

An Application of Tensors in the Stochastic Reaction Diffusion Master Equation

Md Mustafijur Rahman

0000-0002-6874-0801

Texas A&M University - San Antonio,
Department of Computational, Engineering,
and Mathematical Sciences (CEMS),
San Antonio, TX 78224, USA
Email: mrahman2@tamusa.edu

Roger B. Sidje

0000-0001-5353-3642

The University of Alabama,
Department of Mathematics,
Tuscaloosa, AL 35487, USA
Email: roger.b.sidje@ua.edu

Abstract—Various scientific fields are seeking to exploit tensors, which are multi-dimensional arrays that expand on the concept of matrices. Here, we consider an application arising from complex biological systems where the data can be processed and analyzed using tensors. Our modeling framework is the reaction-diffusion master equation (RDME) used to describe the dynamics of biological systems involving both reaction and diffusion processes. It is already notoriously hard to solve the more familiar chemical master equation (CME) that only involves reaction processes. Solving the RDME is even harder because its state space is considerably larger compared to that of the CME, and this further motivates the utilization of tensors. Our study is an illustrative example of how tensor techniques can be used to make predictions on the dynamics of a metapopulation model based on its RDME formulation.

I. INTRODUCTION

TENSORS, fundamentally, are extensions of vectors, and matrices, with the ability to represent and manipulate data in multiple dimensions. Classical linear algebra approaches can circumvent tensors by flattening the data through mechanisms known as matricization or vectorization, although doing so can increase the computational overhead, especially as dimensionality grows. By contrast, specialized tensor methods avoid unfolding the tensor into a matrix format, with the trade-off that the preserved multi-dimensional format drastically complicates the design and implementation of algorithms, not to mention their subsequent understanding and maintenance by others than the original software architects.

Tensor methods encompass the efficient utilization of tensors for the purpose of problem-solving, data analysis, and the extraction of significant patterns from high-dimensional datasets. Tensors find utility not only in the biological application considered here, but also in many other areas, including popular image processing applications (e.g., facial recognition, analysis of musical scores, and identification of cliques within social networks [1], [2]). Together with earlier works on tensors and their decompositions that go back to 1927 [3], we note that more recent literature overviews of tensor techniques for large-scale numerical computations are given in [4], [5], [6], geared towards a scientific computing audience. Of course,

there are way more applications of tensors than we can reasonably cite in a communication such as this one.

In the field of biochemistry, complex networks are shown to emerge as a result of the interactions among many cellular components, including DNA and RNA molecules. Determining the time-varying probability distribution of their states involves solving the chemical master equation (CME) [7]. The stochastic simulation algorithm (SSA) [8] is useful for analyzing the CME. An alternative method for addressing the CME is the finite state projection algorithm (FSP) [9], which amounts to solving differential equations that characterize the evolution of the probability distribution of the underlying Markov chain of the CME over time.

Spatial heterogeneity is a frequent property of many cellular habitats, which are not homogeneous in space. A number of distinct but similar mathematical techniques have been used to model stochastic reaction-diffusion systems in biological cells [10], [11]. For example, Smoluchowski proposed a modeling approach named the spatially continuous diffusion-limited reaction (SCDLR) [12]. Unlike these approaches, the reaction-diffusion master equation (RDME) [13] is obtained when we treat the diffusion at the molecular level as a unique set of reactions in the CME.

The space is partitioned into compartments in the RDME. Every compartment is thought to be thoroughly mixed, and the reactions found within a specific compartment are thought to be consistent with the homogeneous scenario. Additionally, molecules can diffuse (jump) in between neighboring compartments, and here we will model the jump processes as reactions. Similar to the CME but with a much larger dimensionality, the RDME explains the time-dependent probability distribution function (PDF) of the system state. The works in [14], [15], [16] provide instances of the SSA being used in reaction-diffusion systems with the space being one dimensional. The next reaction method (NRM) [17] is an efficient version of the SSA. A software called MesoRD [18] offers a specialized implementation of the NRM, which is designed for diffusive systems and known as the next subvolume method (NSM) [19].

In this contribution rooted in [20], our aim is to illustrate

how a tensor-based approach, specifically a quantized tensor-train decomposition, can be effectively used to manage the potentially high-dimensional state space arising in the RDME of a metapopulation model. Although the RDME shares similarities with the CME, it necessitates greater computational resources due to its considerably larger state space. We employ the tensor technique to derive the marginal distribution of the relevant species. By comparing the tensor technique with the classical FSP, this work aims to highlight the capability of tensors to reduce computational overhead while maintaining accuracy and generalizability across diverse applications.

II. PRELIMINARIES

A. Tensors

We will soon recap how the CME and RDME lead to formulations involving tensors. We begin with some background to define the notation and situate our context. More formally, tensors are multi-dimensional arrays in $\mathbb{K}^{I_1 \times \dots \times I_N}$, where N is the order, i.e., the number of dimensions, which are called modes or ways, and $I_1, \dots, I_N$, are index ranges in each dimension. Tensors extend the notion of vectors and matrices, referred to as 1-way and 2-way tensors, respectively, whereas a 3-way tensor $\mathcal{X} \in \mathbb{R}^{I_1 \times I_2 \times I_3}$, depicted in Fig. 1, is a cuboid (or cube-like) with elements $x_{ijk} \equiv \mathcal{X}(i, j, k)$, $1 \leq i \leq I_1, 1 \leq j \leq I_2, 1 \leq k \leq I_3$. The field $\mathbb{K}$ will be $\mathbb{R}$ in our application but it can be $\mathbb{C}$ in general.

Index fixing can be used to build subarrays. A typical scenario is when all indices are fixed except one, in which case the resulting subarrays are referred to as fibers, c.f. Fig. 2. They can alternatively be thought of as mode- n fibers that are produced when all indices are fixed except the n th one. Another common scenario is when all indices are fixed except two, yielding the examples of slices depicted on Fig. 3 for a 3-way tensor. These examples make clear that there are many different possible orderings, meaning that software developers must agree on common conventions for consistency in their codes. If we stack the resulting fibers or slices, we are performing what is termed vectorization or matricization. In general, tensor unfolding denotes the mechanism by which to flatten a tensor to systematically arrange its entries into a matrix, thereby providing a bridge to circumvent tensors and revert to classical linear algebra operations.

On tensors of equal sizes, simple operations like addition and subtraction are performed component-by-component. However, beyond ordinary matrix-like definitions, there are other tensor-specific operations and decompositions that are much more sophisticated by far.

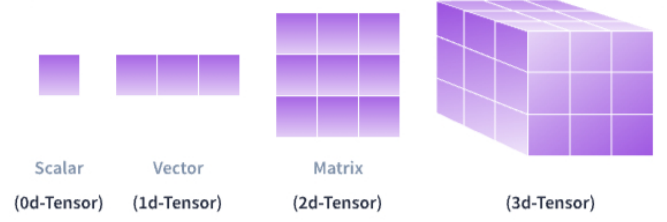


Fig. 1. Tensors of order 0, 1, 2, and 3.

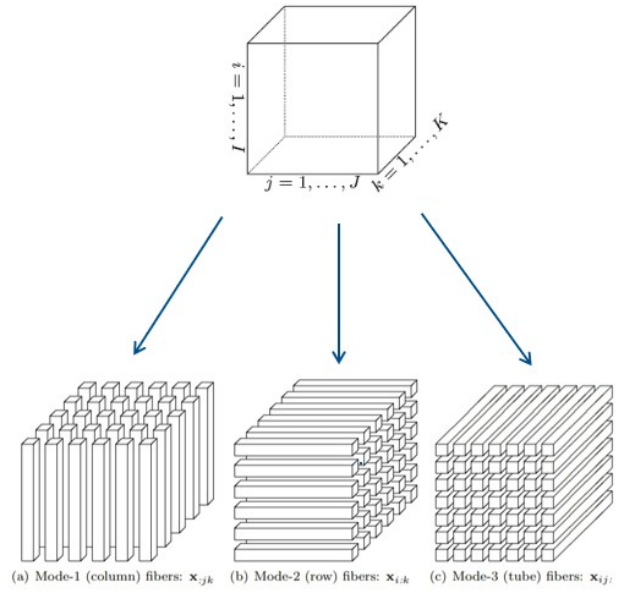


Fig. 2. Different fibers of a 3-mode tensor.

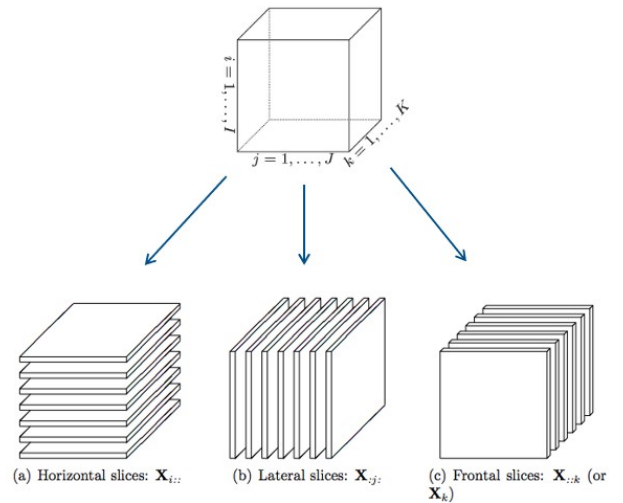


Fig. 3. Different slices of a 3-mode tensor.

B. Higher Order Singular Value Decomposition (HOSVD)

Consider the Singular Value Decomposition (SVD), which is a fundamental matrix factorization that represents any matrix as the product of two orthogonal matrices together with a diagonal matrix of singular values, encapsulating the inherent geometric characteristics of the data. Higher Order Singular Value Decomposition (HOSVD) generalizes this concept to multi-dimensional tensors, by decomposing them into a core tensor and a collection of orthogonal matrices corresponding to each mode. Below we reproduce a characterization of the HOSVD [21], mainly to illustrate how concepts that are familiar with matrices can quickly become drastically harder to grasp when reformulated through tensors.

Theorem II-B.1: Every tensor $\mathcal{X} \in \mathbb{R}^{I_1 \times I_2 \times \dots \times I_N}$ can be written as

$$\mathcal{X} = \mathcal{G} \times_1 \mathbf{A}^{(1)} \times_2 \mathbf{A}^{(2)} \dots \times_N \mathbf{A}^{(N)},$$

where

- 1) $\mathbf{A}^{(n)} = \begin{pmatrix} A_1^{(n)} & A_2^{(n)} & \dots & A_{I_n}^{(n)} \end{pmatrix}$ is an orthogonal matrix of order $I_n \times I_n$,
- 2) $\mathcal{G} \in \mathbb{R}^{I_1 \times I_2 \times \dots \times I_N}$ of which the subtensors $\mathcal{G}_{i_n=\alpha}$, attained by fixing the n th index to α , have the properties of
 - (i) ordering:

$$\|\mathcal{G}_{i_n=1}\|_F \geq \|\mathcal{G}_{i_n=2}\|_F \geq \dots \geq \|\mathcal{G}_{i_n=I_n}\|_F \geq 0,$$

for all possible values of n ,

- (ii) all-orthogonality: for all possible values of n, α and β , two subtensors $\mathcal{G}_{i_n=\alpha}$ and $\mathcal{G}_{i_n=\beta}$ are orthogonal subject to $\alpha \neq \beta$:

$$\langle \mathcal{G}_{i_n=\alpha}, \mathcal{G}_{i_n=\beta} \rangle = 0 \text{ for } \alpha \neq \beta.$$

- 3) $\mathcal{G} = \mathcal{X} \times_1 \mathbf{A}^{(1)T} \times_2 \mathbf{A}^{(2)T} \dots \times_N \mathbf{A}^{(N)T}$.

The vector $A_i^{(n)}$ is an n -mode singular vector and the Frobenius-norms $\|\mathcal{G}_{i_n=i}\|_F$ are n -mode singular values of $\mathcal{X}$. N denotes the total number of modes of the tensor, and $n \in \{1, \dots, N\}$ refers to an arbitrary mode index. The symbol $\times_n$ denotes the mode- n product, which is an operation where each mode- n fiber of a tensor is multiplied by a matrix, affecting only the n th dimension while keeping others fixed.

C. CANDECOMP/PARAFAC (CP) Decomposition

A popular decomposition in practice is the CP decomposition, which for $\mathcal{X} \in \mathbb{R}^{I_1 \times \dots \times I_N}$, is given by [22],

$$\mathcal{X} = \mathcal{I} \times_1 \mathbf{A}^{(1)} \dots \times_N \mathbf{A}^{(N)} \equiv \llbracket \mathbf{A}^{(1)}, \dots, \mathbf{A}^{(N)} \rrbracket \quad (1)$$

where $\mathcal{I}$ is the identity tensor having ones along the super-diagonal and zeros elsewhere, and each $\mathbf{A}^{(n)} \in \mathbb{R}^{I_n \times R}$ for $n = 1, \dots, N$ is a factor matrix. The notation $\llbracket \cdot \rrbracket$ denotes the CP decomposition as a sum of R rank-one tensors formed by outer products of the columns of the factor matrices. The CP rank $R \in \mathbb{N}$ is left as a parameter that tunes the accuracy of the decomposition. The fact that CP boils down to resorting to matrices explains its appeal and popularity.

According to [23], (1) can be written element-wise as a sum of products of the entries of those matrices

$$\mathcal{X}(i_1, \dots, i_N) = \sum_{r=1}^R \mathbf{A}^{(1)}(i_1, r) \dots \mathbf{A}^{(N)}(i_N, r). \quad (2)$$

Equation (3) below is another compact notation for writing (1) or (2) via the outer product

$$\mathcal{X} = \sum_{r=1}^R \mathbf{a}_r^{(1)} \circ \dots \circ \mathbf{a}_r^{(N)}. \quad (3)$$

by referencing columns of $\mathbf{A}^{(n)} = \begin{pmatrix} \mathbf{a}_1^{(n)} & \dots & \mathbf{a}_R^{(n)} \end{pmatrix} \in \mathbb{R}^{I_n \times R}$, $n = 1, \dots, N$.

D. Tensor Train (TT) Decomposition

Our present numerical implementation is based on another decomposition, namely the tensor train (TT) decomposition. For any $\mathcal{X} \in \mathbb{R}^{I_1 \times \dots \times I_N}$, it can be shown that its entries can be written in the form [24],

$$\mathcal{X}(i_1, \dots, i_N) = \mathcal{G}_1(i_1) \cdot \dots \cdot \mathcal{G}_N(i_N). \quad (4)$$

Here, $\mathcal{G}_n(i_n) \in \mathbb{R}^{r_{n-1} \times r_n}$ are matrices (slices) of the cuboid (3-way) tensors $\mathcal{G}_n \in \mathbb{R}^{r_{n-1} \times r_n \times r_n}$, which are called TT cores. Moreover, “boundary conditions” $r_0 = r_N = 1$ are imposed so that the chain product (4) always yields a scalar. TT ranks are defined as $(r_1, \dots, r_{N-1})$.

An explicit storage of $\mathcal{X}$ would cost $\prod_{n=1}^N I_n = O(I^N)$, where $I = \max_{1 \leq n \leq N} I_n$, so that if $N = 10$ in a CME application and $I \approx 10$, its corresponding matrix (7) would be of immense size $10^{10} \times 10^{10}$ and even much more for the RDME. By contrast, the TT format only stores the TT cores at a total cost of $O(r_{TT}^2 \cdot N \cdot I)$, where $r_{TT} = \max_{1 \leq n \leq N} r_n$. Thus, provided that the maximum rank r_{TT} is not large, the storage of the TT format scales linearly with N , leading to dramatic savings from the explicit format where the storage scales exponentially with N . Our actual code builds on another variant, specifically the quantized tensor train (QTT) format that brings the storage requirements further down to $O(r_{QTT}^2 \cdot N \cdot \log_2 I)$. Although r_{QTT} is not necessarily r_{TT} , this reduction remarkably means a linear cost in the number of modes and logarithmic in the mode sizes.

Our implementation did not have to invent everything from scratch. Rather, we leveraged past research, including [25], [26], and the TT toolbox software [24]. Our contribution here is sharing the findings of our hands-on experience on these topics and their application to the RDME context that we describe next.

E. CME and FSP in tensor format

Consider a chemical reacting system in a spatial domain Ω involving N species $\{S_1, \dots, S_N\}$, represented by the state vector $\mathbf{X}(t) = [X_1(t), \dots, X_N(t)]^T$, where each $X_i(t)$ is a non-negative integer denoting the population of species S_i at time t . Suppose there are M reaction channels, denoted

by $\{R_1, \dots, R_M\}$, and assume that the system is well mixed and in thermal equilibrium. The dynamics of the j th reaction channel R_j is characterized by

- (i) the propensity function $a_j(\mathbf{x})$, which indicates how likely it is that R_j will occur given the system's current state $\mathbf{x}$, and
- (ii) the stoichiometric (or state change) vector $\boldsymbol{\nu}_j = [\nu_{1j}, \dots, \nu_{Nj}]^T$, which specifies the change in the population count of each species that results from one occurrence of R_j .

Thus, $a_j(\mathbf{x})dt$ gives the probability that, given $\mathbf{X}(t) = \mathbf{x}$, one R_j reaction will occur in the next infinitesimal time interval $[t, t + dt]$, and ν_{ij} gives the change in X_i induced by one R_j reaction. This system can be modeled as a Markov chain, with its behavior determined by the Chemical Master Equation (CME) [7]

$$\frac{dP(\mathbf{x}, t)}{dt} = \sum_{j=1}^M \left[a_j(\mathbf{x} - \boldsymbol{\nu}_j) P(\mathbf{x} - \boldsymbol{\nu}_j, t) - a_j(\mathbf{x}) P(\mathbf{x}, t) \right], \quad (5)$$

where the function $P(\mathbf{x}, t)$ denotes the probability that $\mathbf{X}(t)$ will be $\mathbf{x}$. Equation (5) may be written in an equivalent matrix-vector form by enumerating all the states. If there are n possible states, $\mathbf{x}_1, \dots, \mathbf{x}_n$, the CME takes the form of a system of linear ordinary differential equations (ODEs),

$$\dot{\mathbf{P}}(t) = \mathbf{M} \cdot \mathbf{P}(t), \quad (6)$$

where the probability vector $\mathbf{P} = [p_1, \dots, p_n]^T$ is such that each component $p_i \equiv P(\mathbf{x}_i, t) \equiv \text{Prob}\{\mathbf{X}(t) = \mathbf{x}_i\}$, the probability of being at state $\mathbf{x}_i$ at time t , for $i = 1, \dots, n$.

The matrix $\mathbf{M}$ is a sparse $n \times n$ matrix that is populated by the propensities and it represents the transition rate matrix of the Markov chain underlying the CME. Its entries are constructed using

$$\mathbf{M}(i, j) = \begin{cases} a_k(\mathbf{x}_j) & \text{if } \mathbf{x}_i = \mathbf{x}_j + \boldsymbol{\nu}_k \\ -\sum_{k=1}^M a_k(\mathbf{x}_j) & \text{if } i = j \\ 0 & \text{otherwise} \end{cases} \quad (7)$$

The FSP method [9] and improvements such as [27] directly compute an approximation to the solution of the CME. The FSP method solves the CME and estimates the probability vector (PV) of the populations in a chemical reaction system using a truncated state space. For a truncated transition matrix $\mathbf{M}_T$ and initial truncated PV denoted by $\mathbf{P}_T(0)$, the FSP finds the PV at any time $\mathbf{P}_T(t)$ by the following,

$$\dot{\mathbf{P}}_T(t) = \mathbf{M}_T \cdot \mathbf{P}_T(t). \quad (8)$$

Equation (8) is a system of linear ODEs and the solution is given by

$$\mathbf{P}_T(t) = \exp(\mathbf{M}_T t) \mathbf{P}_T(0). \quad (9)$$

The CME matrix $\mathbf{M}$ remains sparse under any order relation imposed on the enumeration of the states, and it is natural

to use the lexicographic order on the state space as a subset of $\mathbb{N}$, the set of nonnegative natural numbers. If we impose the lexicographic order on a state space $\mathbb{X}$ in the form of a window of lower bounds $(lb^{(1)}, \dots, lb^{(N)})$ and upper bounds $(ub^{(1)}, \dots, ub^{(N)})$, i.e., $\mathbb{X}$ consists of states $\mathbf{x} = (x_1, \dots, x_N)^T$ such that $lb^{(s)} \leq x_s \leq ub^{(s)}, \forall s = 1, \dots, N$, then the transition rate matrix can be decomposed into a sum of Kronecker products [28], [29]

$$\mathbf{M} = \sum_{k=1}^M \left(\bigotimes_{s=1}^N \mathbf{S}_k^{(s)} - \mathbf{I} \right) \mathbf{M}_k. \quad (10)$$

Each term in the sum (10) corresponds to a reaction: $\mathbf{S}_k^{(s)}$ is a “shift-diagonal” matrix corresponding to the change in species s when reaction k happens; $\mathbf{I}$ is the identity matrix; and $\mathbf{M}_k$ is the diagonal matrix that stores the values of the propensity function a_k . Precisely, let $I_s = ub^{(s)} - lb^{(s)} + 1$ be the window size for each species, $n = I_1 \cdot \dots \cdot I_N$ is the total number of states in $\mathbb{X}$, and denote $\boldsymbol{\nu}_k = (\nu_k^{(1)}, \dots, \nu_k^{(N)})^T$. The identity matrix $\mathbf{I}$ is then of order n , $\mathbf{S}_k^{(s)} \in \mathbb{R}^{I_s \times I_s}$ is given by

$$\mathbf{S}_k^{(s)} \equiv \begin{cases} \begin{pmatrix} 0 & \dots & 1 & & \\ & \ddots & & \ddots & \\ & & \ddots & & 1 \\ & & & \ddots & \vdots \\ & & & & 0 \end{pmatrix} & \begin{array}{l} \text{if } \nu_k^{(s)} < 0, \text{ shifted} \\ \text{up } |\nu_k^{(s)}| \text{ rows} \end{array} \\ \begin{pmatrix} 0 & & & & \\ \vdots & \ddots & & & \\ 1 & & \ddots & & \\ & \ddots & & \ddots & \\ & & 1 & \dots & 0 \end{pmatrix} & \begin{array}{l} \text{if } \nu_k^{(s)} \geq 0, \text{ shifted} \\ \text{down } \nu_k^{(s)} \text{ rows} \end{array} \end{cases}$$

and

$$\mathbf{M}_k \equiv \text{diag}(a_k(\mathbf{x}_1), \dots, a_k(\mathbf{x}_n)), \quad (11)$$

where $\mathbf{x}_1, \dots, \mathbf{x}_n$ are states of $\mathbb{X}$ in the increasing lexicographic order. Furthermore, by assuming that the propensity functions are separable, i.e.,

$$a_k(\mathbf{x}) = a_k^{(1)}(x_1) \cdot \dots \cdot a_k^{(N)}(x_N), \quad k = 1, \dots, M,$$

where $a_k^{(s)} \geq 0$ are scalar functions, (11) can be rewritten as

$$\mathbf{M}_k = \bigotimes_{s=1}^N \text{diag}(a_k^{(s)}),$$

where, for $s = 1, \dots, N$,

$$\text{diag}(a_k^{(s)}) \equiv \text{diag}(a_k^{(s)}(lb^{(s)}), \dots, a_k^{(s)}(ub^{(s)})).$$

Under these assumptions, the transition matrix (10) can be represented by simple matrices of small sizes,

$$\mathbf{M} = \sum_{k=1}^M \left(\bigotimes_{s=1}^N \mathbf{S}_k^{(s)} - \mathbf{I} \right) \left(\bigotimes_{s=1}^N \text{diag}(a_k^{(s)}) \right). \quad (12)$$

In [26], the lower and upper bounds are in the form $lb^{(s)} = 0$ and $ub^{(s)} = I_s - 1$, yielding

$$\text{diag } a_k^{(s)} = \text{diag} \left(a_k^{(s)}(0), \dots, a_k^{(s)}(I_s - 1) \right).$$

The reason for holding the lower bounds anchored at 0 was to help attenuate some of the complexity of dealing with tensor train decompositions when their sizes changed adaptively. Subsequent improvements in [25] allowed both the lower and upper bounds to slide dynamically. A central aspect in our body of research [25], [20], [26] has been that by keeping the state space in the form of a hyper-rectangle,

$$\mathbf{x} \in [lb^{(1)} : ub^{(1)}] \times \dots \times [lb^{(N)} : ub^{(N)}],$$

both the transition operator $\mathcal{M}$ and the probability distribution $\mathbf{P}$ can be represented as multi-dimensional arrays that allow us to use tensor decompositions as pioneered in [28], [29] to resolve storage issues.

F. RDME

Assume the spatial domain Ω is partitioned into compartments (voxels), denoted V_k , $k = 1, \dots, K$. Within each compartment, species have the ability to react with one another as well as diffuse (jump) over the boundaries to reach adjacent compartments. Let $X_{i,k}(t)$ be the population count of species S_i in compartment V_k at time t . Then each species in the domain is given by the subvector $\mathbf{X}_i(t) = [X_{i,1}(t), \dots, X_{i,K}(t)]$, $i = 1, \dots, N$. Hence, $\mathbf{X}_1$ is the subvector which represents the species 1 in all compartments. Likewise, $\mathbf{X}_2$ represents species 2 in all compartments, and so on. Thus, the state vector of the system is $\mathbf{X} = [\mathbf{X}_1, \dots, \mathbf{X}_N]$. The diffusion propensity function $d_{i,j,k}$ and the state change vector $\mu_{k,j}$ characterize the dynamics of the diffusion of species S_i from compartment V_k to V_j . The vector $\mu_{k,j}$ has a length of K with -1 in the k th position, 1 in the j th position, and 0 elsewhere. Given $\mathbf{X}(t) = \mathbf{x}$, the diffusion master equation (DME) can be written as

$$\frac{dP(\mathbf{x}, t)}{dt} = \sum_{i=1}^N \sum_{k=1}^K \sum_{j=1}^K \left[d_{i,j,k}(\mathbf{x}_i - \mu_{k,j}) P(\mathbf{x}_1, \dots, \mathbf{x}_i - \mu_{k,j}, \dots, \mathbf{x}_N, t) - d_{i,j,k}(\mathbf{x}_i) P(\mathbf{x}, t) \right]. \quad (13)$$

We take the diffusion propensity function $d_{i,j,k}(\mathbf{x}_i) = D/l^2$, for $k = j \pm 1$, where D is the diffusion rate and l is the characteristic length of the compartment, meaning that our setting ultimately assumes that $d_{i,j,k} \equiv d$ is constant. Some other settings could make different choices. If $\mathcal{D}$ is the transition matrix describing the diffusion (or jump) of species between compartments, the equivalent matrix-vector form of (13) can be written as

$$\dot{\mathbf{P}}(t) = \mathcal{D} \cdot \mathbf{P}(t). \quad (14)$$

Combining equations (6) and (14) we get the matrix-vector form of the RDME,

$$\dot{\mathbf{P}}(t) = \mathcal{M} \cdot \mathbf{P}(t) + \mathcal{D} \cdot \mathbf{P}(t). \quad (15)$$

Equation (15) is a system of linear constant coefficient ODEs and gives us more possible states than the CME, and so its corresponding transition matrix is substantially extended to represent species in compartments. However, it should be remembered that it is ultimately a tensor representation of the form (12) that is used in the calculations.

III. STUDY FINDINGS

A. Metapopulation Model

Populations that are segmented into smaller groups linked by migration are studied using metapopulation models. Metapopulation models often employ reaction-diffusion principles to describe spatial dynamics [30]. In a reaction-diffusion metapopulation model, the reaction term depicts the occurrences of births and deaths within each subpopulation, whereas the diffusion term captures the mobility of an individual between nearby subpopulations. By taking into account both diffusion and reaction terms, the model can better capture complex interactions between local population dynamics and spatial dispersal. Using the RDME in a metapopulation model, each subpopulation's population is modeled as a stochastic process with birth, death, and migration events occurring randomly. We examine a simple reaction scheme in this study that conserves the population count in order to empirically provide a proof-of-concept



In the context of our metapopulation model, α species represent normal species (healthy individuals) and β species represent active species (infected individuals) intermingling from area to area. This scheme is also known as the SIS model (Susceptible-Infectious-Susceptible) [31]. Recall that when we analyze this model without considering compartments and diffusion, the stoichiometric matrix is

$$S = \begin{pmatrix} \alpha & \beta & \alpha & \beta \\ 0 & 1 & 1 & 0 \\ 1 & 1 & 0 & 2 \end{pmatrix}$$

in which the first and second row of the matrix stand for reactions (16) and (17) respectively. The two left columns contain the coefficients of the reactants while the other two right columns contain the coefficients of the products. Meaning that from $S = (s_{ij})$, the components of the state change vector associated to a reaction R_k , $k = 1, \dots, M$, are effectively

$$(s_{k,N+1:2N} - s_{k,1:N}) \quad (18)$$

where in this example $N = 2$ and $M = 2$. It is less error prone to set S and programmatically proceed as shown above.

We now consider two neighboring areas (i.e., compartments) where both the healthy and infected individuals can move back and forth. We label the areas as Area 1 and Area 2. Consequently, the scheme can be represented by a reaction-diffusion process where reactions (16) and (17) can take place inside both areas separately and the individuals can move back

and forth between Area 1 and Area 2. An illustration is shown in Fig. 4. (Imagine Area 1 as indoor and Area 2 as outdoor.)

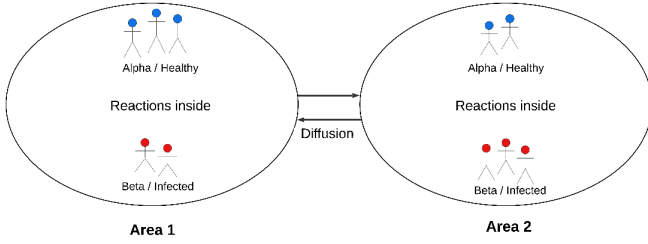


Fig. 4. Illustration of the metapopulation example.

The process is a random process and can be modeled by the RDME consisting of two reactions with two compartments. By our construction, species change forms without creating or destroying themselves, hence the overall population count in the process stays constant.

To make the details clear, we take a small example where we assume that initially, Area 1 has $\alpha = 2$ healthy individuals and $\beta = 1$ infected individual. Area 2 has $\alpha = 1$ healthy individual and $\beta = 2$ infected individuals. With the convention of section II-F, the initial state vector is therefore

$$\left[\overbrace{2, 1}^{\alpha}, \overbrace{1, 2}^{\beta} \right] \quad (19)$$

Area 1 Area 2 Area 1 Area 2

with diacritical annotations in (19) meant to ease the interpretation of the indexing of the components. We set the reaction parameters $c_1 = .30$ and $c_2 = 1.0$, and the diffusion parameters $D = 1.0$ and $l = 10$, yielding $d = D/l^2 = 10^{-2}$. The RDME gives us all the possible states we can get when those individuals react and diffuse randomly, and recall that along a similar spirit to [32, Eq. (4)], diffusion in our context is modeled via stochastic jump processes through additional reactions,

$$(\text{Area 1}) \alpha \xrightleftharpoons[d]{d} \alpha (\text{Area 2}), \quad (20)$$

$$(\text{Area 1}) \beta \xrightleftharpoons[d]{d} \beta (\text{Area 2}). \quad (21)$$

Now that we are considering the system in two areas where diffusion also occurs, the stoichiometric matrix grows much

larger to take the form below.

Area#	Reactants				Products				
	α		β		α		β		
	1	2	1	2	1	2	1	2	
	1	2	1	2	1	2	1	2	
$S =$	0	0	1	0	1	0	0	0	Reactions
	1	0	1	0	0	0	2	0	
	0	0	0	1	0	1	0	0	
	0	1	0	1	0	0	0	2	
	1	0	0	0	0	1	0	0	Diffusions
	0	1	0	0	1	0	0	0	
	0	0	1	0	0	0	0	1	
	0	0	0	1	0	0	1	0	

The annotations here are also aimed at easing the connection of how to map the location of entries with the convention of section II-F. In the matrix, the first two rows indicate reactions (16) and (17) inside Area 1. The third and fourth rows represent the same reactions respectively in Area 2. Diffusions (20) and (21) are characterized by the last four rows. The four left columns contain the coefficients of the reactants and the other four right columns contain the coefficients of the products. Using (18) with $N = 4$ and $M = 8$ gives the state change vectors.

Even starting with a simple initial state as $[2, 1, 1, 2]$ explained in (19), the system explodes to a total of 84 possible states (and quickly much more if we set larger starting values as we alerted before). We solve the associated RDME using the classical FSP equation (9) with a state space truncated to 70 to continue our empirical illustration of the computational building blocks. Taking $t = 10$, we obtain an approximation of the marginal probability distribution of the healthy and infected individuals in the two areas. Results are shown in Fig. 5.

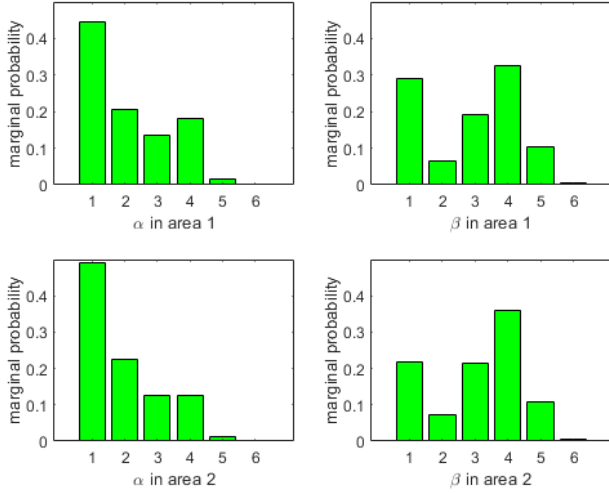


Fig. 5. Marginal probability distribution of the number of healthy and infected individuals in the two areas using the classical FSP at time $t = 10$.

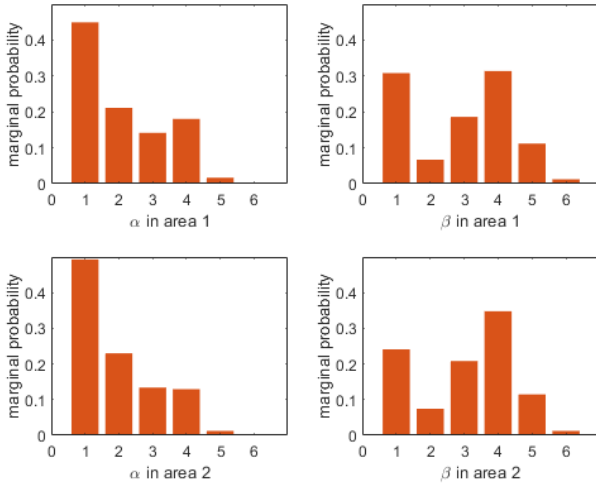


Fig. 6. Marginal probability distribution of the number of healthy and infected individuals in the two areas using tensors at time $t = 10$.

Having obtained a reference solution via the classical FSP, we now attempt to reproduce results using tensors on the same metapopulation model. As explained in the text, the main concept is to express a high-dimensional tensor as an ordered product of smaller tensors that can scale better as problems become more complex than the simple example shown here for illustration and ease of interpretation. For our purposes, the marginal probability distribution of the individuals are found by using the adaptive tensor code in [25], which has already been demonstrated to be effective on much larger problems albeit in the CME context. Fig. 6 shows the computed marginal probability distribution of the healthy and infected individuals in the two areas using the tensor technique in our RDME prototype. Fig. 7 illustrates the accuracy of the implementation

of tensors by showing the near alignment of the marginal distributions obtained using the two alternative approaches.

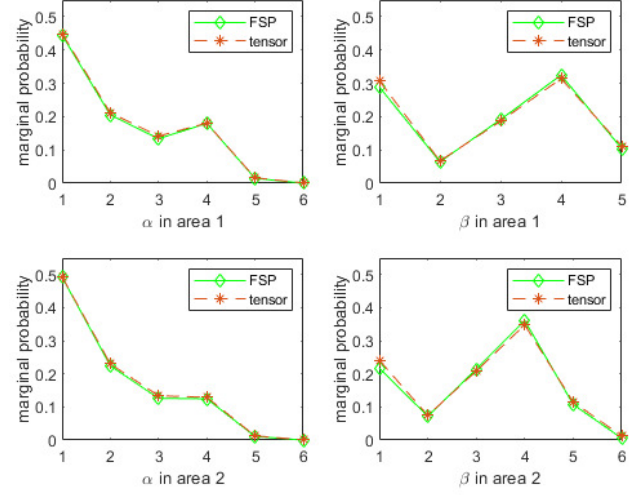


Fig. 7. Comparison of the marginal probability distributions computed using the classical FSP vs tensors.

IV. CONCLUSIONS

We have applied the RDME formulation in a metapopulation model that includes both reaction and diffusion processes. Although our initialization was kept simple to detail the algorithmic steps, it is perhaps worth mentioning that the broader significance of the metapopulation model with reaction-diffusion is to represent the interplay between local population dynamics and migration. It is a way of understanding how populations behave when they are distributed across discrete, interconnected areas and subject to both local environmental factors (birth/death rates) and movement between these areas (diffusion). This model is of value in the literature because it provides insights into how populations persist, grow, or face extinction in fragmented landscapes. The model is essential for addressing real-world issues such as conservation, habitat fragmentation, disease spread, and the management of invasive species, where the dynamics of movement and local interactions play key roles in shaping long-term outcomes.

Therefore, the fact that we have an RDME formulation opens avenues for further computerized simulations. Here, we have utilized the FSP and tensors to approximate the marginal probability distribution. Rather than taking into account all possible states, the FSP operates with a reduced state space. On top of that, the compact representation of the transition matrix via tensors goes hand in hand with tensor train decompositions. Our body of research has shown that tensors scale very well to tackle multi-dimensional CME problems, suggesting promising prospects for their application in the RDME as we illustrated in this case study. Although the tensor approach introduces a significant complexity in the derivation of algorithms, it offers distinct scalability advantages. It excels in

managing high-dimensional problems by efficiently representing extensive state spaces with reduced memory requirements, making it particularly effective in scenarios where the classical matrix-based FSP becomes computationally expensive or infeasible.

REFERENCES

- [1] M. A. O. Vasilescu and D. Terzopoulos, "Multilinear analysis of image ensembles: Tensorfaces," in *Computer Vision—ECCV 2002: 7th European Conference on Computer Vision Copenhagen, Denmark, May 28–31, 2002 Proceedings, Part I* 7. Springer, 2002, pp. 447–460, doi: 10.1007/3-540-47969-4_30.
- [2] E. E. Papalexakis, C. Faloutsos, and N. D. Sidiropoulos, "Parcube: Sparse parallelizable tensor decompositions," in *Machine Learning and Knowledge Discovery in Databases: European Conference, ECML PKDD 2012, Bristol, UK, September 24–28, 2012. Proceedings, Part I* 23. Springer, 2012, pp. 521–536, doi: 10.1007/978-3-642-33460-3_39.
- [3] F. L. Hitchcock, "The expression of a tensor or a polyadic as a sum of products," *Journal of Mathematics and Physics*, vol. 6, no. 1–4, pp. 164–189, 1927. doi: 10.1002/sapm192761164
- [4] L. Grasedyck, D. Kressner, and C. Tobler, "A literature survey of low-rank tensor approximation techniques," *GAMM-Mitteilungen*, vol. 36, no. 1, pp. 53–78, 2013. doi: 10.1002/gamm.201310004
- [5] W. Hackbusch, *Tensor spaces and numerical tensor calculus*. Springer, 2012, vol. 42, doi: 10.1007/978-3-030-35554-8.
- [6] J. Grabek and B. Cyganek, "An impact of tensor-based data compression methods on deep neural network accuracy," in *Position and Communication Papers of the 16th Conference on Computer Science and Intelligence Systems*, ser. Annals of Computer Science and Information Systems, M. Ganzha, L. Maciaszek, M. Paprzycki, and D. Ślęzak, Eds., vol. 26. PTI, 2021. doi: 10.15439/2021F127 p. 3–11. [Online]. Available: <http://dx.doi.org/10.15439/2021F127>
- [7] D. T. Gillespie, "A rigorous derivation of the chemical master equation," *Physica A: Statistical Mechanics and its Applications*, vol. 188, no. 1–3, pp. 404–425, 1992. doi: 10.1016/0378-4371(92)90283-V
- [8] —, "Exact stochastic simulation of coupled chemical reactions," *The journal of physical chemistry*, vol. 81, no. 25, pp. 2340–2361, 1977. doi: 10.1021/j100540a008
- [9] B. Munsky and M. Khammash, "The finite state projection algorithm for the solution of the chemical master equation," *The Journal of Chemical Physics*, vol. 124, no. 4, p. 044104, 2006. doi: 10.1063/1.2145882
- [10] J. S. van Zon, M. J. Morelli, S. Tănase-Nicola, and P. R. ten Wolde, "Diffusion of transcription factors can drastically enhance the noise in gene expression," *Biophysical Journal*, vol. 91, no. 12, pp. 4350–4367, 2006. doi: 10.1529/biophysj.106.086157
- [11] N. Pavin, H. Č. Paljetak, and V. Krstić, "Min-protein oscillations in *Escherichia coli* with spontaneous formation of two-stranded filaments in a three-dimensional stochastic reaction-diffusion model," *Physical Review E*, vol. 73, no. 2, p. 021904, 2006. doi: 10.1103/PhysRevE.73.021904
- [12] M. Smoluchowski, "Mathematical theory of the kinetics of the coagulation of colloidal solutions," *Z. Phys. Chem.*, vol. 92, pp. 129–168, 1917.
- [13] C. Gardiner, K. McNeil, D. Walls, and I. Matheson, "Correlations in stochastic theories of chemical reactions," *Journal of Statistical Physics*, vol. 14, pp. 307–331, 1976. doi: 10.1007/BF01030197
- [14] A. B. Stundzia and C. J. Lumsden, "Stochastic simulation of coupled reaction–diffusion processes," *Journal of Computational Physics*, vol. 127, no. 1, pp. 196–207, 1996. doi: 10.1006/jcph.1996.0168
- [15] M. M. Rahman and R. B. Sidje, "A study of a metapopulation model using the stochastic reaction diffusion master equation," in *Southwest Data Science Conference*. Springer, 2023, pp. 242–253, doi: 10.1007/978-3-031-61816-1_17.
- [16] M. M. Rahman, K. Cook, and R. B. Sidje, "Parallel implementation of the stochastic reaction diffusion master equation," in *Proceedings of the 17th International Conference on Bioinformatics and Computational Biology*, T. Aldwairi, H. Al-Mubaid, and O. Eulenstein, Eds. Cham: Springer, 2025. ISBN 978-3-031-94039-2 Doi: 10.1007/978-3-031-94039-2_10.
- [17] M. A. Gibson and J. Bruck, "Efficient exact stochastic simulation of chemical systems with many species and many channels," *The journal of physical chemistry A*, vol. 104, no. 9, pp. 1876–1889, 2000. doi: 10.1021/jp993732q
- [18] J. Hatne, D. Fange, and J. Elf, "Stochastic reaction-diffusion simulation with MesoRD," *Bioinformatics*, vol. 21, no. 12, pp. 2923–2924, 2005. doi: 10.1093/bioinformatics/bti431
- [19] J. Elf and M. Ehrenberg, "Spontaneous separation of bi-stable biochemical systems into spatial domains of opposite phases," *Systems biology*, vol. 1, no. 2, pp. 230–236, 2004. doi: 10.1049/sb:20045021
- [20] M. M. Rahman, "A numerical study of the stochastic reaction-diffusion master equation using tensors and parallelism with application to biological models," Ph.D. dissertation, The University of Alabama, 2025.
- [21] L. De Lathauwer, B. De Moor, and J. Vandewalle, "A multilinear singular value decomposition," *SIAM Journal on Matrix Analysis and Applications*, vol. 21, no. 4, pp. 1253–1278, 2000. doi: 10.1137/S089547989630569
- [22] T. G. Kolda and B. W. Bader, "Tensor decompositions and applications," *SIAM Review*, vol. 51, no. 3, pp. 455–500, 2009. doi: 10.1137/07070111X
- [23] T. G. Kolda, "Multilinear operators for higher-order decompositions," Sandia National Laboratories (SNL), Albuquerque, NM, and Livermore, CA (United States), Tech. Rep., 2006, doi: 10.2172/923081.
- [24] I. V. Oseledets, "Tensor-train decomposition," *SIAM Journal on Scientific Computing*, vol. 33, no. 5, pp. 2295–2317, 2011. doi: 10.1137/090752286
- [25] T. Dinh and R. B. Sidje, "An adaptive solution to the chemical master equation using quantized tensor trains with sliding windows," *Physical Biology*, vol. 17, no. 6, p. 065014, 2020. doi: 10.1088/1478-3975/aba1d2
- [26] H. D. Vo and R. B. Sidje, "An adaptive solution to the chemical master equation using tensors," *The Journal of Chemical Physics*, vol. 147, no. 4, 2017. doi: 10.1063/1.4994917
- [27] K. Burrage, M. Hegland, S. Macnamara, and R. B. Sidje, "A Krylov-based finite state projection algorithm for solving the chemical master equation arising in the discrete modelling of biological systems," in *Proc. of The AA Markov 150th Anniversary Meeting*, no. 21–37, 2006.
- [28] M. Hegland and J. Garcke, "On the numerical solution of the chemical master equation with sums of rank one tensors," *ANZIAM Journal*, vol. 52, pp. C628–C643, 2010. doi: 10.21914/anziamj.v52i0.3895
- [29] V. Kazeev, M. Khammash, M. Nip, and C. Schwab, "Direct solution of the chemical master equation using quantized tensor trains," *PLoS computational biology*, vol. 10, no. 3, p. e1003359, 2014. doi: 10.1371/journal.pcbi.1003359
- [30] V. Colizza, R. Pastor-Satorras, and A. Vespignani, "Reaction–diffusion processes and metapopulation models in heterogeneous networks," *Nature Physics*, vol. 3, no. 4, pp. 276–282, 2007. doi: 10.1038/nphys560
- [31] R. Anderson, R. May, and B. Anderson, *Infectious Diseases of Humans: Dynamics and Control*. Oxford University Press, 1992, doi: 10.1093/oso/9780198545996.001.0001.
- [32] S. Smith and R. Grima, "Breakdown of the reaction-diffusion master equation with nonelementary rates," *Phys. Rev. E*, vol. 93, p. 052135, May 2016. doi: 10.1103/PhysRevE.93.052135. [Online]. Available: <https://link.aps.org/doi/10.1103/PhysRevE.93.052135>