RATE CONSTANT MATRIX CONTRACTION METHOD FOR STIFF MASTER EQUATIONS WITH DETAILED BALANCE*

SATORU IWATA[†‡], TAIHEI OKI[‡§], AND SHINSAKU SAKAUE[¶‖§]

Abstract. This paper considers master equations for Markovian kinetic schemes that possess the detailed balance property. Chemical kinetics, as a prime example, often yields large-scale, highly stiff equations. Based on chemical intuitions, Sumiya et al. (2015) presented the rate constant matrix contraction (RCMC) method that computes approximate solutions to such intractable systems.

This paper aims to establish a mathematical foundation for the RCMC method. We present a reformulated RCMC method in terms of matrix computation, deriving the method from several natural requirements. We then perform a theoretical error analysis based on eigendecomposition and discuss implementation details caring about computational efficiency and numerical stability. Through numerical experiments on synthetic and real kinetic models, we validate the efficiency, numerical stability, and accuracy of the presented method.

Key words. Chemical kinetics, differential equations, quasi-steady-state approximation, rate constant matrix contraction method

MSC codes. 65C40, 65L04, 65F60

1. Introduction.

A *master equation* is a system of first-order ordinary differential equations (ODEs) describing the time evolution of the quantity or probability distribution on a finite number of states. This paper focuses on master equations for *Markovian kinetic schemes*, which are linear ODEs written in the form of

$$\dot{x}(t) = Kx(t), \tag{1.1}$$

where $x(t)$ is an n-dimensional vector whose ith component $x_i(t)$ represents the quantity or probability of occupying state i at time $t \in \mathbb{R}$ and K is an $n \times n$ matrix whose (i, j) entry K_{ij} with $i \neq j$ is a non-negative value called the (*transition*) *rate constant* from state j to state i. Each diagonal entry in K is the minus of the sum of off-diagonals in the column to which the diagonal belongs, i.e., $K_{jj} = -\sum_{i \neq j} K_{ij}$. This equality implies that the equation satisfies the *conservation law*, i.e., $\sum_{i=1}^{n} x_i(t)$ is preserved throughout time evolution. The matrix K is called a *rate constant matrix*.

A master equation (1.1) is said to satisfy the *detailed balance condition* if there exists a positive vector $\pi = (\pi_i)_n \in \mathbb{R}^n$ such that

$$K_{ij}\pi_j = K_{ji}\pi_i \tag{1.2}$$

holds for each pair of i and j. The condition (1.2) implies $K\pi = \mathbf{0}$. If the system is irreducible (i.e., the supporting graph of K is (strongly) connected), π is a *stationary* (or *equilibrium*) *distribution*, which is a unique distribution such that $x(t)$ converges to π as $t \to \infty$ regardless of the initial distribution $p = (p_i)_i = x(0)$ with $\sum_{i=1}^{n} p_i = \sum_{i=1}^{n} \pi_i$, as expounded in [1].

---

*The author accepted manuscript (AAM).

Funding: This work was supported by JST ERATO Grant Number JPMJER1903 and JST CREST Grant Number JPMJCR24Q2.

[†]Department of Mathematical Informatics, Graduate School of Information Science and Technology, University of Tokyo, Tokyo 113-8656, Japan. E-mail: iwata@mist.i.u-tokyo.ac.jp

[‡]Institute for Chemical Reaction Design and Discovery (ICReDD), Hokkaido University, Sapporo, Hokkaido 001-0021, Japan. E-mail (T. Oki): oki@icredd.hokudai.ac.jp

[§]Center for Advanced Intelligence Project, RIKEN, Tokyo 103-0027, Japan.

[¶]CyberAgent, Tokyo 150-6121, Japan. E-mail: shinsaku.sakaue@gmail.com

[‖]National Institute of Informatics, Tokyo 101-8430, Japan.

Master equations with detailed balance often arise from a monomolecular chemical kinetics simulation of reaction path networks. Within this context, the states represent energetically stable atomic configurations in the potential energy surface, called *equilibrium states* (EQs). If the system is described by a canonical ensemble, the rate constant K_{ij} from state j to state i with $i \neq j$ is given by

$$K_{ij} = \Gamma \frac{k_{\mathrm{B}} T}{h} \exp\left(-\frac{E_{i-j} - E_j}{RT} \right), \tag{1.3}$$

where k_{B} is the Boltzmann constant, h is the Planck constant, R is the gas constant, T is the temperature, Γ is the transmission coefficient, and E_j and $E_{i-j} = E_{j-i}$ are the potential energies of equilibrium state j and the transition state between equilibrium states i and j, respectively [16]. The rate constant matrix K determined by (1.3) satisfies the detailed balance condition (1.2) with the stationary distribution π given by

$$\pi_i = \frac{\exp\left(-\frac{E_i}{RT}\right)}{\sum_{j=1}^{n} \exp\left(-\frac{E_j}{RT}\right)},$$

which is called the *Boltzmann distribution*. Rate constant matrices derived from the microcanonical ensemble also satisfy the detailed balance condition; see [17].

Due to the wide range of reaction time scales, chemical kinetics simulations often yield stiff master equations. In particular, recent developments in automatic chemical reaction pathway search techniques, such as the anharmonic downward distortion following (ADDF) method [14] and the single-component artificial force induced reaction (SC-AFIR) method [13], have led to the emergence of large-scale and extremely stiff master equations, where the ratio of the largest to smallest non-zero entries of K may exceed 10^{150} (see Table 1). Solving such ODEs based on finite difference discretization of differentiation has been recognized as a difficult task. While the solution $x(t)$ to (1.1) can also be obtained from the eigendecomposition of K via the analytic expression $x(t) = \mathrm{e}^{tK} p$ for any initial distribution $p = x(0)$ and time t, computing the eigendecomposition of an extremely ill-conditioned matrix K usually requires floating-point arithmetic with precision at least the spectral condition number of K, which is computationally expensive [21].

To address such intractable equations, Sumiya et al. [16, 17] presented the *rate constant matrix contraction* (RCMC) method for systems described by the canonical or microcanonical ensemble. Based on the quasi-steady-state approximation (QSSA), the RCMC method maintains a set $S \subseteq [n]$ of states regarded to be steady, which monotonically grows from the empty set. For each S, the method computes a vector $q = (q_i)_i \in \mathbb{R}^n$ that approximates the solution $x(t)$ at time t when the states in S and $T \coloneqq [n] \setminus S$ can be regarded to be steady and transient, respectively. The output q preserves the component sum, i.e., $\sum_{i=1}^{n} p_i = \sum_{i=1}^{n} q_i$ holds, and the method runs fast for large-scale, stiff equations because it does not rely on numerical integration, unlike ordinary numerical methods. The RCMC method has been successfully applied to the kinetics-based navigation in the SC-AFIR method [15]. The RCMC method was, however, designed based on chemical intuition, and its mathematical validity and properties have not been discussed thoroughly.

Contributions. This paper improves the RCMC method by establishing mathematical foundations, summarized as follows.

First, we reformulate the RCMC method in terms of matrix computations. The reformulated method applies to any master equation (1.1) whose rate constant matrix

K satisfies the detailed balance condition (1.2). We describe the method by means of matrix operations, while all calculation formulas in the original form [16, 17] are written with direct reference to the matrix entries. In our reformulation, the output q of the method is expressed as a matrix-vector multiplication Vp, where $p \in \mathbb{R}^n$ is a given initial value and $V \in \mathbb{R}^{n\times n}$ is a matrix approximating e^{tK} that can be explicitly constructed from the rate constant matrix K and the set S of steady states. We observe that the matrix V can be determined from natural requirements coming from properties of e^{tK} and the QSSA. Furthermore, we propose a variant of V, which is another natural approximation of e^{tK}. The methods with original and modified V are referred to as Types A and B, respectively. While the matrix V of Type A is easier to compute, the matrix V of Type B is mathematically more tractable as it is indeed the orthogonal projection onto the QSSA subspace.

Second, we theoretically analyze the error between the approximate solution $q = Vp$ and the exact solution $x(t) = \mathrm{e}^{tK}p$. Here, the error is measured in terms of the *π-norm*; see Section 2.1 for definition. The derived error bounds depend on the spectral radii of matrices obtained from K and S, as well as the coefficients in the expansion of p with respect to an eigenbasis of K. Roughly speaking, q becomes a good approximation to $x(t)$ if p consists of eigenvectors of K with eigenvalues far from the spectral radii. We give a better and simpler bound for Type B than for Type A. Our error estimation for Type B leads to a principled way to determine the optimal reference time t^* such that q should be regarded as an approximation of $x(t^*)$.

Third, based on our reformulation and subsequent error analysis, we discuss implementation details of the RCMC method. We provide efficient implementations for calculating the optimal reference time t^* and for the incremental updates of the LU decomposition within the algorithm. We also address concerns regarding numerical stability.

Finally, we conduct numerical experiments using one synthetic dataset and two actual chemical reaction networks to demonstrate the efficiency, numerical stability, and accuracy of the RCMC method. We observe that the RCMC method of both Types A and B using double precision is capable of producing approximate solutions within mere seconds to minutes for master equations with $n \approx 1{,}700$. This is in contrast to the multi-precision eigendecomposition, which took more than one hour for the same matrix. The experimental results substantiate that the actual errors are within our theoretical bounds, subject to issues arising from the machine epsilon. Furthermore, we compare three methods for calculating the optimal reference time t^* and analyze the errors using π- and L_∞-norms.

Previous and Related Work. The RCMC method was originally introduced in [16] as a method to reduce the master equation (1.1) with the rate constant matrix derived from the canonical ensemble (1.3) into a smaller system by combining repeated applications of QSSA and an operation similar to *lumping* [8, 20], which is a classical reduction method for (generally nonlinear) master equations by grouping several states into a single state. The RCMC method in the aforementioned form of computing approximate solutions was subsequently proposed in [17], inheriting the spirit of lumping in its design.

QSSA is classified as a reduction method for generally nonlinear master equations based on timescale analysis. Other representative examples of methods based on timescale analysis include *intrinsic low-dimensional manifold* (ILDM) method [11, 12] and *computational singular perturbation* (CSP) [10]. Both methods rely on the eigendecomposition of the rate constant matrix (or the Jacobian matrix for nonlinear systems), which is prohibitive for large-scale, highly stiff systems we are concerned with.

In contrast, the RCMC method can be executed without performing the eigendecomposition, while our error analysis depends on it. See [18, Chapter 7] for a comprehensive survey of reduction techniques for kinetics equations.

Subsequent Work. After the submission of this paper, the following subsequent work appeared. Hemmi, Iwata, and Oki [6] proposed a fast and numerically stable implementation of the RCMC method, leveraging its connection to the greedy algorithm for submodular function maximization; see Section SM5 in supplemental materials for details on the connection. Harabuchi et al. [5] provided an explicit formulation of the derivative of $q = Vp$ with respect to the entries of K. They demonstrated that these derivatives are useful for analyzing elementary steps that control the reactivity in complicated reaction path networks.

Organization. The rest of this paper is organized as follows. Section 2 affords preliminaries on rate constant matrices, master equations, QSSA, and the RCMC method. Section 3 reformulates the RCMC method in terms of matrix operations. Section 4 provides an error analysis for the RCMC method. Section 5 presents the overall framework of the RCMC method and discusses its implementation details. Section 6 shows the results of numerical experiments. Finally, Section 7 concludes this paper.

2. Preliminaries.

This section provides preliminaries on the master equation (1.1) and rate constant matrices.

Notations. Let $\mathbb{R}$, $\mathbb{R}_{\geq 0}$, and $\mathbb{R}_{>0}$ denote the set of reals, non-negative reals, and positive reals, respectively. For a positive integer n, define $[n] \coloneqq \{1, \ldots, n\}$.

For $I \subseteq [n]$, let $\mathbb{R}^I$ be the $|I|$-dimensional real vector space having coordinates indexed by I. We call $\mathbb{R}^I$ a *coordinate subspace* of $\mathbb{R}^J$ for $I \subseteq J \subseteq [n]$. For $I, J \subseteq [n]$, let $\mathbb{R}^{I \times J}$ be the set of $|I| \times |J|$ matrices whose rows and columns are indexed by I and J, respectively. We simply denote $\mathbb{R}^n \coloneqq \mathbb{R}^{[n]}$, $\mathbb{R}^{I \times n} \coloneqq \mathbb{R}^{I \times [n]}$, $\mathbb{R}^{n \times J} \coloneqq \mathbb{R}^{[n] \times J}$, and $\mathbb{R}^{n \times n} \coloneqq \mathbb{R}^{[n] \times [n]}$. For a vector $y \in \mathbb{R}^n$, let $y_I \in \mathbb{R}^I$ denote the subvector of y indexed by $I \subseteq [n]$, and $A_{IJ} \in \mathbb{R}^{I \times J}$ the submatrix of $A \in \mathbb{R}^{n \times n}$ indexed by rows I and columns J. For brevity, we rewrite $y_{\{i\}}$, $A_{\{i\}J}$, $A_{I\{j\}}$, and $A_{\{i\}\{j\}}$ as y_i, A_{iJ}, A_{Ij}, and A_{ij}, respectively, for $I, J \subseteq [n]$ and $i, j \in [n]$. For $I \subseteq [n]$, we mean by A_{II}^{-1} the inverse of A_{II}. The transpose of a vector a and a matrix A are denoted by $a^\top$ and $A^\top$, respectively.

Let $\mathbf{1}_n$ be the all-one vector in the n-dimensional space and I_n the $n \times n$ identity matrix. For $J \subseteq [n]$, let $\mathbf{1}_J$ and I_J be the all-one vector in $\mathbb{R}^J$ and the identity matrix in $\mathbb{R}^{J \times J}$, respectively. Let $\mathbf{0}$ and O be the zero vector and the zero matrix of an appropriate dimension (size), respectively.

Vectors and matrices are called *non-negative* if their components (entries) are non-negative. For $y \in \mathbb{R}^n$, we mean by $y \geq \mathbf{0}$ that y is a non-negative vector. Similarly, $A \geq O$ for $A \in \mathbb{R}^{n \times m}$ means that A is a non-negative matrix. For a row or column vector y, $\mathrm{diag}(y)$ denotes the diagonal matrix obtained by arranging the components of y on the diagonals.

Let Δ_n denote the $(n-1)$-dimensional probability simplex and Δ_n° be its relative interior, i.e., $\Delta_n = \{y \in \mathbb{R}^n_{\geq 0} \mid \mathbf{1}_n^\top y = 1\}$ and $\Delta_n^\circ = \{y \in \mathbb{R}^n_{>0} \mid \mathbf{1}_n^\top y = 1\}$.

2.1. Rate Constant Matrices.

Let $K \in \mathbb{R}^{n \times n}$ be an $n \times n$ matrix satisfying the following conditions:

(RCM1) $K_{ij} \geq 0$ for $i \neq j$,
(RCM2) $\mathbf{1}_n^\top K = \mathbf{0}^\top$,
(RCM3) the detailed balance condition (1.2) holds for some $\pi \in \Delta_n^\circ$.

We call such K and π a *rate constant matrix* and a *stationary distribution*, respectively (see [20]). A matrix satisfying (RCM1) is called a *Metzler matrix* [4] or an *essentially non-negative matrix* [19]. Setting

$$\Pi \coloneqq \mathrm{diag}(\pi),$$

we can simply write (RCM3) as $K\Pi = \Pi K^\top$, i.e., K is a self-adjoint matrix with respect to the π-*inner product* in the following sense.

For $a, b \in \mathbb{R}^n$, we define the π-*inner product* by

$$\langle a, b \rangle_\pi \coloneqq a^\top \Pi^{-1} b. \tag{2.1}$$

The π-inner product induces the π-*norm* $\|a\|_\pi \coloneqq \sqrt{\langle a, a \rangle_\pi}$. Unless otherwise stated, we regard $\mathbb{R}^n$ as a metric vector space equipped with the π-inner product. Similar arguments to the standard inner product hold for basic notions such as adjoint matrices and positive definiteness in this space. Specifically, the *adjoint matrix* of $A \in \mathbb{R}^{n\times n}$ is the matrix A^* such that $\langle a, Ab \rangle_\pi = \langle A^* a, b \rangle_\pi$ for all $a, b \in \mathbb{R}^n$, and is explicitly written as $A^* = \Pi A^\top \Pi^{-1}$. A matrix $A \in \mathbb{R}^{n\times n}$ is *self-adjoint* if $A = A^*$, i.e., $A\Pi = \Pi A^\top$. Thus, (RCM3) is nothing but a statement that K is self-adjoint. See Sections SM1.1 and SM1.2 in supplemental materials for details of adjoint matrices and positive semi-definiteness on the π-inner product space.

Taking the transpose on both sides of (RCM2), we get $\mathbf{0} = K^\top \mathbf{1}_n = \Pi^{-1} K \Pi \mathbf{1}_n = \Pi^{-1} K \pi$ and thereby $K\pi = \mathbf{0}$. Conversely, a self-adjoint matrix $K \in \mathbb{R}^{n\times n}$ satisfying $K\pi = \mathbf{0}$ meets (RCM2). Thus, $\mathbf{1}_n^\top K = \mathbf{0}^\top$ and $K\pi = \mathbf{0}$ are equivalent under (RCM3).

The eigenvalues of a self-adjoint matrix are real. A self-adjoint matrix $A \in \mathbb{R}^{n\times n}$ is called *positive semi-definite* if $\langle a, Aa \rangle_\pi$ is non-negative for $a \in \mathbb{R}^n$, or equivalently, all eigenvalues of A are non-negative. It follows from the Gershgorin circle theorem that a rate constant matrix K has non-positive eigenvalues, i.e., K is *negative semi-definite*. Let $0 = \lambda_1 \geq \cdots \geq \lambda_n$ be the eigenvalues of K and u_k an eigenvector corresponding to λ_k for $k \in [n]$. Since K is self-adjoint, we can take $u_1, \ldots, u_n$ such that $\langle u_k, u_\ell \rangle_\pi$ is 1 if $k = \ell$ and 0 otherwise. We can set $u_1 = \pi$ by $K\pi = \mathbf{0}$. Note that π satisfies $\|\pi\|_\pi^2 = \pi^\top \Pi^{-1} \pi = \mathbf{1}_n^\top \pi = 1$, i.e., π is normalized. See Section SM1.2 in supplemental materials for details on the semi-definiteness of self-adjoint matrices.

The minus of K is a *Z-matrix*, i.e., every off-diagonal entry is non-negative. In addition, $-K$ has non-negative eigenvalues. Such a matrix is called an *M-matrix* [2, Chapter 6]. Principal submatrices of an M-matrix are M-matrices again, and the inverse of a nonsingular M-matrix is a non-negative matrix. Therefore, $K_{SS}^{-1} \leq O$ holds for $S \subseteq [n]$, provided that K_{SS} is nonsingular.

The matrix $L \coloneqq -K\Pi$ is a (weighted) *graph Laplacian matrix*, i.e., a symmetric Z-matrix such that $L\mathbf{1}_n = \mathbf{0}$. Conversely, given a graph Laplacian matrix $L \in \mathbb{R}^{n\times n}$ and $\pi \in \Delta_n^\circ$, we can construct a rate constant matrix K by $K \coloneqq -L\Pi^{-1}$.

2.2. Master Equations with Detailed Balance. Consider the master equation (1.1) associated with a rate constant matrix K with stationary distribution π. Letting $p \in \mathbb{R}^n$ be an initial value $x(0)$, the solution $x(t)$ of (1.1) can be analytically represented by $x(t) = \mathrm{e}^{tK} p$, where $\mathrm{e}^{tK} \coloneqq \sum_{d=0}^\infty \frac{1}{d!}(tK)^d$. The matrix exponential e^{tK} enjoys the following properties:

(E1) $\mathrm{e}^{tK} \geq O$ for $t \geq 0$,
(E2) $\mathbf{1}_n^\top \mathrm{e}^{tK} = \mathbf{1}_n^\top$,
(E3) e^{tK} is self-adjoint.

We can directly derive (E2) and (E3) from (RCM2) and (RCM3), respectively. Property (E1) follows from the general fact that the matrix exponential of a Metzler matrix is non-negative [19, Section 8.2]. Note that (E2) is equivalent to $\mathrm{e}^{tK}\pi = \pi$ under (E3), which can be shown in the same way as the equivalence of (RCM2) and $K\pi = \mathbf{0}$ under (RCM3). These properties on e^{tK} give rise to the following well-known properties on the solution $x(t)$ of (1.1). We here describe a short proof to clarify from which property of e^{tK} each claim is derived.

PROPOSITION 2.1. *Let $x(t)$ be the solution of (1.1) with $x(0) = p \in \mathbb{R}^n$. If $p \in \Delta_n$, then $x(t) \in \Delta_n$ for all $t \geq 0$. If $p = \pi$, then $x(t) = \pi$ for all $t \geq 0$.*

Proof. By (E1), we have $x(t) = \mathrm{e}^{tK}p \geq \mathbf{0}$ for $p \geq \mathbf{0}$ and $t \geq 0$. By (E2), it holds that $\mathbf{1}_n^\top x(t) = \mathbf{1}_n^\top \mathrm{e}^{tK} p = \mathbf{1}_n^\top p = 1$ if $\mathbf{1}_n^\top p = 1$. By (E2) and (E3), we have $\mathrm{e}^{tK}\pi = \pi$, which implies $x(t) = \mathrm{e}^{tK}\pi = \pi$ for $p = \pi$. ∎

Another important property of (1.1) is the *linearity* of solutions. That is, if $x(t)$ and $y(t)$ are the solutions of (1.1) with initial values $x(0) = p$ and $y(0) = q$, respectively, then $ax(t) + by(t)$ is the solution of (1.1) with initial value $ap + bq$ for $a, b \in \mathbb{R}$. We also remark that the eigenvalues of e^{tK} with $t \geq 0$ are in $[0, 1]$ because K is negative semi-definite.

2.3. Quasi-Steady-State Approximation. QSSA, originated in the work of Bodenstein [3], is a classical tool for reducing the number of variables and stiffness in a (generally nonlinear) master equation. We here describe the method specialized to our linear setting (1.1).

Suppose that we have a bipartition $\{S, T\}$ of $[n]$ separating the components of $x(t)$ into *fast variables* $x_S(t)$ and *slow variables* $x_T(t)$. With this bipartition, the master equation (1.1) can be displayed as

$$\begin{cases} \dot{x}_S(t) = K_{SS}x_S(t) + K_{ST}x_T(t), \\ \dot{x}_T(t) = K_{TS}x_S(t) + K_{TT}x_T(t). \end{cases} \tag{2.2}$$

Now, $x_S(t)$ is approximately regarded as "steady state," i.e., $\dot{x}_S(t) \approx \mathbf{0}$. The QSSA replaces the left-hand side $\dot{x}_S(t)$ of the first equation in (2.2) with $\mathbf{0}$. Namely, letting $\bar{x}(t)$ be a new variable vector, we can formulate the QSSA of (2.2) as

$$\begin{cases} \mathbf{0} = K_{SS}\bar{x}_S(t) + K_{ST}\bar{x}_T(t), \\ \dot{\bar{x}}_T(t) = K_{TS}\bar{x}_S(t) + K_{TT}\bar{x}_T(t). \end{cases} \tag{2.3}$$

The system (2.3) consists of purely algebraic equations and differential equations. Such a system of equations is called a *differential-algebraic equation* (DAE). The algebraic equations constrain the solution to the *QSSA subspace*

$$W := \{y \in \mathbb{R}^n \mid K_{SS}y_S + K_{ST}y_T = \mathbf{0}\}. \tag{2.4}$$

If K_{SS} is nonsingular, $\bar{x}_S(t)$ can be determined from $\bar{x}_T(t)$ by

$$\bar{x}_S(t) = -K_{SS}^{-1}K_{ST}\bar{x}_T(t). \tag{2.5}$$

Substituting (2.5) back into the differential equation in (2.3), we obtain a differential equation involving only $\bar{x}_T(t)$:

$$\dot{\bar{x}}_T(t) = \big(K_{TT} - K_{TS}K_{SS}^{-1}K_{ST}\big)\bar{x}_T(t). \tag{2.6}$$

Algorithm 1 The RCMC method

Input: A rate constant matrix $K \in \mathbb{R}^{n\times n}$, an initial value $p \in \Delta_n$, and an end time $t_{\max} \in \mathbb{R}_{\geq 0}$.

Output: $t^{(0)}, t^{(1)}, \ldots \in \mathbb{R}_{\geq 0}$ and $q^{(0)}, q^{(1)}, \ldots \in \mathbb{R}^n$.

1: $S^{(0)} \coloneqq \emptyset$, $T^{(0)} \coloneqq [n]$, $D^{(0)} \coloneqq K$, $t^{(0)} \coloneqq 0$

▷ Step 1

2: **for** $k = 1, 2, \ldots$ **do**

3: Take $j^{(k)} \in \arg\max\left\{ \left| D^{(k-1)}_{jj} \right| \,\middle|\, j \in T^{(k-1)} \right\}$

4: $S^{(k)} \coloneqq S^{(k-1)} \cup \left\{ j^{(k)} \right\}$ and $T^{(k)} \coloneqq T^{(k-1)} \setminus \left\{ j^{(k)} \right\}$

5: $D^{(k)} \coloneqq D^{(k-1)}_{T^{(k)}T^{(k)}} - \dfrac{D^{(k-1)}_{T^{(k)}j^{(k)}} D^{(k-1)}_{j^{(k)}T^{(k)}}}{D^{(k-1)}_{j^{(k)}j^{(k)}}}$

6: Compute $t^{(k)}$ by (3.1) or (3.2)

7: **if** $t^{(k)} > t_{\max}$ **then**

8: **break**

▷ Step 2

9: For every $k = 0, 1, 2, \ldots$, output $t^{(k)}$ and $q^{(k)} \coloneqq \mathrm{proj}(Vp)$, where V is the matrix (3.3) of a designated type (Type A or B)

The coefficient matrix $D \coloneqq K_{TT} - K_{TS}K_{SS}^{-1}K_{ST}$ in (2.6) is called the *Schur complement* of K_{SS} in K. The Schur complement D is again a rate constant matrix in the sense that it satisfies (RCM1)–(RCM3) with stationary distribution proportional to π_T. Hence, QSSA can be applied repeatedly. From the property of the Schur complement, the equation obtained by applying the QSSA twice with $S = I$ and J coincides with the single application of the QSSA with $S = I \cup J$ for any disjoint subsets $I, J \subseteq [n]$.

3. Rate Constant Matrix Contraction Method. In this section, we describe the RCMC method in terms of matrix operations. Readers interested in the original form of the method given by Sumiya et al. [16, 17] and its correspondence to our form are referred to Section SM2 in supplemental materials.

3.1. Method Description. Algorithm 1 describes the RCMC method. The RCMC method consists of the following two steps:

Step 1: Greedy selection of steady-states (Lines 2–8),

Step 2: Computation of an approximate solutions (Line 9).

In Step 1, the method constructs a growing sequence $\varnothing = S^{(0)} \subseteq S^{(1)} \subseteq \cdots \subseteq [n]$ of steady-states in a greedy fashion, along with the corresponding reference times $0 = t^{(0)} \leq t^{(1)} \leq t^{(2)} \leq \cdots$ such that the elements in $S^{(k)}$ and $T^{(k)} \coloneqq [n] \setminus S^{(k)}$ can be considered "steady" and "transient", respectively, at time $t^{(k)}$. Step 1 also maintains the Schur complement $D^{(k)}$ of $K_{S^{(k)}S^{(k)}}$ in K, which is the coefficient matrix of the equation (2.6) obtained by applying the QSSA with bipartition $\left\{ S^{(k)}, T^{(k)} \right\}$. Then, for every $k = 0, 1, 2, \ldots,$, Step 2 computes a vector $q^{(k)}$ that approximates the solution $x(t)$ of the master equation (1.1) with initial value $x(0) = p$ at time $t^{(k)}$. Details of Steps 1 and 2 are elaborated below.

Details of Step 1. In every kth iteration of Step 1, the algorithm finds the index of a diagonal entry in $D^{(k-1)}$ with the largest absolute value and designates it as the next steady state $j^{(k)}$ (Line 3). While there is no definitive way of selecting the next steady state, this manner stems from a standard heuristic used in QSSA that

considers $1/\left|D^{(k-1)}_{jj}\right|$ as the "timescale" of state $j \in T^{(k-1)}$ and selects the minimum-timescale state as the next steady state [18]. Additional justification of this method, based on error analysis and the connection to submodular function maximization, is discussed in Section SM5 in supplemental materials. If $\arg\max$ is not unique, $j^{(k)}$ can be taken as any maximizer. In practice, ties are usually resolved by selecting the smallest index.

At Line 6, the reference time $t^{(k)}$ should be chosen so that states in $S^{(k)}$ and $T^{(k)}$ can be considered steady and transient, respectively, at time $t^{(k)}$. Following the time-scale heuristics in QSSA, one might set $t^{(k)}$ as

$$t^{(k)} \leftarrow t^{(k)}_{\mathrm{diag}} \coloneqq \frac{1}{\left|D^{(k-1)}_{j^{(k)}j^{(k)}}\right|}. \tag{3.1}$$

The original description of the RCMC method [16, 17] essentially adopts this choice of $t^{(k)}$. In contrast, we will show in Section 4 that the following choice of $t^{(k)}$ is better in terms of the error bound:

$$t^{(k)} \leftarrow t^{(k)}_{\mathrm{eigen}} \coloneqq \frac{\ln 2}{\sqrt{\sigma\bigl(K_{S^{(k)}S^{(k)}}\bigr)\rho\bigl(D^{(k)}\bigr)}}, \tag{3.2}$$

where $\rho(A)$ and $\sigma(A)$ for a square matrix A denote the maximum and minimum absolute values of an eigenvalue of A, respectively.

Details of Step 2. Fix k and let $S = S^{(k)}, T = T^{(k)}$, and $t = t^{(k)}$. The exact solution $x(t)$ can be obtained by multiplying p by the matrix exponential e^{tK}. Here, instead of e^{tK}, Step 2 multiplies p by an alternative matrix $V \in \mathbb{R}^{n\times n}$ that serves as an approximation to e^{tK}. We now introduce two variants of V, referred to as Types A and B, as follows:

$$\begin{aligned}
V_{SS} &= K_{SS}^{-1}K_{ST}V_{TT}K_{TS}K_{SS}^{-1},\\
V_{ST} &= -K_{SS}^{-1}K_{ST}V_{TT},\\
V_{TS} &= -V_{TT}K_{TS}K_{SS}^{-1},\\
V_{TT} &= \begin{cases} \mathrm{diag}\bigl(\mathbf{1}_T^\top M\bigr)^{-1} & (\text{Type A}),\\ M^{-1} & (\text{Type B}), \end{cases}
\end{aligned} \tag{3.3}$$

where

$$M \coloneqq I_T + K_{TS}K_{SS}^{-2}K_{ST} \in \mathbb{R}^{T\times T}. \tag{3.4}$$

As seen in Section 3.2, the vector Vp lies in the QSSA subspace with respect to the bipartition $\{S, T\}$ of $[n]$. Moreover, Vp is guaranteed to be in the probability simplex Δ_n for Type A (Proposition 3.3), whereas it is not for Type B. Thus, after calculating Vp, we compute its (π-norm) projection $\mathrm{proj}(Vp)$ onto Δ_n, i.e.,

$$\mathrm{proj}(Vp) \coloneqq \arg\min\{\, \|q - Vp\|_\pi \mid q \in \Delta_n \,\}. \tag{3.5}$$

See Section SM3 in supplemental materials for the projection algorithm. We will show in Proposition 4.1 that the projection does not enlarge the distance from the exact solution $x(t)$.

While it is not immediately clear, Algorithm 1 of Type A is indeed equivalent to the original RCMC method presented by Sumiya et al. [16, 17]; see Section SM2 in supplemental materials for the correspondence. We see in Section 3.2 that (3.3)

is obtained from natural requirements as an approximation of e^{tK} and the idea of QSSA. The RCMC method of Type A is practically advantageous in computational efficiency and numerical stability (see Section 5). Meanwhile, the RCMC method of Type B has several mathematically desirable properties in comparison with Type A. For example, the matrix V of Type B is the (π-norm) projection onto the QSSA subspace (2.4) with respect to $\{S, T\}$; see Section 3.3. In Section 4, we provide better and simpler error bounds for Type B than Type A.

3.2. Approximating Matrix Exponential. We derive the matrix V in the form of (3.3) from several requirements that should be satisfied for V to approximate e^{tK}.

We introduce additional notions on coordinate subspaces. Since Π^{-1} is diagonal, for any $I \subseteq [n]$, Π_I^{-1} with $\Pi_I \coloneqq \mathrm{diag}(\pi_I)$ is naturally regarded as a metric on $\mathbb{R}^I$. The *adjoint matrix* of $A \in \mathbb{R}^{I\times J}$ with $I, J \subseteq [n]$ is then defined by $A^* \coloneqq \Pi_J A^\top \Pi_I^{-1} \in \mathbb{R}^{J\times I}$. Self-adjointness and positive semi-definiteness are similarly defined for matrices in $\mathbb{R}^{I\times I}$. If $A \in \mathbb{R}^{n\times n}$ is self-adjoint, A_{II} is also self-adjoint and $A_{IJ}^* = A_{JI}$ holds for any $I, J \subseteq [n]$. If $A \in \mathbb{R}^{n\times n}$ is positive semi-definite, so is A_{II}. See Section SM1.3 in supplemental materials for details on metrics of coordinate subspaces.

First, from (E1)–(E3), it is natural to expect the following properties to hold:

(V1) $V \geq O$,
(V2) $\mathbf{1}_n^\top V = \mathbf{1}_n^\top$,
(V3) V is self-adjoint.

Recall that (V2) is equivalent to $V\pi = \pi$ under (V3). Property (V1) is desirable to guarantee $Vp \geq \mathbf{0}$ for $p \geq \mathbf{0}$. As seen in the proof of Proposition 2.1, $\mathbf{1}_n^\top V = \mathbf{1}_n^\top$ and $V\pi = \pi$ correspond to the conservation law and the stationarity of π, respectively. Furthermore, we require V to satisfy the following:

(V4) $\mathrm{Im}\, V$ is contained in the QSSA subspace W (defined by (2.4)) with respect to a bipartition $\{S, T\}$ of $[n]$.

We assume that K_{SS} is nonsingular. This property comes from the idea of QSSA; that is, if V is a good approximation of e^{tK} and QSSA works well with $\{S, T\}$, then $Vp \approx \mathrm{e}^{tK}p = x(t)$ should be close to the QSSA subspace W.

We now design a matrix V so that it satisfies these properties. We first show that the form (3.3) of V is equivalent to (V3) and (V4).

PROPOSITION 3.1. *A matrix $V \in \mathbb{R}^{n\times n}$ satisfies (V3) and (V4) if and only if V is in the form of (3.3) with self-adjoint $V_{TT} \in \mathbb{R}^{T\times T}$.*

Proof. We first rephrase (V4). The statement $Vp \in W$ is written down as

$$\mathbf{0} = K_{SS}(Vp)_S + K_{ST}(Vp)_T = \big(K_{SS}V_{SS} + K_{ST}V_{TS}\big)p_S + \big(K_{SS}V_{ST} + K_{ST}V_{TT}\big)p_T.$$

Since this equation must be satisfied for all $p \in \mathbb{R}^n$, (V4) is equivalent to

$$K_{SS}V_{SS} + K_{ST}V_{TS} = O, \tag{3.6}$$

$$K_{SS}V_{ST} + K_{ST}V_{TT} = O. \tag{3.7}$$

Suppose that V satisfies (V3) and (V4). Solving (3.7) for V_{ST}, we obtain $V_{ST} = -K_{SS}^{-1}K_{ST}V_{TT}$. Now we have $V_{TS} = -V_{TT}K_{TS}K_{SS}^{-1}$ due to (V3) and (RCM3). Substituting this into (3.6), we have $V_{SS} = -K_{SS}^{-1}K_{ST}V_{TS} = K_{SS}^{-1}K_{ST}V_{TT}K_{TS}K_{SS}^{-1}$. Thus, V must be in the form of (3.3). In addition, the self-adjointness of the principal submatrix V_{TT} is necessary for (V3).

Conversely, suppose that a matrix V is in the form of (3.3) with self-adjoint V_{TT}. The self-adjointness (V3) of V can be directly verified by using $K_{ST}^* = K_{TS}$,

$K_{TS}^* = K_{ST}$ and the self-adjointness of K_{SS}^{-1} and V_{TT} to hold. The condition (V4) can be seen by checking (3.6) and (3.7). ∎

By Proposition 3.1, the matrix V satisfying (V3) and (V4) is uniquely determined from V_{TT}. We next rephrase (V1) and (V2) to derive conditions on V_{TT}. Recall that $M \in \mathbb{R}^{T\times T}$ is defined by (3.4).

PROPOSITION 3.2. *For $V \in \mathbb{R}^{n\times n}$ in the form of (3.3), the following hold.*
(1) *The matrix V satisfies (V1) if and only if $V_{TT} \geq O$.*
(2) *The matrix V satisfies (V2) if and only if $\mathbf{1}_T^\top M V_{TT} = \mathbf{1}_T^\top$.*

Proof. (1) The necessity is obvious. We show the sufficiency. Since $-K$ is an M-matrix, $-K_{SS}^{-1}$ is non-negative (see Section 2.1). Additionally, K_{ST} and K_{TS} are also non-negative by (RCM1) and $S \cap T = \varnothing$. Thus, $V \geq O$ follows from $V_{TT} \geq O$.

(2) Using (3.3), we can express the equality $\mathbf{1}_n^\top V = \mathbf{1}_n^\top$ as

$$\mathbf{1}_S^\top K_{SS}^{-1} K_{ST} V_{TT} K_{TS} K_{SS}^{-1} - \mathbf{1}_T^\top V_{TT} K_{TS} K_{SS}^{-1} = \mathbf{1}_S^\top, \tag{3.8}$$

$$-\mathbf{1}_S^\top K_{SS}^{-1} K_{ST} V_{TT} + \mathbf{1}_T^\top V_{TT} = \mathbf{1}_T^\top. \tag{3.9}$$

If (3.9) is satisfied, (3.8) is simplified by using (RCM2) as

$$-\mathbf{1}_T^\top K_{TS} K_{SS}^{-1} = \mathbf{1}_S^\top. \tag{3.10}$$

Thus, $\mathbf{1}_n^\top V = \mathbf{1}_n^\top$ is equivalent to (3.9). Substituting (3.10) into (3.9), we obtain $\mathbf{1}_T^\top = \big(\mathbf{1}_T^\top - \mathbf{1}_S^\top K_{SS}^{-1} K_{ST}\big) V_{TT} = \mathbf{1}_T^\top M V_{TT}$. ∎

The matrices V_{TT} of Types A and B given in (3.3) both satisfy the second condition in Proposition 3.2. Hence, (V2) holds on both types. Regarding the first condition in Proposition 3.2, the Type A matrix V_{TT} satisfies it, since $M = I_T + K_{TS} K_{SS}^{-2} K_{ST}$ is non-negative; this follows from the non-negativity of K_{TS}, K_{ST}, and K_{SS}^{-1}. This leads us to the following proposition.

PROPOSITION 3.3. *Let V be the matrix (3.3) of Type A. Then, $Vp \in \Delta_n$ holds for any $p \in \Delta_n$.*

Unfortunately, M^{-1} is not necessarily non-negative, and hence (V1) is not satisfied for Type B. However, we can project Vp onto Δ_n without making the distance to the exact solution larger. Proposition 3.3 helps us to reduce the computational cost and the complexity of the implementation for Type A, which does not require the projection.

We note that Proposition 3.1 guarantees that the vector Vp lies in the QSSA subspace with respect to $\{S, T\}$. In particular, for connected networks, we have $Vp = \pi$ if $|S| = n - 1$; that is, the final output of the RCMC method (without the projection for Type A) is the stationary distribution π. Indeed, the QSSA subspace with respect to $\{S, T\}$ is a $|T|$-dimensional subspace contining π. Hence, the $(n-1)$st QSSA space is the one-dimensional space spanned by π. Thus, we obtain that Vp is proportional to π, and in fact equals π because $\mathbf{1}_n^\top \pi = 1 = \mathbf{1}^\top Vp$. This holds for both Types A and B.

3.3. Spectral Properties of V. As described in Section 2.2, the matrix exponential e^{tK} with a rate constant matrix K and $t \geq 0$ has eigenvalues in $[0, 1]$. In this section, we see that the matrix V of either type enjoys this spectral property.

For self-adjoint matrices $A, B \in \mathbb{R}^{I\times I}$ with $I \subseteq [n]$, let $A \preceq B$ mean that $B - A$ is positive semi-definite. As with the usual case of symmetric matrices, $\preceq$ forms a partial order on the self-adjoint matrices, called the *Löwner order*, compatible with

the addition. See Section SM1.2 in supplemental materials for details on the Löwner order.

We first show that the two self-adjoint matrices M and V_{TT} are in the following relation of the Löwner order.

LEMMA 3.4. *The matrix V_{TT} of either type satisfies $I_T \preceq M \preceq V_{TT}^{-1}$.*

Proof. Letting $X \coloneqq K_{SS}^{-1}K_{ST}$, we have $X^* = K_{TS}K_{SS}^{-1}$ and $M = I_T + X^*X$. Thus, $I_T \preceq M$ holds. We show $M \preceq V_{TT}^{-1}$. This is clear for Type B since $M = V_{TT}^{-1}$. For Type A, $M - V_{TT}^{-1} = M - \operatorname{diag}\bigl(\mathbf{1}_T^{\top}M\bigr)$ is a rate constant matrix with stationary distribution proportional to π_T, i.e., it satisfies (RCM1)–(RCM3). Therefore, $M - V_{TT}^{-1}$ is negative semi-definite by the Gershgorin circle theorem, as mentioned in Section 2.1. ∎

For a positive semi-definite self-adjoint matrix A, let $A^{\frac{1}{2}}$ denote the *square root* of A, i.e., the unique positive semi-definite matrix whose square is A. Since $A \preceq B$ implies $P^*AP \preceq P^*BP$ for any $P \in \mathbb{R}^{n\times n}$, Lemma 3.4 can be expressed as follows.

LEMMA 3.5. *The matrix V_{TT} of either type satisfies $V_{TT} \preceq V_{TT}^{\frac{1}{2}}MV_{TT}^{\frac{1}{2}} \preceq I_T$.*

The next proposition gives the same bound on the eigenvalues of V as e^{tK}.

PROPOSITION 3.6. *The matrix V of either type satisfies $O \preceq V \preceq I_n$, i.e., all eigenvalues of V are in $[0, 1]$.*

To prove Proposition 3.6, we rewrite the matrix V by using a matrix $\Omega \in \mathbb{R}^{T\times n}$ defined as

$$\Omega_{TS} = -K_{TS}K_{SS}^{-1}, \quad \Omega_{TT} = I_T. \tag{3.11}$$

The adjoint matrix $\Omega^* \in \mathbb{R}^{n\times T}$ of Ω is given by

$$\Omega_{ST}^* = (\Omega_{TS})^* = -K_{SS}^{-1}K_{TS}, \quad \Omega_{TT}^* = (\Omega_{TT})^* = I_T.$$

It can be directly seen that V and M are expressed as $V = \Omega^*V_{TT}\Omega$ and $M = \Omega\Omega^*$, respectively.

Proof of Proposition 3.6. The positive semi-definiteness of V follows from $V = \Omega^*V_{TT}\Omega$ and $V_{TT} \succeq O$. To show $V \preceq I_n$, let (λ, u) be an eigenpair of $V_{TT}^{\frac{1}{2}}MV_{TT}^{\frac{1}{2}}$; $\lambda \leq 1$ holds from Lemma 3.5. By direct calculation, we obtain $V\Omega^*V_{TT}^{\frac{1}{2}}u = \lambda\Omega^*V_{TT}^{\frac{1}{2}}u$, meaning that λ is an eigenvalue of V as well. Since Ω is of row-full rank and V_{TT} is nonsingular, $\operatorname{rank} V = |T|$ holds. That is, V has $|T|$ non-zero eigenvalues, which are exactly the eigenvalues of $V_{TT}^{\frac{1}{2}}MV_{TT}^{\frac{1}{2}}$, and $n-|T|$ zero eigenvalues by the rank-nullity theorem. As any eigenvalue of $V_{TT}^{\frac{1}{2}}MV_{TT}^{\frac{1}{2}}$ is at most 1, $V \preceq I_n$ is obtained. ∎

As for Type B, the eigenvalues of V are either 0 or 1 because $V_{TT}^{\frac{1}{2}}MV_{TT}^{\frac{1}{2}} = I_T$ holds. This means that V is idempotent, i.e., $V^2 = V$. Accordingly, the matrix V of Type B is the orthogonal (with respect to the π-inner product) projection onto $\operatorname{Im} V = \operatorname{Im}\Omega^*$, which is the QSSA subspace with respect to $\{S, T\}$.

4. Error Analysis. In this section, we provide upper bounds on the error between the output of the RCMC method and the analytic solution to the master equation (1.1).

4.1. Results. Let $K \in \mathbb{R}^{n\times n}$ be a rate constant matrix with stationarity distribution $\pi \in \Delta_n^\circ$ and $V \in \mathbb{R}^{n\times n}$ the matrix (3.3) of Type A or B with bipartition $\{S,T\}$ of $[n]$. For an arbitrary $t \in \mathbb{R}_{\geq 0}$, we consider the following form of error:

$$\mathrm{Err}_{S,t}(p) \coloneqq \frac{\left\|\mathrm{proj}(Vp) - \mathrm{e}^{tK}p\right\|_\pi}{\|p\|_\pi} \quad (p \in \Delta_n), \tag{4.1}$$

where $\mathrm{proj}(\cdot)$ is defined by (3.5) and can be omitted for Type A due to Proposition 3.3. We provide bounds on the following two quantities:

- $\mathrm{Err}_{S,t}(p)$ with a given $p \in \Delta_n$,
- the expectation $\mathbb{E}_p[\mathrm{Err}_{S,t}(p)]$, where p is drawn from the uniform distribution on the vertices of Δ_n (i.e. the standard vectors $e_1, \dots, e_n$).

The distribution assumed in the expectation error is natural in that a typical simulation of a kinetic scheme starts from a single state.

Our first step is to bound $\left\|\mathrm{proj}(Vp) - \mathrm{e}^{tK}p\right\|_\pi$ by $\left\|\left(V - \mathrm{e}^{tK}\right)p\right\|_\pi$ as follows.

PROPOSITION 4.1. *For $a \in \Delta_n$ and $q \in \mathbb{R}^n$, we have $\|\mathrm{proj}(q) - a\|_\pi \leq \|q - a\|_\pi$.*

Proof. This inequality is a π-norm variant of the Pythagorean theorem, which follows from the standard Euclidean version as follows. Define $\Pi^{-\frac{1}{2}}\Delta_n \coloneqq \left\{\Pi^{-\frac{1}{2}}y \;\middle|\; y \in \Delta_n\right\}$, which one can easily prove to be convex. Note that we have $\mathrm{proj}(q) = \arg\min\left\{\left\|\Pi^{-\frac{1}{2}}(y-q)\right\|_\pi \;\middle|\; y \in \Delta_n\right\} = \Pi^{\frac{1}{2}}\arg\min\left\{\left\|y' - \Pi^{-\frac{1}{2}}q\right\|_\pi \;\middle|\; y' \in \Pi^{-\frac{1}{2}}\Delta_n\right\}$, i.e., $\Pi^{-\frac{1}{2}}\,\mathrm{proj}(q)$ is the Euclidean projection of $\Pi^{-\frac{1}{2}}q$ onto $\Pi^{-\frac{1}{2}}\Delta_n$. Since $\Pi^{-\frac{1}{2}}a \in \Pi^{-\frac{1}{2}}\Delta_n$, the Pythagorean theorem implies $\left\|\Pi^{-\frac{1}{2}}\,\mathrm{proj}(q) - \Pi^{-\frac{1}{2}}a\right\|_\pi \leq \left\|\Pi^{-\frac{1}{2}}q - \Pi^{-\frac{1}{2}}a\right\|_\pi$, which is nothing but the desired inequality. ∎

Proposition 4.1 ensures that the projection $\mathrm{proj}(Vp)$ does not enlarge the distance between the approximate and exact solutions. Proposition 4.1 together with the spectral properties of V and e^{tK} leads us to the following uniform error bound.

LEMMA 4.2. *For any $p \in \Delta_n$, it holds that $\mathrm{Err}_{S,t}(p) \leq 1$.*

We show Lemma 4.2 in Section 4.2 with the aid of the notion of the *matrix norm*.

We then focus on an error bound specific to a given $p \in \Delta_n$. We first decompose p in terms of the eigenvectors of K.

LEMMA 4.3. *Let $\{u_1, \dots, u_n\}$ be an orthonormal eigenbasis of K. The matrix V of either type satisfies*

$$\frac{\left\|\left(V - \mathrm{e}^{tK}\right)p\right\|_\pi}{\|p\|_\pi} \leq \sum_{k=1}^{n} \frac{|\langle u_k, p\rangle_\pi|}{\|p\|_\pi}\left\|\left(V - \mathrm{e}^{tK}\right)u_k\right\|_\pi \quad (p \in \Delta_n), \tag{4.2}$$

$$\mathbb{E}_p\left[\frac{\left\|\left(V - \mathrm{e}^{tK}\right)p\right\|_\pi}{\|p\|_\pi}\right] \leq \sqrt{\frac{1}{n}\sum_{k=1}^{n}\left\|\left(V - \mathrm{e}^{tK}\right)u_k\right\|_\pi^2}. \tag{4.3}$$

For each eigenvector u of K, the value $\left\|\left(V - \mathrm{e}^{tK}\right)u\right\|_\pi$ is bounded as follows.

LEMMA 4.4. *For an eigenpair (λ, u) of K with $\|u\|_\pi = 1$, we have*

$$\left\|\left(V - \mathrm{e}^{tK}\right)u\right\|_\pi \leq \min\left\{1, \|Vu\|_\pi + \mathrm{e}^{t\lambda}, \|(I_n - V)u\|_\pi + 1 - \mathrm{e}^{t\lambda}\right\}.$$

Proofs of Lemmas 4.3 and 4.4 are also given in Section 4.2. The terms $\|Vu\|_\pi$ and $\|(I_n - V)u\|_\pi$ in (4.2) and (4.3) are bounded by the following theorem. Let $\|A\|_{\infty,\mathrm{off}}$ denote the maximum absolute value of an off-diagonal entry of A.

LEMMA 4.5. *Let D be the Schur complement of K_{SS} in K. For an eigenpair (λ, u) of K with $\|u\|_\pi = 1$, the following two statements hold.*

(1) *As for Type A, we have*

$$\|Vu\|_\pi \le \frac{\rho(D)}{|\lambda|}, \quad \|(I_n - V)u\|_\pi \le \frac{|\lambda| + \sqrt{2\left\|\Pi^{-1}K\right\|_{\infty,\mathrm{off}}|\lambda|}}{\sigma(K_{SS})}.$$

(2) *As for Type B, we have*

$$\|Vu\|_\pi \le \frac{\rho(D)}{|\lambda|}, \quad \|(I_n - V)u\|_\pi \le \frac{|\lambda|}{\sigma(K_{SS})}.$$

Lemma 4.5 is proved in Sections 4.3 and 4.4. It should be noted that the upper bounds on $\|Vu\|_\pi$ and $\|(I_n - V)u\|_\pi$ of either type given in Lemma 4.5 converge to 0 as $\lambda \to -\infty$ and $\lambda \to 0$, respectively. This fact agrees with the following observation: if V approximates e^{tK} well, then

$$\begin{aligned} \|Vu\|_\pi &\approx \left\|\mathrm{e}^{tK}u\right\|_\pi = \left\|\mathrm{e}^{t\lambda}u\right\|_\pi = \mathrm{e}^{t\lambda} \to 0 && (\lambda \to -\infty), \\ \|(I_n - V)u\|_\pi &\approx \left\|(I_n - \mathrm{e}^{tK})u\right\|_\pi = \left\|(1 - \mathrm{e}^{t\lambda})u\right\|_\pi = 1 - \mathrm{e}^{t\lambda} \to 0 && (\lambda \to 0). \end{aligned}$$

Combining Proposition 4.1 and Lemmas 4.2–4.5, we obtain the following.

THEOREM 4.6. *Let $K \in \mathbb{R}^{n\times n}$ be a rate constant matrix with stationary distribution $\pi \in \Delta_n^\circ$, $\{u_1, \dots, u_n\}$ an orthonormal eigenbasis of K with corresponding eigenvalues $\{\lambda_1, \dots, \lambda_n\}$, $\{S, T\}$ a bipartition of $[n]$, D the Schur complement of K_{SS} in K, and $V \in \mathbb{R}^{n\times n}$ the matrix (3.3) of Type A or B. For $t \in \mathbb{R}_{\ge 0}$, it holds that*

$$\mathrm{Err}_{S,t}(p) \le \min\left\{ \sum_{k=1}^n \frac{|\langle u_k, p\rangle_\pi|}{\|p\|_\pi} \min\{1, \alpha_k, \beta_k\}, 1 \right\} \quad (p \in \Delta_n), \tag{4.4}$$

$$\mathbb{E}_p[\mathrm{Err}_{S,t}(p)] \le \min\left\{ \sqrt{\frac{1}{n}\sum_{k=1}^n \min\{1, \alpha_k, \beta_k\}^2}, 1 \right\}$$

where

$$\alpha_k := \frac{\rho(D)}{|\lambda_k|} + \mathrm{e}^{t\lambda_k}, \quad \beta_k := \begin{cases} \frac{|\lambda_k| + \sqrt{2\left\|\Pi^{-1}K\right\|_{\infty,\mathrm{off}}|\lambda_k|}}{\sigma(K_{SS})} + 1 - \mathrm{e}^{t\lambda_k} & (\textit{Type A}), \\ \frac{|\lambda_k|}{\sigma(K_{SS})} + 1 - \mathrm{e}^{t\lambda_k} & (\textit{Type B}) \end{cases}$$

for $k \in [n]$.

Theorem 4.6 claims that the output of the RCMC method for an initial vector $p \in \Delta_n$ becomes a better approximation to the analytic solution $x(t) = \mathrm{e}^{tK}p$ if p comprises few eigencomponents of K with eigenvalues closer to $\sigma(K_{SS})$ or $\rho(D)$. Our numerical experiments in Section 6 show that (4.4) for Type B provides a sharp bound on the actual errors.

Theorem 4.6 also leads to a principled way of determining the appropriate time t^* corresponding to $\{S, T\}$ as follows. Since

$$\mathbb{E}_p[\mathrm{Err}_{S,t}(p)] \le \sqrt{\frac{1}{n}\sum_{k=1}^n \min\{1, \alpha_k, \beta_k\}^2} \le \sqrt{\frac{1}{n}\sum_{k=1}^n f(t)^2} = f(t)$$

with

$$f(t) \coloneqq \max_{\lambda<0} \min\left\{ \frac{\rho(D)}{|\lambda|} + \mathrm{e}^{t\lambda}, \frac{|\lambda|}{\sigma(K_{SS})} + 1 - \mathrm{e}^{t\lambda} \right\} \tag{4.5}$$

holds on Type B, it is a natural idea to set t as the minimizer t^* of (4.5). By calculation (see Section SM4 in supplemental materials), we have

$$t^* = \frac{\ln 2}{\sqrt{\sigma(K_{SS})\rho(D)}}.$$

The formula of $t^{(k)}_{\mathrm{eigen}}$ in (3.2) is derived in this way. While it is possible to consider the optimal time based on the upper bound for Type A, as we will see in Section 6, $\mathrm{Err}_{S,t}(p)$ on Type A also falls below the upper bound for Type B in practice. Therefore, using $t^{(k)}_{\mathrm{eigen}}$ is practical even for Type A.

4.2. Proofs of Lemmas 4.2–4.4. This section gives proofs of Lemmas 4.2–4.4. We denote $V - \mathrm{e}^{tK}$ by E for notational simplicity.

Here, we introduce the notion of the *matrix norm*. For $I, J \subseteq [n]$, consider a matrix $A \in \mathbb{R}^{I\times J}$. The *matrix norm* of A is defined by

$$\|A\|_\pi \coloneqq \max_{v\in\mathbb{R}^J\setminus\{\mathbf{0}\}} \frac{\|Av\|_\pi}{\|v\|_\pi} = \sqrt{\rho(AA^*)} = \sqrt{\rho(A^*A)}. \tag{4.6}$$

Note that different norms are used in the numerator and the denominator in (4.6) unless $I = J$. If $I = J$ and A is self-adjoint, $\|A\|_\pi$ is equal to $\rho(A)$. See Section SM1.4 in supplemental materials for details.

Proof of Lemma 4.2. By Proposition 4.1, it suffices to show $\|E\|_\pi \le 1$. Since E is self-adjoint, we have $\|E\|_\pi = \rho(E)$. Recall from Section 2.2 and Proposition 3.6 that $O \preceq \mathrm{e}^{tK} \preceq I_n$ and $O \preceq V \preceq I_n$ hold. Hence, we have $E = V - \mathrm{e}^{tK} \preceq I_n - O = I_n$ and $E = V - \mathrm{e}^{tK} \succeq O - I_n = -I_n$, meaning that $\rho(E) \le 1$ as required. ∎

Proof of Lemma 4.3. Since $\{u_1, \dots, u_n\}$ is an orthonormal basis of $\mathbb{R}^n$, p can be expanded as

$$p = \sum_{k=1}^n \langle u_k, p\rangle_\pi u_k \tag{4.7}$$

(see Section SM1.1 in supplemental materials). Substituting (4.7) into $\|Ep\|_\pi$, we get

$$\frac{\|Ep\|_\pi}{\|p\|_\pi} = \frac{1}{\|p\|_\pi}\left\| \sum_{k=1}^n \langle u_k, p\rangle_\pi Eu_k \right\|_\pi \le \sum_{k=1}^n \frac{|\langle u_k, p\rangle_\pi|}{\|p\|_\pi}\|Eu_k\|_\pi$$

by the triangle inequality. Thus, (4.2) is proved.

We next consider the expectation error. It follows from Jensen's inequality that

$$\mathbb{E}_p\left[\frac{\|Ep\|_\pi}{\|p\|_\pi}\right]^2 \le \mathbb{E}_p\left[\frac{\|Ep\|_\pi^2}{\|p\|_\pi^2}\right].$$

Since $\|Ep\|_\pi^2$ for $p \in \Delta_n$ can be written as

$$\|Ep\|_\pi^2 = \sum_{k=1}^n \sum_{\ell=1}^n \langle u_k, p\rangle_\pi \langle u_\ell, p\rangle_\pi \langle Eu_k, Eu_\ell\rangle_\pi$$

by (4.7), we have

$$\mathbb{E}_p\left[\frac{\|Ep\|_\pi^2}{\|p\|_\pi^2}\right] = \sum_{k=1}^n\sum_{\ell=1}^n \mathbb{E}_p\left[\frac{\langle u_k,p\rangle_\pi\langle u_\ell,p\rangle_\pi}{\|p\|_\pi^2}\right]\langle Eu_k,Eu_\ell\rangle_\pi. \tag{4.8}$$

Recall that the expectation is taken over the uniform distribution on the vertices $\{e_1,\dots,e_n\}$ of Δ_n, where $e_i\in\mathbb{R}^n$ is the vector whose ith component is 1 and others are 0 for $i\in[n]$. The expectation in the right-hand side of (4.8) is calculated as

$$\mathbb{E}_p\left[\frac{\langle u_k,p\rangle_\pi\langle u_\ell,p\rangle_\pi}{\|p\|_\pi^2}\right] = \frac{1}{n}\sum_{i=1}^n\frac{\langle u_k,e_i\rangle_\pi\langle u_\ell,e_i\rangle_\pi}{\|e_i\|_\pi^2} = \frac{1}{n}\sum_{i=1}^n\frac{(u_k)_i(u_\ell)_i}{\pi_i} = \frac{\langle u_k,u_\ell\rangle_\pi}{n},$$

which is $\frac{1}{n}$ if $k=\ell$ and 0 otherwise. Hence, we have

$$\mathbb{E}_p\left[\frac{\|Ep\|_\pi^2}{\|p\|_\pi^2}\right] = \frac{1}{n}\sum_{k=1}^n\langle Eu_k,Eu_k\rangle_\pi = \frac{1}{n}\sum_{k=1}^n\|Eu_k\|^2$$

and the desired bound (4.3) can be obtained by taking the square root. ∎

Proof of Lemma 4.4. We give three upper bounds on $\|Eu\|_\pi$ which proves (4.2). First, from the self-adjointness of $E=V-\mathrm{e}^{tK}$, we have $\|Eu\|_\pi\le\|E\|_\pi\|u\|_\pi=\rho(E)$. By $O\preceq V,\mathrm{e}^{tK}\preceq I_n$, we get $-I_n\preceq E\preceq I_n$ and hence $\rho(E)\le 1$ holds. Next, we have $\|Eu\|_\pi\le\|Vu\|_\pi+\|\mathrm{e}^{tK}u\|_\pi=\|Vu\|_\pi+\mathrm{e}^{t\lambda}$. Finally, we obtain $\|Eu\|_\pi\le\|(I_n-V)u\|_\pi+\|(I_n-\mathrm{e}^{tK})u\|_\pi=\|(I_n-V)u\|_\pi+1-\mathrm{e}^{t\lambda}$. ∎

4.3. Bounding $\|Vu\|_\pi$. We give a bound on $\|Vu\|_\pi$ for an eigenvector u of K.

LEMMA 4.7. *For an eigenpair (λ,u) of K with $\|u\|_\pi=1$, on both Types A and B, it holds that $\|Vu\|_\pi\le\dfrac{\rho(D)}{|\lambda|}$, where D is the Schur complement of K_{SS} in K.*

Proof. Recall that the matrix V can be expressed as $V=\Omega^*V_{TT}\Omega$, where $\Omega\in\mathbb{R}^{T\times n}$ is defined by (3.11). Then, it holds that

$$\|Vu\|_\pi^2 = \langle u,V^2u\rangle_\pi = \langle u,\Omega^*V_{TT}\Omega\Omega^*V_{TT}\Omega u\rangle_\pi = \langle\Omega u,V_{TT}MV_{TT}\Omega u\rangle_\pi. \tag{4.9}$$

By $K_{SS}u_S+K_{ST}u_T=\lambda u_S$ and $K_{TS}u_S+K_{TT}u_T=\lambda u_T$, we have $Du_T=\lambda\big(u_T-K_{TS}K_{SS}^{-1}u_S\big)=\lambda\Omega u$. Substituting $\Omega u=\frac{1}{\lambda}Du_T$ into (4.9), we get

$$\|Vu\|_\pi^2 = \frac{1}{\lambda^2}\langle Du_T,V_{TT}MV_{TT}Du_T\rangle_\pi = \frac{1}{\lambda^2}\big\|M^{\frac12}V_{TT}Du_T\big\|_\pi$$

and hence $\|Vu\|_\pi\le\frac{1}{|\lambda|}\big\|M^{\frac12}V_{TT}^{\frac12}\big\|_\pi\big\|V_{TT}^{\frac12}\big\|_\pi\|D\|_\pi\|u_T\|_\pi$ holds. By Lemma 3.5, we have $\big\|M^{\frac12}V_{TT}^{\frac12}\big\|_\pi^2=\rho\big(V_{TT}^{\frac12}MV_{TT}^{\frac12}\big)\le 1$ and $\big\|V_{TT}^{\frac12}\big\|_\pi^2=\rho(V_{TT})\le 1$. Since $\|D\|_\pi=\rho(D)$ and $\|u_T\|_\pi\le\|u\|_\pi=1$, we obtain the desired inequality. ∎

4.4. Bounding $\|(I_n-V)u\|_\pi$. We first consider Type B. As described in Section 3.3, the matrix V of Type B is the orthogonal projection onto the QSSA subspace $\operatorname{Im}V$. On the other hand, I_n-V is the orthogonal projection onto $\ker V$. Let

$N \coloneqq I_S + K_{SS}^{-1}K_{ST}K_{TS}K_{SS}^{-1}$. By the Woodbury matrix identity, it holds that

$$\begin{aligned} M^{-1} &= I_T - K_{TS}K_{SS}^{-1}N^{-1}K_{SS}^{-1}K_{ST}, \\ N^{-1} &= I_S - K_{SS}^{-1}K_{ST}M^{-1}K_{TS}K_{SS}^{-1}, \\ M^{-1}K_{TS}K_{SS}^{-1} &= K_{TS}K_{SS}^{-1}N^{-1}, \\ K_{SS}^{-1}K_{ST}M^{-1} &= N^{-1}K_{SS}^{-1}K_{ST}. \end{aligned}$$

Therefore, we have

$$I_n - V = \Psi^* N^{-1}\Psi, \tag{4.10}$$

where $\Psi \in \mathbb{R}^{S\times n}$ is defined as $\Psi_{SS} = I_S$ and $\Psi_{ST} = K_{SS}^{-1}K_{ST}$.

Lemma 4.8. *For an eigenpair (λ, u) of K with $\|u\|_\pi = 1$, the matrix V of Type B satisfies $\|(I_n - V)u\|_\pi \le \dfrac{|\lambda|}{\sigma(K_{SS})}$.*

Proof. By (4.10), we have

$$\|(I_n - V)u\|_\pi^2 = \left\|\Psi^* N^{-1}\Psi u\right\|_\pi^2 = \left\langle \Psi u, N^{-1}\Psi\Psi^* N^{-1}\Psi u\right\rangle_\pi = \left\langle \Psi u, N^{-1}\Psi u\right\rangle_\pi,$$

where we used $\Psi\Psi^* = N$ in the last equality. The vector Ψu can be written as

$$\Psi u = u_S + K_{SS}^{-1}K_{ST}u_T = K_{SS}^{-1}\big(K_{SS}u_S + K_{ST}u_T\big) = \lambda K_{SS}^{-1}u_S.$$

Thus, it holds $\|(I_n - V)u\|_\pi^2 = \lambda^2\left\|N^{\frac{1}{2}}K_{SS}^{-1}u_S\right\|_\pi^2$ and

$$\|(I_n - V)u\|_\pi \le |\lambda|\left\|N^{\frac{1}{2}}\right\|_\pi \left\|K_{SS}^{-1}\right\|_\pi \|u_S\|_\pi \le \frac{|\lambda|}{\sigma(K_{SS})}\left\|N^{\frac{1}{2}}\right\|_\pi.$$

Therefore, it suffices to show $\left\|N^{-\frac{1}{2}}\right\|_\pi \le 1$. Since N is self-adjoint, we have $\left\|N^{-\frac{1}{2}}\right\|_\pi^2 = \rho(N^{-1})$. In addition, $N \succeq I_S$ holds by definition, and thus $N^{-1} \preceq I_S$. Hence, we have $\rho(N^{-1}) \le 1$. ∎

We next consider Type A. For a vector $a \in \mathbb{R}^n$, let $a^* \coloneqq a^\top \Pi^{-1}$.

Lemma 4.9. *Let (λ, u) be an eigenpair of K and*

$$\Delta \coloneqq K\,\mathrm{diag}(u^*) - \mathrm{diag}(u^*)K \in \mathbb{R}^{n\times n}. \tag{4.11}$$

Then, for the matrix V of Type A, we have $\|(I_n - V)u\|_\pi \le \dfrac{|\lambda| + \|\Delta\|_\pi}{\sigma(K_{SS})}$.

Proof. Let $v \coloneqq Vu$. First, we have

$$\begin{aligned} (u - v)_S &= u_S + K_{SS}^{-1}K_{ST}v_T \\ &= -K_{SS}^{-1}K_{ST}(u_T - v_T) + u_S + K_{SS}^{-1}K_{ST}u_T \\ &= -K_{SS}^{-1}K_{ST}(u_T - v_T) + \lambda K_{SS}^{-1}u_S \\ &= \lambda K_{SS}^{-1}u_S - K_{SS}^{-1}K_{ST}(u_T - v_T) \\ &= \Omega_{ST}^*(u_T - v_T) + \lambda K_{SS}^{-1}u_S \end{aligned}$$

and $(u-v)_T = u_T - v_T = \Omega^*_{TT}(u_T - v_T)$. Hence, we have

$$\|u-v\|_\pi \le \big\|\Omega^*(u_T-v_T)\big\|_\pi + \big\|\lambda K_{SS}^{-1}u_S\big\|_\pi \le \big\|\Omega^*(u_T-v_T)\big\|_\pi + \frac{|\lambda|}{\sigma(K_{SS})}.$$

Since $\|\Omega^*(u_T-v_T)\|_\pi = \big\|M^{\frac12}(u_T-v_T)\big\|_\pi$ by $\Omega\Omega^* = M$ and the self-adjointness of $M^{\frac12}$, we obtain

$$\|u-v\|_\pi \le \big\|M^{\frac12}(u_T-v_T)\big\|_\pi + \frac{|\lambda|}{\sigma(K_{SS})}. \tag{4.12}$$

We next bound $\big\|M^{\frac12}(u_T-v_T)\big\|_\pi$. By $O \preceq M^{\frac12}V_{TT}M^{\frac12} \preceq I_T$, we have

$$\big\|M^{\frac12}(u_T-v_T)\big\|_\pi \le \big\|M^{\frac12}V_{TT}^{\frac12}\big\|_\pi \big\|V_{TT}^{-\frac12}(u_T-v_T)\big\|_\pi \le \big\|V_{TT}^{-\frac12}(u_T-v_T)\big\|_\pi. \tag{4.13}$$

By $\mathbf{1}_S^\top K_{SS} + \mathbf{1}_T^\top K_{TS} = \mathbf{0}$, we have

$$V_{TT}^{-1} = \operatorname{diag}\big(\mathbf{1}_T^\top M\big) = I_T + \operatorname{diag}\big(\mathbf{1}_T^\top K_{TS}K_{SS}^{-2}K_{ST}\big) = I_T - \operatorname{diag}\big(\mathbf{1}_S^\top K_{SS}^{-1}K_{ST}\big)$$

and

$$\begin{aligned} V_{TT}^{-\frac12}(u_T-v_T) &= V_{TT}^{-\frac12}\big(u_T - V_{TT}\big(u_T - K_{TS}K_{SS}^{-1}u_S\big)\big) \\ &= V_{TT}^{\frac12}\big(V_{TT}^{-1}u_T - u_T + K_{TS}K_{SS}^{-1}u_S\big) \\ &= V_{TT}^{\frac12}\big(K_{TS}K_{SS}^{-1}u_S - \operatorname{diag}\big(\mathbf{1}_S^\top K_{SS}^{-1}K_{ST}\big)u_T\big). \end{aligned} \tag{4.14}$$

By $(K_{ST})^\top = \Pi_T^{-1}K_{TS}\Pi_S$ and $\big(K_{SS}^{-1}\big)^\top = \Pi_S^{-1}K_{SS}^{-1}\Pi_S$, we get

$$\begin{aligned} \operatorname{diag}\big(\mathbf{1}_S^\top K_{SS}^{-1}K_{ST}\big) &= \operatorname{diag}\big(\big(\mathbf{1}_S^\top K_{SS}^{-1}K_{ST}\big)^\top\big) \\ &= \operatorname{diag}\big(\Pi_T^{-1}K_{TS}K_{SS}^{-1}\Pi_S\mathbf{1}_S\big) = \operatorname{diag}\big(K_{TS}K_{SS}^{-1}\pi_S\big)\Pi_T^{-1}. \end{aligned} \tag{4.15}$$

Recall that the matrix Δ is defined by (4.11). Combining (4.14) and (4.15), we have

$$\begin{aligned} V_{TT}^{-\frac12}(u_T-v_T) &= V_{TT}^{\frac12}\big(K_{TS}K_{SS}^{-1}u_S - \operatorname{diag}\big(K_{TS}K_{SS}^{-1}\pi_S\big)\Pi_T^{-1}u_T\big) \\ &= V_{TT}^{\frac12}\big(K_{TS}K_{SS}^{-1}u_S - \operatorname{diag}(u_T^*)K_{TS}K_{SS}^{-1}\pi_S\big) \\ &= V_{TT}^{\frac12}\big(K_{TS}K_{SS}^{-1}\operatorname{diag}(u_S^*) - \operatorname{diag}(u_T^*)K_{TS}K_{SS}^{-1}\big)\pi_S \\ &= V_{TT}^{\frac12}\big(K_{TS}K_{SS}^{-1}\operatorname{diag}(u_S^*) - K_{TS}\operatorname{diag}(u_S^*)K_{SS}^{-1} \\ &\qquad + K_{TS}\operatorname{diag}(u_S^*)K_{SS}^{-1} - \operatorname{diag}(u_T^*)K_{TS}K_{SS}^{-1}\big)\pi_S \\ &= V_{TT}^{\frac12}\big(K_{TS}K_{SS}^{-1}\big(\operatorname{diag}(u_S^*)K_{SS} - K_{SS}\operatorname{diag}(u_S^*)\big) \\ &\qquad + K_{TS}\operatorname{diag}(u_S^*) - \operatorname{diag}(u_T^*)K_{TS}\big)K_{SS}^{-1}\pi_S \\ &= V_{TT}^{\frac12}\big(-K_{TS}K_{SS}^{-1}\Delta_{SS} + \Delta_{TS}\big)K_{SS}^{-1}\pi_S \\ &= V_{TT}^{\frac12}\Omega\Delta_{[n]S}K_{SS}^{-1}\pi_S, \end{aligned}$$

whose norm is bounded as

$$\big\|V_{TT}^{-\frac12}(u_T-v_T)\big\|_\pi \le \big\|V_{TT}^{\frac12}\Omega\big\|_\pi \|\Delta\|_\pi \big\|K_{SS}^{-1}\pi_S\big\|_\pi. \tag{4.16}$$

By $V_{TT} \preceq M^{-1}$, it holds that

$$\left\|V_{TT}^{\frac{1}{2}}\Omega\right\|_\pi^2 = \rho\big(V_{TT}^{\frac{1}{2}}\Omega\Omega^* V_{TT}^{\frac{1}{2}}\big) = \rho\big(V_{TT}^{\frac{1}{2}} M V_{TT}^{\frac{1}{2}}\big) \le 1 \tag{4.17}$$

and

$$\left\|K_{SS}^{-1}\pi_S\right\|_\pi \le \left\|K_{SS}^{-1}\right\|_\pi \|\pi_S\|_\pi \le \frac{1}{\sigma(K_{SS})}. \tag{4.18}$$

By (4.12), (4.13), and (4.16)–(4.18), we obtain the desired inequality. ∎

To bound $\|\Delta\|_\pi$, we give two lemmas. The first one is a generalization of the inequality that bounds the spectral norm of a matrix by the Frobenius norm.

LEMMA 4.10. *For $A \in \mathbb{R}^{n\times n}$, it holds that $\|A\|_\pi \le \sqrt{\operatorname{tr} A^* A}$.*

Proof. The claim follows from the fact that $\|A\|_\pi^2 = \rho(A^*A)$ is the maximum eigenvalue of A^*A, whereas $\operatorname{tr} A^*A$ is the sum of eigenvalues of A^*A. ∎

Recall that $L = -K\Pi$ is a graph Laplacian matrix. It is well-known that the quadratic form $a^\top La$ of a graph Laplacian matrix with $a \in \mathbb{R}^n$ satisfies $a^\top La = \frac{1}{2}\sum_{i\neq j} L_{ij}(a_i - a_j)^2$. The following lemma is a variant of this identity for K.

LEMMA 4.11. *For $a \in \mathbb{R}^n$, it holds that $\langle a, Ka\rangle_\pi = -\frac{1}{2}\displaystyle\sum_{i\neq j} L_{ij}\left(\frac{a_i}{\pi_i} - \frac{a_j}{\pi_j}\right)^2$.*

Proof. Using $L = -K\Pi$, we obtain

$$\langle a, Ka\rangle_\pi = -a^\top \Pi^{-1} L \Pi^{-1} a = -(\Pi^{-1}a)^\top L(\Pi^{-1}a) = -\frac{1}{2}\sum_{i\neq j} L_{ij}\left(\frac{a_i}{\pi_i} - \frac{a_j}{\pi_j}\right)^2. \qquad ∎$$

We now give an upper bound on $\|\Delta\|_\pi$, which, together with Lemma 4.8, leads to the upper bound on $\|(I_n - V)u\|_\pi$ for Type A in Lemma 4.5.

LEMMA 4.12. *Let (λ, u) be an eigenpair of K with $\|u\|_\pi = 1$ and Δ be the matrix defined by (4.11). Then, it holds that $\|\Delta\|_\pi \le \sqrt{2|\lambda|\left\|\Pi^{-1}K\right\|_{\infty,\mathrm{off}}}$.*

Proof. By Lemma 4.10 and $-\Delta^* = \Delta$, it holds $\|\Delta\|_\pi^2 \le \operatorname{tr}\Delta^*\Delta = -\operatorname{tr}\Delta^2$. The (i,j) entry in Δ is

$$\Delta_{ij} = K_{ij}\left(\frac{u_j}{\pi_j} - \frac{u_i}{\pi_i}\right) = \frac{L_{ij}}{\pi_j}\left(\frac{u_i}{\pi_i} - \frac{u_j}{\pi_j}\right).$$

We thus have

$$\begin{aligned}\|\Delta\|_\pi^2 \le -\operatorname{tr}\Delta^2 &= -\sum_{i,j}\Delta_{ij}\Delta_{ji} = \sum_{i\neq j}\frac{{L_{ij}}^2}{\pi_i\pi_j}\left(\frac{u_i}{\pi_i} - \frac{u_j}{\pi_j}\right)^2 \\ &\le \max_{i\neq j}\frac{|L_{ij}|}{\pi_i\pi_j}\cdot\sum_{i\neq j}|L_{ij}|\left(\frac{u_i}{\pi_i} - \frac{u_j}{\pi_j}\right)^2 = -2\left\|\Pi^{-1}K\right\|_{\infty,\mathrm{off}}\langle u, Ku\rangle_\pi,\end{aligned}$$

where Lemma 4.11 is used in the last equality. Since $\langle u, Ku\rangle_\pi = \lambda$, we obtain the desired inequality. ∎

5. Implementation Details of RCMC Method. In this section, we discuss the implementation details of the RCMC method for fast and stable computation. For more details on the computational complexity of the RCMC method, see [6].

5.1. Incremental Update of LU Decomposition. In the calculation of the matrix-vector multiplication Vp, one needs to solve linear systems with the coefficient matrix $K_{SS} = K_{S^{(k)}S^{(k)}}$, which costs $\mathrm{O}(k^3)$ operations if an LU decomposition of K_{SS} is computed from scratch. Fortunately, we can incrementally update an LU decomposition of K_{SS} for successive k in the following way.

We first discuss the LU decomposition of a self-adjoint positive semi-definite matrix $A \in \mathbb{R}^{n\times n}$ in terms of the π-norm metric. The matrix $B \coloneqq A\Pi^{-1}$ is a symmetric positive semi-definite matrix. Consider a Cholesky decomposition $PCC^\top P^\top$ of B, where $C \in \mathbb{R}^{n\times n}$ is a lower-triangular matrix and $P \in \mathbb{R}^{n\times n}$ is a permutation matrix representing the order of pivoting. Then, $C\big(C^\top\Pi^{-1}\big)$ is an LU decomposition of $P^\top AP$. We regard the row and column spaces of C as the linear spaces equipped with the metrics $P^\top\Pi^{-1}P = \mathrm{diag}(P\pi)^{-1}$ and I_n, respectively. Then, the adjoint matrix of C is given by $C^* = C^\top\Pi^{-1}$ and $CC^* = P^\top AP$ holds. We call this decomposition a *Cholesky decomposition* of A with respect to the π-norm. The matrix C is called a *Cholesky factor* of A.

Throughout the algorithm, we maintain an $n\times k$ matrix $C^{(k)} \in \mathbb{R}^{n\times S^{(k)}}$. Initially, $C^{(0)}$ is set as the $n\times 0$ matrix. At the kth iteration, we update $C^{(k-1)}$ to $C^{(k)}$ by appending a column vector as

$$C^{(k)} \coloneqq \begin{pmatrix} & & \overset{S^{(k-1)}}{} & \overset{j^{(k)}}{} \\ & C^{(k-1)} & \Big| & \mathbf{0} \\ & & & -\sqrt{\frac{\pi_{j^{(k)}}}{\left|D^{(k-1)}_{j^{(k)}j^{(k)}}\right|}}D^{(k-1)}_{T^{(k-1)}j^{(k)}} \end{pmatrix}\begin{matrix} \\ S^{(k-1)} \\ T^{(k-1)}. \end{matrix} \tag{5.1}$$

PROPOSITION 5.1. *The matrix $C^{(k)}$ given by (5.1) is equal to $C_{[n]S^{(k)}}$ for some Cholesky factor C of $-K$.*

Proof. Let $Q^{(k)} \coloneqq -D^{(k)}\Pi_{T^{(k)}} = L_{T^{(k)}T^{(k)}} - L_{T^{(k)}S^{(k)}}L^{-1}_{S^{(k)}S^{(k)}}L_{S^{(k)}T^{(k)}}$ be the Schur complement of $L_{S^{(k)}S^{(k)}}$ in $L = -K\Pi$. Then, we can write (5.1) as

$$C^{(k)} = \left(\begin{array}{c|c} C^{(k-1)} & \begin{matrix}\mathbf{0}\\ \frac{1}{\sqrt{Q^{(k-1)}_{j^{(k)}j^{(k)}}}}Q^{(k-1)}_{T^{(k-1)}j^{(k)}}\end{matrix}\end{array}\right). \tag{5.2}$$

The formula (5.2) is identical to the procedure of computing a Cholesky decomposition $CC^\top$ of L with pivoting order $j^{(1)}, j^{(2)}, \ldots$; see, e.g., [7, Chapter 10]. Thus, it holds that $C^{(k)} = C_{[n]S^{(k)}}$. ∎

By Proposition 5.1, $C^{(k)}_{S^{(k)}S^{(k)}}$ provides a Cholesky factor of $-K_{S^{(k)}S^{(k)}}$.

The RCMC method of Type B additionally involves solving a linear system with the coefficient matrix M defined by (3.4). In Section SM6 in supplemental materials, we describe an efficient algorithm for incrementally updating a Cholesky factor of M.

5.2. Numerical Stabilization. A rate constant matrix K arising from chemical kinetics contains a wide range of numbers in its entries, which may lead to catastrophic cancellation when two like-sign numbers are subtracted in finite precision arithmetic [7]. Therefore, designing an algorithm that avoids subtracting like-sign numbers is crucial to numerical stability.

Fortunately, almost all matrices appearing in Algorithm 1 of Type A have constant sign patterns. For instance, K_{ST}, K_{TS}, and $V_{TT} = \mathrm{diag}\big(\mathbf{1}_T^\top M\big)^{-1}$ are non-negative matrices for any $S = S^{(k)}$ and $T = T^{(k)}$. The Cholesky factor C of $-K$ treated in

Section 5.1 is a lower-triangular matrix whose diagonals and off-diagonals are non-negative and non-positive, respectively, by (5.1) and the fact that $D = D^{(k)}$ is a rate constant matrix. Thus, we have $C_{SS}^{-1} \geq O$ from the inversion formula of triangular matrices. As a result, for any $p \geq \mathbf{0}$, the matrix-vector multiplication Vp on Type A is free from subtractions of like-signed numbers.

Unfortunately, the same property does not apply to the matrix-vector multiplication in Type B due to the absence of the non-negativity of M^{-1}. Nevertheless, in our experiments detailed in Section 6, we have not encountered any numerical instability. Theoretically analyzing the numerical stability for the TypeB RCMC method is left as future work.

Apart from the matrix-vector multiplication, Algorithm 1 performs arithmetic operations during the update of D at Line 5. Unlike the off-diagonal updates, the following diagonal update in Line 5 is the subtraction of two non-positive numbers:

$$D_{ii} \coloneqq D_{ii}^{(k-1)} - \frac{D_{ij}^{(k-1)} D_{ji}^{(k-1)}}{D_{jj}^{(k-1)}} \qquad (i \in T), \tag{5.3}$$

where $j = j^{(k)}$. Indeed, due to $\mathbf{1}_T^\top D = \mathbf{0}^\top$, we can compute D_{ii} from off-diagonals of D as

$$D_{ii} = - \sum_{k \in T \setminus \{i\}} D_{ki} \qquad (i \in T). \tag{5.4}$$

Diagonal entries in D should be updated via (5.4) in place of (5.3) from the viewpoint of numerical stability.

5.3. Faster Reference Time Computation. Computing $t_{\mathrm{eigen}}^{(k)}$ given by (3.2) costs longer time than other operations in Algorithm 1, since it involves the computation of the spectral radii of D and K_{SS}^{-1}. When the running time is a significant concern, for a square matrix A, one can replace $\rho(A)$ and $\sigma(A)$ with

$$\hat{\rho}(A) \coloneqq \min\left\{ \max_i \left\{ \sum_j |A_{ij}| \right\}, \max_j \left\{ \sum_i |A_{ij}| \right\} \right\} \quad \text{and} \quad \hat{\sigma}(A) \coloneqq \frac{1}{\hat{\rho}(A^{-1})},$$

respectively. The value $\hat{\rho}(A)$ is an upper bound on $\rho(A)$ derived by the Gershgorin theorem. Since K_{SS} (with $S = S^{(k)}$) is an M-matrix, we further have

$$\hat{\rho}\big(K_{SS}^{-1}\big) = \min\left\{ \max_{i \in S} \big(K_{SS}^{-1} \mathbf{1}_S\big)_i, \max_{j \in S} \big(\mathbf{1}_S^\top K_{SS}^{-1}\big)_j \right\},$$

which can be computed by solving two linear systems with the coefficient matrix K_{SS}.

6. Experiments. In this section, we report numerical experiments of the RCMC method using synthetic and real-world datasets. All experiments were conducted using C++ and `Eigen`[1] 3.4.0 compiled by `Apple Clang` 14.0.3 on a laptop powered by an Apple M2 CPU (8 Cores) with 24 GB of RAM.

[1] https://eigen.tuxfamily.org/

TABLE 1
Properties of rate constant matrices $K \in \mathbb{R}^{n\times n}$, where $\mathrm{nnz}(K)$ denotes the number of non-zero entries in K.

Data	n	$\mathrm{nnz}(K)$	$\min_j \lvert K_{jj}\rvert$	$\max_j \lvert K_{jj}\rvert$
synth.	6	16	8.1×10^{-21}	3.9×10^{11}
DFG	1,745	9,677	4.3×10^{-159}	3.8×10^{14}
WL	1,759	10,485	1.2×10^{-144}	1.6×10^{16}

6.1. Datasets and Settings. We employ three datasets: one small-scale synthetic and two large-scale real-world data. The synthetic data consists of the following rate constant matrix $K \in \mathbb{R}^{6\times 6}$ (expressed with two significant digits):

$$\begin{pmatrix} -2.0\times10^{9} & 6.9\times10^{9} & 0 & 1.7\times10^{-18} & 0 & 0 \\ 2.0\times10^{9} & -3.9\times10^{11} & 3.3\times10^{11} & 0 & 1.9\times10^{7} & 0 \\ 0 & 3.9\times10^{11} & -3.3\times10^{11} & 0 & 0 & 0 \\ 9.6\times10^{-7} & 0 & 0 & -1.7\times10^{-18} & 0 & 0 \\ 0 & 3.4\times10^{7} & 0 & 0 & -1.9\times10^{7} & 8.1\times10^{-21} \\ 0 & 0 & 0 & 0 & 8.3\times10^{-9} & -8.1\times10^{-21} \end{pmatrix}$$

with $\pi = (9.1\times10^{-13},\ 2.6\times10^{-13},\ 3.1\times10^{-13},\ 0.51,\ 4.8\times10^{-13},\ 0.49)^\top \in \Delta_6^\circ$.

For real-world datasets, we use two chemical reaction networks fetched from the Searching Chemical Action and Network (SCAN) platform [9]. The two datasets, denoted by DFG and WL, represent the synthesis of fluoroglycine and Wöhler's urea synthesis, respectively. Rate constant matrices K are calculated from the free energies of equilibrium and transient states, assuming the canonical ensemble (1.3) with parameters $T = 300\,\mathrm{K}$ and $\Gamma = 1$. We truncate off-diagonal entries K_{ij} and K_{ji} to zero if $\frac{K_{ij}}{\pi_j} = \frac{K_{ji}}{\pi_i} < 10^{-200}$ and eliminate all-zero rows and columns from K. Properties of the resulting rate constant matrices are summarized in Table 1.

We execute the RCMC method (Algorithm 1) of Types A and B to compute approximate solutions $q^{(k)} \approx x\big(t^{(k)}\big)$ for $k = 0, \dots, n$. The initial vector p is fixed to $e_1 = (1, 0, \dots, 0)^\top$. Based on the discussion in Section 5.3, we determine the reference time $t^{(k)}$ in the following three ways:

- "diag": $t^{(k)} = t^{(k)}_{\mathrm{diag}}$ defined by (3.1),
- "eigen": $t^{(k)} = t^{(k)}_{\mathrm{eigen}}$ defined by (3.2),
- "gershgorin": $t^{(k)} = t^{(k)}_{\mathrm{gershgorin}} := \ln 2 / \sqrt{\hat\sigma\big(K_{S^{(k)}S^{(k)}}\big)\hat\rho\big(D^{(k)}\big)}$.

In "eigen," the spectral radii are computed by the Lanczos method[2] using `Spectra`[3] version 1.0.1. Incremental updates for LU decompositions of K_{SS} and M (only for Type B) given in Sections 5.1 and SM6, respectively, are employed. Matrix operations are performed in sparse representation.

Finally, we evaluate the error (4.1), which we call the *π-norm error*, for every k and compare them with the upper bounds given in Theorem 4.6. For calculating error bounds and the exact solution $x\big(t^{(k)}\big)$, we also compute the eigendecomposition

[2] More precisely, for a self-adjoint matrix A, we apply the Lanczos method to the "symmetrized" matrix $B := \Pi^{-\frac{1}{2}} A \Pi^{\frac{1}{2}}$ and retrive the eigendecomposition of A from that of B; see Section SM1.1 for the symmetrization.

[3] https://spectralib.org/

TABLE 2
Running time (in seconds) of eigendecomposition ("EVD") and the RCMC method.

Data	EVD	diag		eigen		gershgorin	
		Type A	Type B	Type A	Type B	Type A	Type B
synth.	0.01	< 1 ms	< 1 ms	< 1 ms	< 1 ms	< 1 ms	< 1 ms
DFG	4092.42	1.28	2.59	72.24	71.48	1.85	3.08
WL	4020.03	1.58	2.86	85.94	86.25	2.21	3.39

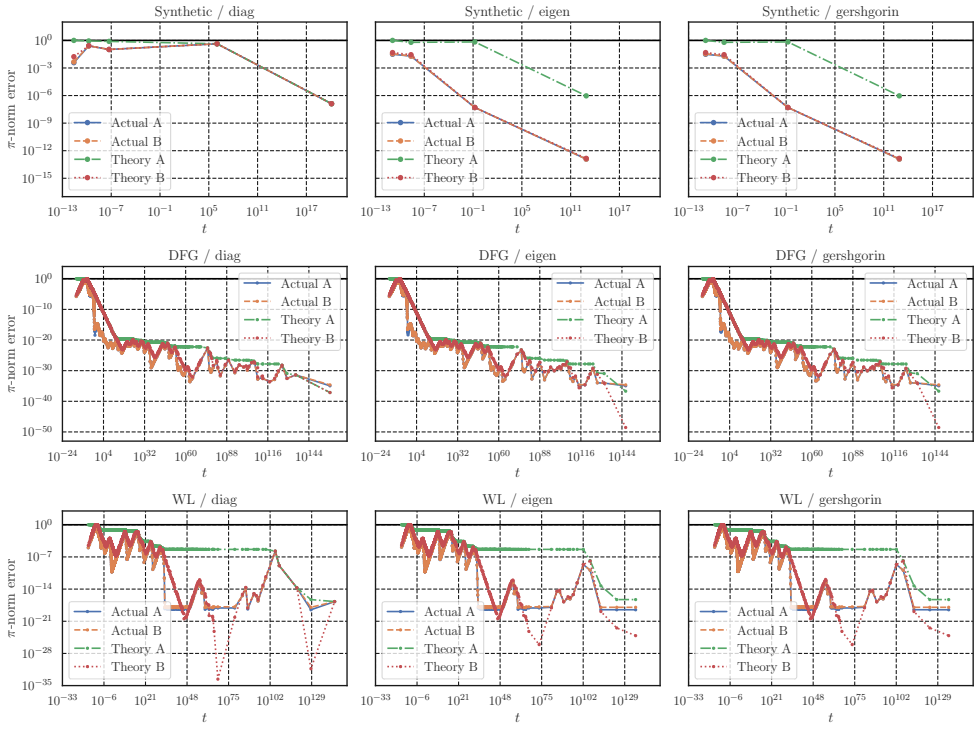

FIG. 1. *Log-log plots of the actual π-norm errors in approximate solutions ("Actual A/B") and their theoretical upper bounds ("Theory A/B"), where A and B indicate the types of the RCMC method. The errors are plotted as points $\left(t^{(k)}, \mathrm{Err}_{S^{(k)},t^{(k)}}(p)\right)$ for each k. Points with $t^{(k)} = 0$ or $+\infty$ are omitted. Actual A and B overlap closely in all plots. In the plots for synthetic data, Theory B also aligns closely with Actual A and B.*

of rate constant matrices by the Lanczos method with 200-digit precision floating-point numbers via the `GNU MPFR` library[4] of version 4.2.0.

6.2. Results. Table 2 shows the running time of the eigendecomposition and the RCMC method. Across all combinations of data, type, and the reference time computation method, the RCMC method ran within 100 seconds. Notably, "diag" and "gershgorin" required at most only 4 seconds. This result is noteworthy, especially when compared to the eigendecomposition, which took over 1 hour for DFG and WL.

[4] https://www.mpfr.org/

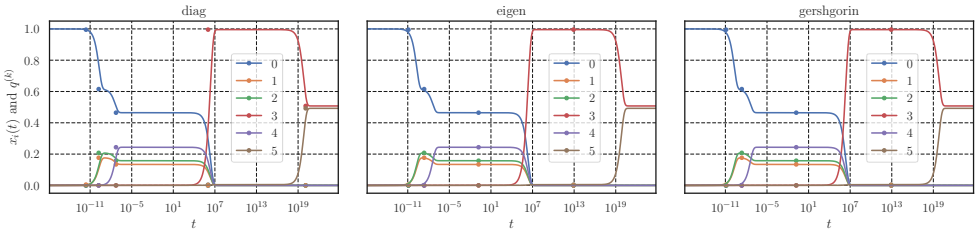

FIG. 2. *Exact solutions (solid curves) and approximate solutions (points) for the synthetic data. Approximate solutions of both Types A and B give visually indistinguishable plots.*

Figure 1 illustrates the comparison of the actual π-norm errors and their theoretical upper bounds. For the synthetic data, the theoretical bounds correctly served as upper bounds on actual errors for all k. For DFG and WL, this was true only for k such that the theoretical bounds for Type B are larger than 10^{-35} and 10^{-18}, respectively. This is because the RCMC method is performed with double-precision floating-point numbers (about 15-digit precision); we have confirmed that the errors fell below the theoretical bounds when recalculated with 50-digit precision. Nevertheless, the approximate solutions for such k computed with double precision align with the exact solutions for the first 12 to 14 digits. Thus, double precision is sufficient for most applications. It is worth mentioning that the double precision was insufficient for the eigendecomposition of K, even when applied to the synthetic data, as the matrix is extremely ill-conditioned. This fact highlights an advantage of the RCMC method in terms of numerical stability.

Type A versus Type B. In all datasets, the output of the RCMC method of Types A and B were almost identical. Notably, the errors of approximate solutions of Type A were consistently lower than the theoretical bounds not only of Type A but also of Type B. Regarding the computational time, Type B applied to DFG and WL additionally took about 1 second to perform the incremental update of an LU decomposition of M. Thus, in practice, employing Type A is sufficient from the viewpoint of running time.

Comparison of reference time computation methods. The output of "eigen" and "gershgorin" are nearly identical in all datasets. Thus, given the computational time shown in Table 2, we recommend the use of "gershgorin". As for "diag," it maintains a relatively small error for DFG and WL but exhibits larger errors to synthetic data. For a detailed analysis, we show plots of the exact and approximate solutions for the synthetic data in Figure 2. In "diag," $t^{(k)}$ are picked right after significant changes in the exact solution. In contrast, both "eigen" and "gershgorin" choose the midpoint (in terms of the logarithmic scale) of periods where the exact solution is nearly constant. Consequently, these methods yield smaller errors between the approximate and exact solutions.

L_∞-norm error. While the π-norm error (4.1) is tractable from the viewpoint of error analysis, it is not commonly used in general contexts. We illustrate in the top and middle rows of Figure 3 the log-linear plots of the π-norm errors and the *L_∞-norm errors*, defined to be $\left\|q^{(k)} - x(t^{(k)})\right\|_\infty$, respectively, focusing on shorter time intervals (with $t^{(k)}$ computed by the "eigen"). Although the π- and L_∞-norm errors typically follow analogous trajectories in the plots, there are moments when the π-norm error is small but the L_∞-norm error is large. To capture the time when

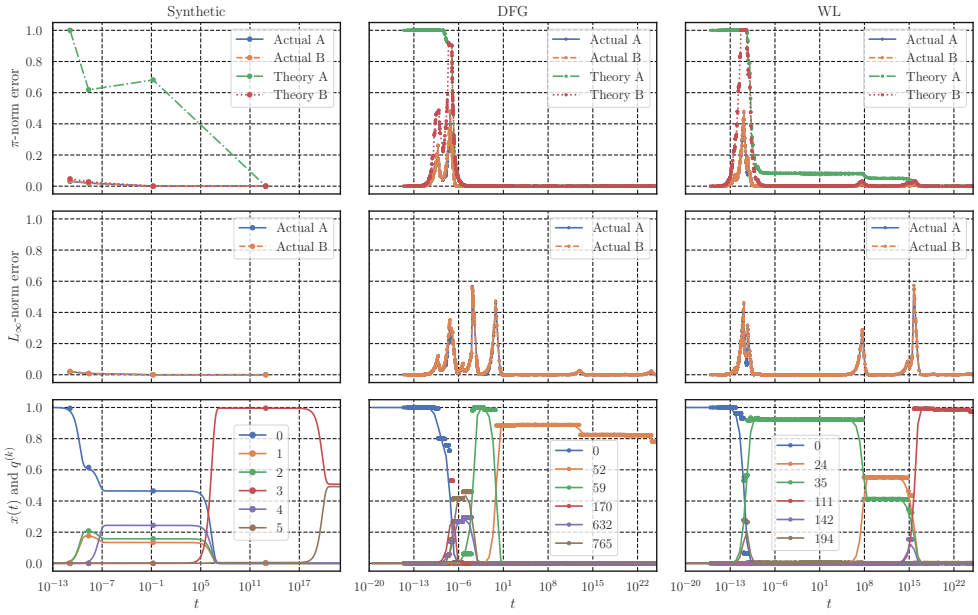

FIG. 3. *The top and middle rows: log-linear plots of the π- and L_∞-norm errors, respectively. "Actual A/B" means the actual errors in approximate solutions, whereas "Theory A/B" in the top row indicates the theoretical upper bounds. "A" and "B" refer to the types of the RCMC method. Actual A and B overlap closely in all error plots. The bottom row: comparison of exact solutions (solid curves) and the approximate solutions (points) computed by the RCMC method of Type B. Only the top six curves are plotted for clarity. In all plots, the times $t^{(k)}$ are computed by "eigen".*

a large L_∞-norm error occurs, we plot the approximate and exact solutions in the bottom row of Figure 3. For DFG and WL, wherein the RCMC method yields dense outputs of approximate solutions, the dynamics of approximate solutions resemble step functions, while the exact solutions form smooth curves. This discrepancy leads to a significant error amplification when the exact solution undergoes substantial fluctuations. Since the RCMC method is based on QSSA, approximating solutions accurately at such transient timings is difficult. Nonetheless, owing to its capability to perform calculations with notable speed and numerical stability, it is deemed suitable for applications that require fast computations of rough approximations of the solutions to master equations, like the kinetics-based navigation in automatic chemical reaction pathway search techniques [14, 15].

7. Conclusion. In this paper, we have explored mathematical aspects of the RCMC method. We have reformulated the RCMC method in terms of matrix computation and conducted theoretical error analysis. Building on this foundation, we have also discussed its efficient and numerically stable implementations. Numerical experiments with both synthetic and real-world datasets have demonstrated computational efficiency and numerical stability, and validated the derived error bounds. Potential future work includes extending the RCMC method to nonlinear master equations, such as those involving bimolecular reactions.

Acknowledgments. The authors thank Satoshi Maeda, Yu Harabuchi, and Wataru Matsuoka of Hokkaido University for fruitful discussions.

REFERENCES

[1] W. J. Anderson, *Continuous-Time Markov Chains*, Springer Series in Statistics, Springer, New York, 1991, https://doi.org/10.1007/978-1-4612-3038-0.

[2] A. Berma and R. J. Plemmons, *Nonnegative Matrices in the Mathematical Sciences*, Classics in Applied Mathematics, SIAM, Philadelphia, 1994.

[3] M. Bodenstein, *Eine Theorie der photochemischen Reaktionsgeschwindigkeiten*, Zeitschrift für Physikalische Chemie, 85 (1913), pp. 329–397, https://doi.org/10.1515/zpch-1913-8512.

[4] L. Farina and S. Rinaldi, *Positive Linear Systems: Theory and Applications*, Pure and Applied Mathematics: A Wiley Series of Texts, Monographs and Tracts, A Wiley-Interscience Publication, New York, 2000.

[5] Y. Harabuchi, T. Yokoyama, W. Matsuoka, T. Oki, S. Iwata, and S. Maeda, *Differentiating the yield of chemical reactions using parameters in first-order kinetic equations to identify elementary steps that control the reactivity from complicated reaction path networks*, The Journal of Physical Chemistry A, 128 (2024), pp. 2883–2890, https://doi.org/10.1021/acs.jpca.4c00204.

[6] S. Hemmi, S. Iwata, and T. Oki, *Fast and numerically stable implementation of rate constant matrix contraction method*, in Recent Developments in Industrial and Applied Mathematics, S. Oishi, H. Okamoto, and K. Hayami, eds., vol. 1 of ICIAM2023 Springer Series, Singapore, 2026, Springer. To appear.

[7] N. J. Higham, *Accuracy and Stability of Numerical Algorithms*, SIAM, Philadelphia, second ed., 2002, https://doi.org/10.1137/1.9780898718027.

[8] J. C. W. Kuo and J. Wei, *Lumping analysis in monomolecular reaction systems. Analysis of approximately lumpable system*, Industrial and Engineering Chemistry Fundamentals, 8 (1969), pp. 124–133, https://doi.org/10.1021/i160029a020.

[9] M. Kuwahara, Y. Harabuchi, S. Maeda, J. Fujima, and K. Takahashi, *Searching chemical action and network (SCAN): an interactive chemical reaction path network platform*, Digital Discovery, (2023), https://doi.org/10.1039/D3DD00026E.

[10] S. H. Lam and D. A. Goussis, *Understanding complex chemical kinetics with computational singular perturbation*, in Proceedings of the 22nd Symposium (International) on Combustion, vol. 22, The Combustion Institute, 1989, pp. 931–941, https://doi.org/10.1016/S0082-0784(89)80102-X.

[11] U. Maas and S. B. Pope, *Implementation of simplified chemical kinetics based on intrinsic low-dimensional manifolds*, in Proceedings of the 24th Symposium (International) on Combustion, vol. 24, The Combustion Institute, 1992, pp. 103–112, https://doi.org/10.1016/S0082-0784(06)80017-2.

[12] U. Maas and S. B. Pope, *Simplifying chemical kinetics: intrinsic low-dimensional manifolds in composition space*, Combustion and Flame, 88 (1992), pp. 239–264, https://doi.org/10.1016/0010-2180(92)90034-M.

[13] S. Maeda, T. Taketsugu, and K. Morokuma, *Exploring transition state structures for intramolecular pathways by the artificial force induced reaction method*, Journal of Computational Chemistry, 35 (2014), pp. 166–173, https://doi.org/10.1002/jcc.23481.

[14] K. Ohno and S. Maeda, *A scaled hypersphere search method for the topography of reaction pathways on the potential energy surface*, Chemical Physics Letters, 384 (2004), pp. 277–282, https://doi.org/10.1016/j.cplett.2003.12.030.

[15] Y. Sumiya and S. Maeda, *Rate constant matrix contraction method for systematic analysis of reaction path networks*, Chemistry Letters, 49 (2020), pp. 553–564, https://doi.org/10.1246/cl.200092.

[16] Y. Sumiya, Y. Nagahata, T. Komatsuzaki, T. Taketsugu, and S. Maeda, *Kinetic analysis for the multistep profiles of organic reactions: significance of the conformational entropy on the rate constants of the Claisen rearrangement*, The Journal of Physical Chemistry A, 119 (2015), pp. 11641–11649, https://doi.org/10.1021/acs.jpca.5b09447.

[17] Y. Sumiya, T. Taketsugu, and S. Maeda, *Full rate constant matrix contraction method for obtaining branching ratio of unimolecular decomposition*, Journal of Computational Chemistry, 38 (2017), pp. 101–109, https://doi.org/10.1002/jcc.24526.

[18] T. Turányi and A. S. Tomlin, *Analysis of Kinetic Reaction Mechanisms*, Springer, Berlin, 2014, https://doi.org/10.1007/978-3-662-44562-4.

[19] R. S. Varga, *Matrix Iterative Analysis*, vol. 27 of Springer Series in Computational Mathematics, Springer, Berlin, second ed., 2000, https://doi.org/10.1007/978-3-642-05156-2.

[20] J. Wei and J. C. W. Kuo, *Lumping analysis in monomolecular reaction systems. Analysis of the exactly lumpable system*, Industrial and Engineering Chemistry Fundamentals, 8 (1969), pp. 114–123, https://doi.org/10.1021/i160029a019.

[21] Q. Ye, *Accurate inverses for computing eigenvalues of extremely ill-conditioned matrices and differential operators*, Mathematics of Computation, 87 (2017), pp. 237–259, https://doi.org/10.1090/mcom/3223.