arXiv:2511.22800v1 [math.PR] 27 Nov 2025

EMBEDDING OF REVERSIBLE MARKOV MATRICES

ELLEN BAAKE, MICHAEL BAAKE, AND JEREMY SUMNER

Abstract. The embeddability of reversible Markov matrices into time-homogeneous Markov semigroups is revisited, with some focus on simplifications and extensions. In particular, we do not demand irreducibility and consider weakly reversible matrices as well as reversible matrices with negative eigenvalues.

1. Introduction

A *Markov matrix* is a square matrix M with non-negative entries and unit row sums, while a *Markov generator* Q has non-negative off-diagonal entries and zero row sums. The latter is also known as a *rate matrix*, and generates a continuous-time semigroup of Markov matrices via $\{\mathrm{e}^{tQ} : t \geqslant 0\}$. Let $\mathcal{M}_d$ denote the convex set of all $d \times d$ Markov matrices; see [26] for general background. It is an old and still only partially answered question [14, 21] to determine which Markov matrices are *embeddable*, which means that they occur in some Markov semigroup of the above type. This is the time-homogeneous situation because e^{tQ} is the solution to the matrix-valued initial value problem $\dot{X} = XQ$ with $X(0) = \mathbb{1}$. This embedding is intimately related with the existence of a real matrix logarithm of M, meaning a real solution R of the equation $M = \mathrm{e}^{R}$, which may or may not be unique. The latter, in turn, may or may not be relevant in a given context. Here, we consider the embedding problem for a large but important class of Markov matrices [24] as follows.

A Markov matrix $M \in \mathcal{M}_d$ is called *reversible* if there exists a strictly positive probability vector $\boldsymbol{p} = (p_1, \ldots, p_d)$ such that the detailed balance equations $p_i M_{ij} = p_j M_{ji}$ hold for all $1 \leqslant i, j \leqslant d$; see [20, Ch. 1] for background. In this case, $\boldsymbol{p}$ must satisfy $\boldsymbol{p}M = \boldsymbol{p}$ and thus be a left eigenvector of M for the eigenvalue 1, that is, $\boldsymbol{p}$ is an equilibrium vector for M. Reversible Markov matrices play an important role both in the theory and in many applications of Markov chains [24]; this is due to the invariance under time reversal, and the fact that the detailed balance property allows for a straightforward and explicit calculation of $\boldsymbol{p}$. One concrete motivation for considering reversible Markov chains comes from the study of phylogenetics, which encompasses the inference of biological evolutionary history from present-day molecular sequence data. Here, the most common approach to statistical parameter inference works with a hierarchy of reversible Markov models with the most general case at the top; see any standard reference such as [27, Ch. 1] for further details.

2020 *Mathematics Subject Classification.* 60J10, 60J27, 15A30.

Key words and phrases. Markov matrices and generators, reversibility, embedding problem.

Questions of embeddability are particularly important in applications because data are usually sampled at discrete time points, but the underlying phenomena often operate in continuous time. When working with Markov matrices, special attention should thus be paid to their embeddability; however, this has sometimes been overlooked. For example, in bioinformatics, one of the famous Dayhoff (or PAM) matrices [12, 13], a family of reversible Markov matrices that describe the evolution of amino acids and are routinely used as scoring matrices for protein sequence alignment (see [15, Ch.6.5] for review), has turned out to not be embeddable [22] — although molecular evolution undoubtedly acts in continuous time.

Since irreducible Markov matrices have a unique eigenvector $\boldsymbol{p} > \boldsymbol{0}$ by the Perron–Frobenius theorem, it is natural to consider the embedding problem first for irreducible reversible matrices. This case has been analysed in [8] under the further restriction to reversible generators, where a fairly complete answer was given, which can still be reformulated in a more algebraic manner and then be simplified. Inevitably, this only covers reversible Markov matrices with positive spectrum. However, as we shall see below, there are reversible Markov matrices with (pairs of) negative eigenvalues, which calls for some extension. Also, there is no need to restrict to irreducible matrices from the beginning, so we will drop this restriction, too.

The paper is organised as follows. In Section 2, we provide some definitions and preparatory results, mainly from a linear algebra perspective. Then, in Section 3, we focus on reversible Markov matrices and their embedding, simplifying the approach of [8] and extending the discussion to reducible, weakly reversible matrices, and to cases with negative eigenvalues.

In Appendix A, we give an interpretation of one embeddable class of Markov matrices (for $d = 3$) with a pair of negative eigenvalues in terms of cyclic processes, which provides a simpler and independent approach to this class of matrices.

2. Preliminaries and preparatory results

The real $d \times d$ matrices are denoted as $\mathrm{Mat}(d, \mathbb{R})$, and the spectrum of a matrix B is written as $\sigma(B)$, which is understood as the set or multi-set of eigenvalues, possibly *including* multiplicities. When the eigenvalues are distinct, the spectrum is called *simple*. Further, we speak of *real* or (strictly) *positive* spectrum when $\sigma(B) \subset \mathbb{R}$ or $\sigma(B) \subset \mathbb{R}_{+}$, respectively, with obvious meaning of the inclusion relation in the multi-set case.

In view of later applications to Markov matrices and generators, we mainly work with row vectors in $\mathbb{R}^d$, denoted by $\boldsymbol{x} = (x_1, \ldots, x_d)$, while we use $\boldsymbol{x}^{\mathsf{T}}$ for column vectors. We call $\boldsymbol{x}$ strictly positive, written as $\boldsymbol{x} > \boldsymbol{0}$, when $x_i > 0$ holds for all $1 \leqslant i \leqslant d$.

Definition 2.1. Let $\boldsymbol{p} = (p_1, \ldots, p_d)$ be a strictly positive probability vector, so $p_i > 0$ for all $1 \leqslant i \leqslant d$ and $\sum_{i=1}^{d} p_i = 1$. Then, two matrices $A, B \in \mathrm{Mat}(d, \mathbb{R})$ form a ***p**-balanced pair* if the exchange relations $p_i A_{ij} = p_j B_{ji}$ hold for all $1 \leqslant i, j \leqslant d$.

No specific relation between $\boldsymbol{p} > \boldsymbol{0}$ and the matrices A and B is assumed, but it is easy to verify that $\boldsymbol{p}$ is a left eigenvector of A (of B) with eigenvalue λ if and only if all row sums of B (of A) are λ, equivalently if and only if $\boldsymbol{1}^{\mathsf{T}} = (1, \ldots, 1)^{\mathsf{T}}$ is a right eigenvector of B (of A)

with eigenvalue λ. The exchange relations can be written in matrix form as

$$B \;=\; D^{-1} A^{\mathsf{T}} D \qquad \text{with} \quad D \;=\; D_{\boldsymbol{p}} \;:=\; \operatorname{diag}(p_1, \dots, p_d). \tag{1}$$

A motivation for this particular form will follow shortly. Note that D possesses a unique matrix square root with positive entries on the diagonal, as does D^{-1}. Clearly, (1) is equivalent with $DB = A^{\mathsf{T}} D$, which can still be used when $\boldsymbol{p} \geqslant \boldsymbol{0}$ and D is no longer invertible.

Remark 2.2. When $\boldsymbol{p} > \boldsymbol{0}$, the mapping $A \mapsto \widetilde{A} := D^{-1} A^{\mathsf{T}} D$ defines an involution on $\operatorname{Mat}(d, \mathbb{R})$ that can be seen as the adjoint map for the right action of A in the bilinear form $\sum_{i=1}^{d} x_i p_i y_i$, in complete correspondence with the transpose of a matrix being the adjoint map for the standard bilinear form $\sum_{i=1}^{d} x_i y_i$. The importance of the new involution will become clear in the context of reversibility for Markov and related processes.

In the context of stationary Markov chains, both in discrete and in continuous time, this involution is induced by the time reversal of the chain; see [26, Ch. 1.9] or [20, Ch. 1.2] for background. Our definition is algebraic in nature and applies to a more general setting, which will pay off later in our analysis. ◊

Of particular interest in Definition 2.1 is the case $B = A$, where the exchange relations are called the equations of *detailed balance*. We use this term in slightly greater generality than usual for ease of accommodating Markov matrices, generators and some of their extensions under a common roof.

Definition 2.3. A matrix $B \in \operatorname{Mat}(d, \mathbb{R})$ is called *reversible for the probability vector* $\boldsymbol{p} > \boldsymbol{0}$, or $\boldsymbol{p}$-reversible for short, if it satisfies the detailed balance equations $p_i B_{ij} = p_j B_{ji}$ for all $1 \leqslant i, j \leqslant d$, or, equivalently, if $B = D^{-1} B^{\mathsf{T}} D$ with the diagonal matrix D from (1). Further, B is simply called *reversible* if it is $\boldsymbol{p}$-reversible for some $\boldsymbol{p} > \boldsymbol{0}$.

For a $\boldsymbol{p}$-reversible B, we see that $\boldsymbol{p}$ is a left eigenvector with eigenvalue λ if and only if $\boldsymbol{1}^{\mathsf{T}}$ is a right eigenvector for λ. Also, since $D^{1/2}$ is symmetric, $\boldsymbol{p}$-reversibility of B implies

$$\bigl(D^{1/2} B D^{-1/2}\bigr)^{\mathsf{T}} \;=\; D^{-1/2} B^{\mathsf{T}} D^{1/2} \;=\; D^{1/2} B D^{-1/2},$$

as well as the analogous equation for positive powers of B, which shows the following important property.

Fact 2.4. *Let $B \in \operatorname{Mat}(d, \mathbb{R})$ be reversible for the probability vector $\boldsymbol{p} > 0$. Then, B is similar to a symmetric matrix, hence diagonalisable, and has real spectrum, $\sigma(B) \subset \mathbb{R}$. Further, B^n is reversible for any $n \in \mathbb{N}$.* □

We will be interested in reversible matrices B that possess a real matrix logarithm with particular properties, that is, in solutions of $B = \mathrm{e}^{R}$ with $R \in \operatorname{Mat}(d, \mathbb{R})$ and possibly further restrictions. Let us first recall Culver's results [10] on the existence and uniqueness of real logarithms, formulated in terms of the (complex) *Jordan normal form* (JNF) of B.

Fact 2.5 (Culver)**.** *A matrix $B \in \mathrm{Mat}(d,\mathbb{R})$ has a real logarithm if and only if B is non-singular and has the property that, in the JNF of B, every elementary Jordan block with a negative eigenvalue occurs with even multiplicity.*

Further, B has a unique real logarithm if and only if all eigenvalues of B are positive real numbers and no elementary Jordan block of the JNF of B occurs more than once. If B is diagonalisable, this is equivalent to B having simple, positive spectrum. □

When B is diagonalisable with positive spectrum but repeated eigenvalues, it has uncountably many real logarithms, as also discussed in [10]. Such a non-uniqueness is relevant for the following reason.

Lemma 2.6. *Let $B \in \mathrm{Mat}(d,\mathbb{R})$ satisfy $B = \mathrm{e}^{Q} = \mathrm{e}^{R}$ with $Q, R \in \mathrm{Mat}(d,\mathbb{R})$, and consider $\varTheta := \{t \in \mathbb{R} : \mathrm{e}^{tQ} = \mathrm{e}^{tR}\}$. Then, either $\varTheta = \mathbb{R}$, which is equivalent with $Q = R$, or $\varTheta \subset \mathbb{R}$ is locally finite, which means that $\varTheta \cap [a, a+r]$ is a finite set for every $a \in \mathbb{R}$ and $r > 0$.*

Proof. Obviously, $Q = R$ implies $\varTheta = \mathbb{R}$, while the converse emerges from $B(t) = \mathrm{e}^{tQ}$ via $Q = \dot{B}(0)$, so the equivalence of $Q = R$ and $\varTheta = \mathbb{R}$ is clear.

Now, observe that $\varTheta$ is a closed subset of $\mathbb{R}$, as follows from the continuity of the matrix exponential maps $t \mapsto \mathrm{e}^{tQ}$ and $t \mapsto \mathrm{e}^{tR}$. Assume that $\varTheta$ contains a sequence $(t_m)_{m\in\mathbb{N}}$ of distinct points that converge in $\mathbb{R}$, so $\lim_{m\to\infty} t_m = t_0$. Then, $t_0 \in \varTheta$ as well, which is to say that we also have $\mathrm{e}^{t_0 Q} = \mathrm{e}^{t_0 R}$.

Now, we can compare the series expansions around t_0, which result in the matrix relation $Q = R + \mathcal{O}(t_m - t_0)$ for every $m \in \mathbb{N}$, interpreted elementwise. As t_m converges towards t_0, this is only possible if $Q = R$. So, when $Q \neq R$, the set $\varTheta$ cannot contain any limit point, and must thus be locally finite, hence discrete and closed. □

Let us return to Culver's result. Indeed, matrices with negative eigenvalues can have a real logarithm, and this situation will still occur in the realm of reversible matrices; see Fact 3.19 below. However, if any of the logarithms is to be reversible, one hits the following constraint from the *spectral mapping theorem* (SMT); compare [23, Thm. 9.4.6].

Fact 2.7. *If $R \in \mathrm{Mat}(d,\mathbb{R})$ is reversible for $\boldsymbol{p} > \boldsymbol{0}$, its exponential e^{R} is $\boldsymbol{p}$-reversible as well, hence diagonalisable, and has positive spectrum, $\sigma(\mathrm{e}^{R}) \subset \mathbb{R}_{+}$.*

Proof. If R is $\boldsymbol{p}$-reversible, it is diagonalisable with $\sigma(R) \subset \mathbb{R}$ by Fact 2.4. All powers of R and their linear combinations are $\boldsymbol{p}$-reversible as well, hence also e^{R}, the latter by a standard continuity argument. Here, e^{R} is diagonalisable (by the same matrix as used for R), and $\sigma(\mathrm{e}^{R}) = \{\mathrm{e}^{\lambda} : \lambda \in \sigma(R)\} \subset \mathbb{R}_{+}$ is clear. □

Let us pause for a comment on the set of $\boldsymbol{p}$-reversible matrices, defined as

$$\mathcal{A}_{\boldsymbol{p}} := \{A \in \mathrm{Mat}(d,\mathbb{R}) : p_i A_{ij} = p_j A_{ji} \text{ holds for all } 1 \leqslant i, j \leqslant d\}, \tag{2}$$

which contains $\mathbb{1}$ and is (topologically) closed. If $A, B \in \mathcal{A}_{\boldsymbol{p}}$, it is clear that also any real linear combination of A and B lies in $\mathcal{A}_{\boldsymbol{p}}$. Moreover, for any $1 \leqslant i, j \leqslant d$, we get

$$p_i (AB)_{ij} \, = \sum_{k=1}^{d} p_i A_{ik} B_{kj} \, = \sum_{k=1}^{d} p_k A_{ki} B_{kj} \, = \sum_{k=1}^{d} p_j B_{jk} A_{ki} \, = \, p_j (BA)_{ji} , \tag{3}$$

which interchanges the matrix order. It then follows that $AB + BA$ is $\boldsymbol{p}$-reversible again, so that $\mathcal{A}_{\boldsymbol{p}}$ is a real Jordan algebra, in line with the results from [9]. Note that we did not need $\boldsymbol{p}$ to be strictly positive in (3), so $\mathcal{A}_{\boldsymbol{p}}$ can also be considered for any $\boldsymbol{p} \geqslant \boldsymbol{0}$ via the detailed balance equations. Put together, we thus have the following result.

Fact 2.8. *For any fixed probability vector $\boldsymbol{p} = (p_1, \ldots, p_d)$, the set $\mathcal{A}_{\boldsymbol{p}}$ from (2), with the product $A \circ B := \frac{1}{2}(AB + BA)$, is a unital Jordan algebra.* □

In fact, $\mathcal{A}_{\boldsymbol{p}}$ being closed under taking squares is equivalent with it being a Jordan algebra, because $A^2 = A \circ A$ and $A \circ B = \frac{1}{2}\bigl((A+B)^2 - A^2 - B^2\bigr)$. In either case, one then also gets closure under taking any positive powers, as one sees from $A^{n+1} = A \circ A^n$ inductively, again for any probability vector $\boldsymbol{p}$. We shall return to this in the context of Markov matrices.

Now, if we are interested in reversible matrices with a reversible real logarithm, Fact 2.7 implies that we need to restrict to matrices with positive spectrum. If $B \in \mathrm{Mat}(d, \mathbb{R})$ has simple, positive spectrum, Fact 2.5 asserts that there is precisely one real logarithm, and this is the principal one. Since such a B is diagonalisable, we have $B = U \mathrm{diag}(\lambda_1, \ldots, \lambda_d) U^{-1}$ for some invertible real matrix U. Then, we simply get the principal logarithm as

$$\log(B) \, = \, U \mathrm{diag}\bigl(\log(\lambda_1), \ldots, \log(\lambda_d)\bigr) U^{-1} .$$

When the spectrum possesses non-trivial multiplicities, there are several real logarithms, but we still have the following result.

Lemma 2.9. *Let $B \in \mathrm{Mat}(d, \mathbb{R})$ be $\boldsymbol{p}$-reversible for $\boldsymbol{p} > \boldsymbol{0}$. If B has positive spectrum, its principal logarithm is a real matrix that is $\boldsymbol{p}$-reversible as well.*

Proof. The principal logarithm of a real matrix with positive spectrum exists and is again real [16, Thm. 1.31]. By [16, Thm. 11.1], it is given by

$$\log(B) \, = \int_0^1 (B - \mathbb{1}) \bigl(t(B - \mathbb{1}) + \mathbb{1}\bigr)^{-1} \, \mathrm{d}t ,$$

where the two matrices under the integral commute.

Now, we have $B^{\mathsf{T}} = D B D^{-1}$ from Definition 2.3, which implies

$$\begin{aligned} \log(B)^{\mathsf{T}} \, &= \int_0^1 \Bigl(\bigl(t(B - \mathbb{1}) + \mathbb{1}\bigr)^{\mathsf{T}} \Bigr)^{-1} (B - \mathbb{1})^{\mathsf{T}} \, \mathrm{d}t \, = \int_0^1 D \bigl(t(B - \mathbb{1}) + \mathbb{1}\bigr)^{-1} (B - \mathbb{1}) D^{-1} \, \mathrm{d}t \\ &= \, D \int_0^1 (B - \mathbb{1}) \bigl(t(B - \mathbb{1}) + \mathbb{1}\bigr)^{-1} \, \mathrm{d}t \, D^{-1} \, = \, D \log(B) D^{-1} \end{aligned}$$

as claimed, where the penultimate step uses the commutativity mentioned above. □

The occurrence of multiple solutions is related to the structure of the *commutant* of the matrix B, which is the matrix ring

$$\mathrm{comm}(B) := \{S \in \mathrm{Mat}(d, \mathbb{R}) : [B, S] = \mathbb{0}\}. \tag{4}$$

When the characteristic polynomial of B is also its minimal one, $\mathrm{comm}(B)$ is Abelian and given by the polynomial ring $\mathbb{R}[B]$. Otherwise, it contains further elements, which complicate matters; see [23, Sec. 12.4] for details. Below, we shall see that some natural conditions will ensure that at most one real logarithm of a given B is reversible.

3. Reversible Markov matrices

Let us now consider the case of Markov matrices, where we begin with a mild extension of the reversibility notion. Note that we deviate from the standard approach in [24, Sec. 1.6] because we do not want to restrict to irreducible Markov matrices.

Definition 3.1. A Markov matrix M is *reversible* if it is reversible in the sense of Definition 2.3. Further, it is called *weakly reversible* if the detailed balance equations $p_i M_{ij} = p_j M_{ji}$ hold for all $1 \leqslant i, j \leqslant d$ and some probability vector $\boldsymbol{p} \geqslant \boldsymbol{0}$, thus admitting zero entries. The corresponding notions are also used for Markov generators.

3.1. **General structure.** If $M \in \mathcal{M}_d$ is weakly reversible, $\boldsymbol{p}$ must be an equilibrium vector of M. Weak reversibility of $M \in \mathcal{M}_d$ is equivalent to the identity $DM = M^{\mathsf{T}} D$ with the matrix D from Eq. (1). Note that D is not invertible if M is only weakly reversible. In relation to $\mathcal{A}_{\boldsymbol{p}}$ from Eq. (2), for any fixed probability vector $\boldsymbol{p}$, one can consider the set of (weakly) $\boldsymbol{p}$-reversible Markov matrices

$$\mathcal{M}_{d,\boldsymbol{p}} := \mathcal{M}_d \cap \mathcal{A}_{\boldsymbol{p}},$$

which is convex and topologically closed. It is also closed under taking squares, under taking arbitrary positive powers, and under the Jordan product $M \circ M' = \frac{1}{2}(MM' + M'M)$. In fact, these three properties are equivalent, because (weak) reversibility is defined by a linear relation, so our earlier argument used after Fact 2.8 still applies.

When M is irreducible, weak reversibility implies reversibility, and reversibility can be established via Kolmogorov's loop criterion; see [20, Sec. 1.5] for details. At least for small state spaces, this is also effective, while the computational effort quickly grows with d; see [6] for details and alternatives.

Example 3.2. Let us consider $\mathcal{M}_2$. Clearly, $\mathbb{1} \in \mathcal{M}_2$ is (weakly) reversible for every probability vector $\boldsymbol{p}$. Otherwise, $\mathbb{1} \neq M \in \mathcal{M}_2$ reads $M = \left(\begin{smallmatrix} 1-a & a \\ b & 1-b \end{smallmatrix}\right)$ with $a, b \in [0,1]$, not both zero. Then, the unique equilibrium vector is $\boldsymbol{p} = \frac{1}{a+b}(b, a)$, and M is $\boldsymbol{p}$-reversible if $ab > 0$ and weakly reversible otherwise.

By Kendall's theorem [21], M is embeddable if and only if $\det(M) > 0$, equivalently $a + b < 1$, and the embedding is unique. Indeed, one then has $M = \mathrm{e}^{Q}$ with the principal matrix logarithm $Q = -\frac{\log(1-a-b)}{a+b}(M - \mathbb{1})$, which is a Markov generator when $a + b < 1$. Moreover, Q is (weakly) reversible if and only if M is (weakly) reversible.

On the other hand, there are reversible matrices that are not embeddable at all, such as $M = \left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)$, and we should expect more complicated cases in $\mathcal{M}_d$ with $d > 2$. ♢

Example 3.3. Let $\boldsymbol{x} \in \mathbb{R}^d$ be a non-negative row vector and let $C_{\boldsymbol{x}} \in \mathrm{Mat}(d, \mathbb{R})$ denote the matrix with d equal rows, each being $\boldsymbol{x}$. Then, $Q_{\boldsymbol{x}} = -x\mathbb{1} + C_{\boldsymbol{x}}$ with $x = \mathrm{tr}(C_{\boldsymbol{x}}) = x_1 + \ldots + x_d$ is a Markov generator and $M_{\boldsymbol{x}} = (1-x)\mathbb{1} + C_{\boldsymbol{x}}$ a Markov matrix, the latter under the condition $x \leqslant \frac{d}{d-1}$; see [2] for details. Such matrices are said to be of *equal-input* type (or to be *parent independent*). Since $\boldsymbol{x} = \boldsymbol{0}$ gives $M_{\boldsymbol{0}} = \mathbb{1}$, which is (weakly) reversible for every probability vector $\boldsymbol{p}$, let us assume that $\boldsymbol{x}$ has at least one positive entry.

Now, $M_{\boldsymbol{x}}$ with non-negative $\boldsymbol{x}$ and $0 < x \leqslant \frac{d}{d-1}$ has a unique equilibrium vector, $\boldsymbol{p} = \frac{\boldsymbol{x}}{x}$, and $M_{\boldsymbol{x}}$ is weakly reversible for this $\boldsymbol{p}$, as we always have $DC_{\boldsymbol{x}} = C_{\boldsymbol{x}}^{\mathsf{T}} D$ with D as in (1), and reversible when $\boldsymbol{x} > \boldsymbol{0}$. So, all equal-input Markov matrices are (weakly) reversible. The embedding problem for this matrix class was studied in detail in [2]. ♢

Example 3.4. Let $M \in \mathcal{M}_d$ with $d \geqslant 2$ be tridiagonal, with $a_i = \mathbb{P}(i \to i{+}1) > 0$ for $1 \leqslant i < d$ on the superdiagonal and $b_i = \mathbb{P}(i \to i{-}1) > 0$ for $1 < i \leqslant d$ on the subdiagonal. This is then the irreducible transition matrix of a simple birth and death process (or random walk) in discrete time with states $\{1, 2, \ldots, d\}$. Here, M is always reversible, because the equations for its equilibrium vector directly equal the detailed balance equations.

For $d = 2$ and $a_1 + b_2 < 1$, we get embeddability from Kendall's theorem, compare [21, 1], but no other case is embeddable because it violates the transitivity property that $M_{ij} > 0$ and $M_{jk} > 0$ also forces $M_{ik} > 0$; see [1, Prop. 2.1] for details. To circumvent this well-known obstruction, one can consider, as is often done, birth and death processes in continuous time, then using tridiagonal rate matrices to describe them. ♢

Recall that $M \in \mathcal{M}_d$ is reversible if it is $\boldsymbol{p}$-reversible for some $\boldsymbol{p} > \boldsymbol{0}$. Such a $\boldsymbol{p}$ must be an equilibrium vector. The existence (and uniqueness) of such a $\boldsymbol{p}$ is clear when M happens to be irreducible, by the Perron–Frobenius theorem [5, Thm. 2.1.4]. Since we do not restrict to irreducible M, we also need the following result, which we could not trace in the literature.

Fact 3.5. *A matrix $M \in \mathcal{M}_d$ has a strictly positive equilibrium vector if and only if M is the direct sum of irreducible Markov matrices, hence a diagonal block matrix with irreducible blocks, possibly after a suitable permutation of the states. The equilibrium vector is unique if and only if M is irreducible itself.*

Proof. Any $M \in \mathcal{M}_d$ has spectral radius 1 and satisfies $M\boldsymbol{1}^{\mathsf{T}} = \boldsymbol{1}^{\mathsf{T}}$, which is strictly positive. If M is irreducible, also the corresponding left eigenvector $\boldsymbol{p}$ is strictly positive, by the Perron–Frobenius theorem. When normalised such that $\boldsymbol{p} \cdot \boldsymbol{1}^{\mathsf{T}} = 1$, this is the equilibrium vector. If $M = \oplus_{i=1}^{s} M^{(i)}$ with $s > 1$ and irreducible $M^{(i)}$ of dimension $d^{(i)}$, every block has its unique equilibrium vector $\boldsymbol{p}^{(i)}$ of length $d^{(i)}$. Then, each $\boldsymbol{0} \oplus \cdots \oplus \boldsymbol{0} \oplus \boldsymbol{p}^{(i)} \oplus \boldsymbol{0} \oplus \cdots \oplus \boldsymbol{0}$ is an equilibrium vector of M, and any convex combination of these s probability vectors is again an equilibrium vector, and among them are many strictly positive ones.

It remains to see why this is the only possibility of $M \in \mathcal{M}_d$ with $\boldsymbol{p}M = \boldsymbol{p} > \boldsymbol{0}$. Thus, we need a criterion for the existence of a strictly positive left eigenvector. Since a strictly positive right eigenvector exists, we can invoke [5, Thm. 2.3.14], which implies that $\lambda = 1$ has strictly positive left *and* right eigenvectors if and only if, after some permutation, $M = \bigoplus_{i=1}^{s} M^{(i)}$ with all $M^{(i)}$ irreducible. This is equivalent to all communication classes being both basic and final, which brings the Frobenius normal form to this diagonal block form. □

A reversible Markov matrix, up to permutations of the states, must then be such a block-diagonal matrix with irreducible blocks, and this is the reason why most treatments of reversibility restrict to one block, and hence to irreducible matrices. In the context of the embedding problem and the relation to weakly reversible extensions, this is neither necessary nor advantageous.

Remark 3.6. When $M \in \mathcal{M}_d$ is irreducible and reversible, it must be reversible for the unique equilibrium vector $\boldsymbol{p}$ of M, which is strictly positive. When M is reducible and reversible for some $\boldsymbol{p} > \boldsymbol{0}$, this is not the only equilibrium vector. Since $M = \bigoplus_{i=1}^{s} M^{(i)}$ as in the proof of Fact 3.5, each irreducible block $M^{(i)}$ must be reversible on its own. Consequently, M must be (weakly) reversible for *each* of its equilibrium vectors, which gives some freedom to choose a suitable one when comparing with nearby matrices.

Note that the converse need not hold: A weakly reversible matrix M can be a direct sum of a reversible and a non-reversible one, where both can still be irreducible. Then, there is also a strictly positive equilibrium vector, but M can neither be reversible with respect to this one, nor with respect to any other strictly positive one. ◊

At this point, we have gained access to the following classes of matrices.

Example 3.7. Consider the convex set of *doubly stochastic* matrices within $\mathcal{M}_d$, which are the Markov matrices with both unit row sums and unit column sums. By the Birkhoff–von Neumann theorem, they form a convex polytope with the $d!$ permutation matrices as its vertices (or extremal points), so each doubly stochastic matrix is a convex combination of (some of) these.

If M is any doubly stochastic matrix, $\boldsymbol{1}$ is a left eigenvector for the leading eigenvalue $\lambda = 1$, so $\boldsymbol{p} = \frac{1}{d}\boldsymbol{1}$ is an equilibrium vector, which gives $D = \mathbb{1}$. All *symmetric* Markov matrices are $\boldsymbol{p}$-reversible, with $\widetilde{M} = D^{-1} M^{\mathsf{T}} D = M^{\mathsf{T}} = M$. Moreover, by Remark 3.6, they are the only reversible ones among the doubly stochastic matrices, though there can be further weakly reversible matrices that are not symmetric. ◊

3.2. **Embedding.** Let us now take a closer look at embeddability, where we start with a precise version of the notion alluded to in the introduction.

Definition 3.8. A (weakly) reversible Markov matrix M is called *embeddable* if it satisfies $M = \mathrm{e}^{Q}$ for some Markov generator Q. Further, M is *reversibly embeddable* if $M = \mathrm{e}^{Q}$ holds for a Markov generator Q that is itself (weakly) reversible.

For embeddability, we thus need a real logarithm that is also a rate matrix. To obtain a practically useful criterion, let us begin with the following observation.

Fact 3.9. *Let $M \in \mathcal{M}_d$ be a Markov matrix with positive spectrum, $\sigma(M) \subset \mathbb{R}_+$, and let $m \leqslant d$ be the degree of its minimal polynomial. Then, M has at least one real logarithm, namely its principal matrix logarithm,*

$$\log(\mathbb{1} + A) \,=\, \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n} A^n,$$

where $A = M - \mathbb{1}$ is a rate matrix with spectral radius $\varrho_A < 1$. Here, $\log(\mathbb{1} + A)$ is a matrix with zero row sums that is a polynomial in A of degree $m-1$ with real coefficients.

Further, when $\sigma(M) \subset \mathbb{R}_+$ is simple, $\log(\mathbb{1} + A)$ is the only real matrix logarithm of M.

Proof. Under the assumptions, we have $\sigma(M) \subset (0, 1]$, hence $\sigma(A) \subset (-1, 0]$, and $\varrho_A < 1$ is clear. This implies the convergence of the series. Since M has unit row sums, A has zero row sums, as do all positive powers of A. As this property is preserved in the limit by continuity, $\log(\mathbb{1} + A)$ has zero row sums as well.

Since the minimal polynomials of M and A have the same degree, the claim on the polynomial form is a consequence of the Cayley–Hamilton theorem; compare [17, Thm. 3.3.1].

The final uniqueness statement is a consequence of Culver's theorem (Fact 2.5). □

It is clear that a reversible Markov matrix has real spectrum, because it is similar to a symmetric matrix. Such a matrix can still have negative eigenvalues. If all negative eigenvalues have even multiplicity, a real matrix logarithm exists and embeddability may occur, but not with a reversible generator, as the following result shows.

Lemma 3.10. *Let $M \in \mathcal{M}_d$ and let $\boldsymbol{p} > \boldsymbol{0}$ be a probability vector. Assume that a $\boldsymbol{p}$-reversible real matrix logarithm R of M exists, so $M = \mathrm{e}^R$. Then, M is $\boldsymbol{p}$-reversible with $\sigma(M) \subset \mathbb{R}_+$. Further, no other $\boldsymbol{p}$-reversible real logarithm of M exists, and $R = \log(\mathbb{1}+A)$ is the principal matrix logarithm from Fact 3.9.*

Proof. When $M = \mathrm{e}^R$ with a $\boldsymbol{p}$-reversible R, Fact 2.7 implies the first claim.

It is clear that the rate matrix $A = M - \mathbb{1}$ is $\boldsymbol{p}$-reversible, as are the powers A^n for $n \in \mathbb{N}$. Consequently, $L = \log(\mathbb{1} + A)$ from Fact 3.9 is a real logarithm of M with zero row sums that is also $\boldsymbol{p}$-reversible, in line with Lemma 2.9.

In the generic case that M has simple spectrum, there is precisely one real logarithm of M by Culver's theorem (Fact 2.5), which must be the existing principal one, so we get $R = L = \log(\mathbb{1} + A)$.

In general, we know that M and L commute, and both are $\boldsymbol{p}$-reversible matrices that are simultaneously diagonalisable, with positive and with real spectrum, respectively. Moreover, their diagonal forms match, in the sense that $\lambda \in \sigma(M)$ and $\log(\lambda) \in \sigma(L)$ share the same (algebraic and geometric) multiplicity. This implies that they have the same commutant, $\mathrm{comm}(M) = \mathrm{comm}(L)$, as defined in (4); see [23] for background and details.

Now, let Q be any real solution of $M = \mathrm{e}^Q$. Then, Q must commute with M and hence also with L. As e^Q is invertible, we get

$$\mathbb{1} \,=\, \mathrm{e}^L \mathrm{e}^{-Q} \,=\, \mathrm{e}^{L-Q},$$

and $L - Q$ is a real logarithm of $\mathbb{1}$. If Q is also $\boldsymbol{p}$-reversible, then so is $L - Q$, which thus is diagonalisable with real spectrum. By the SMT, as 0 is the only real logarithm of 1, the spectrum of the diagonalisable matrix $L - Q$ must consist of 0s entirely, which implies $L - Q = \mathbb{0}$. We thus get $Q = L$, which also applies to the matrix R from the assumption. □

Remark 3.11. Let us note that $M = \mathrm{e}^Q$ with Q a Markov generator implies that M and Q share the same set of equilibrium vectors. Clearly, $\boldsymbol{p}Q = \boldsymbol{0}$ implies $\boldsymbol{p}M = \boldsymbol{p}$. The converse direction depends on the fact that any eigenvalue of a Markov generator is either 0 or has a strictly negative real part, see [1, Prop. 2.3], thus can never be purely imaginary, hence not of the form $k2\pi\mathrm{i}$ for $k \neq 0$. Since the algebraic and the geometric multiplicity of $1 \in \sigma(M)$ is the same [3, Fact 2.2], the spectrum of Q must contain 0 with the identical (algebraic and geometric) multiplicity by the SMT. Consequently, any (real) eigenvector of M with $\boldsymbol{x}M = \boldsymbol{x}$ must also satisfy $\boldsymbol{x}Q = \boldsymbol{0}$, which includes all equilibrium vectors. In particular, in this situation, M has a strictly positive one if and only if Q does.

Let us note that the above argument critically depends on Q being a rate matrix. Indeed, it can already fail if Q has only zero row sums, but violates non-negativity of its off-diagonal elements (such as $Q = \frac{2\pi}{\sqrt{3}}(Q_+ - Q_-)$ with $Q_\pm$ from Example 3.18 below, which has spectrum $\sigma(Q) = \{0, \pm\mathrm{i}\}$). Interestingly, Fact 3.5 now also has the consequence that a reversible and embeddable Markov matrix, possibly after a suitable permutations of the states, must be the direct sum of totally positive (hence in particular primitive) block matrices, because the exponential of an irreducible Markov generator is a totally positive matrix. ♢

Let us continue with a straightforward consequence of Lemma 3.10.

Proposition 3.12. *If $M \in \mathcal{M}_d$ is a reversible Markov matrix with simple, positive spectrum, the following statements are equivalent, where $A = M - \mathbb{1}$ as before.*

(1) *M is embeddable.*
(2) *M is reversibly embeddable.*
(3) *$Q = \log(\mathbb{1} + A)$ is a Markov generator.*

Proof. Under the assumption, there is precisely one real matrix logarithm, namely the principal one from Fact 3.9, which is Q as stated. This Q is reversible with the same $\boldsymbol{p}$ as M.

Consequently, M is embeddable if and only if this Q is a Markov generator. If so, it is automatically a reversible generator, and we are done. □

When the spectrum fails to be simple, further real logarithms exist, and we shall see explicit cases below in Examples 3.18 and 3.20. However, another consequence of Lemma 3.10 in conjunction with Remark 3.6 is the following.

Corollary 3.13. *A reversible Markov matrix $M = \mathbb{1} + A \in \mathcal{M}_d$ with positive spectrum is reversibly embeddable if and only if $Q = \log(\mathbb{1} + A)$ is a Markov generator.* □

When applied to the class of symmetric matrices, compare Example 3.7, this gives the following generalisation of [1, Thm. 6.5].

Corollary 3.14. *A symmetric* $M = \mathbb{1} + A \in \mathcal{M}_d$ *is symmetrically embeddable if and only if* $\sigma(M) \subset \mathbb{R}_+$ *and the symmetric matrix* $Q = \log(\mathbb{1} + A)$ *is a Markov generator.* □

3.3. **Concrete conditions.** Let us next derive a simple criterion for reversible embeddability, that is, for $\log(\mathbb{1} + A)$ to be a Markov generator. Let $M \in \mathcal{M}_d$ be reversible with positive spectrum, possibly with multiplicities, and assume that the degree of its minimal polynomial is m, where $m \leqslant d$. Then, by Fact 3.9, there are real numbers $\alpha_1, \ldots, \alpha_{m-1}$ such that

$$L \, = \, \log(\mathbb{1} + A) \, = \, \sum_{k=1}^{m-1} \alpha_k A^k, \tag{5}$$

where no term with $A^0 = \mathbb{1}$ occurs because L has zero row sums. Since $\sigma(M) \subset (0, 1]$, we can write the eigenvalues of M in descending order as $\lambda_0 = 1 > \lambda_1 > \cdots > \lambda_{m-1} > 0$ and those of A as $\mu_0 = 0 > \mu_1 > \cdots > \mu_{m-1}$, where $\mu_i = \lambda_i - 1$. The corresponding eigenvalues of L are the unique real logarithms $0 = \log(\lambda_0) > \log(\lambda_1) > \cdots > \log(\lambda_{m-1})$. The SMT now gives $m - 1$ conditions for the eigenvalues. They can be written in matrix form as

$$\begin{pmatrix} \mu_1 & \mu_1^2 & \cdots & \mu_1^{m-1} \\ \mu_2 & \mu_2^2 & \cdots & \mu_2^{m-1} \\ \vdots & & & \vdots \\ \mu_{m-1} & \mu_{m-1}^2 & \cdots & \mu_{m-1}^{m-1} \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_{m-1} \end{pmatrix} = \begin{pmatrix} \log(\lambda_1) \\ \log(\lambda_2) \\ \vdots \\ \log(\lambda_{m-1}) \end{pmatrix}. \tag{6}$$

We note that the number of equations is one lower than in the approach of [8]. The matrix on the left has determinant $\bigl(\prod_i \mu_i\bigr) \prod_{k>\ell} (\mu_k - \mu_\ell)$ and is non-singular because the μ_i with $1 \leqslant i \leqslant m-1$ are non-zero and distinct. So, the coefficients α_i are uniquely determined and real; see [2, Sec. 5] for details and an explicit formula for the inverse of this variant of the Vandermonde matrix. We thus have established the following result.

Proposition 3.15. *A reversible Markov matrix* $M \in \mathcal{M}_d$ *with a minimal polynomial of degree* $m \leqslant d$ *is reversibly embeddable if and only if the following two conditions are satisfied.*

1. *The spectrum of* M *is positive,* $\sigma(M) \subset \mathbb{R}_+$.
2. *The coefficients* $\alpha_1, \ldots, \alpha_{m-1}$ *that solve Eq.* (6) *are such that* $L = \log(\mathbb{1} + A)$ *from Eq.* (5) *is a Markov generator.*

In this case, L *is unique in the sense that no other reversible real logarithm of* M *exists.* □

By [21, Prop. 7], for any $2 \leqslant d \in \mathbb{N}$, the set $\mathcal{E}_d$ of embeddable Markov matrices is relatively closed within the set $\{M \in \mathcal{M}_d : \det(M) > 0\}$. In contrast, the set of reversible Markov matrices is *not* relatively closed, because the limit of a converging sequence of reversible Markov matrices, provided that also the corresponding probability vectors converge, might only be weakly reversible.

Likewise, we know from [11, Thm. 7] that the set of embeddable Markov matrices with simple spectrum is relatively open and dense in $\mathcal{E}_d$. In contrast, this is no longer true for embeddable reversible Markov matrices, because they have real spectrum but can never have simple negative eigenvalues. Nevertheless, we still get the following result.

Theorem 3.16. *If $M \in \mathcal{M}_d$ is reversible for some $\boldsymbol{p} > \boldsymbol{0}$, the following are equivalent.*

(1) *M is embeddable and has positive spectrum.*
(2) *M is embeddable with a $\boldsymbol{p}$-reversible Markov generator.*
(3) *M has positive spectrum and its principal logarithm, $L = \log(\mathbb{1} + A)$ with $A = M - \mathbb{1}$, is a Markov generator.*

In this case, L is the only reversible matrix logarithm of M.

Beyond this case, when $d \geqslant 3$, a reversible Markov matrix can still be embeddable, but never with a reversible Markov generator.

Proof. The implication (2) $\Rightarrow$ (1) is clear from Fact 2.7, while (2) $\Leftrightarrow$ (3) follows from Lemma 3.10 by standard arguments.

Next, we need to show that (1) $\Rightarrow$ (2). To this end, assume that $M \in \mathcal{M}_d$ with $\sigma(M) \subset \mathbb{R}_+$ is $\boldsymbol{p}$-reversible, and that $M = \mathrm{e}^Q$ for some Markov generator Q. If Q is $\boldsymbol{p}$-reversible itself, there is nothing to show, so assume that Q is not reversible. By [11, Thm. 7], we know that there are embeddable Markov matrices with simple spectrum in any neighbourhood of M within $\mathcal{M}_d$. In fact, there are also reversible ones among them, as follows from the following deformation argument, which is taken from the proof of [6, Thm. 3.1].

Let us first assume that M is irreducible. Then, if M is reversible, we can modify any of its rows by multiplying it with an arbitrary non-negative number such that the sum of the non-diagonal elements remains $\leqslant 1$, followed by modifying the diagonal element such that the row sum is again 1. Similarly, we can multiply any column with a number in $(0, 1)$, so no row sum can exceed 1, and then re-adjust all diagonal elements for unit row sums.

As follows from Kolmogorov's loop criterion, compare [20, 6], none of these operations destroys reversibility, as long as the resulting equilibrium vector (which may change) remains positive, which is certainly the case for sufficiently small deformations. This gives enough freedom to find, in any neighbourhood of M, a reversible matrix M' with simple spectrum that is also embeddable. Then, the corresponding Markov generator is unique and given by the principal matrix logarithm of M',

More generally, M reversible implies via Fact 3.5 that M is the direct sum of irreducible Markov matrices. By permuting the states appropriately, we may assume that M has diagonal block form. We can now employ the above deformation argument to each block separately. Since we have some freedom to choose $\boldsymbol{p} > \boldsymbol{0}$, as explained in Remark 3.6, we obtain the corresponding conclusion: in any neighbourhood of M, we can find an embeddable, reversible matrix M' with simple spectrum and a nearby probability vector. For the latter, without loss of generality, we may assume that none of its entries is smaller than $\epsilon := \frac{1}{2} \min_i p_i > 0$, say. The Markov generator for M' is again the principal matrix logarithm, which has the same equilibrium vectors as M'; compare Remark 3.11.

Now, we can select a converging sequence of embeddable reversible Markov matrices with simple spectrum that converges to M such that also the corresponding equilibrium vectors converge to some $\boldsymbol{p} > \boldsymbol{0}$, which is possible by a standard compactness argument involving the constant $\epsilon > 0$. All of these matrices are embeddable via their principal logarithm, by Proposition 3.12, which is another converging sequence, this time of Markov generators. The limit is still a Markov generator, and it is reversible. Then, it must be the principal logarithm L of M by Corollary 3.13, which proves the claim.

The uniqueness of L follows from Proposition 3.12 in conjunction with Lemma 3.10 and Remark 3.6, because L must be $\boldsymbol{p}$-reversible and thus also reversible for any other equilibrium vector of M.

The last claim is made explicit in Example 3.18 below, which establishes the statement in line with Lemma 3.10. □

3.4. **Multiple embeddings.** With the above results, Proposition 3.15 becomes the right tool to test all reversible Markov matrices with positive spectrum for embeddability. When the spectrum is degenerate, we do not get uniqueness. Let us look at this case more closely now. In general, when a reversible Markov matrix has more than one real logarithm, non-reversible ones must come in pairs. Indeed, let $M = \widetilde{M} = D^{-1}M^{\intercal}D$ for some probability vector $\boldsymbol{p} > \boldsymbol{0}$ with $D = D_{\boldsymbol{p}}$ as above, where $\boldsymbol{p}$ then is an equilibrium vector for M. Now, if $M = \mathrm{e}^{R}$ for some $R \in \mathrm{Mat}(d, \mathbb{R})$ and $\widetilde{R} = D^{-1}R^{\intercal}D$, we get

$$M \, = \, \widetilde{M} \, = D^{-1}\big(\mathrm{e}^{R}\big)^{\intercal}D \, = \, \exp\big(D^{-1}R^{\intercal}D\big) \, = \, \mathrm{e}^{\widetilde{R}},$$

where $\widetilde{R}$ is a Markov generator if and only if R is one. Indeed, if R is a Markov generator, it must satisfy $\boldsymbol{1}R^{\intercal} = \boldsymbol{0}$ and share the equilibrium vector $\boldsymbol{p}$ with M, so $\boldsymbol{p}R = \boldsymbol{0}$, by Remark 3.11. Then, $\widetilde{R}$ has non-negative off-diagonal entries because R does and D is a non-negative matrix, and one gets $\boldsymbol{p}\widetilde{R} = \boldsymbol{1}R^{\intercal}D = \boldsymbol{0}D = \boldsymbol{0}$. Likewise, $\boldsymbol{1}\widetilde{R}^{\intercal} = \boldsymbol{p}RD^{-1} = \boldsymbol{0}D^{-1} = \boldsymbol{0}$, so $\widetilde{R}$ is a rate matrix. The converse direction follows from $R = D^{-1}\widetilde{R}^{\intercal}D$ in the same way.

So, let us assume that R is a rate matrix. Then, we either have $\widetilde{R} = R$, which is thus reversible, or we get a $\boldsymbol{p}$-balanced pair with $\widetilde{R} \neq R$, where each is the time reversal of the other. Clearly, $R + \widetilde{R}$ is $\boldsymbol{p}$-reversible, as $\widetilde{.}$ is an involution. Then, if $[\widetilde{R}, R] = \mathbb{0}$, we get

$$M^2 \, = \, \mathrm{e}^{R}\mathrm{e}^{\widetilde{R}} \, = \, \mathrm{e}^{R+\widetilde{R}},$$

and Fact 2.7 implies that M^2 is a reversible Markov matrix with positive spectrum. Now, M^2 has a unique square root with positive eigenvalues by [19, Cor. 6], say M'. If M has positive spectrum itself, we get $M = M'$; otherwise, M and M' are different square roots of M^2. One consequence is the following.

Lemma 3.17. *Let $M \in \mathcal{M}_d$ be reversible, and embeddable via a $\boldsymbol{p}$-balanced pair of Markov generators, R and $\widetilde{R}$, with $[\widetilde{R}, R] = \mathbb{0}$ and $R \neq \widetilde{R}$. Then, $R + \widetilde{R}$ is a $\boldsymbol{p}$-reversible Markov generator and $M^2 = \mathrm{e}^{2R} = \mathrm{e}^{2\widetilde{R}}$ is also reversibly embedded as $M^2 = \exp(R + \widetilde{R})$. Further, M^2 has positive spectrum and possesses a unique Markov square root with positive spectrum. This square root is reversibly embedded via the Markov generator $\frac{1}{2}(R + \widetilde{R})$.* □

Note that this result also applies if we start from a Markov matrix with negative eigenvalues, which can be embeddable when each negative eigenvalue appears with even multiplicity.

Clearly, $d = 3$ is the smallest dimension where this can occur, and the embedding for such matrices was only settled relatively recently in [7]; see [3, Sec. 3.1] for a simplified derivation.

Example 3.18. Set $\epsilon = \mathrm{e}^{-\pi\sqrt{3}} \approx 0.163034$ and consider the symmetric Markov matrix

$$M \, = \, \frac{1}{3}\begin{pmatrix} 1-2\epsilon & 1+\epsilon & 1+\epsilon \\ 1+\epsilon & 1-2\epsilon & 1+\epsilon \\ 1+\epsilon & 1+\epsilon & 1-2\epsilon \end{pmatrix},$$

which is reversible with $\boldsymbol{p} = \frac{1}{3}(1,1,1)$, hence $D_{\boldsymbol{p}} = \frac{1}{3}\mathbb{1}$. Here, M cannot be reversibly embeddable, because its spectrum is $\sigma(M) = \{1, -\epsilon, -\epsilon\}$, so it lies outside the case covered by Corollary 3.14. Still, it is embeddable as $M = \mathrm{e}^{Q_+} = \mathrm{e}^{Q_-}$ with the two Markov generators

$$Q_+ \, = \, \frac{2\pi}{\sqrt{3}}\begin{pmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \\ 1 & 0 & -1 \end{pmatrix} \quad \text{and} \quad Q_- \, = \, \frac{2\pi}{\sqrt{3}}\begin{pmatrix} -1 & 0 & 1 \\ 1 & -1 & 0 \\ 0 & 1 & -1 \end{pmatrix},$$

see [1, 3] and references therein for details. An independent approach is given in Appendix A. Clearly, $Q_+ = Q_-^{\mathsf{T}}$, so they form a $\boldsymbol{p}$-balanced pair, and it is easy to check that they commute, $[Q_+, Q_-] = \mathbb{0}$, which means that Lemma 3.17 applies.

Now, $Q_+ + Q_-$ is a $\boldsymbol{p}$-reversible Markov generator of equal-input type (see Example 3.3), and we get $M^2 = \exp(Q_+ + Q_-)$ with spectrum $\{1, \epsilon^2, \epsilon^2\}$. The unique square root of M^2 with positive spectrum $\{1, \epsilon, \epsilon\}$ is the matrix M' that emerges from M by changing ϵ to $-\epsilon$. Note that $M^2 = \mathrm{e}^{2Q_+} = \mathrm{e}^{2Q_-}$ is also reversibly embedded via $Q_+ + Q_-$, so we have an example with positive spectrum and three embeddings.

More generally, the equal-input matrices from [3, Prop. 3.8] have a double negative eigenvalue and are reversible. They are also embeddable, at least with one $\boldsymbol{p}$-balanced pair $Q_\pm$ of Markov generators, though there may be several pairs, as derived in [3] and in Appendix A. One can check that they always satisfy $[Q_+, Q_-] = \mathbb{0}$, so we are in the same situation as above. As was derived in [3, Sec. 3.1], this is the only situation with a double negative eigenvalue that can occur for $d = 3$. ◊

To continue with multiple negative eigenvalues, let us consider $B = -\mathbb{1} \in \mathrm{SL}(2, \mathbb{R})$, which is reversible for $\boldsymbol{p} = \bigl(\frac{1}{2}, \frac{1}{2}\bigr)$ via $\widetilde{B} = B^{\mathsf{T}} = B$. The matrix B has multiple real logarithms by Fact 2.5, which we now determine completely. Clearly, if $B = \mathrm{e}^R$, the spectrum of $R \in \mathrm{Mat}(2, \mathbb{R})$ must be $\sigma(R) = \{\pm(2m+1)\pi\mathrm{i}\}$ for some arbitrary, but fixed $m \in \mathbb{Z}$. We may thus write $R = (2m+1)I$, where I is a real matrix with $\sigma(I) = \{\pm\mathrm{i}\}$ and $I^2 = -\mathbb{1}$. In particular, $\det(I) = 1$ and $I \in \mathrm{SL}(2, \mathbb{R})$. Also, one gets

$$\exp\bigl((2m+1)\pi I\bigr) \, = \, \cos\bigl((2m+1)\pi\bigr)\mathbb{1} + \sin\bigl((2m+1)\pi\bigr) I \, = \, -\mathbb{1},$$

and we have reduced our problem to finding all real square roots of $-\mathbb{1}$.

Set $I = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ and consider $I^2 = -\mathbb{1}$. As one can easily verify, $b = c = 0$ does not lead to any real solution, so we need $a + d = 0$ together with $bc = -(a^2 + 1) < 0$, which gives

$$I = \begin{pmatrix} a & b \\ -\frac{a^2+1}{b} & -a \end{pmatrix}. \tag{7}$$

Each such I together with I^{T} gives a $\boldsymbol{p}$-balanced pair, with $I \neq I^{\mathsf{T}}$. Moreover, one has $[I, I^{\mathsf{T}}] = \mathbb{0}$ if and only if $a = 0$ together with $b = \pm 1$, in which case one has $I^{\mathsf{T}} = I^3$. Put together, we have the following.

Fact 3.19. *The matrix $B = -\mathbb{1} \in \mathrm{SL}(2, \mathbb{R})$ is $\boldsymbol{p}$-reversible for $\boldsymbol{p} = \left(\frac{1}{2}, \frac{1}{2}\right)$, but has no reversible real logarithm. Instead, its real logarithms come in $\boldsymbol{p}$-balanced pairs $\{R, R^{\mathsf{T}}\}$ with $R = (2m + 1)\pi I$ for some $m \in \mathbb{Z}$ and I as in (7), with $a, b \in \mathbb{R}$ and $b \neq 0$.*

They commute if and only if $a = 0$ and $b = \pm 1$. □

Now, we need to see whether the analogous situation occurs among Markov matrices. The non-commutative case cannot happen for $d = 3$, as discussed in Example 3.18. Fortunately, it suffices to search within the class of doubly stochastic matrices from Example 3.7, which allows an approach based on permutation symmetries. Since we need some cyclic structure in the rate matrix (to obtain complex eigenvalues), but also non-commutativity, we replace the cyclic group C_3, which underlies our previous example as detailed in Appendix A, by the dihedral group

$$D_4 = \langle a, b : a^4 = b^2 = (ab)^2 = e \rangle,$$

where e denotes the neutral element. Concretely, we use the permutations $a = (1234)$ and $b = (12)(34)$. Then, starting from the standard 4D permutation representation, we consider the algebra generated by these matrices, and make an ansatz for our generator that is mildly quadratic in the generating elements, namely $Q = \alpha(a - e) + \beta(b - e) + \gamma(a^2 - e)$, and determine the parameters such that the spectrum becomes $\sigma(Q) = \{0, -2\sqrt{3}\pi, \pm\pi\mathrm{i} - 2\sqrt{3}\pi\}$. This means $\alpha = \frac{2\pi}{\sqrt{3}}$, $\beta = \frac{\pi}{\sqrt{3}}$ and $\gamma = \frac{3\pi}{2\sqrt{3}}$, and gives the following.

Example 3.20. Consider the Markov generator

$$Q = \frac{\pi}{2\sqrt{3}} \begin{pmatrix} -9 & 6 & 3 & 0 \\ 2 & -9 & 4 & 3 \\ 3 & 0 & -9 & 6 \\ 4 & 3 & 2 & -9 \end{pmatrix},$$

which is doubly stochastic, but not symmetric, so $\widetilde{Q} = Q^{\mathsf{T}} \neq Q$. One can check that

$$[Q, Q^{\mathsf{T}}] = \frac{4\pi^2}{3} \begin{pmatrix} 1 & 0 & -1 & 0 \\ 0 & -1 & 0 & 1 \\ -1 & 0 & 1 & 0 \\ 0 & 1 & 0 & -1 \end{pmatrix} \neq \mathbb{0},$$

which differs from the situation in Example 3.18. On the other hand, one can compute the matrix exponentials of Q and Q^{T}, which gives

$$M \,=\, \mathrm{e}^{Q} \,=\, \mathrm{e}^{Q^{\mathsf{T}}} \,=\, \frac{1}{4} \begin{pmatrix} 1-\varepsilon & 1-\varepsilon & 1+3\varepsilon & 1-\varepsilon \\ 1-\varepsilon & 1-\varepsilon & 1-\varepsilon & 1+3\varepsilon \\ 1+3\varepsilon & 1-\varepsilon & 1-\varepsilon & 1-\varepsilon \\ 1-\varepsilon & 1+3\varepsilon & 1-\varepsilon & 1-\varepsilon \end{pmatrix}$$

with $\varepsilon = \mathrm{e}^{-2\sqrt{3}\pi} \approx 1.87785 \cdot 10^{-5}$, where $\varepsilon = \epsilon^2$ with the constant ϵ from Example 3.18 and $\det(M) = \varepsilon^3 = \mathrm{e}^{-6\sqrt{3}\pi} \approx 6.62194 \cdot 10^{-15}$. Here, M is symmetric, hence reversible for $\boldsymbol{p} = \frac{1}{4}\mathbf{1}$, with spectrum $\sigma(M) = \{1, \varepsilon, -\varepsilon, -\varepsilon\}$. We thus have an example of a reversible Markov matrix that is embeddable via a $\boldsymbol{p}$-balanced pair of non-commuting Markov generators.

Now, M^2 has positive spectrum, and reads

$$M^2 \,=\, \frac{1}{4} \begin{pmatrix} 1+3\varepsilon^2 & 1-\varepsilon^2 & 1-\varepsilon^2 & 1-\varepsilon^2 \\ 1-\varepsilon^2 & 1+3\varepsilon^2 & 1-\varepsilon^2 & 1-\varepsilon^2 \\ 1-\varepsilon^2 & 1-\varepsilon^2 & 1+3\varepsilon^2 & 1-\varepsilon^2 \\ 1-\varepsilon^2 & 1-\varepsilon^2 & 1-\varepsilon^2 & 1+3\varepsilon^2 \end{pmatrix},$$

which is not only embeddable via $2Q$ and $2Q^{\mathsf{T}}$, so

$$M^2 \,=\, \mathrm{e}^{2Q} \,=\, \mathrm{e}^{2Q^{\mathsf{T}}} \,=\, \mathrm{e}^{Q}\mathrm{e}^{Q^{\mathsf{T}}} \,=\, \mathrm{e}^{Q^{\mathsf{T}}}\mathrm{e}^{Q} \,\neq\, \mathrm{e}^{Q+Q^{\mathsf{T}}},$$

but also via the principal matrix logarithm of M^2, which reads

$$L \,=\, \sqrt{3}\pi \begin{pmatrix} -3 & 1 & 1 & 1 \\ 1 & -3 & 1 & 1 \\ 1 & 1 & -3 & 1 \\ 1 & 1 & 1 & -3 \end{pmatrix} \,\neq\, Q + Q^{\mathsf{T}}.$$

Note that L is a constant-input matrix, which is a special case of the equal-input matrices considered in Example 3.3. So, we also have an example of a symmetric (and hence reversible) Markov matrix with positive spectrum that has three embeddings, one with a reversible rate matrix, and two other ones via a (non-commuting) $\boldsymbol{p}$-balanced pair. $\Diamond$

As our examples show, a reversible Markov matrix M can have non-reversible embeddings as well as multiple embeddings, and the generators of a balanced pair may or may not commute. This has to be determined on an individual basis, where multiple solutions can only occur for sufficiently small values of the determinant of M.

3.5. **Outlook.** Many practically occurring Markov matrices are reversible, and if an estimated one fails to be reversible, one can determine the nearest reversible one, at least numerically [25]. If a specific family of reversible matrices is investigated, additional tools might simplify its embeddability question, as we saw for the equal-input matrices [2]. There are certainly other important matrix classes that could and should be analysed along these lines.

However, what we have discussed above can only be the first step, as it considers the time-homogeneous case. More generally, it is of interest whether any given reversible Markov matrix appears in a continuous-time Markov flow, as one obtains from the initial value problem $\dot{X} = XQ$ with $X(0) = \mathbb{1}$ with a time-dependent family $Q = Q(t)$ of Markov generators. As long as $[Q(t), Q(s)] = \mathbb{0}$ holds for all $t, s \geqslant 0$, no new cases emerge. But when they fail to commute, an entirely new chapter is opened; see [18, 4] and references therein for a more detailed discussion. This extension certainly deserves further investigations in the future.

Appendix A. Poisson process approach to Example 3.18

Consider the matrix

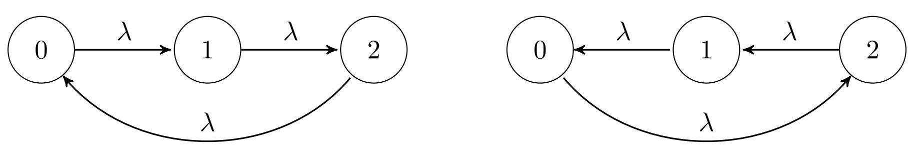
$$M(\delta) := \frac{1}{3}\begin{pmatrix} 1-2\delta & 1+\delta & 1+\delta \\ 1+\delta & 1-2\delta & 1+\delta \\ 1+\delta & 1+\delta & 1-2\delta \end{pmatrix},$$

for $0 < \delta \leqslant \frac{1}{2}$, where $M(\delta)$ is Markov, and the Markov generators

$$Q_{+}(\lambda) = \lambda \begin{pmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \\ 1 & 0 & -1 \end{pmatrix} \quad \text{and} \quad Q_{-}(\lambda) = \lambda \begin{pmatrix} -1 & 0 & 1 \\ 1 & -1 & 0 \\ 0 & 1 & -1 \end{pmatrix}$$

with $\lambda > 0$. We know from Example 3.18 that $\lambda = 2\pi/\sqrt{3}$ gives $\mathrm{e}^{Q_{+}(\lambda)} = \mathrm{e}^{Q_{-}(\lambda)} = M(\epsilon)$ with $\epsilon = \mathrm{e}^{-\pi\sqrt{3}}$. More generally, it follows from [3] that $M(\delta)$ is embeddable for $0 < \delta \leqslant \epsilon$, but not embeddable for $\delta > \epsilon$. The corresponding continuous-time processes $\big(X_t^{+}\big)_{t\geqslant 0}$ and $\big(X_t^{-}\big)_{t\geqslant 0}$ possess the transition graphs

(8)



and have an interpretation as clockwise or anti-clockwise cycles.

For the further derivation, we need a split of the power series of the exponential function as $\mathrm{e}^{x} = f_0(x) + f_1(x) + f_3(x)$ with

$$f_\ell(x) = \sum_{m=0}^{\infty} \frac{x^{3m+\ell}}{(3m+\ell)!} \qquad \text{for} \quad \ell \in \{0,1,2\}. \tag{9}$$

These functions are well defined and given as follows.

Lemma A.1. *The power series of the functions f_ℓ from Eq. (9) define entire functions. They are given by*

$$\begin{aligned} f_0(x) &= \frac{1}{3}\Big(\mathrm{e}^{x} + 2\mathrm{e}^{-x/2}\cos\frac{\sqrt{3}x}{2}\Big), \\ f_1(x) &= \frac{1}{3}\Big(\mathrm{e}^{x} - \mathrm{e}^{-x/2}\Big(\cos\frac{\sqrt{3}x}{2} + \sqrt{3}\sin\frac{\sqrt{3}x}{2}\Big)\Big), \\ f_2(x) &= \frac{1}{3}\Big(\mathrm{e}^{x} - \mathrm{e}^{-x/2}\Big(\cos\frac{\sqrt{3}x}{2} - \sqrt{3}\sin\frac{\sqrt{3}x}{2}\Big)\Big). \end{aligned}$$

Proof. The claim on analyticity is standard and justified by the absolute convergence of all series involved on all of $\mathbb{C}$. To compute them, define $\omega = \mathrm{e}^{2\pi \mathrm{i}/3} = -\frac{1}{2} + \frac{\mathrm{i}}{2}\sqrt{3}$, where one has $\overline{\omega} = \omega^2$ together with $1 + \omega + \overline{\omega} = 0$. Now, observe that, for $k, \ell \in \{0,1,2\}$, one has

$$f_\ell(\omega^k x) \, = \, \omega^{k\ell} f_\ell(x),$$

hence $\mathrm{e}^{\omega^k x} = f_0(x) + \omega^k f_1(x) + \overline{\omega}^k f_2(x)$. This allows the following computations,

$$\begin{aligned}
f_0(x) \, &= \, \frac{1}{3}\big(\mathrm{e}^x + \mathrm{e}^{\omega x} + \mathrm{e}^{\overline{\omega} x}\big) \, = \, \frac{1}{3}\Big(\mathrm{e}^x + \mathrm{e}^{-x/2}\big(\mathrm{e}^{\mathrm{i}\sqrt{3}x/2} + \mathrm{e}^{-\mathrm{i}\sqrt{3}x/2}\big)\Big), \\
f_1(x) \, &= \, \frac{1}{3}\big(\mathrm{e}^x + \overline{\omega}\mathrm{e}^{\omega x} + \omega\mathrm{e}^{\overline{\omega} x}\big) \, = \, \frac{1}{3}\Big(\mathrm{e}^x - \frac{1}{2}\big(\mathrm{e}^{\omega x} + \mathrm{e}^{\overline{\omega} x}\big) - \frac{\mathrm{i}}{2}\sqrt{3}\big(\mathrm{e}^{\omega x} - \mathrm{e}^{\overline{\omega} x}\big)\Big), \\
f_2(x) \, &= \, \frac{1}{3}\big(\mathrm{e}^x + \omega\mathrm{e}^{\omega x} + \overline{\omega}\mathrm{e}^{\overline{\omega} x}\big) \, = \, \frac{1}{3}\Big(\mathrm{e}^x - \frac{1}{2}\big(\mathrm{e}^{\omega x} + \mathrm{e}^{\overline{\omega} x}\big) + \frac{\mathrm{i}}{2}\sqrt{3}\big(\mathrm{e}^{\omega x} - \mathrm{e}^{\overline{\omega} x}\big)\Big),
\end{aligned}$$

from which the stated expressions follow via Euler's identity, $\mathrm{e}^{\mathrm{i}y} = \cos(y) + \mathrm{i}\sin(y)$. □

We note in passing that, using trigonometric identities, the expressions for f_1 and f_2 could be simplified to

$$\cos\frac{\sqrt{3}x}{2} \pm \sqrt{3}\sin\frac{\sqrt{3}x}{2} \, = \, \sin\Big(\frac{\pi}{6} \mp \frac{\sqrt{3}x}{2}\Big),$$

but this form would be less suitable for the coming computations.

Now, let $(S_t)_{t\geqslant 0}$ be an $\mathbb{N}_0$-valued Poisson process with parameter (or intensity) λ, which means that S_t is Poisson distributed with $\mathbb{P}(S_t = k) = \frac{(\lambda t)^k}{k!}\,\mathrm{e}^{-\lambda t}$. Then, the distributions of $X_t^{\pm}$ agree with those of $\pm S_t$ mod 3, and we get

$$\begin{aligned}
p_0(\lambda) \, &:= \, \mathbb{P}\big(X_1^+ = 0\big) \, = \, \mathbb{P}\big(X_1^- = 0\big) \, = \, \mathbb{P}\big(S_1 \in 3\mathbb{N}_0\big) \\
&= \, \mathrm{e}^{-\lambda}\sum_{m=0}^{\infty}\frac{\lambda^{3m}}{(3m)!} \, = \, \mathrm{e}^{-\lambda} f_0(\lambda) \, = \, \frac{1}{3}\Big(1 + 2\mathrm{e}^{-3\lambda/2}\cos\frac{\sqrt{3}\lambda}{2}\Big).
\end{aligned}$$

Analogously, we also get

$$\begin{aligned}
p_1(\lambda) \, &= \, \mathbb{P}\big(X_1^+ = 1\big) \, = \, \mathbb{P}\big(X_1^- = 2\big) \, = \, \mathbb{P}\big(S_1 \in 3\mathbb{N}_0 + 1\big) \\
&= \, \mathrm{e}^{-\lambda} f_1(\lambda) \, = \, \frac{1}{3}\Big(1 - \mathrm{e}^{-3\lambda/2}\Big(\cos\frac{\sqrt{3}x}{2} + \sqrt{3}\sin\frac{\sqrt{3}x}{2}\Big)\Big)
\end{aligned}$$

and

$$\begin{aligned}
p_2(\lambda) \, &= \, \mathbb{P}\big(X_1^+ = 2\big) \, = \, \mathbb{P}\big(X_1^- = 1\big) \, = \, \mathbb{P}\big(S_1 \in 3\mathbb{N}_0 + 2\big) \\
&= \, \mathrm{e}^{-\lambda} f_2(\lambda) \, = \, \frac{1}{3}\Big(1 - \mathrm{e}^{-3\lambda/2}\Big(\cos\frac{\sqrt{3}x}{2} - \sqrt{3}\sin\frac{\sqrt{3}x}{2}\Big)\Big).
\end{aligned}$$

Inserting $\lambda = 2\pi/\sqrt{3}$ gives $p_0(\lambda) = \frac{1}{3}(1 - 2\epsilon)$ and $p_1(\lambda) = p_2(\lambda) = \frac{1}{3}(1+\epsilon)$, which provides an independent derivation of the claim in Example 3.18.

In fact, $\lambda = 2\pi/\sqrt{3}$ is the *smallest* $\lambda > 0$ with $p_1(\lambda) = p_2(\lambda)$ and thus the relation $\mathbb{P}\big(X_1^+ = \ell\big) = \mathbb{P}\big(X_1^- = \ell\big)$ for $\ell \in \{0,1,2\}$, which is equivalent to $\mathrm{e}^{Q^+(\lambda)}$ being a symmetric Markov matrix. This symmetry precisely appears for $\lambda_k = 2(2k+1)\pi/\sqrt{3}$ with $k \in \mathbb{N}_0$, then giving $\delta_k = \mathrm{e}^{-(2k+1)\pi\sqrt{3}}$, with $\delta_0 = \epsilon$. This sequence corresponds to the one derived in [3] via a minimisation procedure, and defines a decreasing sequence with $\delta_k > \delta_{k+1}$ for all

$k \in \mathbb{N}_0$ and 0 as its limit point. Recall from [3, Rem. 3.9] that $M(\delta)$ is embeddable for all $\delta \leqslant \delta_0$, with Markov generators that derive from the $Q^{\pm}(\delta_0)$ via a simple deformation. Then, when one reaches δ_k, another new pair of commuting generators emerges, and this process continues. Consequently, the number of embeddings of $M(\delta)$ increases without bound as $\delta \searrow 0$ or, equivalently, as the traces of the constant-input matrices $M(\delta)$ approach 1.

Acknowledgements

This work was supported by the German Research Foundation (DFG), within the CRC 1283/2 (2021 - 317210226) at Bielefeld University.

References

[1] M. Baake and J. Sumner, Notes on Markov embedding, *Lin. Alg. Appl.* **594** (2020) 262–299; `arXiv:1903.08736`.

[2] M. Baake and J. Sumner, On equal-input and monotone Markov matrices, *J. Appl. Probab.* **54** (2022) 460–492; `arXiv:2007.11433`.

[3] M. Baake and J. Sumner, Embedding of Markov matrices for $d \leqslant 4$, *J. Math. Biol.* **89** (2024) 23:1–45; `arXiv:2311.02596`.

[4] M. Baake and J. Sumner, On the algebra of equal-input matrices in time-inhomogeneous Markov flows, *Docum. Math.* **30** (2025) 455–474; `arXiv:2404.11222`.

[5] A. Berman and R.J. Plemmons, *Nonnegative Matrices in the Mathematical Sciences*, corr. reprint, SIAM, Philadelphia, PA (1994).

[6] P.H. Brill, C.H. Cheung, M. Hlynka and Q. Jiang, Reversibility checking for Markov chains, *Commun. Stoch. Anal.* **12** (2018) 129–135; `arXiv:1806.10154`.

[7] P. Carette, Characterizations of embeddable 3×3 stochastic matrices with a negative eigenvalue, *New York J. Math.* **1** (1995) 120–129.

[8] J. Chen, A solution to the reversible embedding problem for finite Markov chains, *Stat. Probab. Lett.* **116** (2016) 122–130; `arXiv:1605.03502`.

[9] L. Cooper and J. Sumner, Uniformization stable Markov models and their Jordan algebraic structure, *SIAM J. Matrix Anal. Appl.* **44** (2023) 1822–1851; `arXiv:2105.03558`.

[10] W.J. Culver, On the existence and uniqueness of the real logarithm of a matrix, *Proc. Amer. Math. Soc.* **17** (1966) 1146–1151.

[11] E.B. Davies, Embeddable Markov matrices, *Electr. J. Probab.* **15** (2010) 1474–1486 (article 47); `arXiv:1001.1693`.

[12] M.O. Dayhoff, R.V. Eck, and C.M. Park, A model of evolutionary change in proteins, in: *Atlas of Protein Sequence and Structure. Vol. 5*, ed. M.O. Dayhoff, National Biomedical Research Foundation, Washington, D.C. (1972), pp. 89–99.

[13] M.O. Dayhoff, R.M. Schwartz, and B.C. Orcutt, A model of evolutionary change in proteins, in: *Atlas of Protein Sequence and Structure. Vol. 5, Suppl. 3*, ed. M.O. Dayhoff, National Biomedical Research Foundation, Washington, D.C. (1978), pp. 345–352.

[14] G. Elfving, Zur Theorie der Markoffschen Ketten, *Acta Soc. Sci. Fennicae* **A2** (1937) 1–17.

[15] W.J. Ewens and G.R. Grant, *Statistical Methods in Bioinformatics*, 2nd ed., Springer, New York (2005).

[16] N.J. Higham, *Functions of Matrices: Theory and Computation*, SIAM, Philadelphia, PA (2008).
[17] R.A. Horn and C.R. Johnson, *Matrix Analysis*, 2nd ed., Cambridge University Press, Cambridge (2013).
[18] S. Johansen, The Bang-Bang problem for stochastic matrices, *Z. Wahrscheinlichkeitsth. Verw. Geb.* **26** (1973) 191–195.
[19] C.R. Johnson, K. Okubo and R. Reams, Uniqueness of matrix square roots and an application, *Lin. Alg. Appl.* **323** (2001) 51–60.
[20] F.P. Kelly, *Reversibility and Stochastic Networks*, rev. ed., Cambridge University Press, Cambridge (2011).
[21] J.F.C. Kingman, The imbedding problem for finite Markov chains, *Z. Wahrscheinlichkeitsth. verw. Geb.* **1** (1962) 14–24.
[22] C. Kosiol and N. Goldman, Different versions of the Dayhoff rate matrix, *Mol. Biol. Evol.* **22** (2005) 19–199.
[23] P. Lancaster and M. Tismenetsky, *The Theory of Matrices*, 2nd ed., Academic Press, Orlando, FL (1985).
[24] D.A. Levin and Y. Peres, *Markov Chains and Mixing Times*, 2nd ed., AMS, Providence, RI (2017).
[25] A.J.N. Nielsen and M. Weber, Computing the nearest reversible Markov chain, *Numer. Lin. Alg. Appl.* **22** (2025) 483–499.
[26] J.R. Norris, *Markov Chains*, Cambridge University Press, Cambridge (1997), reprint (2005).
[27] Z. Yang, *Molecular Evolution: A Statistical Approach*, Oxford University Press, Oxford (2014).

Technische Fakultät, Universität Bielefeld, Postfach 100131, 33501 Bielefeld, Germany
Email address: ebaake@techfak.uni-bielefeld.de

Fakultät für Mathematik, Universität Bielefeld, Postfach 100131, 33501 Bielefeld, Germany
Email address: mbaake@math.uni-bielefeld.de

School of Natural Sciences, Discipline of Mathematics, University of Tasmania,
PO Box 807, Sandy Bay, TAS 7006, Australia
Email address: Jeremy.Sumner@utas.edu.au