Feasibility and spatial heterogeneity of lineage conversion processes

Carlos Mendonca[1] Shloka Janapaty[1] Michael Tulskikh[1]

[1]Department of Applied Mathematics,
Columbia University

Dec 11, 2023

Who are we?

Carlos Mendonca

- CC
- Interested in developmental and molecular embryogenesis.

Shloka Janapaty

- SEAS
- Interested in theoretical ecology, biogeochemistry, and dynamical systems.

Michael Tulskikh

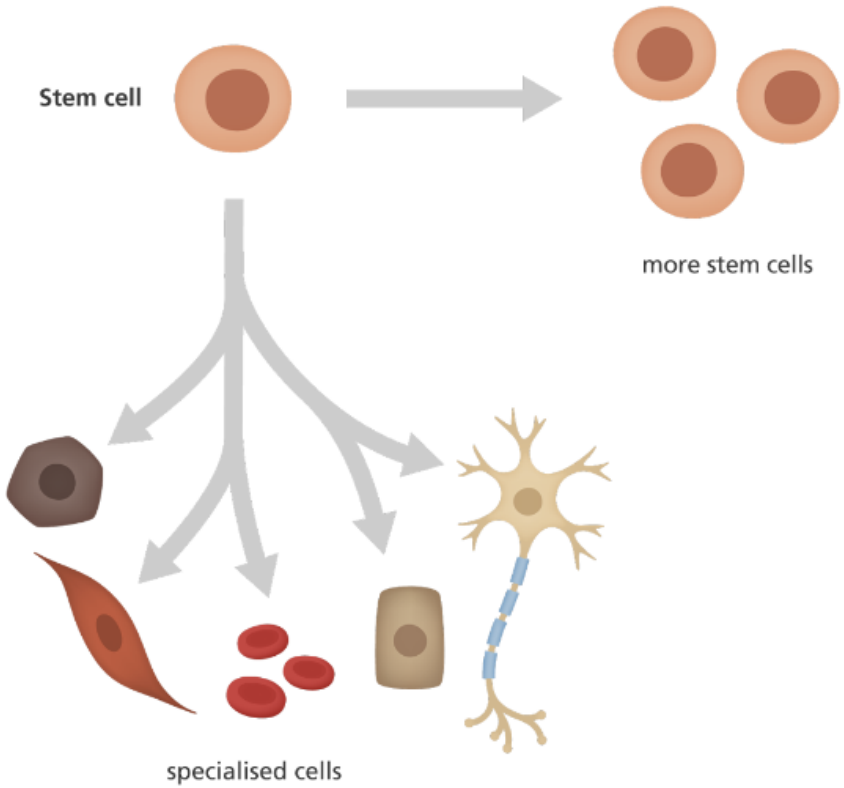
- SEAS
- Interested in PDEs and differential geometry.

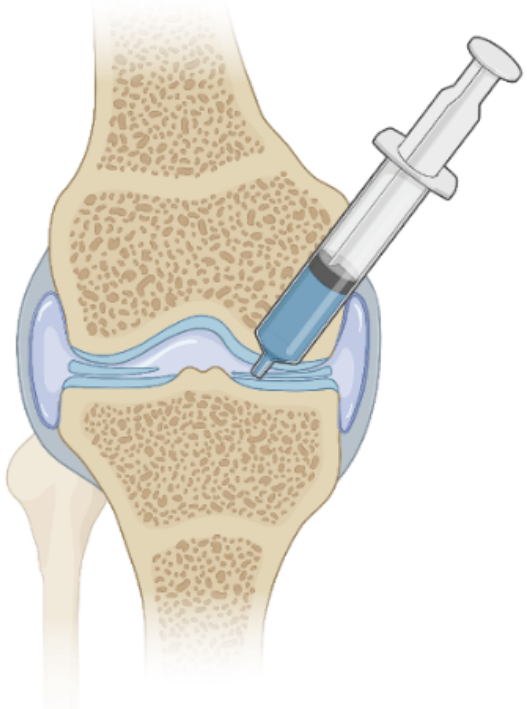
Primary References

1. Joo, J. I., Zhou, J. X., Huang, S., & Cho, K. H. (2018). **Determining relative dynamic stability of cell states using boolean network model.** *Nature Scientific Reports*, 8(1), 12077.
2. Gonze, D., Lahti, L., Raes, J., Faust, K. (2017). **Multi-stability and the origin of microbial community types.** *The ISME journal*, 11(10), 2159-2166.
3. Lu, M., & Hedin, L. O. (2019). **Global plant–symbiont organization and emergence of biogeochemical cycles resolved by evolution-based trait modelling.** *Nature Ecology & Evolution*, 3(2), 239-250.

What is a Stem Cell?

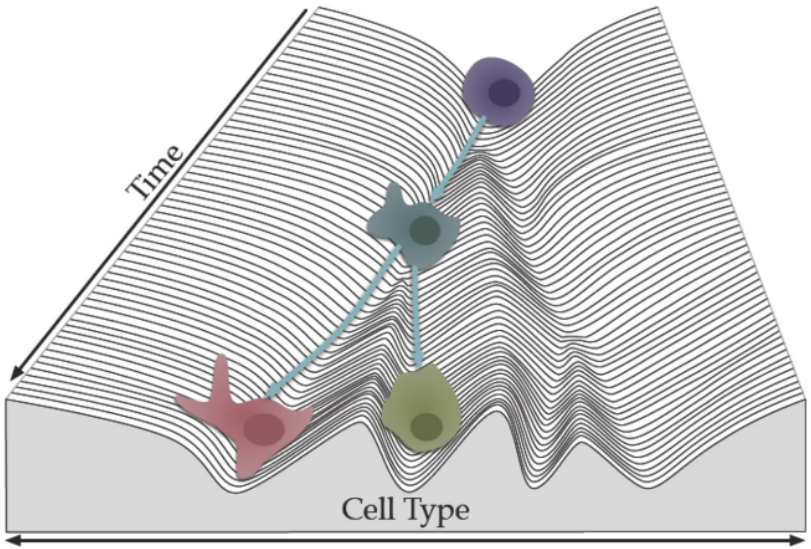
- Ability to self-renew and proliferate
- May differentiate into specialized cell types

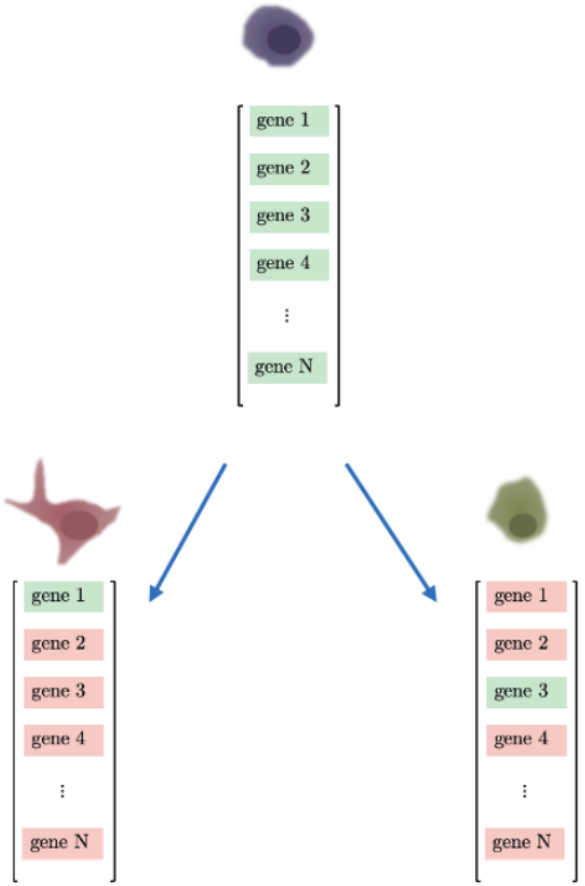




Effective therapy for cartilage repair in osteoarthritis

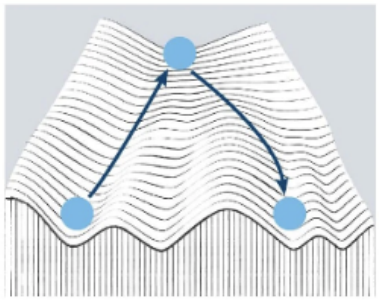
Quantitative description of differentiation

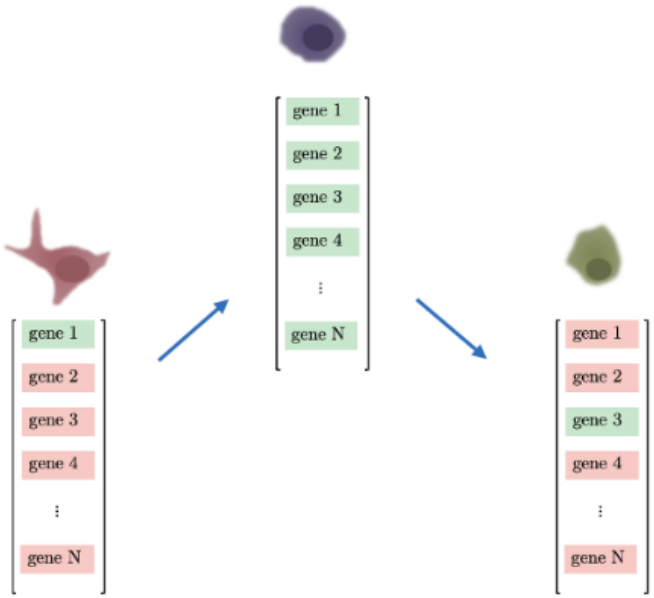




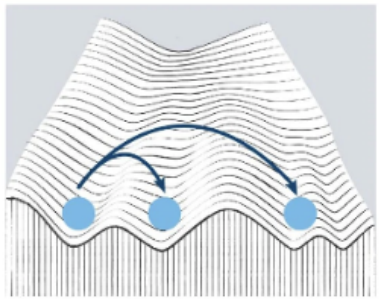
Bypassing pluripotency

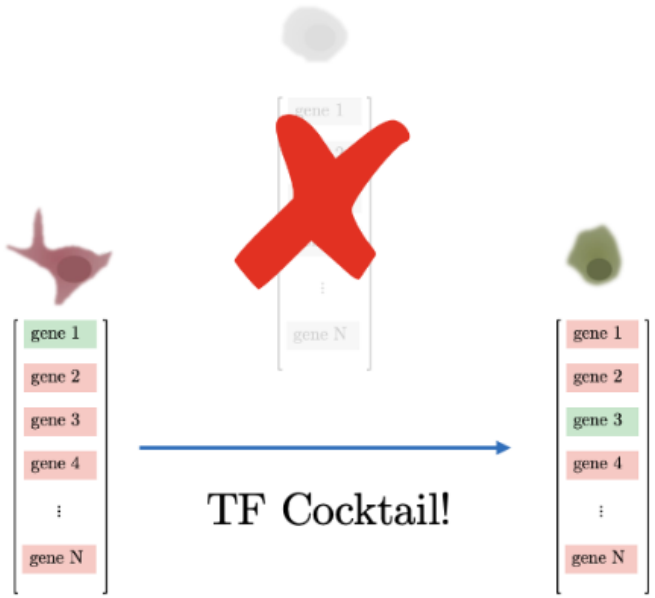
Pluripotent Reprogramming



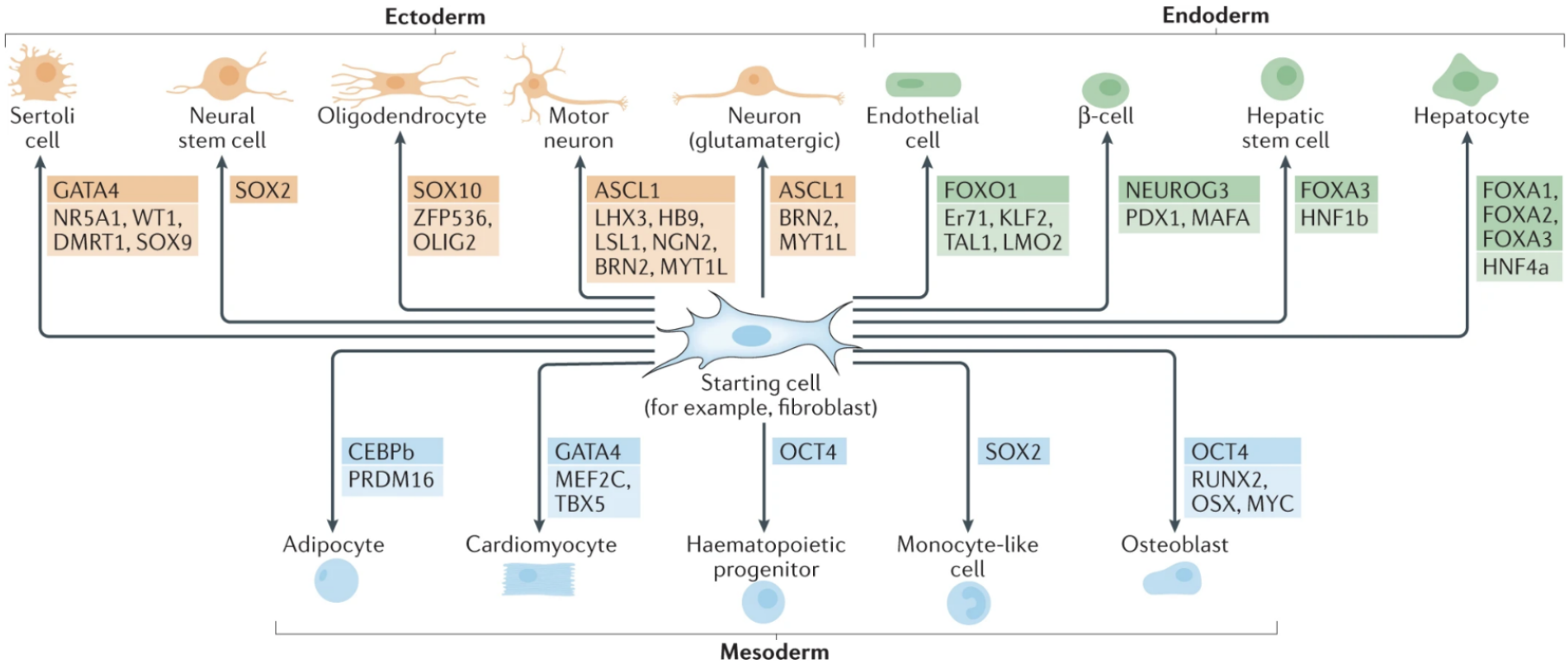


Lineage Conversion





Lineage Conversion



Wang et al. (2021)

Lineage Conversion

Advantages

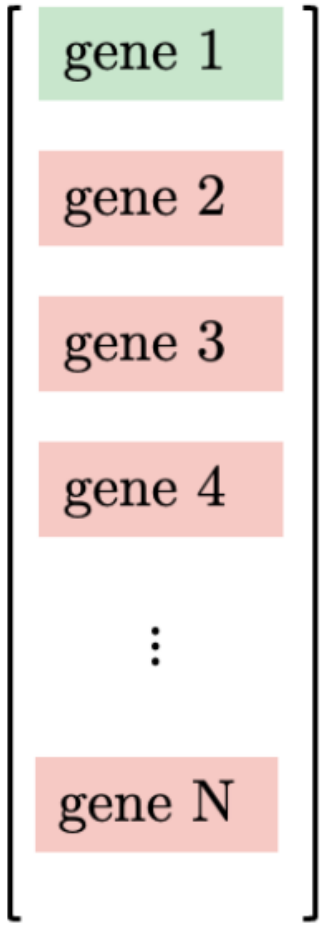
- Patient specific cell generation
- Bypasses the pluripotent state, thereby decreasing risks of off-target tumorigenesis

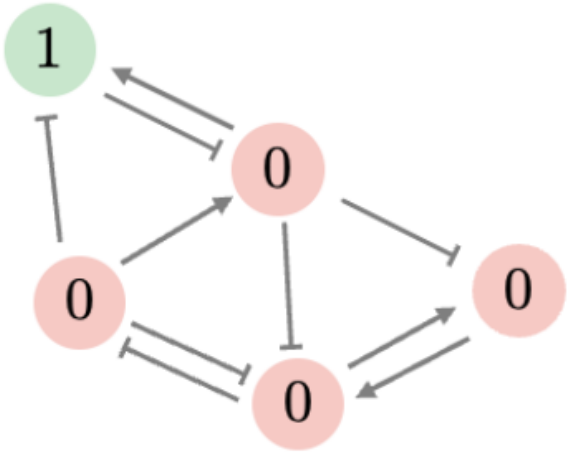
Disadvantages

- Limited mechanistic understanding and laborious search for transcription factor cocktail
- Can lead to heterogeneous target tissues

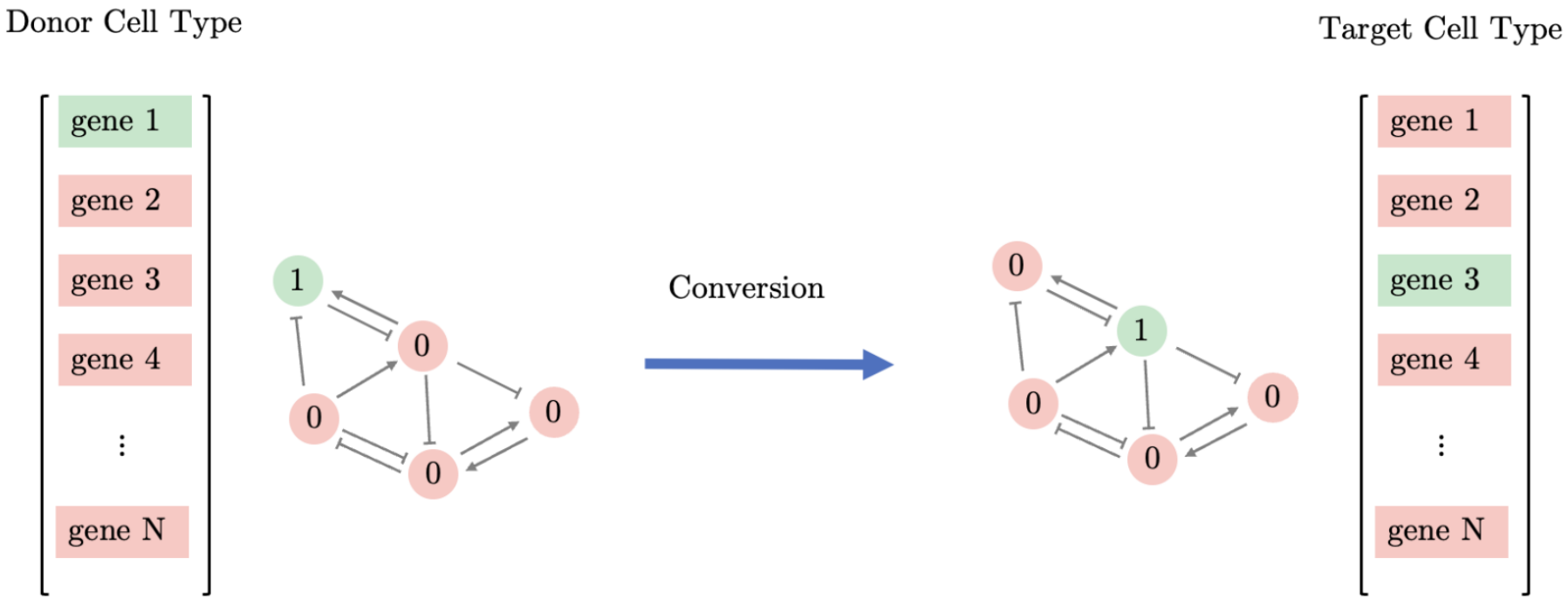
1. How can we better inform the search for novel lineage conversion pathways?
2. Can we develop simple models of differentiation and conversion to characterize spatial heterogeneity of target tissues?

Network of master regulatory genes



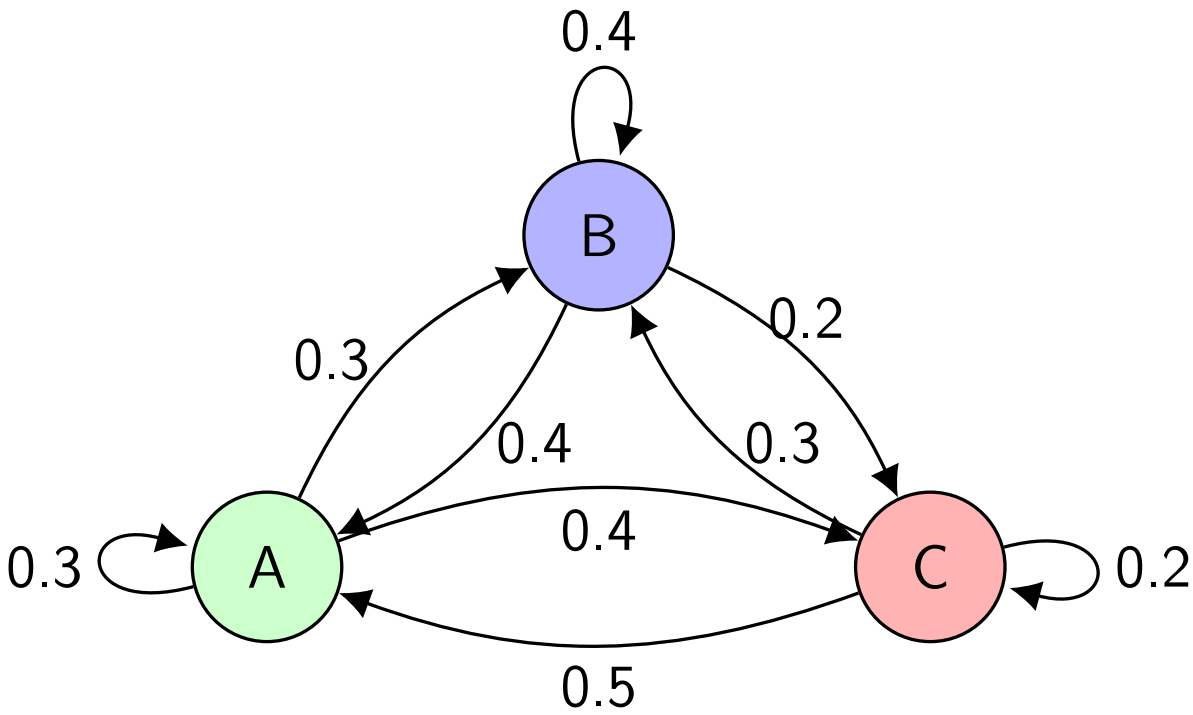


Lineage conversion is transition between stable states



Example of a Probabilistic Discrete Dynamical System

- Consider a simple system with states {A, B, C}.
- Transition probabilities:



Understanding Steady States in Dynamical Systems

- **Steady State Definition:** A state where probabilities do not change over time.
- **Mathematical Representation:** If x is the state vector and P the transition matrix, a steady state x satisfies $\mathrm{x} = P\mathrm{x}$.
- **Example:** For a transition matrix

$$P = \begin{pmatrix} 0.3 & 0.4 & 0.5 \\ 0.3 & 0.4 & 0.3 \\ 0.4 & 0.2 & 0.2 \end{pmatrix},$$

find a vector $\mathrm{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$ such that $\mathrm{x} = P\mathrm{x}$.

Computing Steady States

- **Steady State Equation:** $x = Px$.
- **Solving the System:**
  - Reformulate as $(P - I)x = 0$, where I is the identity matrix.
  - Solve the linear system to find x.
- **Normalization:** Ensure that the elements of x sum to 1, since they represent probabilities.
- **Example Continued:**
  - Solve $(P - I)x = 0$ for the given P.
  - Normalize x to find the steady-state distribution.

Matrix Generation

The matrix characterising the GRN should have:

1. M stable states corresponding to M cell fate or tissue type
2. Biochemically logical relationships between regulatory genes

There is no matrix which satisfies both (1) and (2).

Hybrid Matrix Generation

Find a set of $(n \times n)$ matrices $A \in B$ that have the n required stable states $\vec{y_n}$

$$\lim_{n\to\infty} A^n \vec{x_1} = \vec{y_1}$$
$$\lim_{n\to\infty} A^n \vec{x_2} = \vec{y_2}$$
$$\vdots$$
$$\lim_{n\to\infty} A^n \vec{x_n} = \vec{y_n}$$

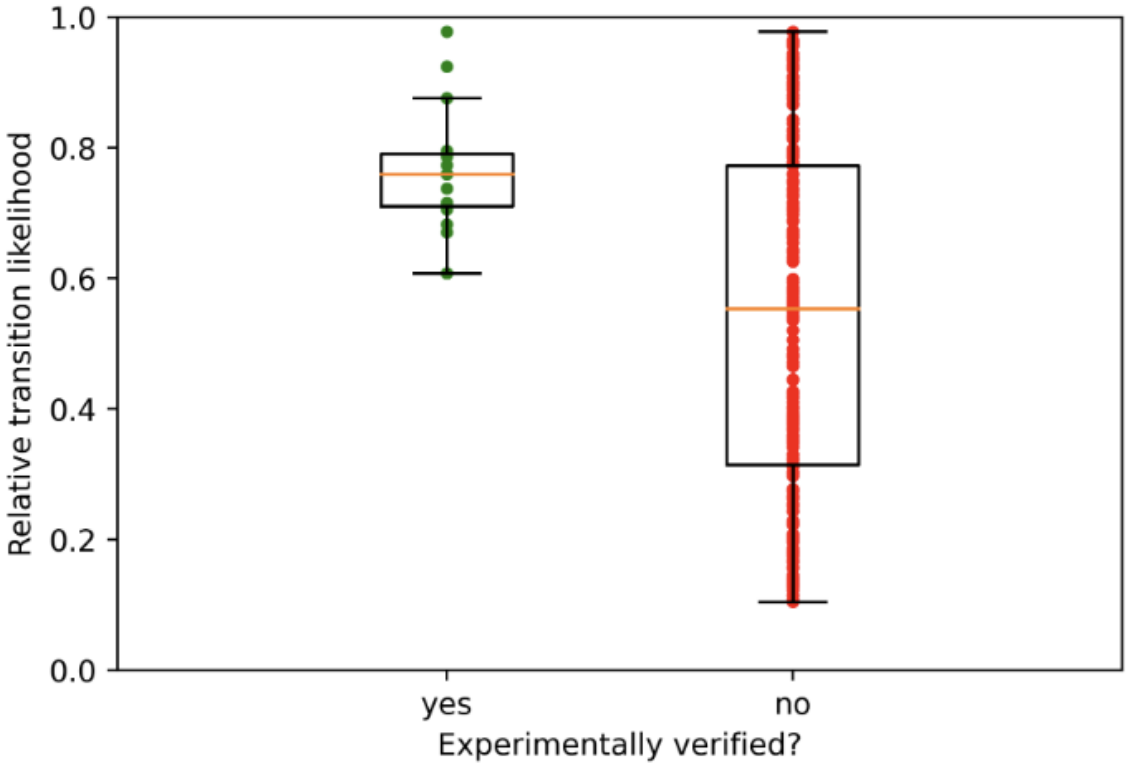
Define a $(n \times n)$ matrix C from empirical biochemical studies logic sourced from the literature

$$C = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix}$$

Choose $D \in B$ by minimizing distance to C

$$D \equiv \min(d(A, C))$$

Verification from Biological Literature

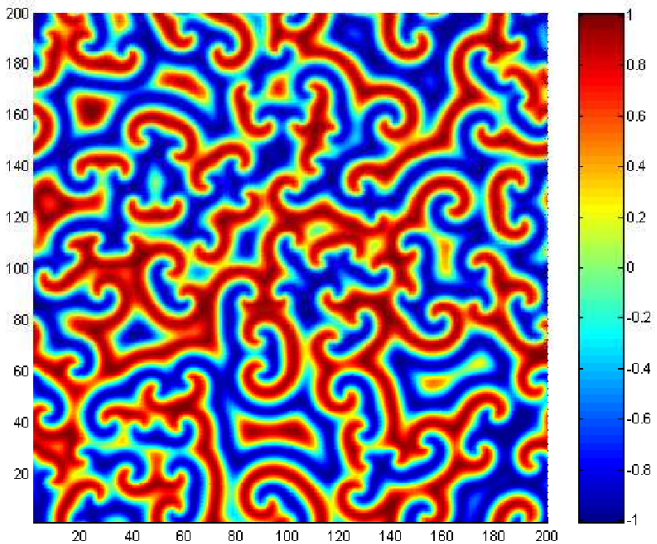


Introduction to Reaction-Diffusion Models

- **Concept:** Models that describe how the concentration of one or more substances distributed in space changes under the influence of two processes: reaction (chemical) and diffusion (spatial).
- **Applications:** Widely used in various fields such as physics, chemistry, biology, and ecology.
- **Analysis:** Used to understand how small perturbations in a system evolve over time, as well as how certain patterns arise.

Mathematical Formulation of Reaction-Diffusion Models

- **General Form:** $\frac{\partial u}{\partial t} = D\nabla^2 u + f(u)$, where:
  - u represents the concentration of a substance.
  - D is the diffusion coefficient.
  - $\nabla^2 u$ is the Laplacian, representing diffusion.
  - $f(u)$ represents the reaction component.
- **Example:** The Ginzburg–Landau equation $\frac{\partial A}{\partial t} = \nabla^2 A + A - A|A|^2$

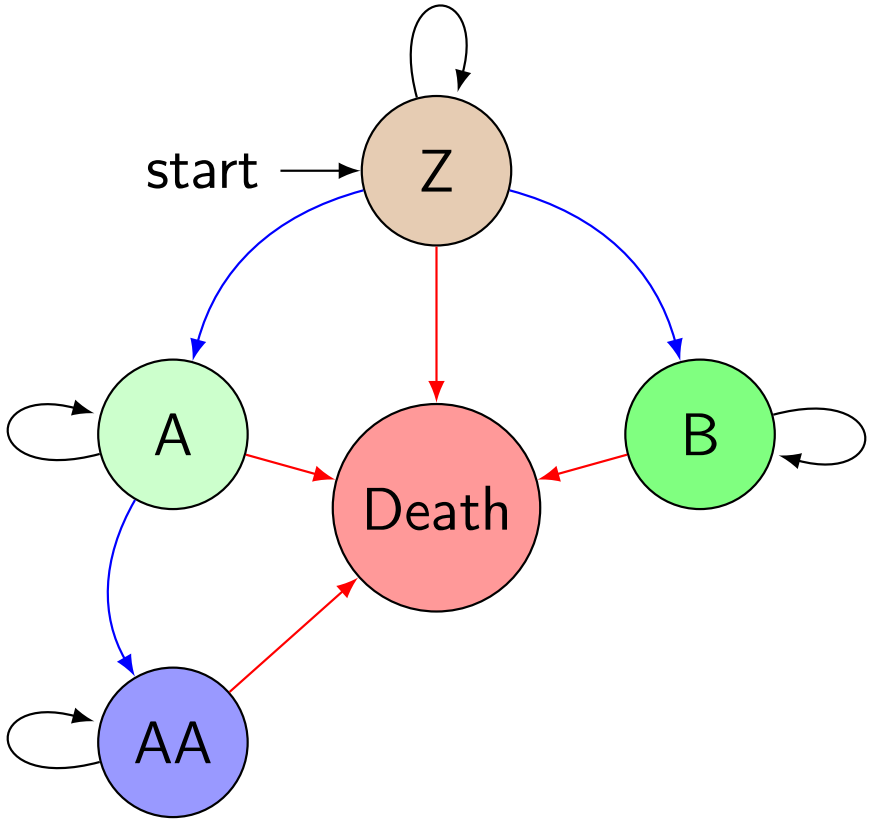


Reaction-Diffusion Models in Biological Systems

Relevance in Biological Systems:

- Explain spatial and temporal patterns in biological systems, such as cellular processes.
- Model the dynamics of ecological systems and population distributions.

Case 1: Differentiation, Apoptosis [1]



[1] Morris 2016, *Development*

Model 1 Equations

$$\frac{\partial Z}{\partial t} = \text{Growth} - \text{Death} - \text{Diff (Loss)} + \nabla^2 \tag{1}$$

$$\frac{\partial A}{\partial t} = \text{Growth} - \text{Death} + \text{Diff (Gain)} - \text{Diff (Loss)} + \nabla^2 \tag{2}$$

$$\frac{\partial B}{\partial t} = \text{Growth} - \text{Death} + \text{Diff (Gain)} + \nabla^2 \tag{3}$$

$$\frac{\partial AA}{\partial t} = \text{Growth} - \text{Death} + \text{Diff (Gain)} + \nabla^2 \tag{4}$$

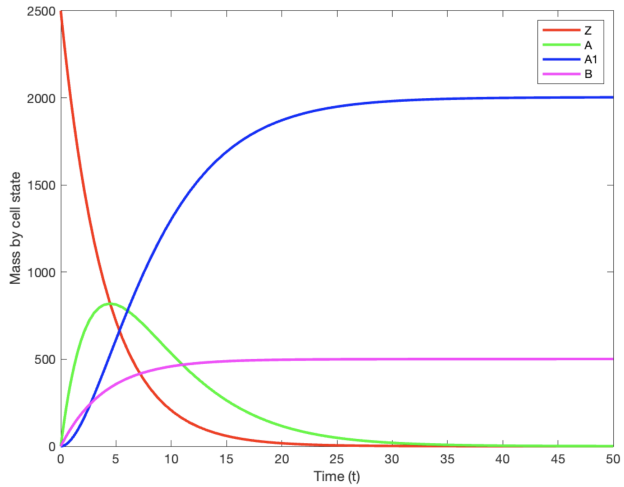
Model 1 Equations

$$\frac{\partial Z}{\partial t} = g_Z - d_Z - (p_{ZA} + p_{ZB})Z + D_Z \nabla^2 Z \tag{1}$$

$$\frac{dA}{dt} = g_A - d_A + p_{ZA} Z - p_{AAA} A + D_A \nabla^2 A \tag{2}$$

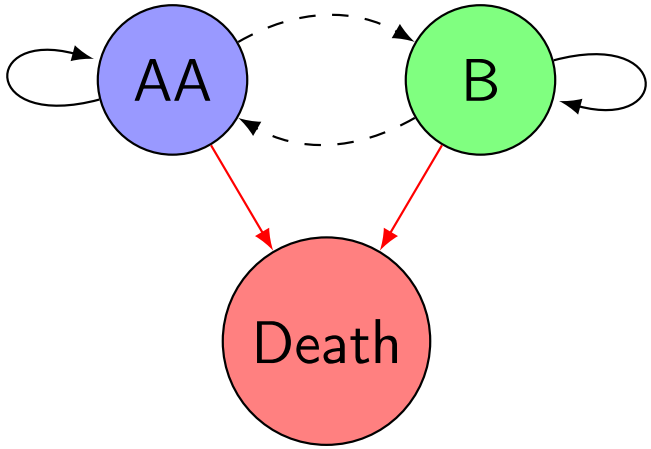
$$\frac{\partial B}{\partial t} = g_B - d_B + p_{ZB} Z + D_B \nabla^2 B \tag{3}$$

$$\frac{\partial AA}{\partial t} = g_{AA} - d_{AA} + p_{AAA} A + D_{AA} \nabla^2 AA \tag{4}$$



Time evolution of (1) differentiation, (2) apoptosis in Model 1

Case 2: Lineage Conversion



Model 2 Equations

$$\frac{\partial AA}{\partial t} = \text{Growth} - \text{Death} - \text{LC (Loss)} + \text{LC (Gain)} + \nabla^2 \tag{1}$$

$$\frac{\partial B}{\partial t} = \text{Growth} - \text{Death} + \text{LC (Gain)} - \text{LC (Loss)} + \nabla^2 \tag{2}$$

$$\frac{\partial t_{AA}}{\partial t} = -\text{Decay} + \nabla^2 \tag{3}$$

$$\frac{\partial t_B}{\partial t} = -\text{Decay} + \nabla^2 \tag{4}$$

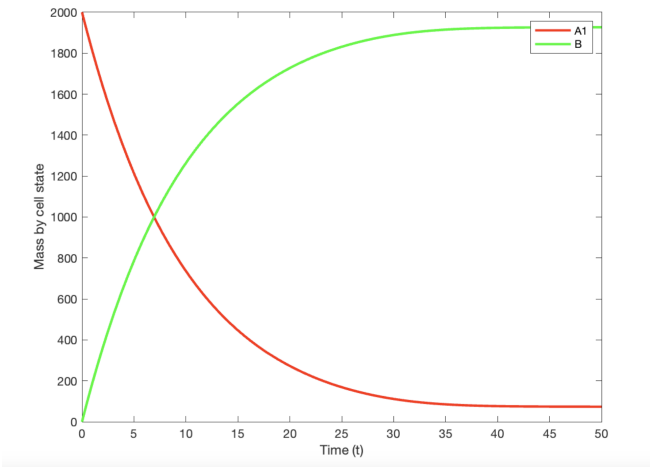
Model 2 Equations

$$\frac{\partial AA}{\partial t} = g_{AA} - d_{AA} - \mu_{AA}\frac{t_{AA}}{k_{AA}+t_{AA}}B + \mu_B\frac{t_B}{k_B+t_B}AA + D_{AA}\nabla^2 AA \quad (1)$$

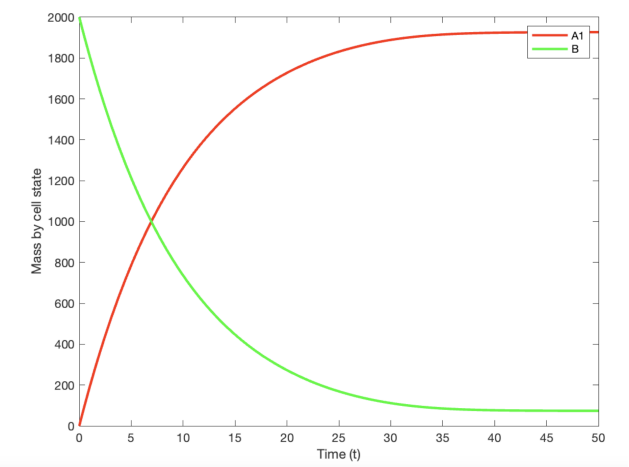
$$\frac{\partial B}{\partial t} = g_B - d_B - \mu_B\frac{t_B}{k_B+t_B}AA + \mu_{AA}\frac{t_{AA}}{k_{AA}+t_{AA}}B + D_B\nabla^2 B \quad (2)$$

$$\frac{\partial t_{AA}}{\partial t} = -q_{AA}t_{AA} + D_{tAA}\nabla^2 t_{AA} \quad (3)$$

$$\frac{\partial t_B}{\partial t} = -q_B t_B + D_{tB}\nabla^2 t_B \quad (4)$$

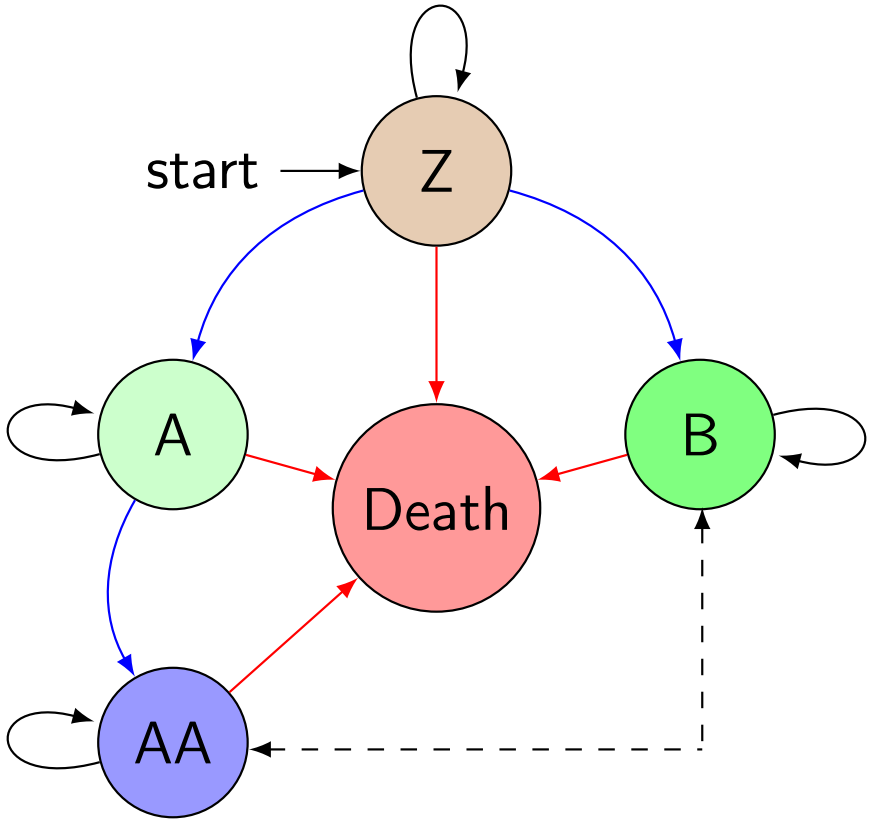


(a) Lineage Conv: AA to B



(b) Lineage Conv: B to AA

Case 3: Differentiation, Apoptosis, Lineage Conversion



Model 3 Equations

$$\frac{\partial Z}{\partial t} = \text{Growth} - \text{Death} - \text{Diff (Loss)} + \nabla^2 \quad (1)$$

$$\frac{\partial A}{\partial t} = \text{Growth} - \text{Death} + \text{Diff (Gain)} - \text{Diff (Loss)} + \nabla^2 \quad (2)$$

$$\frac{\partial B}{\partial t} = \text{Growth} - \text{Death} + \text{Diff (Gain)} + \text{LC (Gain)} - \text{LC (Loss)} + \nabla^2 \quad (3)$$

$$\frac{\partial AA}{\partial t} = \text{Growth} - \text{Death} + \text{Diff (Gain)} + \text{LC (Gain)} - \text{LC (Loss)} + \nabla^2 \quad (4)$$

$$\frac{\partial t_{AA}}{\partial t} = -\text{Decay} + \nabla^2 \quad (5)$$

$$\frac{\partial t_B}{\partial t} = -\text{Decay} + \nabla^2 \quad (6)$$

Model 3 Equations

$$\frac{\partial Z}{\partial t} = g_Z - d_Z - (p_{ZA} + p_{ZB})Z + D_Z \nabla^2 Z \quad (1)$$

$$\frac{dA}{dt} = g_A - d_A + p_{ZA} Z - p_{AAA} A + D_A \nabla^2 A \quad (2)$$

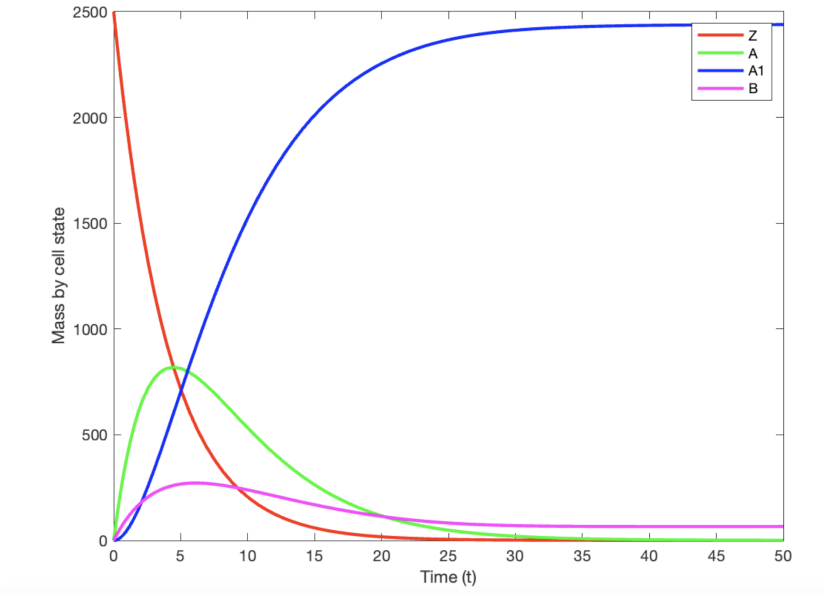
$$\frac{\partial B}{\partial t} = g_B - d_B + p_{ZB} Z + \mu_B \frac{t_B}{k_B + t_B} AA - \mu_{AA} \frac{t_{AA}}{k_{AA} + t_{AA}} B + + D_B \nabla^2 B \quad (3)$$

$$\frac{\partial AA}{\partial t} = g_{AA} - d_{AA} + p_{AAA} A + \mu_{AA} \frac{t_{AA}}{k_{AA} + t_{AA}} B - \mu_B \frac{t_B}{k_B + t_B} AA + D_{AA} \nabla^2 AA \quad (4)$$

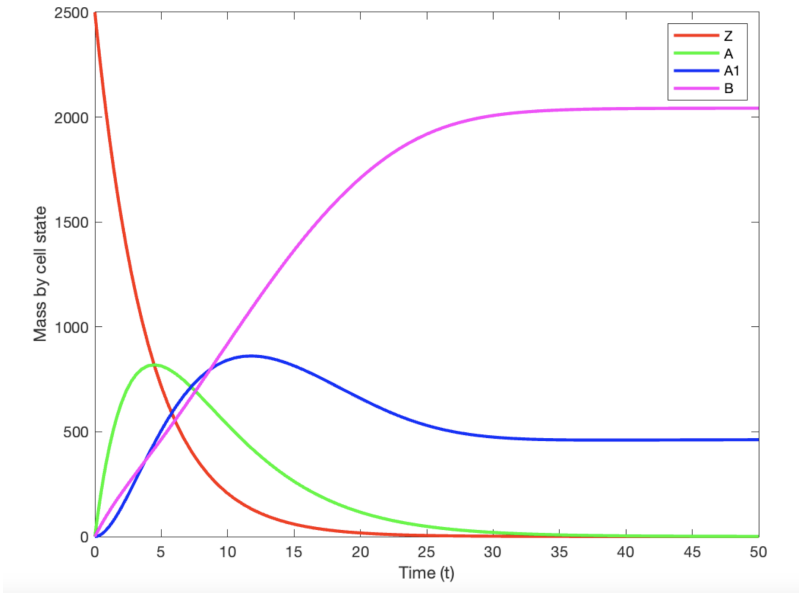
$$\frac{\partial t_{AA}}{\partial t} = - q_{AA} t_{AA} + D_{t\,AA} \nabla^2 t_{AA} \quad (5)$$

$$\frac{\partial t_B}{\partial t} = - q_B t_B + D_{t\,B} \nabla^2 t_B \quad (6)$$

Case 3: Time Evolution



(a) B to AA Lineage Conversion



(b) AA to B Lineage Conversion

(1) Differentiation, (2) Apoptosis, (3) Lineage Conversion

Code Demo: A Simple Model of Spatial Heterogeneity

Algorithm Simulation Pseudocode

1: $dim_x, dim_y, c_{tf}, c_{cell}, states \leftarrow 10, 10, 0.3, 0.3, 5$
2: $tissue, index_x, index_y \leftarrow$ initialize tissue and random indices
3: $tissue(index_x, index_y) \leftarrow 2, tissue_c map \leftarrow$ color map definition
4: initialize video and visualization settings
5: initialize *tf _grid* with zeros and set initial conditions
6: **for** $t \leftarrow 1$ to 10 **do**
7: **for** $row \leftarrow 1$ to dim_x **do**
8: **for** $col \leftarrow 1$ to dim_y **do**
9: $curr, y0, t\_range, [time, y], nei \leftarrow$ current state and ODE solution
10: **for** $k \leftarrow 1$ to 4 **do**
11: diffusion in the neighborhood
12: **end for**

Algorithm Simulation Pseudocode Continued

for $k \leftarrow 1$ to 4 **do**
  for $l \leftarrow 5$ to 6 **do**
    if $nei[k,:] \neq 0$ **then**
      $tf\_grid[row, col][l] \leftarrow c_{tf} \cdot (nei[k,l] - tf\_grid[row, col][l])$
    end if
  end for
  for $l \leftarrow 1$ to 4 **do**
    if $nei[k,:] \neq 0$ **then**
      $tf\_grid[row, col][l] \leftarrow c_{cell} \cdot (nei[k,l] - tf\_grid[row, col][l])$
    end if
  end for
end for
$tf\_grid[row, col] \leftarrow y[11,:] = 0$

Our Futures

Carlos
Applying to graduate programs in biology

Shloka
Applying to PhD programs in ecology

Michael
Moving to Armenia

Introduction to IIF Method

The Implicit Integration Factor (IIF) method is designed to solve stiff systems in reaction-diffusion equations, where some terms act on much shorter time scales and cause rapid changes. It uses explicit methods to solve the linear diffusion terms, and implicit methods to handle nonlinear reactions.

- Stiffness is a common challenge in reaction-diffusion systems.
- IIF method offers an efficient solution for biological applications.

Derivation of IIF Method

Consider a semi-discrete system of the form:

$$\frac{\partial u}{\partial t} = cu + f(u), \quad t > 0, \quad u(0) = u_0$$

Where:

- c is a constant representing diffusion.
- $f(u)$ is a nonlinear function representing reaction.

We first apply an integrating factor e^{-ct}, and then integrate:

$$u(t_{n+1}) = u(t_n)e^{c\Delta t} + \int_0^{\Delta t} e^{-c\tau} f(u(t_n + \tau))\, d\tau \tag{7}$$

Implicit Scheme

We want to approximate $e^{-c\tau} f(u(t_n + \tau))$ using an interpolation polynomial depending on t_{n+1} in order to construct a scheme that treats the stiff part implicitly, but integrates the diffusion term exactly. To do this, first define:

$$g(\tau) = e^{-c\tau} f(u(t_n + \tau))$$

And then take a Lagrange polynomial of order r-1 to be:

$$p(\tau) = \sum_{i=1}^{r-2} e^{ic\Delta t} f(u_{n-i}) \prod_{\substack{j=-1 \\ j \neq i}}^{r-2} \frac{\tau + j\Delta t}{(j-i)\Delta t}, \quad 0 \leq \tau \leq \Delta t.$$

Calculations

Given that $g(0) = f(u_n)$ and $g(\Delta t) = e^{-c\Delta t} f(u_{n+1})$, we can simplify the first order approximation of $g(\tau)$:

$$p(\tau) = \frac{1}{\Delta t}\left[f(u_n)(\Delta t - \tau) + e^{-c\Delta t} f(u_{n+1})\tau\right], \quad 0 \leq \tau \leq \Delta t.$$

We can repeat this process with further values of $g(c\Delta t)$ to derive higher order approximations. Once this is completed, we see that 7 can be rewritten as:

$$u_{n+1} = e^{c\Delta t} u_n + e^{c\Delta t} \int_0^{\Delta t} p(\tau)\, d\tau. \tag{8}$$

Calculations

Evaluating the integral, we get from Nie (2005) that

$$u_{n+1} = e^{c\Delta t} u_n + \Delta t \left(a_{n+1,1} f(u_{n+1}) + \sum_{i=0}^{r-2} a_{n-i} f(u_{n-i}) \right).$$

where a_i is defined by:

$$a_{n-i} = \frac{e^{(i+1)c\Delta t}}{\Delta t} \int_0^{\Delta t} \prod_{\substack{j=-1 \\ j \neq i}}^{r-2} \frac{\tau + j\Delta t}{(j-i)\Delta t} \, d\tau, \quad -1 \leq i \leq r-2.$$

And these can be solved computationally once the initial conditions, time steps, and desired level of accuracy is known.

Pseudocode for IIF Method

Algorithm Solve Reaction-Diffusion System Using IIF

Require: $u_0, c, \Delta t, T, r$

Ensure: u_{n+1} at each time step until final time T

$u \leftarrow u_0$ ▷ Initialize with initial condition

for $n \leftarrow 0$ to $T/\Delta t - 1$ **do**

    Compute $e^{c\Delta t}$ and other constants

    Define $g(\tau) \leftarrow e^{-c\tau} f(u(t_n + \tau))$

    Approximate $g(\tau)$ with Lagrange polynomial $p(\tau)$

    for $i \leftarrow -1$ to $r - 2$ **do**

        Compute coefficients a_{n-i} using the integral formula

    end for

    Compute integral of $p(\tau)$ from 0 to Δt

    $u_{n+1} \leftarrow e^{c\Delta t} u_n + e^{c\Delta t} \int_0^{\Delta t} p(\tau)\, d\tau$

end for

Conclusion

The IIF method:

- Enhances efficiency by decoupling diffusion and nonlinear reaction.
- Is robust and accurate, maintaining stability even in stiff systems.
- Has computational advantages, especially for large-scale problems.