

# A Software Tool for the Compact Solution of the Chemical Master Equation

Tuğrul Dayar<sup>1(✉)</sup> and M. Can Orhan<sup>2</sup>

<sup>1</sup> Department of Computer Engineering, Bilkent University,  
06800 Bilkent, Ankara, Turkey  
tugrul@cs.bilkent.edu.tr

<sup>2</sup> Kanava Technologies, 06800 Ankara, Turkey  
m.canorhan@gmail.com

**Abstract.** The problem of computing the transient probability distribution of countably infinite multidimensional continuous-time Markov chains (CTMCs) arising in systems of stochastic chemical kinetics is addressed by a software tool. Starting from an initial probability distribution, time evolution of the probability distribution associated with the CTMC is described by a system of linear first-order ordinary differential equations, known as the chemical master equation (CME). The solver for the CME uses the time stepping implicit backward differentiation formulae (BDF). Solution vectors in BDF can be stored compactly during transient analysis in one of the Hierarchical Tucker Decomposition, Quantized Tensor Train, or Transposed Quantized Tensor Train formats.

**Keywords:** Continuous-time Markov chain  
Chemical master equation · Backward differentiation  
Compact vector · Kronecker decomposition

## 1 The Problem

Letting the initial probability distribution vector of the infinitesimal generator matrix  $Q$  underlying a multidimensional continuous-time Markov chain (CTMC) be denoted by  $\pi_0$ , the transient probability distribution vector  $\pi_t \in \mathbb{R}_{\geq 0}^{1 \times |\mathcal{R}|}$  of  $Q$  at time  $t \in \mathbb{R}_{\geq 0}$  satisfies [12]

$$\frac{d\pi_t}{dt} = \pi_t Q, \quad \pi_t \mathbf{e} = 1. \quad (1)$$

Here,  $\mathcal{R}$  is the reachable state space of the CTMC and  $\mathbf{e}$  is a vector of 1's.

When the CTMC arises in the area of systems of stochastic chemical kinetics, (1) is referred to as the chemical master equation (CME) [4]. In this case, there are a finite number of dimensions and transitions, but  $\mathcal{R}$  is almost always countably infinite. Therefore, a CTMC  $\{S(t), t \geq 0\}$  having  $H$  dimensions such that  $S(t) = (S_1(t), \dots, S_H(t))$  and  $\Pr(S(t) = \mathbf{i}) = \Pr(S_1(t) = i_1, \dots, S_H(t) = i_H)$

can be used with the state vector  $\mathbf{i} = (i_1, \dots, i_H)$ . The state space of dimension  $h$  is given by  $\mathcal{S}^{(h)} = \mathbb{Z}_{\geq 0}^h$  for  $h = 1, \dots, H$ , and when there are no unreachable states, we have  $\mathcal{R} = \times_{h=1}^H \mathcal{S}^{(h)}$  and  $K$  transition classes in which transition class  $k$  is represented by the pair  $(\alpha_k(\mathbf{i}), \mathbf{v}^{(k)})$  for  $k = 1, \dots, K$ . Here,  $\alpha_k(\mathbf{i}) \in \mathbb{R}_{\geq 0}$  is the transition rate function specifying the transition rate from state  $\mathbf{i} \in \mathcal{R}$  to state  $(\mathbf{i} + \mathbf{v}^{(k)}) \in \mathcal{R}$  and  $\mathbf{v}^{(k)} \in \mathbb{Z}^{1 \times H}$  is the state change vector specifying the successor state of the transition, with  $v_h^{(k)}$  denoting the change in state variable  $i_h \in \mathcal{S}^{(h)}$  due to a class  $k$  transition [3]. That  $\mathcal{R}$  is equal to the product state space is a property of the models under consideration; but, this can be relaxed with the help of a well known hierarchical state space structuring approach [1].

A Kronecker representation for models in this area, which has separable state dependent transition rate functions in the form

$$\alpha_k(\mathbf{i}) = \phi_k \prod_{h=1}^H \alpha_k^{(h)}(i_h),$$

can be obtained by letting the transition matrix of dimension  $h$  with state space  $\mathcal{S}^{(h)}$  for  $h = 1, \dots, H$  and transition class  $k = 1, \dots, K$  be denoted by  $Q_k^{(h)} \in \mathbb{R}_{\geq 0}^{|\mathcal{S}^{(h)}| \times |\mathcal{S}^{(h)}|}$  and given entrywise as

$$Q_k^{(h)}(i_h, j_h) = \begin{cases} \alpha_k^{(h)}(i_h) & \text{if } j_h = i_h + v_h^{(k)} \\ 0 & \text{otherwise} \end{cases} \quad \text{for } i_h, j_h \in \mathcal{S}^{(h)}.$$

Then

$$Q = Q_O + Q_D, \quad Q_O = \sum_{k=1}^K \phi_k \bigotimes_{h=1}^H Q_k^{(h)}, \quad Q_D = - \sum_{k=1}^K \phi_k \bigotimes_{h=1}^H \text{diag}(Q_k^{(h)} \mathbf{e}).$$

Next, we introduce a tool to solve the initial value problem associated with the system of linear first-order ordinary differential equations (ODEs) in (1) [12].

## 2 A Software Tool

We present a software tool [2] for the transient analysis of countably infinite multidimensional CTMCs introduced in the previous section in a sequential setting. Details regarding the tool may be obtained from its user manual. Time is discretized into smaller time steps and the solver for the CME [10] uses the implicit backward differentiation formulae (BDF). BDF methods are a class of implicit multistep methods to solve stiff ODEs [12]. Stiffness generally manifests itself when reaction rates occur at different time scales, and this is the case for many realistic systems. The  $o$ -step BDF method, denoted BDF $o$ , keeps approximations of solutions at  $o$  previous time steps and computes the solution at the current time step by solving a linear system. BDF $o$  methods have local truncation error proportional to the  $o$ th power of the step size, and therefore, are said to be of order  $o$ . The particular solver initializes the first  $o$  backward differences

with the embedded Runge–Kutta method due to Fehlberg, written  $\text{RKF}k-1(k)$ , which is an order  $k$  method without the error estimate [12].

At each time step,  $n$ , the reachable state space,  $\mathcal{R}$ , is truncated by using a well defined aggregation operator on the prediction vector of BDFo [11] to obtain  $\mathcal{R}_n$  [10]. Solution vectors can be stored compactly during transient analysis using one of the Hierarchical Tucker Decomposition (HTD) [5], Quantized Tensor Train (QTT) [8], or Transposed Quantized Tensor Train (QT3) [7] formats. Compact vectors in HTD format can work with a truncated generator matrix represented as a sum of Kronecker products of small molecule matrices, whereas those in QTT/QT3 format can work with a low-rank approximation of the truncated generator matrix in the same format [10].

The solution of the linear system at each time step in BDF is performed by the Jacobi iteration [12] using the Newton-Schulz method [9] to compute reciprocals of diagonal elements of the coefficient matrix for HTD and the density matrix renormalization group (DMRG) method for QTT/QT3 [7]. It is possible to use fixed and adaptive rank control strategies with compact vectors in HTD format. There are  $\text{HTD}_A$ ,  $\text{HTD}_M$ , and  $\text{HTD}_F$  variants of the BDFo solver in which (adaptive, adaptive), (fixed, adaptive), and (fixed, fixed) rank bounds are used in (Jacobi, Newton-Schulz) methods [10].

Next we show examples of results that can be obtained with the tool.

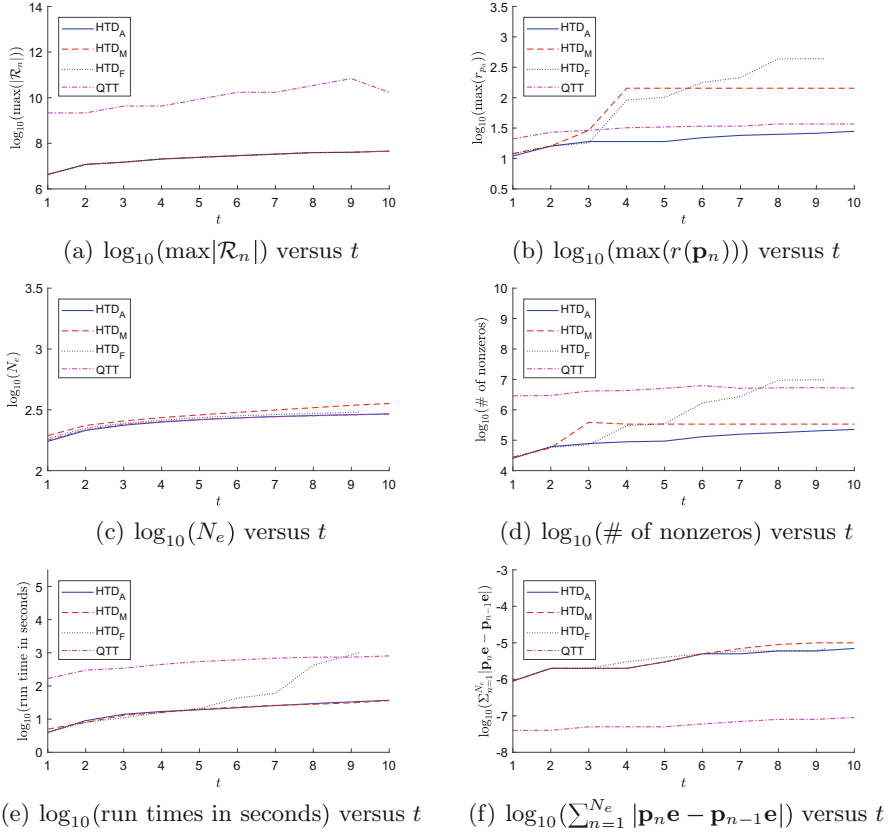
### 3 An Example

We consider a cascade model [6] that has five molecules each corresponding to a different dimension with the transition classes in Table 1. Here,  $H = 5$ ,  $\mathbf{i} = (i_1, i_2, i_3, i_4, i_5)$ ,  $K = 10$ ,  $a, b, c, \mu \in \mathbb{R}_{>0}$ , and  $\mathbf{e}_h$  is the  $h$ th principal axis vector. We let  $a = 0.7$ ,  $b = 1$ ,  $c = 5$ , and  $\mu = 0.07$  as in [6].

The cascade model is analyzed using BDF5 with an accuracy tolerance of  $10^{-9}$  and the indicated compact vector formats starting from the initial distribution  $\boldsymbol{\pi}_0(10, 10, 10, 10, 10) = 1$  for final time values  $t \in \{1, \dots, 10\}$ . A maximum run time of 1,000 seconds is imposed on the experiments performed on an Intel

**Table 1.** Transition classes of the cascade model

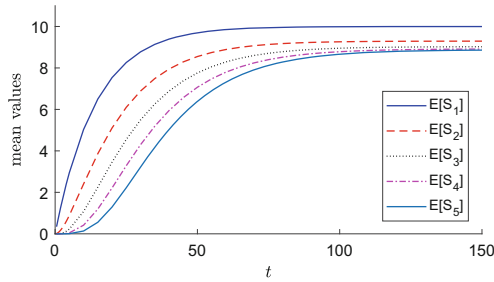
$k$	$\phi_k$	$\alpha_k^{(1)}(i_1)$	$\alpha_k^{(2)}(i_2)$	$\alpha_k^{(3)}(i_3)$	$\alpha_k^{(4)}(i_4)$	$\alpha_k^{(5)}(i_5)$	$\mathbf{v}^{(k)}$
1	$a$	1	1	1	1	1	$\mathbf{e}_1^T$
2	$\mu$	$i_1$	1	1	1	1	$-\mathbf{e}_1^T$
3	$b$	$\frac{i_1}{bi_1+c}$	1	1	1	1	$\mathbf{e}_2^T$
4	$\mu$	1	$i_2$	1	1	1	$-\mathbf{e}_2^T$
5	$b$	1	$\frac{i_2}{bi_2+c}$	1	1	1	$\mathbf{e}_3^T$
6	$\mu$	1	1	$i_3$	1	1	$-\mathbf{e}_3^T$
7	$b$	1	1	$\frac{i_3}{bi_3+c}$	1	1	$\mathbf{e}_4^T$
8	$\mu$	1	1	1	$i_4$	1	$-\mathbf{e}_4^T$
9	$b$	1	1	1	$\frac{i_4}{bi_4+c}$	1	$\mathbf{e}_5^T$
10	$\mu$	1	1	1	1	$i_5$	$-\mathbf{e}_5^T$



**Fig. 1.** Various measures associated with BDF5 for the cascade model

Core i7 2.6 GHz processor with 16 Gigabytes main memory under Linux. We let  $\mathbf{p}_n$  denote the transient probability distribution vector computed at time step  $n$ ,  $\max(|\mathcal{R}_n|)$  denote the maximum truncated state space size,  $\max(r(\mathbf{p}_n))$  denote the maximum rank associated with compact solution vectors,  $N_e$  denote the total number of time steps taken up to  $t$  (if  $t$  is reached within 1,000 seconds), and  $\sum_{n=1}^{N_e} |\mathbf{p}_n \mathbf{e} - \mathbf{p}_{n-1} \mathbf{e}|$  express the total state space truncation error, which has been shown to be in the same order as the relative error in the solution. We do not report the results with QT3 since they did not fare well. The results in Fig. 1 indicate that relative errors of at most  $10^{-7}$  and  $10^{-5}$  are obtained respectively with QTT and adaptive rank controlled HTD formats within 1,000 seconds in all problems. Furthermore, memory and time requirements of HTD<sub>A</sub> are at least an order of magnitude better than those with QTT.

We depict in Fig. 2 the mean number of molecules when BDF5 with HTD<sub>A</sub> is used to analyze the cascade model starting from the initial distribution  $\boldsymbol{\pi}_0(0, 0, 0, 0, 0) = 1$ . We remark that all results are obtained in at most 3,783 seconds with relative errors in  $[5 \times 10^{-7}, 10^{-3}]$  using a maximum truncated state space size of 58,786,560 and a maximum of 2,301,678 nonzeros.



**Fig. 2.** Mean values with BDF5 using HTD<sub>A</sub> for the cascade model

**Acknowledgement.** Part of this work is supported by the Alexander von Humboldt Foundation through the Research Group Linkage Programme. The research of M. Can Orhan is carried out during his PhD studies at Bilkent University and supported by The Scientific and Technological Research Council of Turkey under grant 2211-A. We thank the referees whose comments led to an improved manuscript.

## References

1. Buchholz, P., Dayar, T., Kriege, J., Orhan, M.C.: On compact solution vectors in Kronecker-based Markovian analysis. *Perform. Eval.* **115**, 132–149 (2017)
2. CompactTransientSolver software (2017). <http://www.cs.bilkent.edu.tr/~tugrul/software.html>
3. Dayar, T.: Analyzing Markov Chains using Kronecker Products: Theory and Applications. Springer, New York (2012). <https://doi.org/10.1007/978-1-4614-4190-8>
4. Goutsias, J., Jenkinson, G.: Markovian dynamics on complex reaction networks. *Phys. Rep.* **529**(2), 199–264 (2013)
5. Hackbusch, W.: Tensor Spaces and Numerical Tensor Calculus. Springer, Heidelberg (2012). <https://doi.org/10.1007/978-3-642-28027-6>
6. Hegland, M., Burden, C., Santoso, L., MacNamara, S., Booth, H.: A solver for the stochastic master equation applied to gene regulatory networks. *J. Comput. Appl. Math.* **205**(2), 708–724 (2007)
7. Kazeev, V., Khammash, M., Nip, M., Schwab, C.: Direct solution of the chemical master equation using quantized tensor trains. *PLoS Comput. Biol.* **10**(3), e1003359 (2014)
8. Khoromskij, B.N.:  $O(d \log N)$ -Quantics approximation of  $N - d$  tensors in high-dimensional numerical modeling. *Constructive Approximation* **34**(2), 257–280 (2011)
9. Kressner, D., Tobler, C.: htucker – A Matlab toolbox for tensors in hierarchical Tucker format. Technical Report 2012–02, Mathematics Institute of Computational Science and Engineering, Lausanne (2012)
10. Orhan, M.C.: On the Numerical Analysis of Infinite Multi-dimensional Markov Chains. PhD Thesis, Department of Computer Engineering, Bilkent University, Ankara (2017)
11. Shampine, L.F., Reichelt, M.W.: The MATLAB ODE suite. *SIAM J. Sci. Comput.* **18**(1), 1–22 (1997)
12. Stewart, W.J.: Introduction to the Numerical Solution of Markov Chains. Princeton University Press, Princeton (1994)