East Asian Journal on Applied Mathematics
doi: 10.4208/eajam.230414.311214a

Vol. **5**, No. 1, pp. 75-84
February 2015

Computing Solutions of the Yang-Baxter-like Matrix Equation for Diagonalisable Matrices

J. Ding[1,*] and Noah H. Rhee[2]

[1] *Department of Mathematics, The University of Southern Mississippi, USA; School of Mathematical Sciences, Yangzhou University, China.*
[2] *Department of Mathematics, University of Missouri - Kansas City, Kansas City, MO 64110-2499, USA.*

Received 23 April 2014; Accepted (in revised version) 31 December 2014

Abstract. The Yang-Baxter-like matrix equation $AXA = XAX$ is reconsidered, where A is any complex square matrix. A collection of spectral solutions for the unknown square matrix X were previously found. When A is diagonalisable, by applying the mean ergodic theorem we propose numerical methods to calculate those solutions.

AMS subject classifications: 15A24, 47A35
Key words: Matrix equation, mean ergodic theorem, diagonalisable matrix.

1. Introduction

In this article, we reconsider the matrix equation

$$AXA = XAX \,, \tag{1.1}$$

where both A and X are constant complex square matrices of the same size $(n \times n)$. Eq. (1.1) has been called *Yang-Baxter-like*, after the Yang-Baxter equation for two-dimensional integrable models in statistical mechanics [1, 13]. The *parameter-dependent* equation

$$A(u)B(u+v)A(v) = B(u)A(u+v)B(v), \tag{1.2}$$

where A and B are rational matrix functions of their arguments, obviously reduces to Eq. (1.1) when A and B are constant matrices. The size of A and B in applied problems is typically not large — e.g. the matrices that were considered in Refs. [1, 13] are only 4×4, namely

$$\begin{bmatrix} 1+u & 0 & 0 & 0 \\ 0 & u & 1 & 0 \\ 0 & 1 & u & 0 \\ 0 & 0 & 0 & 1+u \end{bmatrix} \text{ and } \begin{bmatrix} a(u) & 0 & 0 & d(u) \\ 0 & b(u) & c(u) & 0 \\ 0 & c(u) & b(u) & 0 \\ d(u) & 0 & 0 & a(u) \end{bmatrix},$$

*Corresponding author. *Email addresses:* `Jiu.Ding@usm.edu` (J. Ding), `RheeN@umkc.edu` (N. H. Rhee)

respectively. The Yang-Baxter equation has been applied for decades by physicists and mathematicians in many areas such as group theory, braid groups and knot theory [7, 10, 14]. In contrast, the Yang-Baxter-like matrix equation (1.1) has not attracted much attention from matrix theorists, perhaps due to its nonlinearity or their lack of background in knot theory and braid groups (or quantum mechanics). Even for matrices of small size, it has been difficult to find all of the solutions. Some solutions have been obtained, but mostly via direct computation from the polynomial equations corresponding to multiplying out the matrix equation in specific cases. There is still no systematic approach to the existence and computation of solutions of Eq. (1.1) in general, but the numerical method proposed here yields spectral solutions for any matrix size.

The concept of braids was introduced in 1925 by Emil Artin, and a braid group with n strands B_n is a group where the multiplication of a braid s to another braid t corresponds to gluing s onto the bottom of t. It follows that every braid s has a unique inverse braid t ($st = ts = e$, where e is the unit braid such that strands are preserved. It is fundamental that there are elementary braids $s_1, s_2, \cdots, s_{n-1}$, where the i-th strand of s_i goes over to the right of the $(i+1)$-th strand of s_i for each i, to generate the whole braid group B_n. Furthermore, these braids satisfy the Yang-Baxter-like relation

$$s_{i+1}s_is_{i+1} = s_is_{i+1}s_i \tag{1.3}$$

for each $i = 1, \cdots, n-2$, and also $s_is_j = s_js_i$ for any i and j where $|i-j| > 1$. This relation (1.3) evidently has the same form as the matrix equations (1.1) and (1.2). The matrix equations may also be viewed as word equations — cf. Ref. [8] and references therein for more detail. Given a uniquely divisible group G where every element has an n-th root for any positive integer n, a word equation has the form $W(X,A) = B$, where W is a finite word consisting of the unknown element X and the known element A of G and B is a given element in G. Under certain conditions, a solution can be obtained in terms of radicals. Eq. (1.1) written as $W(X,A) = AXA - XAX = 0$ is a word equation. However, since the class of all square matrices is not a group under matrix multiplication, the general setting of word equations does not apply unless we restrict our consideration to invertible matrices of the same size say. (Even then, not every invertible matrix has a root, as is often required in solution techniques for solving word equations.) Thus computing solutions of the Yang-Baxter-like equation (1.1) in practice is generally a challenging problem, and different techniques should be employed.

In first considering the existence and computation of solutions to the Yang-Baxter-like matrix equation (1.1) of arbitrary size, we obtained some numerical solutions when A is a nonsingular quasi-stochastic matrix such that A^{-1} is stochastic [3]. Recently, we have proven some general existence results for an arbitrary square matrix A, by finding a collection of solutions of Eq. (1.1) in terms of spectral projections associated with all of its eigenvalues [4]. More solutions were found in Ref. [5] for some classes of matrix A with special Jordan canonical forms, based on a general result for commuting solutions, but there has not been any actual numerical computation of such spectral solutions. In this article, we show that the solutions found in Ref. [4] can be computed by means of the mean ergodic theorem if A is diagonalisable.

The set of all eigenvalues of A is called the *spectrum* of A, and usually denoted by $\sigma(A)$. The maximum of the absolute values of all eigenvalues of any matrix is called the *spectral radius*, usually denoted by $\rho(A)$. Any eigenvalue of A with absolute value equal to the spectral radius is called a *dominant eigenvalue*. An eigenvalue λ is said to be *semisimple* if its algebraic multiplicity and geometric multiplicity are equal and this common multiplicity is simply called the *multiplicity* of λ. A collection of solutions of Eq. (1.1) found in Ref. [4] is briefly reviewed in Section 2. In Section 3, we discuss how to find these solutions using the mean ergodic theorem. Some numerical examples are then discussed in Section 4, and our conclusions are in Section 5.

2. A Collection of Solutions

Let A^H, $N(A)$ and $R(A)$ denote the conjugate transpose, the null space and the range of A, respectively. The index $\nu(\lambda)$ of a complex number λ with respect to the matrix A is the smallest nonnegative integer j such that $N((A-\lambda I)^{j+1}) = N((A-\lambda I)^j)$, where I is the identity matrix. Note that λ is an eigenvalue of A if and only if $\nu(\lambda) \geq 1$, and it is a semisimple eigenvalue if and only if $\nu(\lambda) = 1$. The space $N((A-\lambda I)^{\nu(\lambda)})$ is called the eigenspace of A corresponding to λ if $\nu(\lambda) = 1$ and the generalised eigenspace if $\nu(\lambda) \geq 2$. The following lemma is crucial in our present approach to find solutions to Eq. (1.1) — cf. Refs. [6,11] for its proof.

Lemma 2.1. *Let P_i be the spectral projection matrix onto $N((A-\lambda_i I)^{\nu(\lambda_i)})$ along $R((A-\lambda_i I)^{\nu(\lambda_i)})$ for each $\lambda_i \in \sigma(A)$. Then*

(a) *$P_i^2 = P_i$ and $P_i P_j = 0$ for $i \neq j$.*

(b) *$AP_i = P_i A$. If λ_i is semisimple, then $AP_i = P_i A = \lambda_i P_i$.*

(c) *There exist bases $\{x_1, \cdots, x_{d_i}\}$ of $N((A-\lambda_i I)^{\nu(\lambda_i)})$ and $\{y_1, \cdots, y_{d_i}\}$ of $N((A^H-\bar{\lambda}_i I)^{\nu(\lambda_i)})$ such that $x_j^H y_k = \delta_{jk}$ for $1 \leq j,k \leq d_i$, and $P_i = \sum_{j=1}^{d_i} x_j y_j^H$, where d_i is the dimension of $N((A-\lambda_i I)^{\nu(\lambda_i)})$.*

(d) *$\sum_{\lambda_i \in \sigma(A)} P_i = I$.*

The following theorem was proved in Ref. [4].

Theorem 2.1.

(a) *For each $\lambda_i \in \sigma(A)$ let P_i be the spectral projection matrix in Lemma* 2.1. *Then the matrix AP_i is a solution of* (1.1).

(b) *Let $\lambda_1, \cdots, \lambda_s$ denote all distinct eigenvalues of A. Then the sum of any number of matrices among $AP_1, \cdots, AP_s$ is a solution of* Eq. (1.1).

Remark 2.1. Since eigenfunctions are key elements in the physical problems related to the Yang-Baxter equation, such as the completely integrable quantum system, the spectral solutions of Eq. (1.1) are of practical value, and notably often those associated with peripheral eigenvalues — cf. Refs. [10,14] for various models of mathematical physics.

The solutions identified in the above theorem can be computed if the spectral projection matrices P_i corresponding to the eigenvalue λ_i of A are found. In general, P_i can be computed either by using Lagrange-Hermite interpolating polynomials or by finding bi-orthonormal bases for $N((A-\lambda_i I)^{\nu(\lambda_i)})$ and $N((A^H-\bar{\lambda}_i I)^{\nu(\lambda_i)})$ — cf. Ref. [4] for details. In particular, if A is diagonalisable such that all of its eigenvalues are semisimple and hence $\nu(\lambda_i) = 1$ for all i, then P_i can be computed either by using Lagrange interpolating polynomials (because all eigenvalues are semisimple) or by finding bi-orthonormal bases for $N((A-\lambda_i I))$ and $N((A^H-\bar{\lambda}_i I))$. We demonstrate here that the solutions $AP_i = \lambda_i P_i$ can be computed via the mean ergodic theorem for matrices, if A is diagonalisable. This avoids the necessity of finding all the other eigenvalues or the eigenvectors of the bi-orthonormal bases, greatly reducing the computational time.

3. Computing the Solutions by the Mean Ergodic Theorem

In this section, we introduce the mean ergodic theorem for matrices and show how it can be used to compute solutions of Eq. (1.1) when A is diagonalisable. First of all, we need the concept of the *Cesáro limit*.

Definition 3.1. For a given square matrix B, we form the average matrices

$$B_m = \frac{1}{m}\sum_{i=0}^{m-1} B^i$$

of all nonnegative powers of B. Then the Cesáro limit P of B is defined as

$$P = \lim_{m\to\infty} B_m\,, \tag{3.1}$$

if this limit exists. If the *Cesáro limit P* of B exists, B is said to be mean ergodic.

Various mean ergodic theorems are important ingredients of ergodic theory, originating from the so-called von Neumann's mean ergodic theorem for measure preserving transformations. Ref. [6] contains several representative ergodic theorems for nonnegative matrices and infinite dimensional positive operators. The following well-known result [2,11] on the mean ergodicity of a matrix is sufficient here.

Lemma 3.1. *The Cesáro limit P of B in* (3.1) *exists if and only if either $\rho(B) < 1$ or $\rho(B) = 1$ and all the dominant eigenvalues are semisimple. In addition, $P \neq 0$ if and only if $1 \in \sigma(B)$. Furthermore, when $1 \in \sigma(B)$, we have the following results.*

(a) *P is a projection ($P^2 = P$) onto $N(B-I)$ along $R(B-I)$.*

(b) *$BP = PB = P$.*

(c) *There exist bases $\{x_1,\cdots,x_k\}$ of $N(B-I)$ and $\{y_1,\cdots,y_k\}$ of $N(B^H-I)$ such that $x_i^H y_j = \delta_{ij}$ for $1 \le i,j \le k$, and $P = \sum_{i=1}^k x_i y_i^H$, where k is the multiplicity of the eigenvalue 1 of B.*

If $\rho(B) = 1$ and 1 is the only dominant eigenvalue of B and is also semisimple, then Lemma 3.1 can be strengthened as follows [11].

Lemma 3.2. *Suppose $\rho(B) = 1$. If 1 is the only dominant eigenvalue of B and is also semisimple, then all the conclusions of Lemma* 3.1 *are valid with*

$$P = \lim_{m\to\infty} B^m. \tag{3.2}$$

Suppose that A is diagonalisable such that all its eigenvalues are semisimple. Let λ be an eigenvalue of A. There are two cases.

- If λ is a dominant eigenvalue, we let $B = \lambda^{-1}A$. Then the eigenvalue λ of A is mapped to the eigenvalue 1 of B, which is a dominant eigenvalue of B. Depending on whether or not 1 is the only dominant eigenvalue of B, one can use either Eq. (3.2) or Eq. (3.1) for the computation of the corresponding P. Since the eigenvectors of a matrix are invariant under the scaling of the matrix, the projection matrix of P of Lemma 3.1 is the projection matrix P corresponding to λ in Lemma 2.1. So by Theorem 2.1 (a) λP is a solution to Eq. (1.1).
- If λ is not a dominant eigenvalue of A, we choose a non-eigenvalue number α such that λ is the nearest eigenvalue of A to α. Then $1/(\lambda-\alpha)$ is a dominant eigenvalue of the matrix $(A-\alpha I)^{-1}$. So if we let $B = (\lambda-\alpha)(A-\alpha I)^{-1}$, then B has 1 as its dominant eigenvalue, which was mapped from the eigenvalue λ of A. Again depending on whether 1 is the only dominant eigenvalue of B or not, one can use either Eq. (3.2) or Eq. (3.1) for the computation of the corresponding matrix P. Since the eigenvectors of a matrix are invariant under shifting, inverse operation and scaling of the matrix, again λP is a solution to Eq. (1.1).

Consequently, the computation of the solution is reduced to finding the projection matrix P in Lemma 3.1, which is the Cesáro limit of a proper diagonalisable matrix B with spectral radius 1. If 1 is the only dominant eigenvalue of B, from Lemma 3.2 the projection $P = \lim_{m\to\infty} B^m$ and P can be computed efficiently as follows:

$$\text{Let } P^{(0)} = B \text{ and } P^{(k)} = \left(P^{(k-1)}\right)^2 \text{ for } k = 1, 2, \cdots. \tag{3.3}$$

Clearly, $P = \lim_{k\to\infty} P^{(k)}$ and the rate of convergence is quadratic. If 1 is not the only dominant eigenvalue of B we can use (3.1), but it may require too much computation. However, if the dominant eigenvalues of B are the h-th roots of unity with $h \geq 2$, we have the following result.

Theorem 3.1. *Let B be a matrix with $\rho(B) = 1$. Further, assume that dominant eigenvalues of B are semisimple and the h-th roots of unity with $h \geq 2$. Let $B = B_0$ and $C = (I + B + B^2 + \cdots + B^{h-1})/h$, and iterate $B_k = B_{k-1}^2$ and $P^{(k)} = B_k C$ for $k = 1, 2, \cdots$. Then*

$$P = \lim_{k\to\infty} P^{(k)},$$

where P is the projection matrix in Lemma 3.1.

Proof. Let us write the dominant eigenvalues of B, the h-th roots of unity, by

$$1, \omega, \omega^2, \cdots, \omega^{h-1}, \text{ where } \omega = e^{2\pi i/h}.$$

Let $J = S^{-1}BS$ be a Jordan form of B. We may assume that $J = \mathrm{diag}(J_1, J_2)$, where J_1 is associated with dominant eigenvalues and J_2 with interior eigenvalues. If we let m_j be the algebraic multiplicity of ω^{j-1} for $j = 1, \cdots, h$, we can write $J_1 = \mathrm{diag}(I_{m_1}, \omega I_{m_2}, \cdots, \omega^{h-1} I_{m_h})$, where I_{m_j} is the $m_j \times m_j$ identity matrix for $j = 1, 2, \cdots, h$. Then the j-th Jordan block of C is

$$\frac{1}{h}\left(1 + \omega^{j-1} + \omega^{2(j-1)} + \cdots + \omega^{(h-1)(j-1)}\right) I_{m_j},$$

which is I_{m_1} if $j = 1$ but 0_{m_j} if $j = 2, \cdots, h$, where 0_{m_j} is the $m_j \times m_j$ zero matrix. Hence we have the Jordan blocks $\mathrm{diag}(I_{m_1}, 0_{m_2}, \cdots, 0_{m_h})$ of C corresponding to J_1, and still $\mathrm{diag}(I_{m_1}, 0_{m_2}, \cdots, 0_{m_h})$ of C for the Jordan blocks of $P^{(k)}$ corresponding to J_1. Note that the Jordan block form of B_k corresponding to J_2 goes to zero because J_2 is associated with the eigenvalues of B, which are less than 1 in absolute value. So the Jordan form of $P^{(k)}$ goes to the zero matrix except for the top $m_1 \times m_1$ block, which is I_{m_1}. It follows that $P = \lim_{k\to\infty} P^{(k)}$. ☐

Remark 3.1. Clearly, the rate of convergence of the algorithm in Theorem 3.1 is quadratic. Further, Theorem 3.1 remains valid if the dominant eigenvalues of B form a subset of the h-th roots of unity.

Remark 3.2. When the size of Eq. (1.1) is large, one eigenvalue λ of A is near a cluster point of the spectrum of A. In this case, a chosen α may be closer to another eigenvalue such that 1 is not a dominant eigenvalue of $B = (\lambda - \alpha)(A - \alpha I)^{-1}$, so the direct iteration of B may not converge. However, if this happens we can choose a new α with distance to λ half of the original one, and repeat the process until the iteration converges.

4. Numerical Examples

We now present three examples illustrating the computed solutions of Eq. (1.1), using the schemes of the previous section.

Example 4.1. Let

$$A = \begin{bmatrix} 1 & 0 & 0 & 3 \\ 3/2 & 1 & 3/2 & 0 \\ 4/3 & 2/3 & 2/3 & 4/3 \\ 0 & 0 & 2 & 2 \end{bmatrix}.$$

One can verify that A is diagonalisable and its eigenvalues are $\lambda_1 = 4$, $\lambda_2 = -0.2184$, $\lambda_3 = 0.4425 + 1.1536i$, and $\lambda_4 = 0.4425 - 1.1536i$. If we let $B = \lambda_1^{-1} A$, then using Eq. (3.3) we compute P_1, so according to Theorem 2.1 (a)

$$\lambda_1 P_1 = \begin{bmatrix} 0.6780 & 0.2712 & 1.2203 & 1.8305 \\ 0.6780 & 0.2712 & 1.2203 & 1.8305 \\ 0.6780 & 0.2712 & 1.2203 & 1.8305 \\ 0.6780 & 0.2712 & 1.2203 & 1.8305 \end{bmatrix}$$

is a solution of Eq, (1.1).

Now we find a solution corresponding to λ_2, which is a non-dominant eigenvalue of A. On choosing $\alpha = -0.2$, we let $B = (\lambda_2 - \alpha)(A - \alpha I)^{-1}$. Then as above we can compute P_2, and again from Theorem 2.1 (a)

$$\lambda_2 P_2 = \begin{bmatrix} 0.0901 & 0.1171 & -0.2141 & 0.0069 \\ -0.1608 & -0.2092 & 0.3823 & 0.0123 \\ 0.0406 & 0.0528 & -0.0964 & 0.0031 \\ -0.0366 & -0.0476 & 1.2203 & -0.0028 \end{bmatrix}$$

is a solution of Eq. (1.1). Similarly, we can find a solution $\lambda_3 P_3$ corresponding to λ_3 (for example using $\alpha = 0.4 + 1.2i$), which is

$$\begin{bmatrix} 0.1160+0.7779i & -0.1942+0.3846i & -0.5031-0.6576i & 0.5813-0.5049i \\ 0.4914-0.7880i & 0.4690-0.1970i & -0.0513+0.9763i & -0.9091+0.0087i \\ 0.3074-0.1071i & 0.1714+0.0493i & -0.2286+0.2553i & -0.2501-0.1975i \\ -0.3207-0.1000i & -0.1118-0.1461i & 0.3464-0.0713i & 0.0861+0.3174i \end{bmatrix}.$$

Of course, a solution $\lambda_4 P_4$ corresponding to λ_4 is the conjugate of $\lambda_3 P_3$. According to Theorem 2.1 (b), the sum of any number of solutions from $\{\lambda_1 P_1, \lambda_2 P_2, \lambda_3 P_3, \lambda_4 P_4\}$ is also a solution, so we have located $2^4 = 16$ solutions.

Example 4.2. Let

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 2 & 0 & 1 & 0 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 0 \end{bmatrix}.$$

One can verify that $\lambda_1 = 2$, $\lambda_2 = -2$, $\lambda_3 = 1$ and $\lambda_4 = -1$. Even though λ_1 is a dominant eigenvalue, the powers of $B = \lambda_1^{-1} A$ do not converge to P_1, because λ_2 is also a dominant eigenvalue of A. However, since the dominant eigenvalues are two square roots of unity, the algorithm in Theorem 3.1 applies. Thus we can compute P_1 efficiently, and

$$\lambda_1 P_1 = \begin{bmatrix} 0.3333 & 0.3333 & 0.3333 & 0.3333 \\ 0.6667 & 0.6667 & 0.6667 & 0.6667 \\ 0.6667 & 0.6667 & 0.6667 & 0.6667 \\ 0.3333 & 0.3333 & 0.3333 & 0.3333 \end{bmatrix}$$

is a solution of (1.1). Similarly, by letting $B = \lambda_2^{-1}A$ we can compute

$$\lambda_2 P_2 = \begin{bmatrix} -0.3333 & 0.3333 & -0.3333 & 0.3333 \\ 0.6667 & -0.6667 & 0.6667 & -0.6667 \\ -0.6667 & 0.6667 & -0.6667 & 0.6667 \\ 0.3333 & -0.3333 & 0.3333 & -0.3333 \end{bmatrix}.$$

Now let us find solutions corresponding to non-dominant eigenvalues. First, we consider a solution corresponding to λ_3. By choosing $\alpha = 0$, we form $B = \lambda_3 A^{-1}$. Then λ_3 is transformed into 1, which is a dominant eigenvalue of B, but λ_4 is also transformed into -1, another dominant eigenvalue of B. The situation is therefore exactly same as above, so again using the algorithm in Theorem 3.1 we obtain

$$\lambda_3 P_3 = \begin{bmatrix} 0.3333 & 0.1667 & -0.1667 & -0.3333 \\ 0.3333 & 0.1667 & -0.1667 & -0.3333 \\ -0.3333 & -0.1667 & 0.1667 & 0.3333 \\ -0.3333 & -0.1667 & 0.1667 & 0.3333 \end{bmatrix},$$

and similarly using $B = \lambda_4 A^{-1}$ (noting that A^{-1} is already available) we compute

$$\lambda_4 P_4 = \begin{bmatrix} -0.3333 & 0.1667 & 0.1667 & -0.3333 \\ 0.3333 & -0.1667 & -0.1667 & 0.3333 \\ 0.3333 & -0.1667 & -0.1667 & 0.3333 \\ -0.3333 & 0.1667 & 0.1667 & -0.3333 \end{bmatrix}.$$

From Theorem 2.1 (b), the sum of any number of solutions from $\{\lambda_1 P_1, \lambda_2 P_2, \lambda_3 P_3, \lambda_4 P_4\}$ is also a solution, so we have again located $2^4 = 16$ solutions.

Example 4.3. The class of matrices

$$A = \begin{bmatrix} a_1 & 0 & 0 & a_3 \\ 0 & a_2 & a_4 & 0 \\ 0 & a_5 & a_2 & 0 \\ a_6 & 0 & 0 & a_1 \end{bmatrix},$$

for finding all the eight-vertex type two-state solutions of the Yang-Baxter-like equation in a different format.

The eigenvalues of A are

$$\lambda_1 = a_1 + \sqrt{a_3 a_6}\,, \quad \lambda_2 = a_1 - \sqrt{a_3 a_6}\,, \quad \lambda_3 = a_2 + \sqrt{a_4 a_5}\,, \quad \lambda_4 = a_2 - \sqrt{a_4 a_5}\,.$$

If we let $a_1 = a_3 = a_4 = a_5 = a_6 = 1$ and $a_2 = 0$, then $\lambda_1 = 2, \lambda_2 = 0, \lambda_3 = 1$, and $\lambda_4 = -1$. From the eigenvalue distribution, it is clear that Eq. (3.3) may be used to compute all corresponding spectral solutions using an appropriate B. For $\lambda_2 = 0$, the corresponding spectral solution is the trivial zero solution. The other spectral solutions are

$$\lambda_1 P_1 = \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix},$$

$$\lambda_3 P_3 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0.5 & 0.5 & 0 \\ 0 & 0.5 & 0.5 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

and

$$\lambda_4 P_4 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & -0.5 & 0.5 & 0 \\ 0 & 0.5 & -0.5 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

From Theorem 2.1 (b), the sum of any number of solutions from $\{\lambda_1 P_1, \lambda_3 P_3, \lambda_4 P_4\}$ is also a solution, so once again we have located $2^3 = 8$ solutions.

5. Conclusion

Based on eigen-projections, a collection of solutions of the Yang-Baxter-like matrix equation (1.1) for a given matrix A were obtained in Ref. [4]. If the matrix A is diagonalisable, in this article we have proposed numerical methods for computing these solutions via the mean ergodic theorem. The convergence rates of the developed numerical methods for calculating the involved projections are quadratic. Another Yang-Baxter-like matrix equation is of the form

$$ABC = CBA \tag{5.1}$$

— cf. Refs. [10,12,14] for related physical applications. Eq. (5.1) is a generalisation of the commutability property $AB = BA$ of two matrices A and B, and an abstract linear algebra version of a relation in the physics literature — viz. In the context of an associative algebra U with unit element e, an invertible element R of the tensor product $U \otimes U$ has the format

$$\phi_{12}(R)\phi_{13}(R)\phi_{23}(R) = \phi_{23}(R)\phi_{13}(R)\phi_{12}(R)\,,$$

where $\phi_{12}, \phi_{13}, \phi_{23}$ are algebra morphisms from $U \otimes U$ to $U \otimes U \otimes U$, defined for all $u, v \in U$ by

$$\phi_{12}(u,v) = u \otimes v \otimes e, \quad \phi_{13}(u,v) = u \otimes e \otimes v\,, \quad \phi_{23}(u,v) = e \otimes u \otimes v\,.$$

A systematic investigation of Eq. (5.1) for its solutions and their efficient computation is envisaged in our future research on linear algebra with application to the physical sciences.

Acknowledgments

We thank two referees for their constructive suggestions.

References

[1] R.J. Baxter, *Partition function of the eight-vertex lattice model*, Ann. Phys. **70**, 193-228 (1972).
[2] J. Ding and N. Rhee, *On the equality of algebraic and geometric multiplicities of matrix eigenvalues*, Applied Math Letters **24**, 2211-2215 (2011).
[3] J. Ding and N. Rhee, *A nontrivial solution to a stochastic matrix equation*, East Asian J. Applied Math. **2**, 277-284 (2012).
[4] J. Ding and N. Rhee, *Spectral solutions of the Yang-Baxter-like matrix equation*, J. Math. Anal. Appl. **402**, 567-573 (2013).
[5] J. Ding, N. Rhee and C. Zhang, *Further solutions of a Yang-Baxter-like matrix equation*, East Asian J. Applied Math. **3**, 352-362 (2013).
[6] J. Ding and A. Zhou, *Nonnegative Matrices, Positive Operators, and Applications*, World Scientific (2009).
[7] F. Felix, *Nonlinear Equations, Quantum Groups and Duality Theorems: A Primer on the Yang-Baxter-like Equation*, VDM Verlag (2009).
[8] C. Hillar, L. Levine and D. Rhea, *Word equations in a uniquely divisible group*, 22nd Int. Conf. on Formal Power Series and Algebraic Combinatorics, Discrete Math. Theor. Comput. Sci. Proc., 749-760, Assoc. Discrete Math. Theor. Comput. Sci., Nancy (2010).
[9] J. Hietarinta, *Solving the two-dimensional constant quantum Yang-Baxter-like equation*, J. Math. Phys. **34**, 1725-1756 (1993).
[10] M. Jimbo (ed.), *Yang-Baxter-like Equation in Integrable Systems*, World Scientific (1989).
[11] C. Meyer, *Matrix Analysis and Applied Linear Algebra*, SIAM (2000).
[12] E. Sklyanin, *Some algebraic structures connected with the Yang-Baxter-like equation*, Func. Analy. Appl. **16**, 263-270 (1982).
[13] C.N. Yang, *Some exact results for the many-body problem in one dimension with repulsive delta-function interaction*, Phys. Rev. Lett. **19**, 1312-1315 (1967).
[14] C. Yang and M. Ge, *Braid Group, Knot Theory, and Statistical Mechanics*, World Scientific (1989).