The Gillespie Algorithm: Stochastic Modeling of Molecular Reaction Networks[*]

Jiayu Guo Sinan Ma Anna Szydlowski

2024-11-27

Abstract

The Gillespie Algorithm provides a framework for simulating stochastic molecular reaction networks, capturing the inherent randomness of chemical and biochemical processes. Unlike deterministic methods, the algorithm accurately models systems with low molecular counts, where fluctuations significantly influence reaction dynamics. This project explores the theoretical foundation of the Gillespie Algorithm, including its basis in the master equation, and demonstrates its application in simulating two molecular systems: the photodimerization of thymine, a process linked to UV-induced DNA damage, and the radioactive decay of isotopes. Through these simulations, the algorithm's capability to model stochastic trajectories is highlighted, offering insights into reaction mechanisms and the role of randomness in molecular processes. The implications of these findings extend to a variety of disciplines, including molecular biology, environmental monitoring, and nuclear medicine.

[*] We would like to thank Dr. Omidali A. Jazifor for his invaluable guidance throughout this project. The source code for simulations can be found on our GitHub repository.

Contents

1 Introduction **4**

2 Topic Review **4**

2.1 Background 4

3 Theory **5**

3.1 Master Equation 5

3.2 Gillespie Algorithm 6

4 Application **7**

4.1 Photodimerization Simulation 8

4.1.1 Observations 8

4.1.2 Biological and Physical Significance 10

4.2 Radioactive Decay Simulation 10

4.2.1 Observations 11

4.2.2 Biological and Physical Significance 13

5 Conclusion **14**

6 References **16**

A Appendix **17**

A.1 R Code for Photodimerization Simulation (section 4.1) 17

A.1.1 Simulation 17

A.1.2 Visulization for Figure 1 19

A.2 R code for Radioactive Decay Simulation (section 4.2) 19

A.2.1 Simulation 19

A.2.2 Visulization for Figure 2 21

A.2.3 Visulization for Figure 3 21

1 Introduction

The Gillespie Algorithm was presented in 1976 by Daniel T. Gillespie (Gillespie [2]). This algorithm was developed to answer the question, “If a fixed volume V contains a spatially uniform mixture of N chemical species which can inter-react through M specified chemical reaction channels, then given the numbers of molecules of each species present at some initial time, what will these molecular population levels be at any later time?” (Gillespie [3]). Given the trajectory of this molecular reaction is a stochastic process, this algorithm can simulate various molecular equation systems, such as radioactive decay models or decaying-dimerization reaction sets (Pineda-Krch and Cannoodt [10]).

Biochemical reactions in living cells are inherently stochastic. Reactions occur at random intervals, and the state of the system at any given time determines the probability of a particular chemical reaction occurring next. This stochasticity arises from the random thermal motion of molecules and the small numbers of reactant molecules involved. When molecule counts are low (often in the range of dozens or hundreds), random fluctuations in molecule collisions and interactions become significant, making reaction timing and outcomes probabilistic rather than deterministic. Consequently, deterministic approaches fail to accurately capture the dynamics of such systems, necessitating stochastic simulation methods like the Gillespie Algorithm.

In the following sections, we will first provide an overview of the Gillespie Algorithm and its historical development in the Section 2. Next, the theoretical underpinnings of the algorithm, including its basis in the master equation and stochastic modeling principles, will be detailed in the Section 3. We will then demonstrate the application of the Gillespie Algorithm to two molecular systems in the Section 4: the photodimerization of thymine, a DNA damage process induced by UV radiation (Martinez-Fernandez and Improta [7]), and the stochastic decay of radioactive isotopes (Mouret et al. [8]). These simulations will be performed using the R programming language (R Core Team [11]). Finally, our Section 5 will summarize the findings, discuss the algorithm’s broader implications, and explore potential future applications.

2 Topic Review

2.1 Background

In this project, we focus on the original version of the Gillespie Algorithm. This algorithm generates statistically correct trajectories to simulate chemical or biochemical reaction systems with high efficiency. By leveraging stochastic simulation, the Gillespie Algorithm enables researchers to model dynamics that cannot be solved analytically using the master equation, particularly for complex reaction networks. Its application is especially valuable in cases where deterministic methods fail to capture the inherent randomness of biochemical processes, such as systems with low molecular counts or highly interconnected reaction pathways.

3 Theory

3.1 Master Equation

Generally we use the distribution function $P(S,t)$ to describe a many particle system at time t, where $P(S,t)$ is the probability for the system to be in the state S at time t.

The equation governing the evolution of $P(S,t)$ is called the master equation. During the time interval $[t, t+dt)$, $P(S,t)$ is changed due to two types of process:

1. The particles in the state S leave due to some reaction $S \to S'$.
2. The particles in other state S' enter the state S state due to some reaction $S' \to S$.

Combining the two factors, we obtain the master equation with the form (Kampen [5])

$$\frac{\partial}{\partial t} P(S,t) = \sum_{S'} P(S',t) R(S',S) - \sum_{S'} P(S,t) R(S,S')$$

where $R(S,S')$ represents the reaction rate from the state S to the state S'.

The idea of master equation is inherently stochastic. Firstly $P(S,t)$ itself is a probabilistic description of the system. Next, in most of the cases, the reaction/process is stochastic. Therefore the master equation describes the ensemble average of all possible systems, here `ensemble` is the collection of all possible states with the same initial condition and macroscopic parameters. Following some certain initial distribution, we are not guaranteed to reach the same distribution at time t.

In the article *Cellular growth and division in the Gillespie algorithm* (Lu et al. [6]), the authors discussed the system of well-stirred mixture of N chemical species. Since the system is well-stirred, the distribution of particles is always uniform in space. The state of the system can be described using a N-vector $Y = (X_1, ..., X_N)$, where X_i is the number of molecules of type i in the system. Assume there are $\mu = 1, ..., M$ elementary reaction channels R_μ. Let c_μ be the reaction rate, i.e., the probability that a random combination of molecules from channel R_μ selected at the moment t react in the interval $[t, t+dt)$ with probability $c_\mu dt + o(dt)$. Let $h_\mu(Y)$ be the total number of possible distinct combinations of molecules for a channel R_μ when the system is in state Y, and $\alpha_\mu = (\alpha_{1,\mu}, ..., \alpha_{N,\mu})$ is a constant stoichiometric vector prescribing the change in the state of the system after the reaction R_μ has occurred.

For example, in a second order reaction $A + B \to AB$, we have

$$\frac{d}{dt}X_{AB} = V\frac{d}{dt}[AB] = V * c'[A][B] = \frac{c'}{V} * X_A X_B = c * h(Y)$$

where c' is the true reaction constant measured in chemistry, V is the volume of the system and $[i] = \frac{X_i}{V}$ is the number concentration of species i.

With the above setting-up, The master equation for this system becomes

$$\frac{\partial}{\partial t}P(Y, t|Y_0, t_0) = \sum_{\mu=1}^{M} c_\mu h_\mu(Y - \alpha_\mu)P(Y - \alpha_\mu, t|Y_0, t_0) - \sum_{\mu=1}^{M} c_\mu h_\mu(Y)P(Y, t|Y_0, t_0)$$

here $P(Y, t|Y_0, t_0)$ is the probability of the system being in state Y at time t given it is in state Y_0 at time t_0.

3.2 Gillespie Algorithm

The master equation is a coupled linear ordinary equation for $|S_Y|$ dependent variables $P(Y, t|Y_0, t_0)$, where $|S_Y|$ is the size of the state space. For a system with n particles, $|S_Y|$ is of order $O(n^N)$. Therefore it is almost impossible to solve the equation analytically, especially when the number of possible elementary reactions M is large.

A practical alternative to solve the master equation is the stochastic simulation approach devised by Gillespie (Gillespie [3]). Using the idea that $P(Y, t|Y_0, t_0)$ describes the ensemble average of all possible micro-states given initial state Y_0 at time t_0, the Gillespie algorithm generates an ensemble of sample trajectories of the system with statistics which asymptotically converges to the solution of the corresponding Master Equation.

In the Gillespie algorithm, we update the state of the system by determining

(i) the time τ to the next reaction

(ii) which reaction R_μ will occur next.

Assuming that each reaction R_μ is independent with rate $a_\mu = h_\mu c_\mu$, the first occurrence τ_μ of the reaction R_μ follows the exponential distribution with rate/propensity a_μ. The first occurrence time of some reaction is then $\tau = \min_\mu \tau_\mu$, it follows the exponential distribution with an overall rate/propensity $A = \sum_\mu a_\mu$. Given that a reaction occurs, the probability that it is the μ'-th reaction is $P(\mu = \mu') = \frac{a_{\mu'}}{A}$. Therefore the the original Gillespie recipe follows the following steps:

1) Input values c_μ, $\mu = 1, ..., M$ and initial state $Y_0 = (x_1, ..., x_N)$ at t_0.

2) Compute the current propensities $a_\mu = h_\mu(Y)c_\mu$, $\mu = 1, ..., M$ and $A = \sum_\mu a_\mu$

3) Generate uniform random numbers $u_1, u_2 \in [0, 1)$

4) Compute the time interval τ until the next reaction according to distribution $\exp(A)$, i.e., $\tau = -\ln u_1/A$

5) Find the channel of the next reaction μ, i.e., take μ to be the integer for which $\sum_{\nu=1}^{\mu-1} a_\nu < u_2 * A \leq \sum_{\nu=1}^{\mu} a_\nu$

6) Update time $t \to t + \tau$, and adjust Y in accordance with the particular reaction R_μ, i.e., update $Y \to Y + a_\mu$, and proceed to step 2 until t reaches presetting terminal time T_f and obtain a final state Y_f at T_f.

After generating sufficiently many N_{Traj} such trajectories, we can estimate $P(Y, T_f|Y_0, t_0) = \frac{\#Y_{f,Traj}=Y}{N_{Traj}}$.

The Gillespie algorithm is straightforward and easy to implement. However, it is only an approximate rather than an exact stochastic algorithm. A very large number of trajectories is needed to reach a reasonable accuracy. The complexity of the algorithm is very high and the accuracy is not guaranteed due to randomness.

4 Application

This section explores the practical application of the Gillespie Algorithm by simulating two molecular systems. The first simulation, Photodimerization Simulation, models the photodimerization of thymine, a molecular reaction relevant to DNA damage caused by ultraviolet (UV) radiation. The second simulation, Radioactive Decay Simulation, examines the stochastic decay of radioactive isotopes, a fundamental process in nuclear physics. Both simulations reference the `GillespieSSA` package (Pineda-Krch and Cannoodt [10]), a R implementation of the Gillespie Algorithm that facilitates the generation of statistically accurate trajectories for molecular reaction systems. By consulting this package, we can simulate complex stochastic processes and explore the dynamics of these reaction systems.

4.1 Photodimerization Simulation

We can simulate the dimerization of DNA pyrimidines, a molecular reaction which can lead to skin cancer. Bipyrimidine photodimerization occurs when UV radiation alters the chemical bonds of two consecutive pyrimidine bases, with two possible outcomes: the formation of the cyclobutene pyrimidine dimer (CPD) if a cycloaddition between the C5-C6 double bond of the pyrimidines occurs, or the formation of pyrimidine (6-4) pyrimidone (64-PP) if a covalent bond forms between C4 and C6 of the pyrimidines (Martinez-Fernandez and Improta [7]).

Therefore, this simulation involves four species: pyrimidine monomers Thymine 1 (T_1), Thymine 2 (T_2), and dimer photoproducts CPD and $64PP$; and two reactions channels: c_1 and c_2.

$$T_1 + T_2 \xrightarrow{c_1} CPD$$

$$T_1 + T_2 \xrightarrow{c_2} 64PP$$

We will simulate the dimerization of thymine using propensity based on the formation yield of CPD (~37 per 10^6 normal bases) and 64-PP (~2 per 10^6 normal bases) found by exposing skin cells to a UVB dose of 0.2 J/cm^2 (Mouret et al. [8]). As the amount of thymine bases in DNA varies from person to person, we will simulate thymine photodimerization of 1000 pairs of thymine.

4.1.1 Observations

As shown in Figure 1, both CPD and 64-PP formation follow logarithmic growth, which aligns with expectations based on specific UV radiation conditions and previously reported quantum yields (Martinez-Fernandez and Improta [7]; Mouret et al. [8]). The number of reactions per unit time levels off as pyrimidine pairs decrease, which is consistent with the propensity functions, as the propensity of 64-PP formation is much significantly than the propensity of CPD formation (Mouret et al. [8]). This plateau suggests that, as the number of available reactants decreases, the rate of reaction slows down. Due to the stochastic nature of molecular reactions, this simulation demonstrates how the Gillespie algorithm reflects fluctuations of photodimerization as it models the probabilistic process of each individual reaction event. These fluctuations are especially important in low-molecule systems, where randomness significantly impacts the overall reaction dynamics, providing a more accurate representation of the molecular behavior compared to deterministic models.

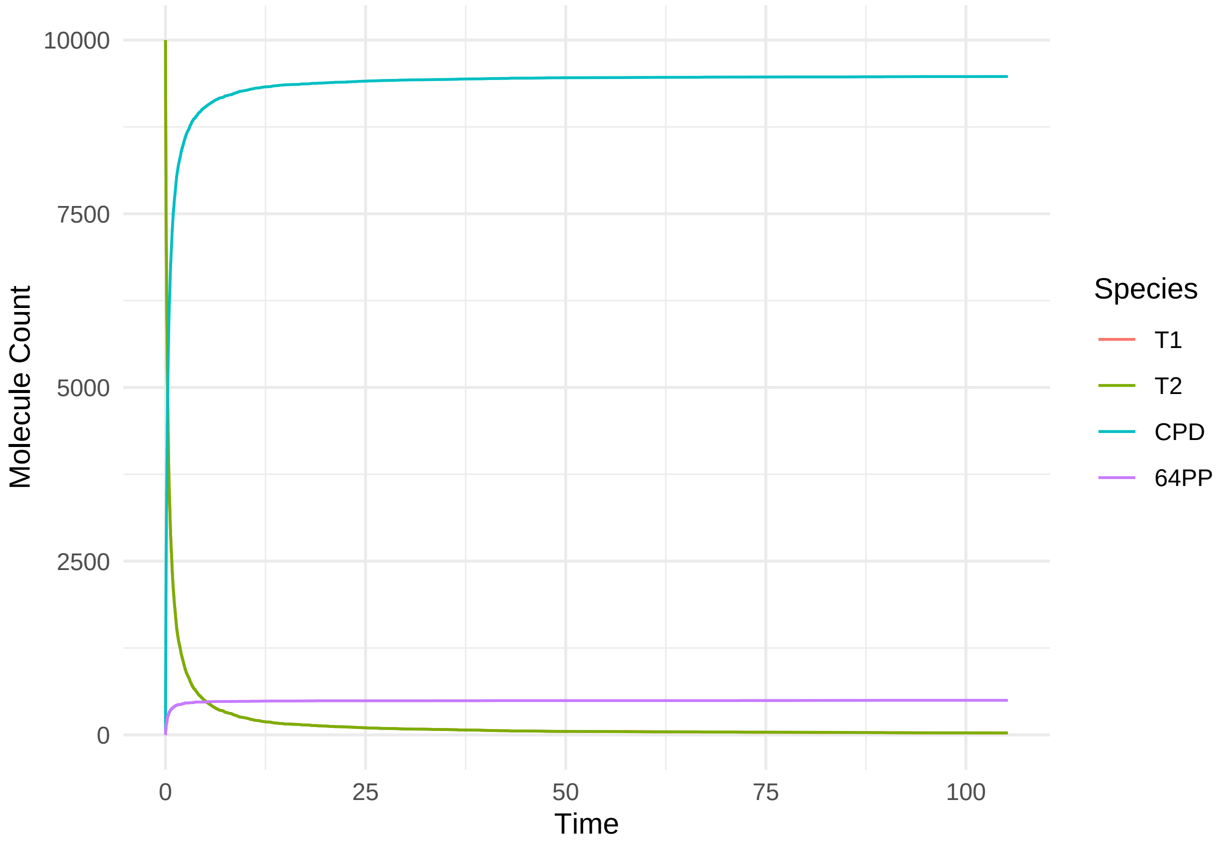

Figure 1: Thymine Photodimerization Reaction Set Simulation

4.1.2 Biological and Physical Significance

4.1.2.1 Dimerization As pyrimidine photodimerization transforms two pyrimidine monomers into a dimer, this simulation has accurately represented the formation of two possible dimers, CPD and 64-PP, from pairs of thymine monomers. Dimerization is also known to occur between pairs of cytosine, as well as between thymine and cytosine, with varying rates dependent on the monomers involved as well as the level of UV radiation (Mouret et al. [8]).

4.1.2.2 Implications and Real-World Applications The Gillespie algorithm offers profound implications for understanding molecular reactions, particularly in systems where stochastic effects are prominent, such as the dimerization. This process, highlighted in Figure 1, is widely studied, as it can lead to mutations potentially resulting in skin cancer, when not repaired. The Gillespie algorithm allows us to realistically model these reactions, as it demonstrates the probabilistic nature of DNA damage, capturing fluctuations in molecular interactions that deterministic models cannot, providing a deeper understanding of molecular events, insights into mutation rates, and the impact of environmental factors like UV radiation on genetic integrity. The algorithm's ability to model these stochastic processes is crucial in areas such as genetics, where random molecular events can significantly influence disease outcomes. This algorithm can be further adjusted to show molecular reactions with various molecular species and reaction channels. Using the Gillespie algorithm to model molecular reactions has wide-ranging implications in fields such as genetics, cancer research, pharmacology, environmental science, and biotechnology. By capturing the inherent randomness in molecular processes, this approach can provide valuable insights into complex biological systems, enabling better prediction of outcomes, such as in disease development, medicine, and environmental impact studies.

4.2 Radioactive Decay Simulation

This simulation models the stochastic decay of a single species (R), which represents a radioactive substance such as radon, radium, or plutonium. The decay process is governed by the following reaction:

$$R \xrightarrow{c} 0$$

where c is the decay rate constant that defines the likelihood of decay per unit time.

The Gillespie algorithm is used to simulate this system, capturing the inherent randomness of radioactive decay events. Unlike deterministic approaches, the Gillespie algorithm tracks individual decay events over time, providing an accurate representation of the stochastic nature of the process. The goal of this simulation is to observe the time evolution of the remaining radioactive molecules (R) and visualize their decay over time.

4.2.1 Observations

The simulation results are presented in two figures, each offering unique insights into the behavior of the decay process. Figure 2 displays the decay on a linear scale, illustrating the overall exponential decline of the radioactive substance over time. Figure 3 presents the same data on a logarithmic scale, highlighting the exponential nature of the process by linearizing the decay trajectory.

4.2.1.1 General Explanation of Exponential Decay The radioactive decay of R follows the well-known exponential decay law:

$$R(t) = R_0 e^{-ct}$$

where R_0 is the initial quantity of the substance, c is the decay constant, and t is time (Henriksen and Maillie [4]). As R decreases, the propensity function $a(R) = c \cdot R$ ensures that the decay rate becomes slower over time. This results in a high number of decay events at the beginning of the simulation, which gradually diminishes as fewer molecules remain.

While the decay trajectory conforms to deterministic exponential decay theory, the Gillespie algorithm introduces stochastic variations. These arise because the time between successive reactions is governed by random exponential waiting times, and the occurrence of each reaction is probabilistic. As the number of R molecules declines, stochastic fluctuations become more apparent due to the reduced scale of the system.

4.2.1.2 Analysis of Radioactive Decay Simulation in Linear Scale (Figure 2) Figure 2 illustrates the exponential decay of R over time. The steep initial decline in the curve corresponds to the high decay rate at the beginning of the simulation when R is at its maximum. As R decreases, the curve flattens, reflecting the slower decay rate. This behavior exemplifies the exponential decay dynamics, where the number of remaining molecules reduces rapidly at first and progressively lessens as the substance depletes.

The gradual flattening of the curve also underscores the diminishing probability of decay events as time progresses. This effect is a direct result of the propensity function $a(R) = c \cdot R$, which decreases linearly with R. The linear-scale plot effectively illustrates the overall trend of radioactive decay over time, making it particularly suitable for understanding the macroscopic behavior of the process.

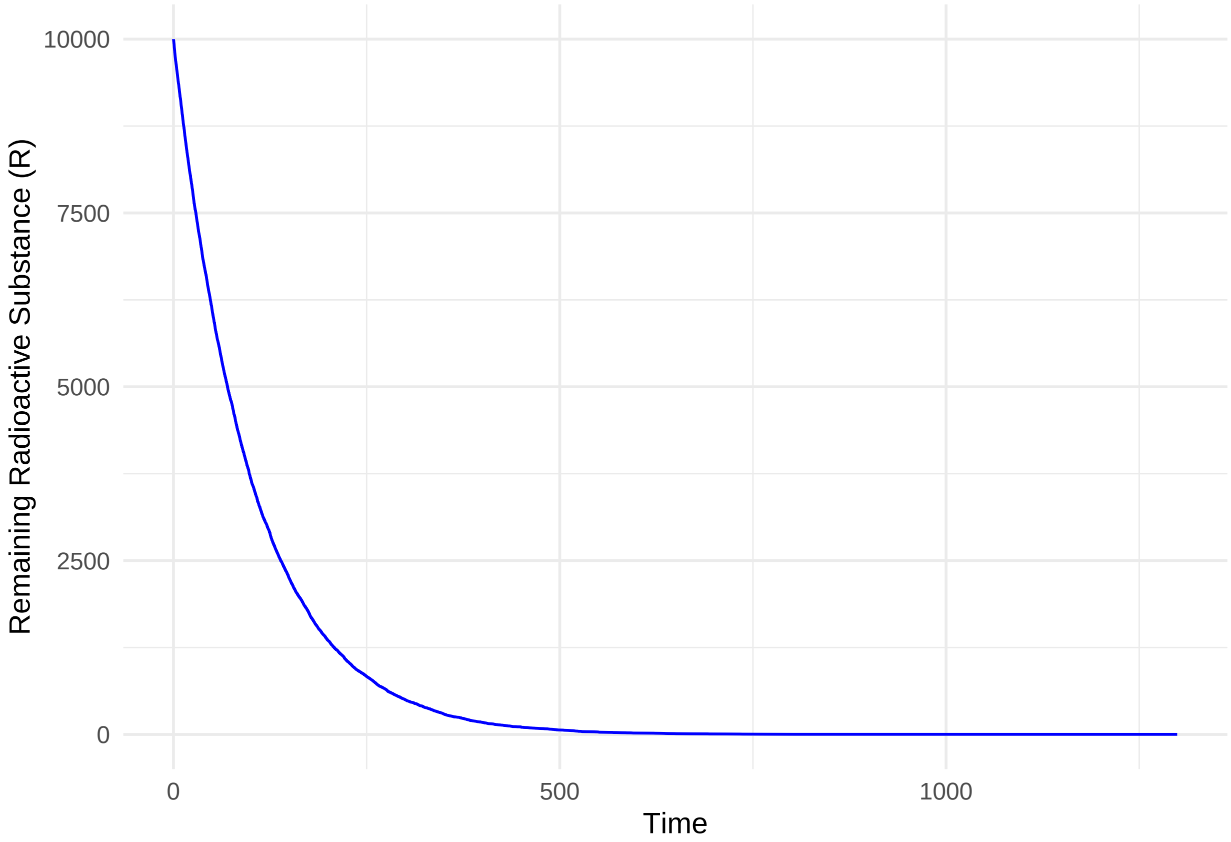

Figure 2: Radioactive Decay Simulation in Linear Scale

4.2.1.3 Analysis of Radioactive Decay Simulation in Logarithmic Scale (Figure 3) Figure 3 depicts the same decay trajectory on a logarithmic scale, offering a complementary perspective. On this scale, the exponential decay law is linearized, and the trajectory appears as a straight line. The slope of the line corresponds to the negative decay constant, $-c$, providing a quantitative validation of the theoretical model.

The logarithmic representation also highlights deviations from perfect linearity, which become more noticeable as the molecule count approaches zero. These deviations are the result of stochastic fluctuations inherent in the Gillespie algorithm, particularly in the later stages when R is low, and individual decay events have a greater impact on the overall trend. This logarithmic plot is especially useful for analyzing the detailed kinetics of the decay process and for confirming the exponential nature of the decay.

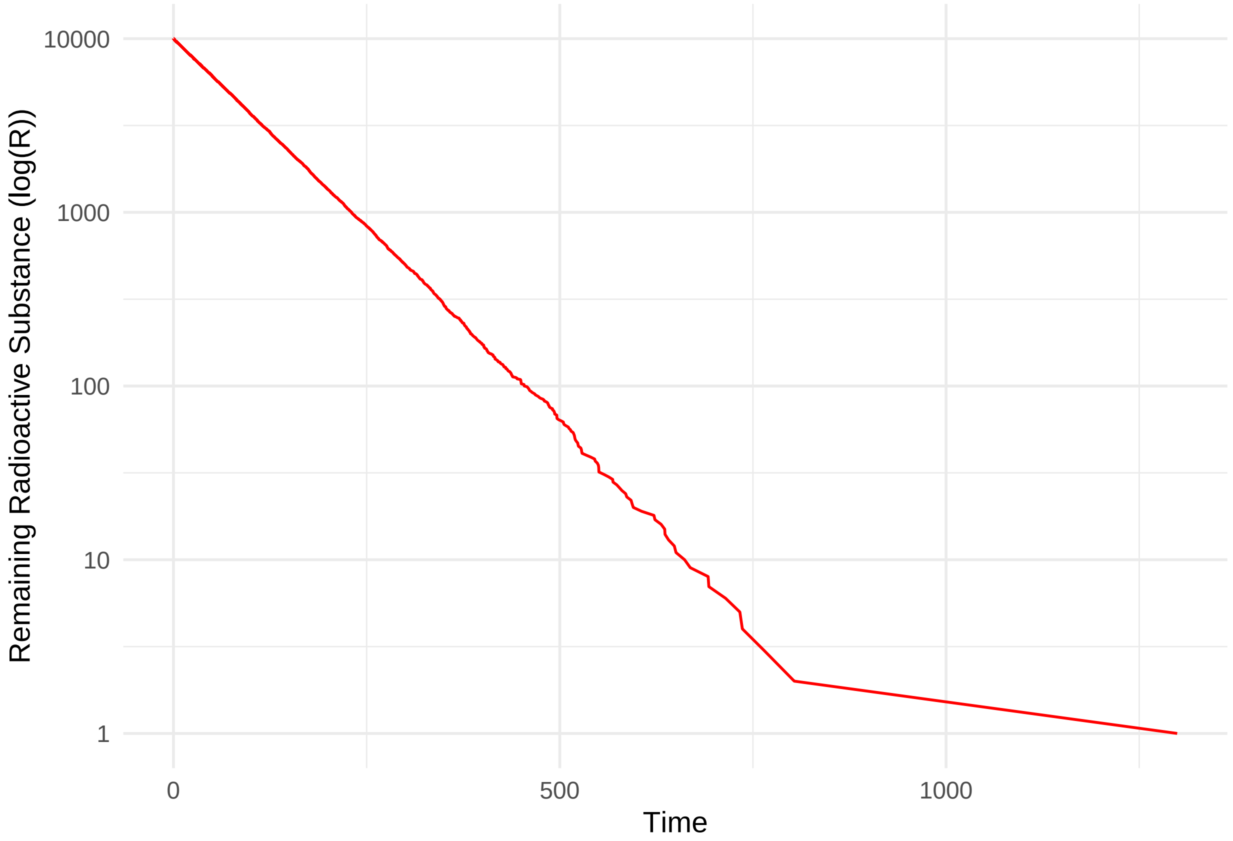

Figure 3: Radioactive Decay Simulation in Logarithmic Scale

4.2.2 Biological and Physical Significance

4.2.2.1 Radioactive Half-Life The exponential decay observed in the simulation aligns with the concept of half-life, a critical measure in the study of radioactive processes. The half-life ($t_{1/2}$) is defined as the time required for half of the radioactive substance to decay (Henriksen and Maillie [4]). It can be calculated as:

$$t_{1/2} = \frac{\ln(2)}{c}$$

For the decay constant $c = 0.01$ used in this simulation, the half-life is approximately $t_{1/2} \approx 69.3$ time units. This simulation demonstrates how the decay rate slows down over time as the radioactive substance depletes, with the time between successive decay events increasing in proportion to the decreasing molecule count.

4.2.2.2 Implications and Real-World Applications Radioactive decay is a fundamental phenomenon in nuclear physics and has wide-ranging applications across multiple scientific disciplines. For instance, in nuclear medicine, the controlled decay of radioactive isotopes is utilized in radiotherapy to deliver targeted doses of radiation for cancer treatment (Britannica [1]). Similarly, in environmental science, the decay of isotopes such as carbon-14 is monitored to assess contamination levels and potential health risks (OpenMedScience [9]).

The stochastic approach modeled here provides insights into systems with low molecule counts, where randomness plays a significant role. For example, in radiological dosimetry, small-scale decay dynamics can impact the precision of dose delivery. In environmental monitoring, understanding the stochastic nature of decay processes aids in the interpretation of measurement data, particularly in situations where low concentrations of isotopes are involved.

In the broader context of molecular reaction networks, the Gillespie algorithm's ability to capture both macroscopic trends and stochastic variations makes it an invaluable tool for studying systems with inherent randomness. By accurately simulating individual reaction events, it bridges the gap between theoretical models and real-world observations, offering a deeper understanding of complex kinetic processes.

5 Conclusion

The Gillespie Algorithm represents a significant role in the numerical simulation of stochastic molecular reaction networks. Its ability to generate statistically trajectories of reaction systems provides insights into processes governed by inherent randomness, especially in systems with low molecule counts. This project has illustrated the versatility and power of the algorithm through two applications: the photodimerization of thymine and the radioactive decay of isotopes.

The thymine photodimerization simulation illustrated the algorithm's capacity to model complex reaction networks, offering a detailed view of the formation of UV-induced DNA lesions such as CPDs and 64-PPs. This has direct implications for understanding the molecular origins of UV-induced mutagenesis and skin cancer. Similarly, the radioactive decay simulation demonstrated the algorithm's precision in modeling exponential decay processes and capturing stochastic fluctuations, a crucial feature for understanding radioactive isotopes in fields such as nuclear medicine and environmental science.

While the Gillespie Algorithm excels at modeling stochastic dynamics, it is not without limitations. The computational cost of simulating systems with large molecule counts or numerous reaction channels can become prohibitive, and its reliance on random number generation introduces variability that requires a sufficiently large number of trajectories for robust statistical analysis. Nevertheless, its utility in studying molecular systems with inherent randomness cannot be overstated.

The implications of this work extend beyond the specific applications discussed. By providing a framework for understanding and simulating stochastic reaction systems, the Gillespie Algorithm continues to play a pivotal role in advancing our knowledge of molecular biology, chemical kinetics, and physical chemistry. Future developments in computational efficiency and hybrid methods may further expand its applicability, enabling the simulation of even more complex systems with greater precision.

6 References

[1] Encyclopaedia Britannica. *Applications of Radioactivity*. Accessed: 2024-11-27. n.d. URL: https://www.britannica.com/science/radioactivity/Applications-of-radioactivity.

[2] D. T. Gillespie. “A general method for numerically simulating the stochastic time evolution of coupled chemical reactions”. In: *Journal of Computational Physics* 22.4 (1976), pp. 403–434. DOI: 10.1016/0021-9991(76)90041-3.

[3] D. T. Gillespie. “Exact Stochastic Simulation of Coupled Chemical Reactions”. In: *The Journal of Physical Chemistry* 81.25 (1977), pp. 2340–2361. DOI: 10.1021/j100540a008. URL: https://doi.org/10.1021/j100540a008.

[4] T. Henriksen and D. H. Maillie. “Radioactive Decay Laws”. In: *Radiation and Health*. United States: CRC Press, 2002, pp. 21–26. DOI: 10.4324/9780203166352-8.

[5] N. G. Van Kampen. *Stochastic Processes in Physics and Chemistry*. Amsterdam: North-Holland, 1992.

[6] T. Lu et al. “Cellular growth and division in the Gillespie algorithm”. In: *Syst Biol (Stevenage)* 1.1 (June 2004), pp. 121–128.

[7] L. Martinez-Fernandez and R. Improta. “Photoinduced Dimerization in DNA: Singlet Excited State Reaction Paths”. In: *DNA Photodamage: From Light Absorption to Cellular Responses and Skin Cancer*. Royal Society of Chemistry, 2022, pp. 17–36. ISBN: 9781839164017. DOI: 10.1039/9781839165703-00017. URL: https://hdl.handle.net/20.500.14243/499141.

[8] S. Mouret et al. “Cyclobutane pyrimidine dimers are predominant DNA lesions in whole human skin exposed to UVA radiation”. In: *Proceedings of the National Academy of Sciences* 103.37 (2006), pp. 13765–13770. DOI: 10.1073/pnas.0604213103.

[9] OpenMedScience. *Understanding the Radioactive Decay of Carbon-14 and Its Applications*. Accessed: 2024-11-27. n.d. URL: https://openmedscience.com/understanding-the-radioactive-decay-of-carbon-14-and-its-applications/.

[10] M. Pineda-Krch and R. Cannoodt. *GillespieSSA: Gillespie’s Stochastic Simulation Algorithm (SSA)*. R package version 0.6.2. 2022. URL: https://cran.r-project.org/package=GillespieSSA.

[11] R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. Vienna, Austria, 2023. URL: https://www.R-project.org/.

A Appendix

A.1 R Code for Photodimerization Simulation (section 4.1)

A.1.1 Simulation

```
# Work environment set-up
library(reshape2)
library(ggplot2)
# Define reaction rates
c1 <- (37 / 1e6) * 10             # Formation of CPD
c2 <- (2 / 1e6) * 10              # Formation of 64-PP
# Initial state
state <- c(T1 = 10000, T2 = 10000, CPD = 0, `64PP` = 0)
# Define propensity functions
propensities <- function(state) {
  c(
    c1 * state["T1"] * state["T2"], # Reaction 1
    c2 * state["T1"] * state["T2"]  # Reaction 2
  )
}
# Gillespie simulation function
simulate_gillespie <- function(state, max_time) {
  # Initialize time and output data frame
  time <- 0
  output <- data.frame(
    Time = numeric(0),
    T1 = numeric(0),
    T2 = numeric(0),
    CPD = numeric(0),
    `64PP` = numeric(0)
  )
```

```
# Record the initial state
initial_row <- c(Time = time, state)
output <- rbind(output, as.data.frame(t(initial_row), stringsAsFactors = FALSE))

while (time < max_time) {
  # Calculate propensities
  props <- propensities(state)
  total_prop <- sum(props)

  # Stop if no reactions can occur
  if (total_prop == 0) break

  # Time to next reaction
  delta_t <- rexp(1, rate = total_prop)
  time <- time + delta_t

  # Select which reaction occurs
  reaction <- sample(length(props), 1, prob = props)

  # Update state based on reaction
  if (reaction == 1) {
    state["T1"] <- state["T1"] - 1
    state["T2"] <- state["T2"] - 1
    state["CPD"] <- state["CPD"] + 1
  }
  if (reaction == 2) {
    state["T1"] <- state["T1"] - 1
    state["T2"] <- state["T2"] - 1
    state["64PP"] <- state["64PP"] + 1
  }

  # Record the new state
```

```
    new_row <- c(Time = time, state)
    output <- rbind(output, as.data.frame(t(new_row), stringsAsFactors = FALSE))
  }

  # Convert columns to numeric for consistency
  output[] <- lapply(output, as.numeric)
  return(output)
}
```

A.1.2 Visulization for Figure 1

```
# Run the simulation
max_time <- 100
simulation <- simulate_gillespie(state, max_time)

# Reshape data for plotting
simulation_long <- melt(simulation, id.vars = "Time")
# Plot results
ggplot(simulation_long, aes(x = Time, y = value, color = variable)) +
  geom_line() +
  labs(
    x = "Time",
    y = "Molecule Count",
    color = "Species"
  ) +
  theme_minimal()
```

A.2 R code for Radioactive Decay Simulation (section 4.2)

A.2.1 Simulation

```r
# Define parameters
c <- 0.01  # Decay rate constant
state <- c(R = 10000)  # Initial radioactive substance
max_time <- 1000

# Define propensity function
propensities <- function(state) {
  c * state["R"]
}

# Gillespie simulation function
simulate_gillespie <- function(state, max_time) {
  time <- 0
  output <- data.frame(Time = numeric(0), R = numeric(0))
  output <- rbind(output, data.frame(Time = time, R = state["R"]))

  while (time < max_time) {
    # Calculate propensity
    prop <- propensities(state)
    if (prop == 0) break  # Stop if no reactions can occur

    # Time to next reaction
    delta_t <- rexp(1, rate = prop)
    time <- time + delta_t

    # Update state (decay reaction)
    state["R"] <- state["R"] - 1

    # Record the state
    output <- rbind(output, data.frame(Time = time, R = state["R"]))
  }
  return(output)
}

# Run the simulation
simulation <- simulate_gillespie(state, max_time)
```

A.2.2 Visulization for Figure 2

```r
# Load ggplot2 for visualization
library(ggplot2)

# Linear-scale plot
ggplot(simulation, aes(x = Time, y = R)) +
  geom_line(color = "blue") +
  labs(
    x = "Time",
    y = "Remaining Radioactive Substance (R)"
  ) +
  theme_minimal()
```

A.2.3 Visulization for Figure 3

```r
# Logarithmic-scale plot
ggplot(simulation, aes(x = Time, y = R)) +
  geom_line(color = "red") +
  scale_y_log10() +
  labs(
    x = "Time",
    y = "Remaining Radioactive Substance (log(R))"
  ) +
  theme_minimal()
```