Author Manuscript

Pattern formation in dense populations studied by inference of nonlinear diffusion-reaction mechanisms

Siddhartha Srivastava[1,3] **& Krishna Garikipati**[*1,2,3]

[1]*Department of Mechanical Engineering, University of Michigan, United States*
[2]*Department of Mathematics, University of Michigan, United States*
[3]*Michigan Institute for Computational Discovery & Engineering, University of Michigan, United States*

Abstract

Reaction-diffusion systems have been proposed as a model for pattern formation and morphogenesis. The Fickian diffusion typically employed in these constructions model the Brownian motion of particles. The biological and chemical elements that form the basis of this process, like cells and proteins, occupy finite mass and volume and interact during migration. We propose a Reaction-Diffusion system with MS formulation to construct the diffusive flux. This formulation relies on inter-species force balance and provides a more realistic model for interacting elements. We also present a variational system inference-based technique to extract these models from spatiotemporal data for these processes. We show that the inferred models can capture the characteristics of local Turing instability that instigates the pattern formation process. Moreover, the equilibrium solutions of the inferred models form similar patterns to the observed data.

1 Introduction

Biological patterns and morphogenesis are emergent phenomena with discernibly coherent structures forming at larger length scales, spawning out of local mechanochemical interactions of natural materials. A large class of PDEs have been proposed to explain this phenomenon. These models have been reviewed in a previous communication[1]. Among these models, the Reaction-Diffusion (RD) model is one of the most widely studied in literature[2]. The seminal work of Alan Turing[3] laid out the mathematical understanding of how specific spatial modes are destabilized in an RD system, leading to a heterogeneous equilibrium solution. The physical understanding of this mechanism regards the morphogens as short-range activators coupled with long-range inhibitors. A local positive perturbation in the activator leads to a localized growth in activators and inhibitors. Meanwhile, the faster diffusion of the inhibitors leads to a decay in activators and inhibitors in the neighboring locations. In this work, we take a mathematical outlook on the RD system, albeit guided by our understanding of the physical system. In particular, we focus on the diffusion aspect of this model.

A lot of development in the study of RD systems, starting from Turing's original work[3], has been with Fickian diffusion with varying self-diffusivities of different components[2,4]. Some studies[5,6] have also considered cross-diffusion as a generalization to the diagonal dominant diffusion. Fickian diffusion[7] posits that the flux of a chemical component is directly proportional to the negative gradient of its concentration. At the microscopic scale, this transport models the *Brownian motion* that is the stochastic collision-driven motion of particles. The cumulative migratory behavior of these particles translates to a concentration gradient equalizing flow of the density. This linear relationship of flux and concentration gradient works reasonably well for approximating diffusion processes in binary mixtures and under dilute limits[8]. On the other hand, extending this model to more complicated scenarios entails including correction terms in the diffusivity coefficient[9,10,11].

*Corresponding author at: Department of Mechanical Engineering, University of Michigan, United States. *E-mail address*: krishna@umich.edu (K. Garikipati).

This is the author manuscript accepted for publication and has undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1002/nme.7475

This article is protected by copyright. All rights reserved.

Author Manuscript

Often, it is not even possible to describe the behavior with these corrections while also maintaining thermodynamic admissibility. Duncan and Toor[12] designed an experiment with a ternary mixture of gases that showcased behavior like osmotic diffusion, reverse diffusion and diffusion barrier where the apparent driving force seems to be along the direction of gradient; such phenomena are not captured by Fickian diffusion. The Maxwell-Stefan (MS) approach to multicomponent diffusion is based on the kinetic theory and molecular force balances[13]. It has been applied to various physical systems such as intracellular organization of biomolecules[14], protein adsorption[15,16], flow of solutes across polymer membranes[17] and dispersion of atmospheric aerosols[18]. This mechanism can better describe the interaction of elements in specific biological systems, for instance, cellular migration in tissue formation.

In this work, we have developed an MS diffusion-based RD system. A similar model has been developed by Herberg et.al.[19] where they have considered a mass-conserving transport with reactive flow; however, they have not considered pattern-forming reactive dynamics. We also provide a mathematical framework to obtain these models in a data-driven fashion. We believe this approach will also help alleviate the problem in the experimental measurements of MS diffusivity coefficients[20]. Our approach leverages Variational System Identification (VSI) techniques that adopt a regression-based methodology. This methodology involves evaluating a set of potential operators within partial differential equations (PDEs) using available data. Subsequently, the parameter estimation problem is reformulated as a nonlinear regression problem. This transformation can be achieved through the construction of operators, either in a strong form of PDE, such as finite difference representations as seen in the Sparse Identification of Nonlinear Dynamics (SINDy) approach[21], or in a weak or variational form of the PDE as is done in VSI methods. Both of these operator representations impose rigorous demands on the quality and quantity of data necessary for accurate operator construction. However, using weak forms offers distinct advantages, including the appearance of boundary conditions within the formulation and reduced prerequisites concerning spatial derivatives. PDE-constrained optimization techniques have also been employed for parameter identification[22]. These techniques, unlike VSI, minimize the difference in the PDE's predicted solution and the data field. Therefore, they enjoy the advantage of being much more robust against noisy data, but at a computational cost. However, in the study of pattern formation, the equilibrium solution has a strong sensitivity to initial conditions, making PDE-constrained optimization a poor choice for such problems. In the case of PDE-constrained optimization, this problem is further exacerbated when the forward solution of models does not converge due to poor guesses for parameters. The mathematically simplified nature of the regression approach, particularly when working with weak forms, renders it amenable to a wide array of operator selection techniques. These include regularization, stepwise regression, and genetic algorithms, facilitating the identification of a parsimonious model (PDE with minimal parameters) from an extensive set of potential operators[22]. Similar techniques have been previously studied across diverse domains, including constitutive theories[22,23],epidemiology[24], and pattern formation[25].

We have made the following contributions in this work:

- We have proposed a new model for studying pattern formation using an MS-based RD system.
- We show that MS-based RD systems can capture a rich set of dynamics that generalizes the ones produced by Fick-based RD systems.
- We have formulated a VSI framework for extracting the RD model from spatiotemporal data.
- Our approach captures the salient features of the Turing instability, and inferred models can produce similar patterns to the original data.

Author Manuscript

- We show that temporally refined data near the onset of instability results in better recovery of stability diagrams. Meanwhile, including longer time ranges in the data results in better recovery of patterns

In the next section, we will introduce the theory for the transport of multicomponent systems and Maxwell-Stefan diffusion. In Sec. 3, we develop the VSI-based computational framework for inferring MS-based RD systems from spatiotemporal data of concentration fields. In Sec. 4, we show pattern formation with MS-based RD systems and employ our data-driven inference approach for identifying these systems. Finally, we present the concluding remarks in Sec. 5.

2 Transport of multicomponent system

Let $\Omega \subset \mathbb{R}^m$ be an open bounded domain with boundary $\partial\Omega$ with outer normal field $\boldsymbol{n}$. We consider a mixture with $n \geq 2$ components. Each constituent has a molar mass $M_k > 0$, individual mass density $\rho_k \geq 0$ and a velocity $\boldsymbol{u}_i$ satisfying $\boldsymbol{u}_i \cdot \boldsymbol{n} = 0$ on the boundary $\partial\Omega$. The mass balance of each component is given by:

$$\frac{\partial \rho_i}{\partial t} + \nabla \cdot \rho_i \boldsymbol{u}_i = M_i r_i \quad \text{in } \Omega \text{ for } t \geq 0 \tag{1}$$

where r_i is the rate of production of i^{th} species. Note that r_i is *positivity-preserving* i.e. if $\rho_i = 0$ then $r_i \geq 0$. The total mass balance is obtained by summing over the balance equation of each component:

$$\frac{\partial \rho}{\partial t} + \nabla \cdot \rho \boldsymbol{u} = \sum_i M_i r_i, \quad \text{with } \boldsymbol{u} \cdot \boldsymbol{n} = 0 \tag{2}$$

where $\rho = \sum_i \rho_i$ is the total mass. The barycentric velocity, $\boldsymbol{u}$, is estimated as $\rho \boldsymbol{u} = \sum_i \rho_i \boldsymbol{u}_i$. Transport is also studied in terms of other quantities like mass faction $y_i = \rho_i/\rho$, molar concentration, $c_i = \rho_i/M_i$, and molar factions $x_i = c_i/c_{\text{tot}}$ where the total concentration is estimated as $c_{\text{tot}} = \sum_i c_i$. The transport of these quantities can be obtained from the mass transport equation. For instance, combining (1), (2) and the definition of mass faction gives:

$$\rho(\partial_t y_i + \boldsymbol{u} \cdot \nabla y_i) + \nabla \cdot \overline{\boldsymbol{J}}_i = M_i r_i - y_i \sum_i M_i r_i, \quad \text{with } \overline{\boldsymbol{J}}_i \cdot \boldsymbol{n} = 0 \tag{3}$$

where $\overline{\boldsymbol{J}}_i = \rho_i(u_i - u)$. The transport of molar concentration is obtained as follows:

$$\partial_t c_i + \nabla \cdot c_i \boldsymbol{u}_i = r_i \tag{4}$$

The transport of total concentration is obtained by adding the transport of individual concentrations:

$$\partial_t c_{\text{tot}} + \nabla \cdot c_{\text{tot}} \boldsymbol{v} = \sum_i r_i, \quad \text{with } \boldsymbol{v} \cdot \boldsymbol{n} = 0 \tag{5}$$

where molar averaged velocity, $\boldsymbol{v}$ is evaluated as $c_{\text{tot}} \boldsymbol{u} = \sum_i c_i \boldsymbol{u}$. Finally the transport of molar

Author Manuscript

factions are obtained from (4) and (5):

$$c_{\text{tot}}(\partial_t x_i + \boldsymbol{v} \cdot \nabla x_i) + \nabla \cdot \boldsymbol{J} = r_i - x_i \sum_i r_i, \quad \text{with } \boldsymbol{J}_i \cdot \boldsymbol{n} = 0 \tag{6}$$

where $\boldsymbol{J}_i = c_i(\boldsymbol{u}_i - \boldsymbol{v})$. It should be remarked that previous studies on non-reactive systems[8] assume c_{tot} to be constant over time. It can be observed from (5) that this requires: (1) $\sum_i r_i = 0$, i.e. there is no net production of material in the system, (2) $\nabla \cdot c_{\text{tot}}\boldsymbol{v} = 0$ that can be achieved by fixing a frame of reference moving with molar averaged velocity. meanwhile, for reactive flows[19], it is assumed that ρ is constant over time, requiring (1) $\sum_i M_i r_i = 0$ i.e. there is no net production of mass in the system, (2) $\nabla \cdot \rho\boldsymbol{u} = 0$ that can be achieved by fixing a frame of reference moving with barycentric velocity.

For this work, we will study the transport of molar factions and fix $c_{\text{tot}} = 1$. It should be noted that in this work, molar factions and molar concentration will have the same numerical value in the units of c_{tot}. As mentioned earlier, we will impose a constant c_{tot} by setting $\sum_i r_i = 0$ that we refer to as the *no-net production* property. Moreover, we fix the frame of reference with respect to molar averaged velocity, and $\boldsymbol{v} = 0 \implies \sum_i J_i = 0$. From here on, we will also refer to flux as $J_i = c_i\boldsymbol{u}_i$. With these remarks, the system under consideration is rewritten as follows:

$$\partial_t c_i + \nabla \cdot \boldsymbol{J}_i = r_i, \quad \text{such that } \sum_i r_i = 0 \tag{7}$$

The first prominent equation of diffusion is Fick's law. This law states that for a single-component system, the diffusion flux is proportional to the negative gradient of the concentration.

$$\boldsymbol{J} = -D\nabla c \tag{8}$$

Therefore, we get that the time derivative of the concentration is the negative of the divergence of the flux:

$$\frac{\partial c}{\partial t} = -\nabla \cdot \boldsymbol{J} = \nabla \cdot D\nabla c \tag{9}$$

This theory can be phenomenologically extended to the multicomponent case by assuming flux to be given as a linear function of the gradient fields as follows:

$$\boldsymbol{J}_i = -\sum_{j=1}^{n} D_{ij}\nabla c_j \tag{10}$$

where the diffusivity tensor, $\boldsymbol{D} \equiv \boldsymbol{D}([c_1, \cdots c_n])$ consists of non-zero off-diagonal terms:

$$\boldsymbol{D} = \begin{bmatrix} D_{11} & D_{12} & \cdots & D_{1n} \\ D_{21} & D_{22} & \cdots & D_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ D_{n1} & D_{n2} & \cdots & D_{nn} \end{bmatrix} \tag{11}$$

This results in the following RD model:

$$\partial_t c_i = \sum_j \nabla \cdot D_{ij}\nabla c_j + r_i \tag{12}$$

Author Manuscript

However, it has been proven that this phenomenology for diffusive fluxes is thermodynamically inconsistent[26]. In particular, (1) mass conservation is only satisfied if all diffusivities are the same. (2) The second law is satisfied only in very special cases. In the next section, we present Maxwell-Stefan diffusion that alleviates these deficiencies.

2.1 Maxwell-Stefan diffusion

The MS equations rely on inter-species force balances. More precisely, it is assumed that the thermodynamical driving force d_i is in local equilibrium with the total friction force. The friction force on 1 mole of particle i due to particle j is modeled as

$$f_{ij} = \frac{1}{\mathscr{D}_{ij}} c_j (u_i - u_j) \tag{13}$$

where the $\mathscr{D}_{ij} > 0$ has a physical significance of an inverse drag coefficient and $\mathscr{D}_{ij} = \mathscr{D}_{ji}$ is a natural condition due to force balance. We remark that the formulations presented in the literature show some variations in the model for net frictional forces, for instance, Krishna and Wesselingh[8] consider $f_{i,j} \propto x_j(u_i - u_j)$, meanwhile Herberg et.al.[19] consider $f_{i,j} \propto y_j(u_i - u_j)$.

The net force balance on particles results in the following:

$$d_i = -\sum_{j \neq i} \frac{c_i c_j}{\mathscr{D}_{ij}} (u_i - u_j) = -\sum_{j \neq i} \frac{c_j J_i - c_i J_j}{\mathscr{D}_{ij}} \tag{14}$$

Note that this relation satisfies $\sum_i d_i = 0$. There have been various models for driving forces proposed for specific systems[27]. In this work we will assume that $d_i = c_i \nabla \mu_i$, with the standard chemical potentials $\mu_i = \log(\gamma_i c_i)$ for the potentials. Here $\gamma_i > 0$ are the so-called activity coefficients. This results in $d_i = \nabla c_i$. These relations are then used to estimate J_i in terms of the thermodynamic forces.

The special case of equimolar transport is well-studied in literature[8]. Mathematically, it implies that $\sum_i J_i = 0$ and c_{tot} is constant. Consequently, $\sum_i \nabla c_i = \nabla c_{\text{tot}} = 0$. This allows us to rewrite the system with $n-1$-independent variables, $(c_1, \cdots c_{n-1})$:

$$\boldsymbol{B} \begin{bmatrix} J_1 \\ \vdots \\ J_{n-1} \end{bmatrix} = - \begin{bmatrix} \nabla c_1 \\ \vdots \\ \nabla c_{n-1} \end{bmatrix} \tag{15}$$

where

$$B_{ij} = \begin{cases} -c_i \left(\frac{1}{\mathscr{D}_{ij}} - \frac{1}{\mathscr{D}_{1n}} \right) & i \neq j \\ \frac{1}{\mathscr{D}_{in}} + \sum_{k \neq i} c_k \left(\frac{1}{\mathscr{D}_{ik}} - \frac{1}{\mathscr{D}_{in}} \right) & i = j \end{cases}$$

where we have used the facts: $(1) \sum_1^n J_i = 0$, $(2) \sum_1^n c_i = c_{\text{tot}} = 1$, and $(3)\ \sum_1^n \nabla c_i = 0$

Example: Consider a non-reactive ternary system. In this case, we have two independent equations

Author Manuscript

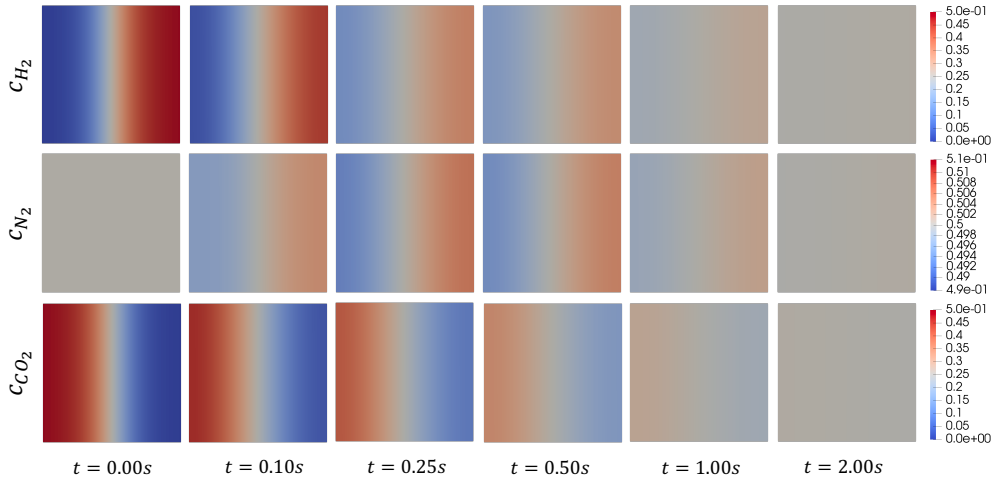


Figure 1: Evolution of molar concentrations c_{H_2} and c_{N_2} over $t \in [0, 2]$.

where the $\boldsymbol{B}$ matrix is given as:

$$\boldsymbol{B} = \begin{bmatrix} \frac{1}{\mathcal{D}_{13}} + c_2\left(\frac{1}{\mathcal{D}_{12}} - \frac{1}{\mathcal{D}_{13}}\right) & -c_1\left(\frac{1}{\mathcal{D}_{12}} - \frac{1}{\mathcal{D}_{13}}\right) \\ -c_2\left(\frac{1}{\mathcal{D}_{12}} - \frac{1}{\mathcal{D}_{23}}\right) & \frac{1}{\mathcal{D}_{23}} + c_1\left(\frac{1}{\mathcal{D}_{12}} - \frac{1}{\mathcal{D}_{23}}\right) \end{bmatrix}$$

Noting that $\det \boldsymbol{B} = (\mathcal{D}_{13}\mathcal{D}_{23})^{-1} + c_2\mathcal{D}_{23}^{-1}\left(\mathcal{D}_{12}^{-1} - \mathcal{D}_{13}^{-1}\right) + c_1\mathcal{D}_{13}^{-1}\left(\mathcal{D}_{12}^{-1} - \mathcal{D}_{23}^{-1}\right)$

$$\boldsymbol{B}^{-1} = \frac{1}{\det \boldsymbol{B}} \begin{bmatrix} \frac{1}{\mathcal{D}_{23}} + c_1\left(\frac{1}{\mathcal{D}_{12}} - \frac{1}{\mathcal{D}_{23}}\right) & c_1\left(\frac{1}{\mathcal{D}_{12}} - \frac{1}{\mathcal{D}_{13}}\right) \\ c_2\left(\frac{1}{\mathcal{D}_{12}} - \frac{1}{\mathcal{D}_{23}}\right) & \frac{1}{\mathcal{D}_{13}} + c_2\left(\frac{1}{\mathcal{D}_{12}} - \frac{1}{\mathcal{D}_{13}}\right) \end{bmatrix} \tag{16}$$

The evolution equation for this reduced system is given as:

$$\frac{\partial c_i}{\partial t} = \sum_{j=1}^{2} \nabla \cdot [\boldsymbol{B}^{-1}]_{ij} \nabla c_j \tag{17}$$

Given the similarity between the forms of MS and Fickian diffusivities from here on, we will denote with $D_{ij} = [\boldsymbol{B}^{-1}]_{ij}$, noting that these coefficients are now considered to be functions of the molar concentration.

2.2 Case study for non-reactive equimolar transport

Duncan and Toor's experiment[12] provides a canonical illustration for the generalizability of MS formulation over Fickian fluxes. The experimental set-up consisted of two-bulb diffusion cells with an ideal ternary gas mixture of (1) H_2 (2) N_2 (3) CO_2. , where bulb 1 consisted of equal moles of H_2 and N_2 mixture and bulb 2 consisted of equal moles of N_2 and CO_2 mixture. The mixtures in the bulbs were allowed to diffuse into each other via a capillary. The experimental conditions allow for a non-reactive equimolar flow. We first consider a numerical idealization of this experiment on a square domain with H_2 and N_2 on the right half of the domain and CO_2 and N_2 on the left half of the domain. We employ the reduced formulation Eq. (17) with c_1 and c_2 denoting the H_2 and N_2

Author Manuscript

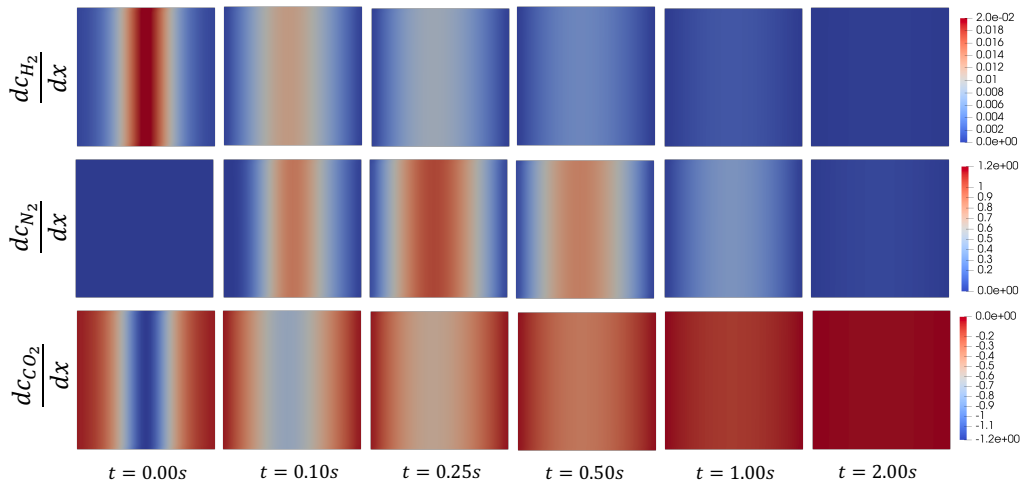


Figure 2: Evolution of gradients of molar concentrations c_{H_2} and c_{N_2} over $t \in [0, 2]$.

molar concentrations, respectively. Let $\Omega = [-0.5, 0.5]^2 \in \mathbb{R}^2$ with following initial conditions:

$$\begin{aligned} c_1(x, y, t = 0) &= 0.5/(1 + \exp(-20x)) \\ c_2(x, y, t = 0) &= 0.5 \end{aligned} \tag{18}$$

We consider the binary diffusion coefficients provided by Boudin et.al.[28]: $\mathscr{D}_{12} = 0.833$, $\mathscr{D}_{13} = 0.680$ and $\mathscr{D}_{23} = 0.168$. The simulation results for the molar concentrations are presented in Fig. 1, and the gradients of the molar concentrations are presented in Fig. 2, where only the horizontal components of the gradients are presented as the fields are uniform in the vertical direction. We observe that the molar concentrations of H_2 and CO_2 behave in a familiar way, where diffusion has an equalizing effect of removing gradients in the molar concentration (see rows 1 and 3 of Fig. 2) that results in uniform equilibrium steady state (see rows 1 and 3 of Fig. 1). However, N_2 shows an anomalous behavior. We first focus on the x-concentration gradient of N_2, Fig. 2) and observe that there is no gradient initially. However, it evolves to form a peak at $t = 0.25$, and then this peak decays to 0 at t=2. In terms of the concentration, c_{N_2}, this phenomenon appears as starting from a uniform field, N_2 first flows to the right, creating a gradient field and then flows back towards the left to a uniform equilibrium. This phenomenon we simulated through MS-diffusion cannot be explained through Fickian diffusion. In particular, at $t = 0$, N_2 starts to redistribute even though $\nabla c_{N2} = 0$. This phenomenon is dubbed as osmotic diffusion that, in the case of Fickian diffusion, demands $D_{N_2} \to \infty$. In $t \in (0, 0.25)$, the diffusion of N_2 is in a direction opposite to that dictated by its driving force in a Fickan model; this is known as reverse diffusion and requires $D_{N_2} < 0$. At $t = 0.25$, the diffusion mechanism switches from reverse diffusion to the usual Fickian diffusion. At this time, we observe a non-zero gradient field ($\nabla c_{N_2} \neq 0$); however, $J_{N_2} = 0$. In Fickian diffusion, this will require $D_{N_2} = 0$. This example shows that MS-diffusion can capture a richer set of diffusive mechanisms. In Sec. 4.1.1, we will discuss the relevance of these mechanisms on pattern formation.

3 Variational System Inference

In this section, we will develop the framework for data-driven inference of the parameterized parabolic PDE to obtain the coefficients of MS diffusion and functional forms of production rates. We will

Author Manuscript

first consider the variational formulation of Eq. (12). As discussed above, the equivalent diffusivity coefficients, D_{ij}, for MS diffusion, are nonlinear functions of the molar concentration $(c_1, \cdots c_{n-1})$. We will start with the general nonlinear form of diffusion coefficients and reaction terms. In Sec. 3.2, we will assume a parameterized form for the production rate.

3.1 Finite element formulation for RD model

We will consider $\boldsymbol{c} \in \boldsymbol{\mathcal{V}}(\Omega) \times L_2([0,T])$ where the hilbert space, $\boldsymbol{\mathcal{V}}(\Omega) = H^1(\Omega) \times_{\text{n-1 times}} \cdots \times H^1(\Omega)$. For the infinite-dimensional initial and boundary value problem (IBVP) with homogeneous Neumann boundary conditions on $\partial\Omega$, the weak form can be stated as: for all admissible test functions $\boldsymbol{w} \equiv (w_1, \cdots w_{n-1}) \in \boldsymbol{\mathcal{V}}$ find $\boldsymbol{c}(\boldsymbol{x}, t)$ such that $\boldsymbol{c}(\boldsymbol{x}, t=0) = \overline{\boldsymbol{c}}(\boldsymbol{x})$ and for all $i \in \{1, \cdots n-1\}$

$$\int_\Omega \frac{\partial c_i}{\partial t} w_i d\Omega = -\sum_j \int_\Omega D_{ij} \nabla w_i \cdot \nabla c_j d\Omega + \int_\Omega r_i w_i d\Omega \tag{19}$$

3.1.1 Time discretization

For a finite time evolution $t \in [0,T]$, we consider finite timepoints $[t_0, t_1, \cdots, t_{n_T}]$ with $t_0 = 0$, $t_{n_T} = T$ and $t_k < t_{k+1}$. The timestep size is evaluated as $\Delta t_k = t_{k+1} - t_k$. We consider the backward Euler scheme for updating c_i

$$\int_\Omega \frac{c_i - c_i^{t_k}}{\Delta t_k} w_i d\Omega = -\sum_j \int_\Omega D_{ij} \nabla w_i \cdot \nabla c_j d\Omega + \int_\Omega r_i w_i d\Omega \tag{20}$$

3.1.2 Finite element approximation

For finite-dimensional fields $\boldsymbol{c}^h, \boldsymbol{w}^h$ the weak form is as follows: find $\boldsymbol{c}^h \in \boldsymbol{\mathcal{V}}^h \subset \boldsymbol{\mathcal{V}}$ such that for all $\boldsymbol{w}^h \in \boldsymbol{\mathcal{V}}^h \subset \boldsymbol{\mathcal{V}}$, the finite-dimensional (Galerkin) weak form of the problem is satisfied. The variation components w_i^h and trial solutions c_i^h are defined using a finite number of basis functions,

$$\begin{aligned} c_i^h &= \sum_{k=1}^{n_b} a_{i,k} N_k \\ w_i^h &= \sum_{k=1}^{n_b} b_{i,k} N_k \end{aligned} \tag{21}$$

where n_b is the dimensionality of the $H^{1,h}(\Omega)$, and and N_k represents the global basis functions. Substituting w^h and c^h in Eq. (20), accounting for arbitrariness of w_i^h, and decomposing the integration over Ω by a sum over subdomains Ω_e leads to following set of residual equations at each timestep for $i \in \{1, \cdots n-1\}$ and $l \in \{1, \cdots n_b\}$:

$$R_{i,l,s} = \sum_{e=1}^{n_e} \sum_{k=1}^{n_{b_e}} \int_{\Omega_e} \left(\frac{a_{i,e(k)} - a_{i,e(k)}^{t_{s-1}}}{\Delta t_s} N_k N_l d\Omega + \sum_j D_{ij} a_{j,e(k)} \nabla N_{e(k)} \cdot \nabla N_l - r_i N_l \right) d\Omega \tag{22}$$

Author Manuscript

The solution to the PDE at each timestep is obtained by solving the above set of $n \times n_b$ equations, $R_{i,l} = 0$ for all i and l. Note that these equations are nonlinear in general as D_{ij} and r_j are typically functions of $\boldsymbol{c}$.

3.2 Identification of governing parabolic PDEs in weak form

We are provided with the data as a spatiotemporal field for the multicomponent system, $\boldsymbol{c}(\boldsymbol{x}, t)$ with $(\boldsymbol{x}, t) \in \Omega \times [0, T]$ where $\Omega \subset \mathbb{R}^n$ and $[0, T]$ are the spatial domain and time-period of interest. We will assume that irrespective of the source of data, be it experimental images or high-fidelity simulations, it is possible to obtain the finite element projection of the data field as $c_i^{d,h} = \sum_k a_{i,k}^d N_k$ where we are denoting the data-field as $\boldsymbol{c}^d$.

The exact parametric form of PDE can be obtained by considering a set of possible reaction operators. In this work, we consider polynomial operators for reactive terms.

$$r_i(\boldsymbol{c}; \boldsymbol{\theta}) = \sum_{k=0}^{n_o} \sum_{|\boldsymbol{\alpha}|=k} \theta_{i,\alpha} \boldsymbol{c}^{\boldsymbol{\alpha}} \tag{23}$$

where the multi-index notation with $\boldsymbol{\alpha} \equiv (\alpha_1, \cdots, \alpha_{n-1})$ with $|\boldsymbol{\alpha}| = \alpha_1 + \cdots + \alpha_{n-1}$, and $\boldsymbol{c}^{\boldsymbol{\alpha}} = c_1^{\alpha_1} c_2^{\alpha_2} \cdots c_{n-1}^{\alpha_{n-1}}$. And, n_o defines some cut-off order for the polynomial approximation.

These polynomial forms are a popular choice in modeling pattern-forming RD models such as Schnackenberg, Gray Scott and FitzHugh-Nagumo models. It should also be noted that the *no-net production* is implicitly built in this ansatz. This can be observed by fixing the reactive dynamics of the n^{th} component as $r_n(\boldsymbol{c}; \boldsymbol{\theta}) = \sum_{k=0}^{n_o} \sum_{|\boldsymbol{\alpha}|=k} \theta_{n,\alpha} \boldsymbol{c}^{\boldsymbol{\alpha}}$ with $\theta_{n,\alpha} = -\sum_{i=1}^{n-1} \theta_{i,\alpha}$. This results in the required condition that $\sum_{i=1}^{n} r_i = 0$. The condition for *positivity-preserving* is more challenging to achieve. A sufficient condition for this property to hold is that we only consider polynomials of the form $\boldsymbol{c}^{(1,1,\cdots,1)+\boldsymbol{\alpha}}$ forcing $r_i(\boldsymbol{c}, \boldsymbol{\theta}) = 0$ if any of the components are zero. However, this is a very restrictive condition. Since it is sufficient and not necessary, we will not enforce *positivity-preserving* property in this work and restrict our attention to models and respective IbVPs that naturally remain in the positive quadrant.

With this ansatz in place, we can formulate the optimization problem for PDE estimation. The set of optimization variables are given as $\boldsymbol{\Theta} = \{\mathcal{D}_{ij}\}_{i \neq j \in 1, \cdot n} \cup \{\theta_{i,\boldsymbol{\alpha}}\}_{i \in \{1, \cdot n-1\}, |\boldsymbol{\alpha}| \leq n_o}$. The optimization problem is posed as:

$$\boldsymbol{\theta}^* = \underset{\boldsymbol{\theta} \in \boldsymbol{\Theta}}{\arg\min} \left|\left| \boldsymbol{R}(\boldsymbol{\theta} | \boldsymbol{c}^d) \right|\right|^2, \quad \text{such that } \mathcal{D}_{ij} \geq 0 \quad \forall i, j \tag{24}$$

where $\boldsymbol{R}(\boldsymbol{\theta}|\boldsymbol{c}^d)$ is an $(n-1) \times n_b \times n_T$-dimensional vector indexed as $R_{i,l,s}$ with $(i, l, s) \in \{1, \cdots n-1, \} \times \{1, \cdots n_b\} \times \{1, \cdots n_T\}$ where:

$$R_{i,l,s}(\boldsymbol{\theta}|\boldsymbol{c^d}) = \sum_e \sum_{k=1}^{n_b} \int_{\Omega_e} \left(\frac{a_{i,k}^{d,t_s} - a_{i,k}^{d,t_{s-1}}}{\Delta t_s} N_k N_l d\Omega + \sum_j D_{ij}(\boldsymbol{c}^d; \boldsymbol{\theta}) a_{j,k}^{d,t_s} \nabla N_k \cdot \nabla N_l - r_i(\boldsymbol{c}^d; \boldsymbol{\theta}) N_l \right) d\Omega \tag{25}$$

Author Manuscript

Table 1: Inferred parameters for a 2-component MS diffusion with different IBVPs

Parameter	Original	Inferred (IBVP 1)	Inferred (IBVP 2)	Inferred (IBVP 3)
$\mathscr{D}_{12}$	0.833	0.83298162	0.83305782	0.83300214
$\mathscr{D}_{23}$	0.680	0.68001752	0.67998865	0.24130419
$\mathscr{D}_{13}$	0.168	0.16799677	0.1679995	0.05961649

3.2.1 Inference with MS diffusion

We first employ this VSI approach on MS diffusion. In particular, we explore the equimolar non-reactive transport example of Duncan and Toor's experiment (see Eq. (18)). From here on, we will refer to H_2, N_2 and CO_2 as species 1, 2, and 3, respectively. We are interested in inferring reduced 2-component MS equation Eq. (16) parameterized with $\boldsymbol{\Theta} = \{\mathscr{D}_{12}, \mathscr{D}_{23}, \mathscr{D}_{13}\}$. We consider 3 different IBVPs with IBVP1 same as Sec. 2.2 and IBVP2 and 3 defined as:

$$\begin{aligned}
\text{IBVP 2}: \quad & c_1(x, y, t=0) = 0.25 + 0.1\cos(10x) + 0.1\cos(10y) \\
& c_2(x, y, t=0) = 0.25 + 0.1\sin(10x) + 0.1\sin(10y) \\
\text{IBVP 3}: \quad & c_1(x, y, t=0) = 0.5/(1 + \exp(-20x)) \\
& c_2(x, y, t=0) = 0.5/(1 + \exp(20x))
\end{aligned} \tag{26}$$

The IBVPs are also provided in Fig. 3 (left panel). We infer the parameters for 2-component MS diffusion by solving Eq. (24), and the inferred parameters are presented in Tab. 1. We observe that for the case of IBVP 1 and IBVP 2, we can recover the PDE to a reasonable precision. However, that is not the case in IBVP 3. We can also consider the sensitivity of the residual to the parameters by evaluating the contours for a range of parameter space (shown in the right panel Fig. 3). We can observe from the bottom-right contour plot ($\mathscr{D}_{23}, \mathscr{D}_{13}$) for IBVP 3 that multiple minima are present along a line. This results in a degeneracy in the inferred parameters due to the chosen initial conditions. One may notice that in IBVP 3, $c_1 + c_2 \approx 1$ almost everywhere. It follows that:

- For a ternary system, if $c_1 + c_2 = 1$ at any given point in time then $\frac{\partial(c_1+c_2)}{\partial t} = 0$ i.e. if $c_1 + c_2 = 1$ at some time, then it remains 1 for all times. Consequently there is no information of interactions of the third species, resulting in degeneracy in the inferred frictional forces that involves this species,
- If $\boldsymbol{c}^d$ is such that it satisfies the above condition and $(\mathscr{D}_{12}^*, \mathscr{D}_{23}^*, \mathscr{D}_{13}^*)$ are such that Eq. (17) is satisfied, then choosing $(\mathscr{D}_{12}^*, \beta\mathscr{D}_{23}^*, \beta\mathscr{D}_{13}^*)$ for any will also satisfy Eq. (17).

The first point is easy to verify by adding the two components of Eq. (17):

$$\frac{\partial(c_1 + c_2)}{\partial t} = \nabla.\left(\frac{1 - c_1 - c_2}{\det \boldsymbol{B}}\left(\frac{\nabla c_1}{\mathscr{D}_{13}} - \frac{\nabla c_2}{\mathscr{D}_{23}}\right) + \frac{1}{\mathscr{D}_{12}\det \boldsymbol{B}}\nabla(c_1 + c_2)^2\right)$$

Noting that $c_1 + c_2 = c_{\text{tot}} = 1$ and $\nabla(c_1 + c_2)^2 = \nabla c_{\text{tot}}^2 = 1$, we get that $\frac{\partial(c_1+c_2)}{\partial t} = 0$

To show the second point, we first observe that $\frac{\partial c_1}{\partial t} = -\frac{\partial c_2}{\partial t}$, therefore it is only necessary to consider evolution of one of the species. Without loss of generality, we consider the evolution of c_1

$$\frac{\partial c_1}{\partial t} = \nabla \cdot \left(\frac{1}{\det \boldsymbol{B}}\left(\left(\frac{1}{\mathscr{D}_{13}} - \frac{1}{\mathscr{D}_{23}}\right)\nabla c_1^2 + \frac{\nabla c_1}{\mathscr{D}_{23}}\right)\right)$$

here we eliminated c_2 using the additional requirement $c_2 = 1 - c_1$. If the data, $\boldsymbol{c}^d$ satisfies this property, one can write the following inference consistency error, $e(\mathscr{D}_{12}, \mathscr{D}_{23}, \mathscr{D}_{13}|c_1^d)$ for the data c_1^d :

$$e(\mathscr{D}_{12}, \mathscr{D}_{23}, \mathscr{D}_{13}|c_1^d) = \left|\left| \frac{\partial c_1^d}{\partial t} - \nabla \cdot \left(\frac{1}{\det \boldsymbol{B}(\mathscr{D}_{12}, \mathscr{D}_{23}, \mathscr{D}_{13}, c_1^d)} \left(\left(\frac{1}{\mathscr{D}_{13}} - \frac{1}{\mathscr{D}_{23}} \right) \nabla (c_1^d)^2 + \frac{\nabla c_1^d}{\mathscr{D}_{23}} \right) \right) \right|\right|$$

where

$$\det \boldsymbol{B} = \frac{1 - c_1}{\mathscr{D}_{12}\mathscr{D}_{23}} - \frac{c_1}{\mathscr{D}_{13}\mathscr{D}_{12}}$$

Here the inference consistency error, e, provides a measure of how closely the data field, c_1^d and the parameters $(\mathscr{D}_{12}, \mathscr{D}_{23}, \mathscr{D}_{13})$ correspond to each other. In particular, if $(\mathscr{D}_{12}^*, \mathscr{D}_{23}^*, \mathscr{D}_{13}^*)$ is the parameter that generated the solution c_1^d then $e(\mathscr{D}_{12}^*, \mathscr{D}_{23}^*, \mathscr{D}_{13}^*|c_1^d)) = 0$. We also observe that for any $\beta > 0$:

$$\begin{aligned} e(\mathscr{D}_{12}, \beta\mathscr{D}_{23}, \beta\mathscr{D}_{13}|c_1^d) &= \left|\left| \frac{\partial c_1^d}{\partial t} - \nabla \cdot \left(\frac{1}{\det \boldsymbol{B}(\mathscr{D}_{12}, \beta\mathscr{D}_{23}, \beta\mathscr{D}_{13}, c_1^d)} \left(\left(\frac{1}{\beta\mathscr{D}_{13}^*} - \frac{1}{\beta\mathscr{D}_{23}^*} \right) \nabla (c_1^d)^2 + \frac{\nabla c_1^d}{\beta\mathscr{D}_{23}^*} \right) \right) \right|\right| \\ &= \left|\left| \frac{\partial c_1^d}{\partial t} - \nabla \cdot \left(\frac{1}{\det \boldsymbol{B}(\mathscr{D}_{12}, \mathscr{D}_{23}, \mathscr{D}_{13}, c_1^d)} \left(\left(\frac{1}{\mathscr{D}_{13}^*} - \frac{1}{\mathscr{D}_{23}^*} \right) \nabla (c_1^d)^2 + \frac{\nabla c_1^d}{\mathscr{D}_{23}^*} \right) \right) \right|\right| = e(D_{12}^*, D_{23}^*, D_{13}^*) = 0 \end{aligned}$$

This degeneracy can be generalized to n-component systems where if any of the components is uniformly zero at any given time, then the binary coefficients corresponding to that component are degenerate. Using these arguments for the three IBVPs, we see that this degeneracy is absent in IBVPs 1 and 2 since we start with a mixture with a non-zero concentration of each component. Therefore, the inference is more robust for these two cases.

3.3 Computational framework for inference

The numerical examples presented in this work have been posed and solved in two dimensions using the finite element method programmed using the automated computational framework of *Fenics* package[29,30]. All the simulations presented in this work are done on a 2D square domain with structured mesh with 5000 linear triangular elements. The nonlinear variational problem in the forward solution is solved using a Newton solver. The nonlinear optimization for minimizing Eq. (24) is done using *L-BFGS-B* method. We use the implementation provided by the *scikit-learn* package[31]. The objective function, $\texttt{obj} = \left|\left|\boldsymbol{R}(\boldsymbol{\theta}|\boldsymbol{c}^d)\right|\right|^2$ is evaluated using the *Fenics* library using the same mesh that the data was generated on. The gradient of the objective function is also supplemented to the optimization scheme, where it is evaluated as $\texttt{grad_obj} = \left|\left|\boldsymbol{R}(\boldsymbol{\theta}|\boldsymbol{c}^d)\right|\right| \nabla_{\boldsymbol{\theta}} \boldsymbol{R}(\boldsymbol{\theta}|\boldsymbol{c}^d)$, where the gradient of the residual is estimated using the automatic differentiation capability of *Fenics* package.

4 Pattern formation in RD model

Turing's work showed that two or more interacting components diffusing at significantly different rates can create spontaneous, self-organizing spatial concentration patterns. In this section, we first present the analysis for MS-based RD systems, followed by a discussion on the inference of these systems. However, the ideas presented here apply to the general n-component model. We illustrate these properties for a ternary mixture represented with the concentration of 2-components, (c_1, c_2),

Author Manuscript

Author Manuscript

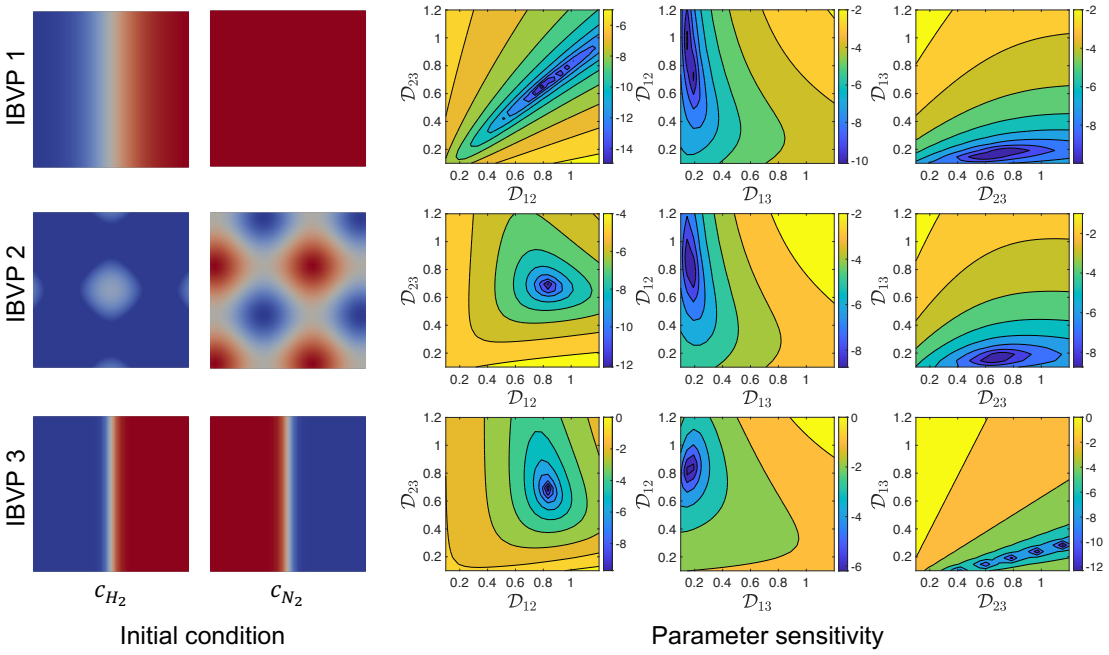


Figure 3: Parameter sensitivity for different IBVPs. Each row represent a different IBVP with the exact form provided in Eqs. (18) and (26). The left panel shows the initial concentration of the dataset. The right panel shows sensitivity contours (left to right) in the $(\mathcal{D}_{23}, \mathcal{D}_{12})$, $(\mathcal{D}_{13}, \mathcal{D}_{12})$, $(\mathcal{D}_{23}, \mathcal{D}_{13})$ planes while fixing $\mathcal{D}_{13}$, $\mathcal{D}_{23}$ and $\mathcal{D}_{12}$, respectively to the original solution. The sensitivity contours are constructed using $\log ||\boldsymbol{R}(\boldsymbol{c}^d|\boldsymbol{\theta})||$

Author Manuscript

with dynamics governed by an RD system composed of MS diffusion and Schnackenberg kinetics for reaction, resulting in the following set of equations:

$$\begin{aligned}\frac{\partial c_1}{\partial t} &= \nabla \cdot D_{11}(c_1, c_2)\nabla c_1 + \nabla \cdot D_{12}(c_1, c_2)\nabla c_2 + R_{10} + R_{11}c_1 + R_{12}c_1^2 c_2 \\ \frac{\partial c_2}{\partial t} &= \nabla \cdot D_{21}(c_1, c_2)\nabla c_1 + \nabla \cdot D_{22}(c_1, c_2)\nabla c_2 + R_{20} + R_{21}c_1 + R_{22}c_1^2 c_2 \end{aligned} \tag{27}$$

Here, the terms $D_{ij}(c_1, c_2) = (\boldsymbol{B}^{-1})_{ij}$ represent the nonlinear diffusion obtained by solving Eq. (16). The reactive dynamics represented by $r_1 = R_{10} + R_{11}c_1 + R_{12}c_1^2 c_2$, and $r_2 = R_{20} + R_{21}c_1 + R_{22}c_1^2 c_2$ is known as the Schnackenberg dynamics[1] and is the basis for many biological patterns. This system contains 3 parameters for diffusion, namely $(\mathscr{D}_{12}, \mathscr{D}_{23}, \mathscr{D}_{13})$, and 6 parameters for reaction, namely $(R_{10}, R_{11}, R_{12}, R_{20}, R_{21}, R_{22})$. It is also worth noting that the dynamics of the concentration of the 3^{rd} species can be determined using the relationship:

$$\frac{\partial c_3}{\partial t} = -\frac{\partial c_1}{\partial t} - \frac{\partial c_2}{\partial t}, \qquad c_3(\boldsymbol{x}, 0) = 1 - c_1(\boldsymbol{x}, 0) - c_2(\boldsymbol{x}, 0) \tag{28}$$

4.1 Turing's analysis with nonlinear diffusion

In this section, we extend Turing's analysis to MS-based RD systems. Turing instability is a linear instability that is observed in some RD systems for specific choices of parameters. To observe this phenomenon, one starts with a homogeneous IBVP, i.e. $\boldsymbol{c}(\boldsymbol{x}, 0) \equiv \boldsymbol{c}_0$ is a constant vector in space. This constant vector $\boldsymbol{c}_0$ is selected such that $r_i(\boldsymbol{c}) = 0$ for all $i \in \{1, \cdots, n\}$, i.e. the initial homogeneous condition is a solution to the system of nonlinear reaction equations. As the solution is homogeneous, Eq. (15) results in the absence of diffusive flux at all points in the domain. Consequently, this homogeneous solution is also a solution to the full RD system, and one should expect that if this homogeneous solution is attained at any time during evolution, then the system will remain at this solution for all future times. However, for certain choices of parameters, if this solution is perturbed from its equilibrium, these perturbations may grow, causing the system to evolve to a non-homogeneous equilibrium state. A typical observation of this phenomenon is presented in Fig. 4 where a numerical solution to Eq. (27) is presented on $\Omega = [-5, 5]^2$ with parameters $(\mathscr{D}_{12}, \mathscr{D}_{13}, \mathscr{D}_{23}) = (0.833, 0.680, 0.168)$ and $(R_{10}, R_{11}, R_{12}, R_{20}, R_{21}, R_{22}) = (0, -20, 7200, 0, 20, -7200)$. The simulation is started at the homogeneous state $\boldsymbol{c}_0 = (0.0527, 0.0527)$ with a perturbation of $10^{-4}\sin(||x||)/(1e-3+||x||)$. We can observe that this homogeneous state is a solution to the Schnackenberg equation for the prescribed parameters. Moreover, the magnitude of perturbation is small in comparison to the homogeneous solution. In the simulations, we first focus on the evolution of c_1 and c_2, noting that the fields rapidly move towards a non-homogeneous solution in the first few time steps, forming concentric rings of localized high and low concentrations. The changes in these concentration fields become slower as time progresses, and beyond a certain time, there is no discernible change in these concentration fields. The inhomogeneous nature of these equilibrium fields is what we refer to as patterns in RD systems. Turing[3] explained the onset of this instability that we observed at $t = 0$ using the stability of the linearized RD system.

Noting that the initial condition is a homogeneous equilibrium solution with some added perturbation, we consider the following general form of perturbed solution about an equilibrium point:

$$c_i(\boldsymbol{x}) = \overline{c}_i + \widetilde{c}_i(\boldsymbol{x})$$

where $\overline{\boldsymbol{c}} = (\overline{c}_1, \cdots, \overline{c}_{n-1})$ represent the equilibrium solution and the $\widetilde{c}_i$ represent the perturbations

Author Manuscript

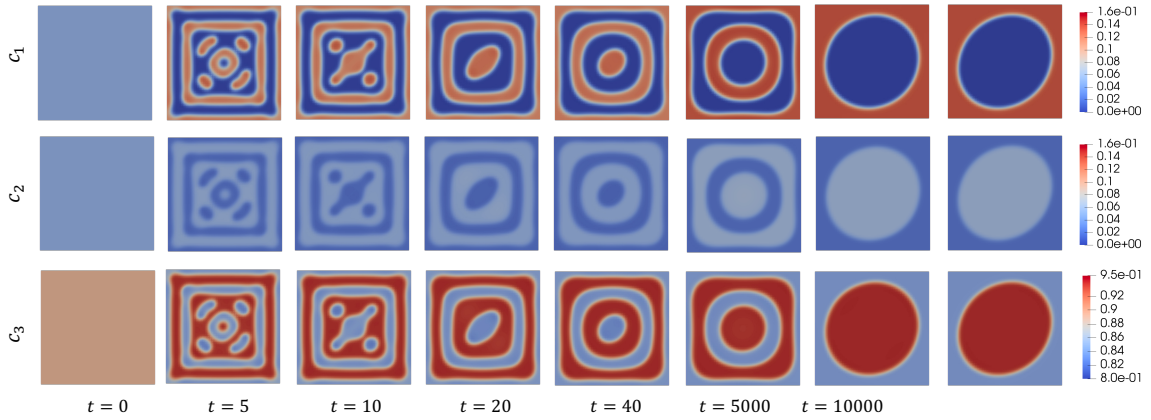


Figure 4: A numerical solution of Eq. (27) on $\Omega = [-5, 5]^2$ with parameters $(\mathscr{D}_{12}, \mathscr{D}_{13}, \mathscr{D}_{23}) = (0.833, 0.680, 0.168)$ and $(R_{10}, R_{11}, R_{12}, R_{20}, R_{21}, R_{22}) = (0, -20, 7200, 0, 20, -7200)$. The simulation is started with the unstable equilibrium solution $\boldsymbol{c}(\boldsymbol{x}, 0) = (0.0527, 0.0527)$ with a perturbation of $10^{-4}\sin(||x||)/(1e-3+||x||)$

about the i^{th} component equilibrium solution. Substituting this form in Eq. (12) gives

$$\frac{\partial \widetilde{c}_i}{\partial t} = \sum_j D_{ij}(\overline{\boldsymbol{c}})\Delta\widetilde{c}_j + \frac{\partial r_i}{\partial c_j}(\overline{\boldsymbol{c}})\widetilde{c}_j \tag{29}$$

where we used the fact that $\overline{\boldsymbol{c}}$ is an equilibrium point i.e. for all $i \in \{1, \cdots, n-1\}$, we have $\frac{\partial \overline{c}_i}{\partial t} = \sum_j D_{ij}(\overline{\boldsymbol{c}})\Delta\overline{c}_j + r_i(\overline{\boldsymbol{c}})$. Moreover, the homogeneous solution is gradient-free, i.e. for all $i \in \{1, \cdots, n-1\}$ and $i \in \{1, \cdots, n_d\}$, we have $\frac{\partial \overline{c}_j}{\partial x_i} = 0$. Eq. (29) approximates the dynamics of perturbation growth when the solution is *close* to a homogeneous equilibrium solution. For the special case of Schnackenberg dynamics, this equation gives:

$$\begin{aligned}
\frac{\partial \widetilde{c}_1}{\partial t} &= \overline{D}_{11}\Delta\widetilde{c}_1 + \overline{D}_{12}\Delta\widetilde{c}_2 + \overline{R}_{11}\widetilde{c}_1 + \overline{R}_{12}\widetilde{c}_2 \\
\frac{\partial \widetilde{c}_2}{\partial t} &= \overline{D}_{21}\Delta\widetilde{c}_1 + \overline{D}_{22}\Delta\widetilde{c}_2 + \overline{R}_{21}\widetilde{c}_1 + \overline{R}_{22}\widetilde{c}_2
\end{aligned} \tag{30}$$

where:

$$\begin{aligned}
\overline{R}_{11} &= R_{11} + 2R_{12}\overline{c}_1\overline{c}_2 \\
\overline{R}_{12} &= R_{12}\overline{c}_1^2 \\
\overline{R}_{21} &= R_{21} + 2R_{22}\overline{c}_1\overline{c}_2 \\
\overline{R}_{22} &= R_{22}\overline{c}_1^2
\end{aligned}$$

We consider solutions of the following form:

$$\widetilde{c}_\alpha = \sum_k T_{\alpha,k} e^{i\boldsymbol{k}\cdot\boldsymbol{x}} e^{\omega_k t} \tag{31}$$

Author Manuscript

with wave vector $\boldsymbol{k}$ and frequency ω_k for the plane wave mode k. Substituting in Eq. (30):

$$\left(|\boldsymbol{k}|^2\overline{D}_{11} - \overline{R}_{11} + \omega_k\right) T_{1,k} + \left(|\boldsymbol{k}|^2\overline{D}_{12} - \overline{R}_{12}\right) T_{2,k} = 0$$
$$\left(|\boldsymbol{k}|^2\overline{D}_{21} - \overline{R}_{21}\right) T_{1,k} + \left(|\boldsymbol{k}|^2\overline{D}_{22} - \overline{R}_{22} + \omega_k\right) T_{2,k} = 0$$

For non-trivial $T_{1,k}$ and $T_{2,k}$, the frequency and wave number must satisfy:

$$0 = w_k^2 + b\omega_k + c \tag{32}$$

where

$$b = |\boldsymbol{k}|^2(\overline{D}_{11} + \overline{D}_{22}) - (\overline{R}_{11} + \overline{R}_{22})$$
$$c = \left(\left(|\boldsymbol{k}|^2\overline{D}_{11} - \overline{R}_{11}\right)\left(|\boldsymbol{k}|^2\overline{D}_{22} - \overline{R}_{22}\right) - \left(|\boldsymbol{k}|^2\overline{D}_{12} - \overline{R}_{12}\right)\left(|\boldsymbol{k}|^2\overline{D}_{21} - \overline{R}_{21}\right)\right)$$

This quadratic equation has the solutions, $\omega_k = 1/2(-b \pm \sqrt{b^2 - 4c})$. We will refer to these solutions as the branches. The form in Eq. (31) reveals that if $\text{Real}(\omega_k) > 0$ for any value of k and any branch, then the corresponding mode will grow over time. This growth is typically expected only close to the homogeneous equilibrium solution, $\overline{\boldsymbol{c}}$. As these modes grow, the state departs from the homogeneous solution and may stabilize at non-homogeneous equilibrium points that we refer to as *'patterns'*. One way of studying this growth in the linearized regime is via stability diagrams. In this visualization, we plot $\text{Real}(\omega_k)$ for the different branches for a range of wave numbers k. If any region of the figure lies in the first quadrant, then the system is linearly unstable. Moreover, the values of k for which $\text{Real}(\omega_k) > 0$ determine the modes that will grow. In particular, one may identify the different modes in the perturbation and use this visualization to determine the unstable ones. The unstable branch of the system discussed in Fig. 4 is presented in Fig. 7. In this case, all modes are observed to be unstable. In another case study presented in Sec. 4.2, we get a stability diagram (shown in Fig. 5) where not all modes are unstable. It should be remarked that the patterns are formed when the system's state evolves beyond a small neighborhood of homogeneous equilibrium solution. Therefore, this stability analysis is only valid at the onset of the linear instability and not at all subsequent times. It should also be observed that different nonlinear systems can have the same onset to instability, i.e. two different systems may have the same instability diagram. For instance, consider the following two RD systems with the same linearized reactions but different diffusivities.

$$\begin{aligned}
&\text{Case 1}: \quad \begin{bmatrix}\overline{R}_{11} & \overline{R}_{12}\\ \overline{R}_{21} & \overline{R}_{22}\end{bmatrix} = \begin{bmatrix}-1 & 1\\ -1 & 1\end{bmatrix}, \quad \begin{bmatrix}\overline{D}_{11} & \overline{D}_{12}\\ \overline{D}_{21} & \overline{D}_{22}\end{bmatrix} = \begin{bmatrix}1 & 0\\ 0 & 1/100\end{bmatrix}\\
&\text{Case 2}: \quad \begin{bmatrix}\overline{R}_{11} & \overline{R}_{12}\\ \overline{R}_{21} & \overline{R}_{22}\end{bmatrix} = \begin{bmatrix}-1 & 1\\ -1 & 1\end{bmatrix}, \quad \begin{bmatrix}\overline{D}_{11} & \overline{D}_{12}\\ \overline{D}_{21} & \overline{D}_{22}\end{bmatrix} = \begin{bmatrix}505/1000 & 99(\sqrt{2}+1)/200\\ 99(\sqrt{2}-1)/200 & 505/1000\end{bmatrix}
\end{aligned} \tag{33}$$

Solving Eq. (32) for both these cases results in the same frequency-wave number relationship:

$$\omega_k = \frac{1}{2}\left(-\frac{101}{50}|\boldsymbol{k}|^2 \pm \sqrt{\left(2 + \frac{99}{100}|\boldsymbol{k}|^2\right)^2 - 4}\right)$$

Therefore, these two cases behave similarly at the homogeneous equilibrium state. It is also worth noting that in case 1, the components are diffusing at significantly different rates; meanwhile, in case 2, the instability arises due to cross-diffusion terms. Regarding physical mechanisms of diffusion,

Author Manuscript

these cases represent two entirely different approaches to the onset of pattern formation. These mechanisms have been previously studied in literature[5,6] in the context of Fickian diffusion. In the next section, we will discuss the physical mechanisms in the MS-based RD model.

4.1.1 Physical interpretation of Turing instability in MS-based RD systems

Turing instabilities were initially studied for a Fickian diffusion-based RD system[3]. In this case, diffusivities D_{ij} are independent of concentrations. The simplicity of the Fick-based RD system allows for an activator-inhibitor interpretation of the morphogens that lead to instability and, consequently, steady-state patterns. To understand this interpretation, we can consider the linearized system in Eq. (30). For this brief excursion into Fick-based RD systems, we will ignore the species, c_3, that in the case of MS-based diffusion is determined as $c_3(c_1, c_2) = 1 - c_1 - c_2$ (conservation of total molar concentration). Moreover, we will ignore the dependence of linearized coefficients $\overline{D}_{ij}$ and $\overline{R}_{ij}$ on the equilibrium homogeneous molar concentrations, $(\bar{c}_1, \bar{c}_2)$. In this setting, we will refer to $\overline{R}_{11}, \overline{R}_{22} > 0$ as auto-activation, $\overline{R}_{11}, \overline{R}_{22} < 0$ as auto-inhibition, $\overline{R}_{12}, \overline{R}_{21} > 0$ as cross-activation, and $\overline{R}_{12}, \overline{R}_{21} < 0$ as cross-inhibition. Moreover, for diffusion, thermodynamic admissibility requires that self-diffusion $\overline{D}_{11}, \overline{D}_{22} > 0$ meanwhile cross-diffusion $(\overline{D}_{12}, \overline{D}_{21})$ can be positive or negative. One of the simplest cases with instability in this activator-inhibitor setting is in the absence of cross-diffusion $(\overline{D}_{12} = \overline{D}_{21} = 0)$ with fast inhibitor (species 1) and slow activator (species 2) $(\overline{D}_{11} >> \overline{D}_{22})$. Here, inhibitors are species with auto and cross-inhibitions $(\overline{R}_{11}, \overline{R}_{12} < 0)$ and activators are species with auto and cross-activations $(\overline{R}_{22}, \overline{R}_{21} > 0)$. An example of this scenario is presented in case 1 of Eq. (33). In this scenario, if there is a perturbation that results in a locally higher concentration of activator, the activator results in a local increase of the inhibitor. However, inhibitors having a faster spreading rate result in a long-range inhibition. Consequently, pushing the system towards a non-homogeneous solution with low and high concentrations. This process is often referred to as short-range activation and long-range inhibition. It should also be noted that this onset to instability doesn't require different self-diffusions as in case 1 of Eq. (33). It is shown in case 2 of Eq. (33) that the same $(\omega_k - \boldsymbol{k})$ relationship can be obtained with species with the same self-diffusion but different cross-diffusions. In this scenario, the different spreading rate in the activator-inhibitor model is achieved through this cross-diffusion mechanism.

It is evident from the above discussion that various diffusive mechanisms can lead to pattern formation. It is, therefore, important to have a physically accurate description of diffusion mechanisms in real-life systems. In Sec. 2.2, we considered Duncan and Toor's experiment that serves as a case study where Fickian diffusion is inadequate to describe transport in a ternary system. MS diffusion describes a more general class of diffusive mechanisms as discussed in Sec. 2.2. At the same time, MS-diffusion remains thermodynamically consistent[32] for positive coefficients, $\mathscr{D}_{ij} > 0$. The form of diffusion provided in Eq. (16) shows a concentration-dependent form. It can be observed that at low concentrations i.e. $c_1, c_2 << 1$, Eq. (16) gives:

$$\boldsymbol{B}^{-1} \approx \begin{bmatrix} \mathscr{D}_{13} & 0 \\ 0 & \mathscr{D}_{23} \end{bmatrix}$$

Therefore, if the frictional forces between species (1-3) and species (2-3), given as $\mathscr{D}_{13}^{-1}$ and $\mathscr{D}_{23}^{-1}$ are very different, we may obtain the fast and slow diffusing species that sometimes form the basis for the activator-inhibitor model. Even at higher densities, the relative values of frictional forces are essential in understanding the diffusion mechanisms; for instance, at the onset of instability in the

Author Manuscript

Table 2: Original and inferred parameters for MS Diffusion and Schnackenberg reaction system.

	$\mathscr{D}_{12}$	$\mathscr{D}_{23}$	$\mathscr{D}_{13}$	R_{10}	R_{11}	R_{12}	R_{20}	R_{21}	R_{22}
Exact A	5.00e0	5.00e0	5.00e-2	1.00e-1	-2.00e+1	7.20e+3	1.00e0	0.00e0	-7.20e+3
Model A1	5.00e0	5.00e0	8.59e-2	1.35e-1	-2.77e+1	9.99e+3	7.65e-1	9.69e0	-9.99e+3
Exact B	8.33e-1	6.80e-1	1.68e-1	0.00e0	-2.00e+1	7.20e+3	0.00e0	2.00e+1	-7.20e+3
Model B1	1.19e0	9.43e-1	2.33e-1	5.82e-4	-2.78e+1	9.99e+3	3.34e-4	2.78e+1	-9.99e+3
Model B2	1.17e0	9.44e-1	2.33e-1	2.32e-4	-2.78e+1	9.99e+3	1.30e-4	2.78e+1	-9.99e+3
Model B3	1.16e0	9.44e-1	2.33e-1	1.25e-4	-2.78e+1	9.99e+3	5.33e-5	2.78e+1	-9.99e+3
Model B4	3.58e0	9.80e-1	3.26e-1	4.55e-2	-2.86e+1	9.99e+3	-4.01e-2	2.85e+1	-9.99e+3

example presented in Fig. 4, the coefficients of linearized RD system Eq. (30) evaluates to be:

$$\begin{bmatrix}\overline{R}_{11} & \overline{R}_{12}\\ \overline{R}_{21} & \overline{R}_{22}\end{bmatrix} = \begin{bmatrix}39.9860 & 19.9930\\ -39.9860 & -19.9930\end{bmatrix}, \quad \begin{bmatrix}\overline{D}_{11} & \overline{D}_{12}\\ \overline{D}_{21} & \overline{D}_{22}\end{bmatrix} = \begin{bmatrix}0.1755 & -0.0302\\ -0.0017 & 0.6869\end{bmatrix} \tag{34}$$

Here, we observe that the self-diffusions are of similar magnitude while there is the presence of cross-diffusion terms. These examples show that MS-based RD systems can exhibit equivalent diffusion tensors as Fickian diffusion at the onset of instability. However, MS diffusion remains thermodynamically consistent for all molar concentrations in the permissible range. More interestingly, the concentration-dependent diffusion implies that a spatially and temporally varying diffusion is observed as concentration evolves after being perturbed from the unstable equilibrium. For instance, in some regions, the system may evolve with slow and fast diffusing species, while in others, it may evolve with cross-diffusion. This out-of-equilibrium dynamics is significantly more complex than Fick-based RD systems. On the other hand, this rich dynamics complicates the inverse problem of inferring parameters in MS-based RD systems due to the nonlinearities introduced by the diffusion process in the optimization problem of Eq. (24).

4.2 Data-driven inference of governing equations for pattern formation

In this section, we present two numerical studies to demonstrate the application of the data-driven system inference method outlined in Sec. 3.2 to pattern formation. The coefficients of the exact and inferred models are presented in Tab. 2.

4.2.1 Data generation

We consider two particular instances of the RD model with MS diffusion. The parameters are prescribed in Table 2. In both cases, we infer a model by identifying the coefficients of the reaction basis terms as they appear in the Schnackenberg model, i.e. $\{c_1, c_2, c_1^2 c_2\}$ for both components. The models are inferred by solving the optimization problem of Eq. (24).

Case A: The model with parameters labeled *Exact A* is solved on the domain $\Omega = [-5, 5]^2$ with $t \in [0, 2]$. The initial conditions are chosen to be the perturbed unstable equilibrium as $\boldsymbol{c}(\boldsymbol{x}, 0) = (0.055, 0.046) + \boldsymbol{\delta}$ where the perturbation, $\boldsymbol{\delta} = 0.01(\cos(10x) + \cos(10y), \sin(10x) + \sin(10y))$.

Case B: The model with parameters labeled *Exact 2* is solved on the domain $\Omega = [-5, 5]^2$ with $t \in [0, 100]$. The initial conditions are chosen to be the perturbed unstable equilibrium as $\boldsymbol{c}(\boldsymbol{x}, 0) = (0.0527, 0.0527) + \boldsymbol{\delta}$ where the perturbation, $\boldsymbol{\delta} = 10^{-4} \sin(||x||)/(0.001 + ||x||)(1, 1)$.

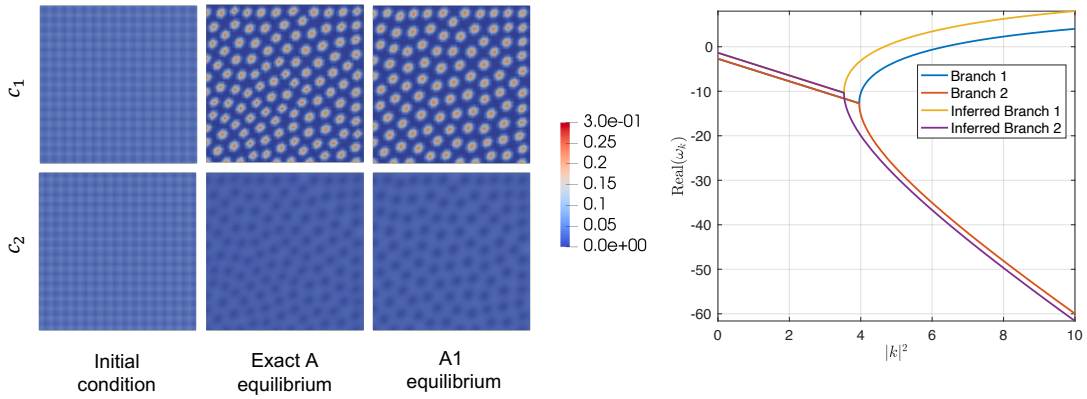


Figure 5: Comparison of original and inferred model for case 1. The left panel shows the state of the original and inferred model at $t = 2$. The right panel shows the stability diagram for the original and inferred model.

4.2.2 Recovery of patterns and stability diagrams

In this subsection, we will look at the results for inference of *Case A*. We first observe the steady state behavior in this case (Fig. 5, left panel) forms *spot*-patterns. The coefficients of the inferred model are provided in Tab. 2 with label, $A1$. We start with considering the reactive part of the inferred model and solve the reaction equations to estimate the homogeneous equilibrium solution, $\boldsymbol{c}(\boldsymbol{x}) = (0.05, 0.05)$. This solution is very close to the exact model's solution, $\boldsymbol{c}(\boldsymbol{x}) = (0.0527, 0.0527)$. Next, we solve the complete RD system with the inferred parameters A1 but with the initial conditions used to generate data. We observe that the equilibrium state of the inferred model (see Fig. 5) also shows a spotted pattern. The pattern formation process in the RD system is sensitive to initial conditions. Therefore, it is not possible to replicate the same steady-state state pattern from different models. However, we obtain similar patterns. We then construct and compare the stability diagram of the exact and the inferred models (see Fig. 5, right panel) and note that they are very similar. A similar stability diagram means that if two models start at the homogeneous unstable equilibrium, the same modes will grow and possibly lead to some steady state. However, the stability diagram doesn't directly determine the final patterns that are formed or the out-of-equilibrium evolution of the solution. This numerical study shows that the VSI technique can capture salient characteristics of the instabilities near the unstable equilibrium point and the long time patterning behavior of the system. There are two key takeaways from this numerical study:

a. Similar patterns can be obtained from different models. This property will impact biological systems (discussed in the concluding remarks).

b. VSI can capture the characteristic features of instability and the long-term behavior or patterning in the RD system. However, VSI could not recover the coefficient of the original system. This phenomenon is attributed to the nonlinear optimization of (24). The minimization problem has multiple local minima, and the algorithm identifies one of them as the solution.

In the following subsection, we will discuss ways to better inform the inference by choosing timestamps at which data is being collected.

Author Manuscript

Author Manuscript

4.2.3 Effect of time slicing on inference

In this subsection, we will consider the inference of case B. We will solve this inference problem four times with data collected at different time points at $t \in \{0, \Delta t, 2\Delta t, \cdots n_T \Delta t\}$ where we consider the following cases:

- *Sub-case B1*: simulation range: $t \in [0, 20]$, time discretization $n_T = 200, \Delta t = 0.1$
- *Sub-case B2*: simulation range: $t \in [0, 50]$, time discretization $n_T = 500, \Delta t = 0.1$
- *Sub-case B3*: simulation range: $t \in [0, 100]$, time discretization $n_T = 1000, \Delta t = 0.1$
- *Sub-case B4*: simulation range: $t \in [0, 1]$, time discretization $n_T = 100, \Delta t = 0.01$

The numerical results for the inferred model are presented in Tab. 2, labeled as B1-4. We again observe that in all cases, B1-4 inferred values differ from the original system. The solutions to the inferred models B1-4 for $t \in [0, 10000]$ are provided in Fig. 6. We reemphasize that cases B1, B2, B3 and B4 were trained with initial time series data up to $t = 20$, $t = 50$, $t = 100$ and $t = 1$ respectively. The forward solutions of the inferred models (presented in Fig. 6) are carried out for a longer time range of $t \in [0, 10000]$ to determine the steady state solution. We observe that all inferred models resemble the square patterns in the original data in the initial few time steps. However, sub-case B1 deviates from the data when it continues to evolve beyond its training range. Interestingly, the equilibrium solutions to sub-case B2 and sub-case B3 resemble the data, illustrating that these models can be extrapolated to longer times than the data with which they were trained. It is also seen that including a more extended range for data gathering B4 to B1 to B2 to B3 results in a better recovery of steady-state solution. However, when we compare the stability diagram of the inferred and exact models (see Fig. 7), we observe that the sub-cases B1, B2, and B3 are overpredicting the growth rates ($\text{Real}(\omega_k)$) for all wavenumbers k. Moreover, the stability diagrams for cases $B1$ to $B3$ are similar (overlapping in the Fig. 7). Meanwhile, the sub-case B4, which contains data for a small range of $t \in [0, 1]$ with refined timesteps of $\Delta t = 0.01$, shows considerable improvement in recovering the stability diagram. There are two key takeaways from this study:

a. Increasing the time range allows VSI to identify model parameters that recover the patterns.

b. Refining time in the initial state close to the unstable equilibrium allows VSI to identify model parameters that recover the stability diagram

5 Concluding remarks

RD systems form a general class of PDEs that model pattern formation in biological systems. These models are typically studied in the context of Fick-based diffusion that defines a linear relationship between flux and the gradient of the concentration field. This idea is generalized for multicomponent transport by constructing flux as a linear combination of concentration gradients of each species. However, this relationship is only valid for low concentrations and, in general, is not thermodynamically consistent. MS diffusion provides a more general framework for multicomponent transport that is valid for arbitrary concentrations. We illustrated in Sec. 2.2 that MS diffusion includes mechanisms for intermolecular force balance in the diffusion flux that allow us to describe phenomena like osmotic and reverse diffusion. Constructing an RD system with this mechanism will lead to a richer class of dynamics while being grounded in physics. We have also considered cases in Sec. 4.1.1, where at low concentrations, one can have slow and fast species that lead to pattern formation with the guiding physical principle of *short-range activators coupled with long-range*

Author Manuscript

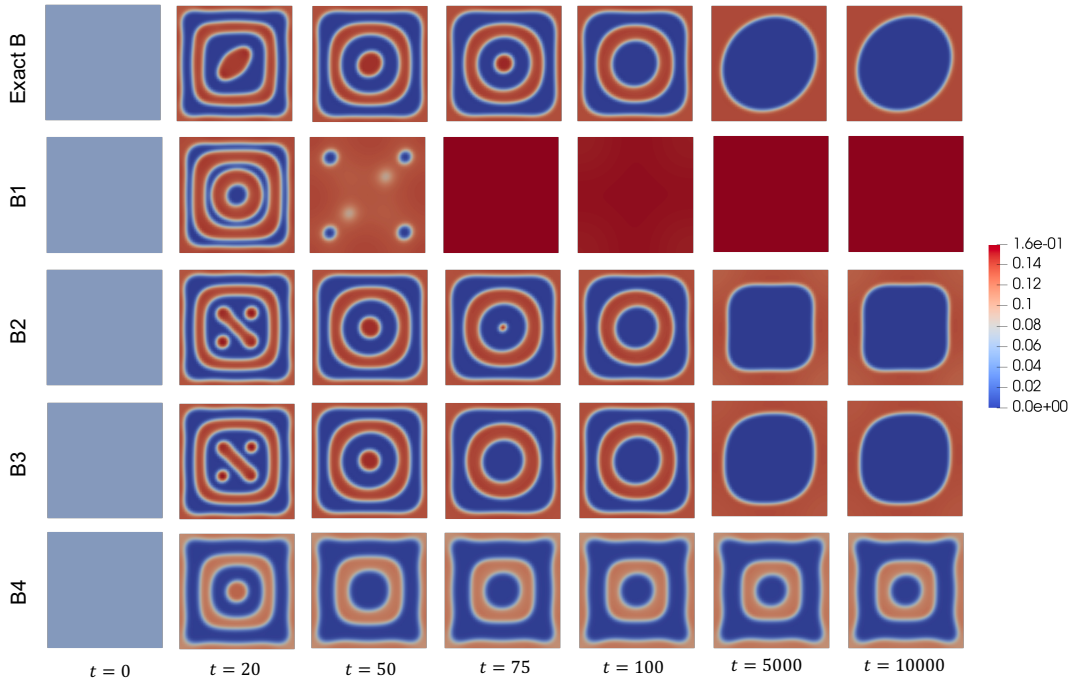


Figure 6: Spatiotemporal evolution of c_1 described by an RD system with parameters provided in Tab. 2.

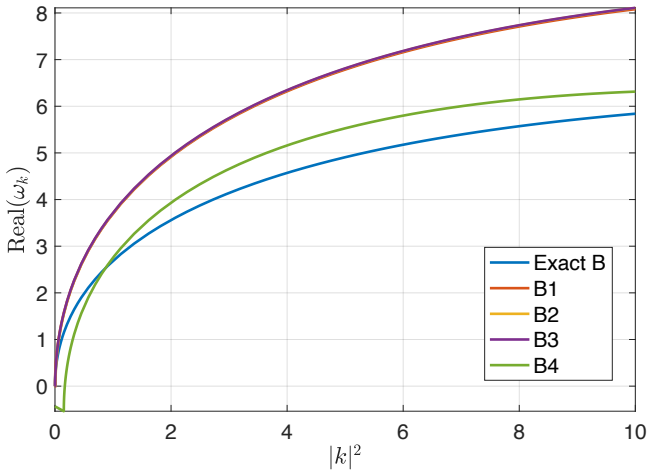


Figure 7: Unstable mode in the stability analysis for the models described in Tab. 2 (Exact B, B1, B2m B3, B4). All diagrams are evaluated at the equilibrium point $\boldsymbol{c}(\boldsymbol{x}) = (0.0527, 0.0527)$. The curves corresponding to B1, B2 and B3 overlap and cannot be distinguished in this plot.

Author Manuscript

inhibitors. We also showed that cross-diffusion can trigger these instabilities for pattern formation. MS-based diffusion allows a natural way to incorporate cross-diffusion at higher concentrations while maintaining thermodynamic consistency.

Meanwhile, simple reaction dynamics can capture a relatively broad class of patterns[2]. In our work, we observed spots and ring patterns with Schnackenberg reaction dynamics. The MS-based RD system can be very quickly coupled with other reaction models studied in pattern formation, like the FitzHugh-Nagumo and the Gray-Scott models, offering a general framework for studying pattern formation while maintaining thermodynamic consistency. One should, however, practice caution when using these general reaction systems. The initial conditions and parameter space of this system must reflect the positivity-preserving property in the reaction dynamics required by the MS-based framework. The other restriction of this framework, the no-net production property, is automatically satisfied by the dynamics of the reduced system. At the same, this property offers more insight into the mechanics of pattern formation in these systems.

No-net production property in MS-based RD system: In a recent study by Braun et.al.[33], they illustrated Turing instabilities and patterning with mass-conserving reaction dynamics.

$$\frac{\partial c_1}{\partial t} = \Delta c_1 + f(c_1, c_2)$$
$$\frac{\partial c_2}{\partial t} = \Delta c_2 - f(c_1, c_2)$$

the implication being that we get source-free transient diffusion for $c_1 + c_2$. A central assumption in MS-based RD systems is that the total molar concentration remains constant when considering the complete $n-$component system. As a consequence, the ternary systems that we have illustrated in Sec. 4.2 follow concentration-conserving reaction dynamics. Therefore, the patterns that we observe in this system are formed by redistribution and interchanging of species in the domain. This contrasts mechanisms studied in the general form of RD dynamics, where a local concentration change may alter the global concentration.

Implications in biological patterning and morphogenesis: RD systems form one of the two primary schools of thought in the genesis of natural pattern formation, the other being mechanochemical systems[1,34]. In the former approach, different species are treated as morphogens (or hypothetical chemicals) that can react and diffuse to form patterns. In many biological systems, these morphogens may interact via physical contact; for instance, the physics of pattern formation in tissues is dictated by the transport of a mixture of cells. Practical examples of such systems include embryogenesis, tissue differentiation, vasculogenesis, neural tube formation, invasive tumor growth[1,34,35,36,37]. Modeling these phenomena must be done considering the high cell densities, warranting the need for more sophisticated modeling of diffusion mechanisms. As we discussed in Sec. 4.1.1, the spatial variations in densities may lead to different mechanisms for diffusion in different parts of the domain.

MS-based RD system offers a general yet simple representation for modeling these complex biological behavior. The initial perturbation dependence of the final pattern also captures the uncertainty and variability we observe in nature. Moreover, the nonlinearity in the inference formulation shows a possible existence of multiple mechanisms (local minima in parameter space) that can explain pattern formation with these models. The class of patterns studied with this approach can be further broadened by including more pattern-forming reactive models like the FitzHugh-Nagumo model to determine the best reactive operators for the data. The VSI approach developed in the work can be integrated with operator selection techniques that we have presented in some previous works[24,25,22],

allowing us to identify a parsimonious representation for reactive dynamics in a data-driven fashion. In future, we will generalize this work to non-equimolar settings[38] that will allow us to model a more general class of biological systems, such as the ones with significant cell proliferation and death.

Author Manuscript

Author Manuscript

Bibliography

1. Garikipati, K. Perspectives on the mathematics of biological patterning and morphogenesis. *Journal of the Mechanics and Physics of Solids* **99**, 192–210 (2017).

2. Kondo, S. & Miura, T. Reaction-diffusion model as a framework for understanding biological pattern formation. *Science* **329**, 1616–1620 (2010). URL `https://www.science.org/doi/abs/10.1126/science.1179047`. `https://www.science.org/doi/pdf/10.1126/science.1179047`.

3. Turing, A. M. The chemical basis of morphogenesis. *Bulletin of mathematical biology* **52**, 153–197 (1990).

4. Green, J. B. & Sharpe, J. Positional information and reaction-diffusion: two big ideas in developmental biology combine. *Development* **142**, 1203–1211 (2015).

5. Madzvamuse, A., Ndakwo, H. S. & Barreira, R. Cross-diffusion-driven instability for reaction-diffusion systems: analysis and simulations. *Journal of mathematical biology* **70**, 709–743 (2015).

6. Vanag, V. K. & Epstein, I. R. Cross-diffusion and pattern formation in reaction–diffusion systems. *Physical Chemistry Chemical Physics* **11**, 897–912 (2009).

7. Gorban, A. N., Sargsyan, H. & Wahab, H. A. Quasichemical models of multicomponent nonlinear diffusion. *Mathematical Modelling of Natural Phenomena* **6**, 184–262 (2010).

8. Krishna, R. & Wesselingh, J. The maxwell-stefan approach to mass transfer. *Chemical Engineering Science* **52**, 861–911 (1997). URL `https://www.sciencedirect.com/science/article/pii/S0009250996004587`.

9. Darken, L. S. Diffusion, mobility and their interrelation through free energy in binary metallic systems. *Trans. Aime* **175**, 184–201 (1948).

10. Vignes, A. Diffusion in binary solutions. variation of diffusion coefficient with composition. *Industrial & Engineering Chemistry Fundamentals* **5**, 189–199 (1966).

11. D'Agostino, C., Mantle, M., Gladden, L. & Moggridge, G. Prediction of binary diffusion coefficients in non-ideal mixtures from nmr data: Hexane–nitrobenzene near its consolute point. *Chemical Engineering Science* **66**, 3898–3906 (2011). URL `https://www.sciencedirect.com/science/article/pii/S0009250911003150`.

12. Duncan, J. B. & Toor, H. An experimental study of three component gas diffusion. *AIChE Journal* **8**, 38–41 (1962).

13. Bothe, D. *On the Maxwell-Stefan Approach to Multicomponent Diffusion*, 81–93 (Springer Basel, Basel, 2011). URL `https://doi.org/10.1007/978-3-0348-0075-4_5`.

Author Manuscript

14. Ramm, B. *et al.* A diffusiophoretic mechanism for atp-driven transport without motor proteins. *Nature Physics* **17**, 850–858 (2021).

15. Yang, K. Dynamic binary protein adsorption in ion-exchange media depicted with a parallel diffusion model derived from maxwell–stefan theory. *Chemical Engineering Science* **139**, 163–172 (2016).

16. Sun, Y. & Yang, K. Analysis of mass transport models based on maxwell–stefan theory and fick's law for protein uptake to porous anion exchanger. *Separation and Purification Technology* **60**, 180–189 (2008).

17. Mazumder, A., Dobyns, B. M., Howard, M. P. & Beckingham, B. S. Theoretical and experimental considerations for investigating multicomponent diffusion in hydrated, dense polymer membranes. *Membranes* **12** (2022). URL `https://www.mdpi.com/2077-0375/12/10/942`.

18. Fowler, K., Connolly, P. J., Topping, D. O. & O'Meara, S. Maxwell–stefan diffusion: a framework for predicting condensed phase diffusion and phase separation in atmospheric aerosol. *Atmospheric Chemistry and Physics* **18**, 1629–1642 (2018).

19. Herberg, M., Meyries, M., Prüss, J. & Wilke, M. Reaction–diffusion systems of maxwell–stefan type with reversible mass-action kinetics. *Nonlinear Analysis* **159**, 264–284 (2017).

20. Giacomazzi, E., Picchia, F. & Arcidiacono, N. A review of chemical diffusion: Criticism and limits of simplified methods for diffusion coefficient calculation. *Combustion Theory and Modelling* **12**, 135–158 (2008).

21. Brunton, S. L., Proctor, J. L. & Kutz, J. N. Discovering governing equations from data by sparse identification of nonlinear dynamical systems. *Proceedings of the National Academy of Sciences* **113**, 3932–3937 (2016). URL `https://www.pnas.org/doi/abs/10.1073/pnas.1517384113`. `https://www.pnas.org/doi/pdf/10.1073/pnas.1517384113`.

22. Wang, Z., Huan, X. & Garikipati, K. Variational system identification of the partial differential equations governing microstructure evolution in materials: Inference over sparse and spatially unrelated data. *Computer Methods in Applied Mechanics and Engineering* **377**, 113706 (2021).

23. Nikolov, D. P. *et al.* Ogden material calibration via magnetic resonance cartography, parameter sensitivity and variational system identification. *Philosophical Transactions of the Royal Society A* **380**, 20210324 (2022).

24. Wang, Z., Carrasco-Teja, M., Zhang, X., Teichert, G. H. & Garikipati, K. System Inference Via Field Inversion for the Spatio-Temporal Progression of Infectious Diseases: Studies of COVID-19 in Michigan and Mexico. *Archives of Computational Methods in Engineering* **28**, 4283–4295 (2021). URL `https://link.springer.com/10.1007/s11831-021-09643-1`.

25. Wang, Z., Huan, X. & Garikipati, K. Variational system identification of the partial differential equations governing the physics of pattern-formation: Inference under varying fidelity and

Author Manuscript

noise. *Computer Methods in Applied Mechanics and Engineering* **356**, 44–74 (2019). URL https://linkinghub.elsevier.com/retrieve/pii/S0045782519304037.

26. Bothe, D. & Étienne Druet, P. On the structure of continuum thermodynamical diffusion fluxes—a novel closure scheme and its relation to the maxwell–stefan and the fick–onsager approach. *International Journal of Engineering Science* **184**, 103818 (2023). URL https://www.sciencedirect.com/science/article/pii/S0020722523000095.

27. Bothe, D. & Dreyer, W. Continuum thermodynamics of chemically reacting fluid mixtures. *Acta Mechanica* **226**, 1757–1805 (2015).

28. Boudin, L., Grec, B. & Salvarani, F. A mathematical and numerical analysis of the maxwell-stefan diffusion equations. *Discrete and Continuous Dynamical Systems-Series B* **17**, 1427–1440 (2012).

29. Logg, A. & Wells, G. N. Dolfin: Automated finite element computing. *ACM Trans. Math. Softw.* **37** (2010). URL https://doi.org/10.1145/1731022.1731030.

30. Logg, A., Mardal, K.-A. & Wells, G. *Automated solution of differential equations by the finite element method: The FEniCS book*, vol. 84 (Springer Science & Business Media, 2012).

31. Pedregosa, F. *et al.* Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research* **12**, 2825–2830 (2011).

32. Krishna, R. & Wesselingh, J. The maxwell-stefan approach to mass transfer. *Chemical Engineering Science* **52**, 861–911 (1997). URL https://www.sciencedirect.com/science/article/pii/S0009250996004587.

33. Brauns, F., Halatek, J. & Frey, E. Phase-space geometry of mass-conserving reaction-diffusion dynamics. *Physical Review X* **10**, 041036 (2020).

34. Murray, J. D., Maini, P. K. & Tranquillo, R. T. Mechanochemical models for generating biological pattern and form in development. *Physics Reports* **171**, 59–84 (1988).

35. Gierer, A. & Meinhardt, H. A theory of biological pattern formation. *Kybernetik* **12**, 30–39 (1972).

36. Czirok, A. Endothelial cell motility, coordination and pattern formation during vasculogenesis. *Wiley Interdisciplinary Reviews: Systems Biology and Medicine* **5**, 587–602 (2013).

37. Khain, E. & Sander, L. M. Dynamics and pattern formation in invasive tumor growth. *Phys. Rev. Lett.* **96**, 188103 (2006). URL https://link.aps.org/doi/10.1103/PhysRevLett.96.188103.

38. Tai, R. C., Chang, H.-K. & Farhi, L. E. Non-equimolar counter-diffusion in ternary gas systems. *Respiration Physiology* **40**, 253–267 (1980).