking particles? Do the rates need to be renormalized and how?

References

[1] A. K. Bhattacharjee, K. Balakrishnan, A. L. Garcia, J. B. Bell, and A. Donev.
Fluctuating hydrodynamics of multispecies reactive mixtures.
J. Chem. Phys., 142(22):224107, 2015.

[2] Radek Erban and S Jonathan Chapman.
Stochastic modelling of reaction–diffusion processes: algorithms for bimolecular reactions.
Physical biology, 6(4):046001, 2009.

[3] Samuel A. Isaacson.
A convergent reaction-diffusion master equation.
The Journal of Chemical Physics, 139(5):–, 2013.

[4] A. Donev, V. V. Bulatov, T. Oppelstrup, G. H. Gilmer, B. Sadigh, and M. H. Kalos.
A First-Passage Kinetic Monte Carlo Algorithm for Complex Diffusion-Reaction Systems.
J. Comp. Phys., 229(9):3214–3236, 2010.

[5] Anton A. Winkler and Erwin Frey.
Validity of the law of mass action in three-dimensional coagulation processes.
Phys. Rev. Lett., 108:108301, 2012.

[6] Irina V Gopich, Alexander A Ovchinnikov, and Attila Szabo.
Long-time tails in the kinetics of reversible bimolecular reactions.
Phys. Rev. Lett., 86(5):922, 2001.

[7] Ariel Balter and Alexandre M Tartakovsky.
Multinomial diffusion equation.
Physical Review E, 83(6):061143, 2011.

[8] TaiJung Choi, Mano Ram Maurya, Daniel M Tartakovsky, and Shankar Subramaniam.
Stochastic operator-splitting method for reaction-diffusion systems.
The Journal of chemical physics, 137(18):184102, 2012.