Comparison of chemical reaction between ODE and probabilistic Gillespie simulations

Jere Arokivi Max Wesamaa Matti Aalto

December 2024

1 Introduction

For our practical assignment we chose Task 3. In the task 3 the idea is to simulate a chemical reaction. In this reaction there are 4 different substances A, B, C and D. The reactions simulated are A + B $\rightarrow$ C and B + C $\rightarrow$ D. The reaction can be modeled with deterministic ODE solutions, but since the experiment is done with a small initial population of molecules, the more informative way to inspect it is to use a probabilistic Gillespie simulation, otherwise known as a stochastic simulation algorithm. We will compare the results from the simulation to a deterministic solution.

Initial molecule populations are 100 for A and B and 0 for C and D. These initial conditions mean that the reaction B + C $\rightarrow$ D cannot begin from the start, but rather requires the first reaction to happen first. The reaction rates are $R_1 = k_1 \cdot A \cdot B$ for the first reaction and $R_2 = k_2 \cdot B \cdot C$ for the second, with rate constants of $k_1 = 0.0013$, and $k_2 = 0.01$.

To capture the stochastic nature of molecular interactions in small populations, we will implement the Gillespie algorithm. The algorithm accounts for random fluctuations and discrete events over time. In addition to this we will perform the simulations over larger populations to observe the behavior of the system in a situation where deterministic ODE models are more accurate. This allows for a better comparison between the stochastic and deterministic approaches.

The analysis will be conducted over a time interval sufficient for the reaction to reach its final state, that is, where all molecules are converted to substance D. Using both approaches will let us evaluate the strengths and weaknesses of both methods.

2 Mathematical Formulation and Methods

2.1 Deterministic Model

The deterministic ODE model is given by:

$$\begin{aligned}
\frac{dA}{dt} &= -k_1 AB, \\
\frac{dB}{dt} &= -k_1 AB - k_2 BC, \\
\frac{dC}{dt} &= k_1 AB - k_2 BC, \\
\frac{dD}{dt} &= k_2 BC,
\end{aligned}$$

with rate constants $k_1 = 0.0013$ and $k_2 = 0.01$. The system was integrated numerically over the time interval $0 \leq t \leq 35$ with a fixed step size $\Delta t = 0.01$.

2.2 Stochastic Model

To capture natural noise, we employed a stochastic simulation approach based on the Gillespie algorithm. For each time step, the expected number of reactions was calculated using the current state:

$$\lambda_1 = k_1 AB, \quad \lambda_2 = k_2 BC.$$

The number of events for each reaction was sampled from a Poisson distribution:

$$N_1 \sim \text{Poisson}(\lambda_1 \Delta t), \quad N_2 \sim \text{Poisson}(\lambda_2 \Delta t).$$

The molecular counts were then updated accordingly. If a negative population would result, the update was discarded for that step.

We ran 100 independent stochastic simulations to explore the distribution of possible trajectories. Each simulation was initiated from the same initial conditions and proceeded with the same $\Delta t = 0.01$ over $t = 0$ to $t = 35$.

3 Results

3.1 Temporal Evolution of Species Populations

Figure 1 presents the deterministic solution alongside the average and individual trajectories obtained from the stochastic simulations. The deterministic solution (dashed lines) provides a single, smooth trajectory for each species. In contrast, individual stochastic trajectories (faint lines) exhibit pronounced variability, particularly during the early portion of the simulation when the number of C and D molecules remains low. As the system evolves and reaction events accumulate, the stochastic trajectories gradually cluster, and the mean of these simulations (solid lines) approaches the deterministic solution.

These findings highlight the fundamental difference between the deterministic and stochastic representations. Whilst the deterministic solution conveys an expected trend, it does not reflect the variability caused by discrete molecular events. The stochastic simulations, however, capture a distribution of possible outcomes. By averaging over multiple stochastic realizations, we find that the mean trajectory aligns closely with the deterministic solution, confirming that the ODE model provides an accurate expectation value in this parameter regime.

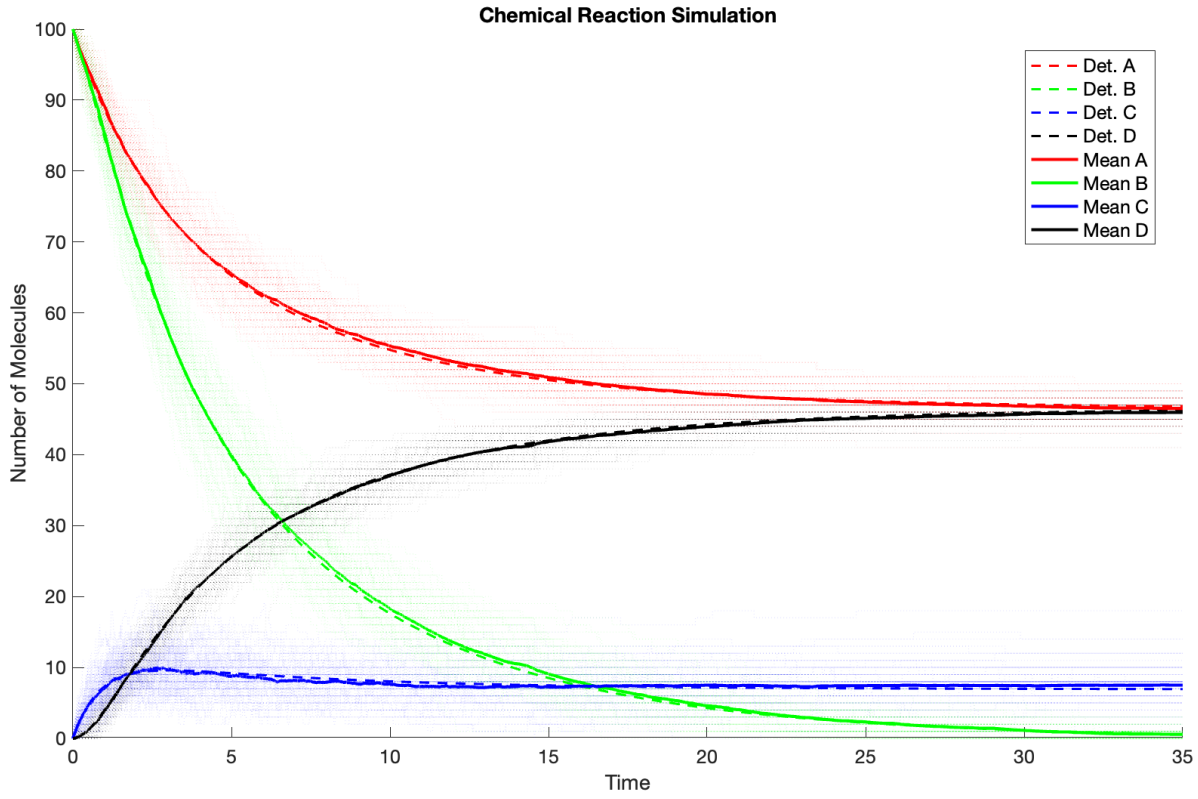

Figure 1: Comparison of the deterministic ODE solution (dashed lines) and the mean of 100 stochastic Gillespie simulations (solid lines) for species A (red), B (green), C (blue), and D (black). Faint lines represent individual stochastic trajectories.

3.2 Stochastic Variability in Final Population Counts

To better understand the long-term variability, we examined the final distributions of each species at $t = 70$ over 100 stochastic simulations. Figure 2 shows histograms of the end-state populations for A, B, C, and D.

The results indicate that both A and D display relatively narrow distributions at the final time point. These findings suggest that the ultimate abundances of A and D are less influenced by random fluctuations. The depletion of A early in the process, as well as the eventual accumulation of D, both reach relatively stable endpoints. Consequently, individual realizations of the stochastic process differ only slightly in the final counts of these species, resulting in low variance.

In contrast, the dynamics of B and C exhibit more pronounced stochastic variability. For B, the distribution is sharply peaked near zero, reflecting the rapid depletion of B as it participates in both reactions. This near-complete consumption occurs consistently across the simulations, resulting in a final distribution heavily skewed toward negligible quantities of B.

The species C presents the most substantial variability at $t = 70$. As an intermediate in the reaction network, C is continuously formed in the first step and consumed in the second. Even small stochastic differences early in the simulation can influence the balance between production and consumption, ultimately propagating into a broad range of final C counts. Hence, while the deterministic model might predict a single favored trajectory for C, the stochastic analysis reveals a broad ensemble of equally possible end states, emphasizing the importance of intrinsic noise in shaping long-term outcomes.

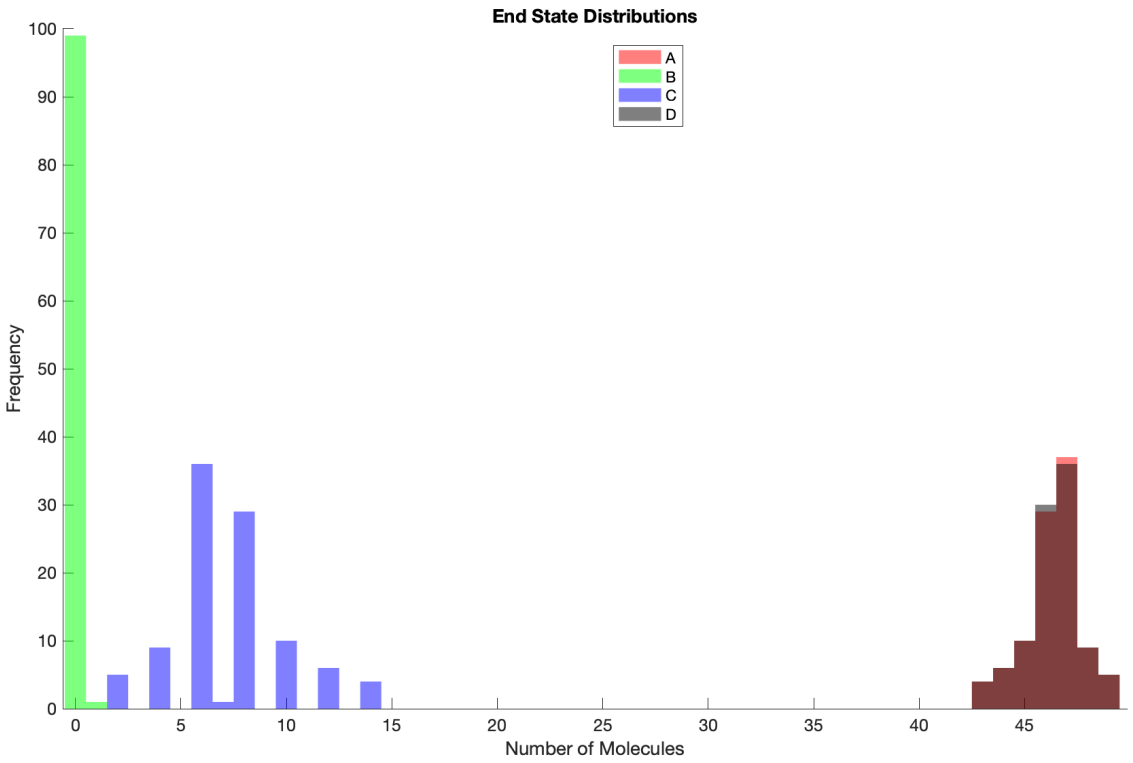

Figure 2: End-state distributions of the four species at $t = 70$ for 100 stochastic simulations. Species A and D show low variability, while B converges toward full depletion. The intermediate C exhibits the greatest spread, emphasizing the role of stochasticity in its final state.

3.3 Scaling Behavior with Increased Initial Populations

As the number of molecules increases from 100 to 1000, the underlying stochastic processes involve a larger number of reaction events. Since these events can often be modeled by Poisson distributions with larger means, we leverage the fact that a Poisson(λ) distribution becomes increasingly well-approximated by a normal distribution $N(\lambda, \lambda)$ as λ grows. This normal approximation implies that while the absolute variance grows, the relative variance decreases, causing fluctuations to average out more smoothly.

Consequently, the stochastic simulations converge more closely and more consistently to the deterministic solution. In other words, as the reaction counts scale up, the Central Limit Theorem ensures the distribution of reaction events becomes more 'normal-like' and relatively less noisy, thus aligning the stochastic predictions more tightly with the deterministic approximation.

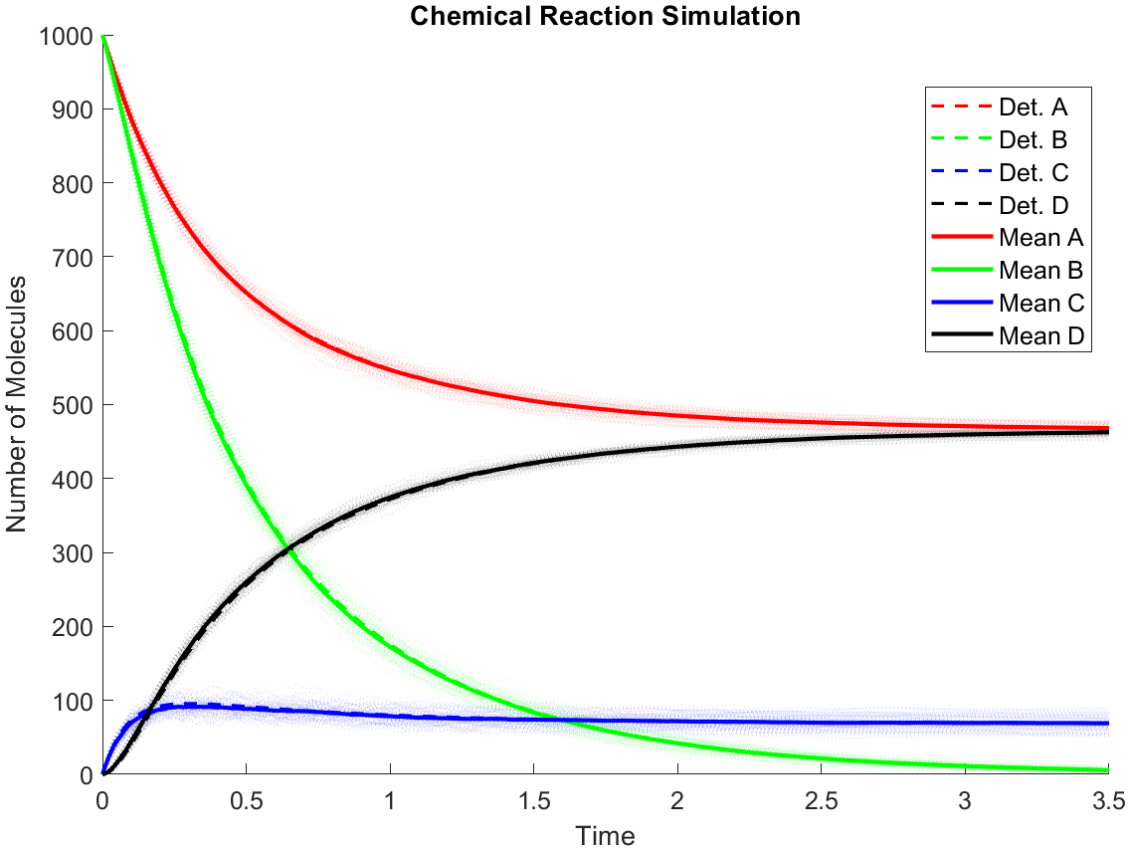

Figure 3: Comparison of deterministic (dashed lines) and stochastic simulations (solid lines) as initial populations increase from 100 to 1000. As the system scales up, relative fluctuations diminish, resulting in stochastic trajectories that closely follow the deterministic predictions.

4 Discussion and Conclusion

The comparison between deterministic ODE modeling and the stochastic Gillespie approach underscores the importance of method selection in analyzing chemical reaction systems. Deterministic models are computationally efficient and straightforward to interpret, making them valuable for large-population or steady-state conditions. However, their deterministic nature inherently overlooks the random fluctuations that can be critical when populations are small or reaction events are relatively rare.

Stochastic simulations, by incorporating randomness at the event level, yield a more deep understanding

of system behavior, including variance, probability distributions of outcomes, and the temporal dynamics of uncertainty. Such insights are particularly relevant in biochemical systems where molecular counts can be low and stochastic effects may influence critical processes

In conclusion, while the deterministic ODE approach provides a good baseline approximation, the stochastic Gillespie method offers a richer representation of the underlying reaction dynamics. For systems with small populations or where variability plays a central role, stochastic modeling is essential. Future investigations might include examining the sensitivity of outcomes to parameter changes.