Correspondence

A Linear State-Space Analysis of the Migration Model in an Island Biogeography System

Abhishek Sinha, Swagatam Das, and B. K. Panigrahi

Abstract—Biogeography deals with the study of the distribution of biodiversity over space and time and has been well studied by naturists and biologists for over the last five decades. Recently, the theory of biogeography has been applied to solve difficult engineering optimization problems in the form of a nature-inspired metaheuristic, known as biogeography-based optimization (BBO) algorithm. In this correspondence paper, we present an in-depth analysis of the linear time-invariant (LTI) system model of immigration and emigration of organisms in an island biogeography system that forms the basis of BBO. We find the bound of the eigenvalues of the general LTI system matrix using the Perron–Frobenius theorem from linear algebra. Based on the bounds of the eigenvalues, we further investigate four important properties of the LTI biogeography system, including the system reasonability with probability distribution vectors, stability, convergence, and nature of the equilibrium state. Our analysis gives a better insight into the dynamics of migration in actual biogeography systems and also helps in the understanding of the search mechanism of BBO on multimodal fitness landscapes.

Index Terms—Biogeography-based optimization (BBO), convergence, equilibrium theory of island biogeography, island biogeography, linear time-invariant (LTI) system, population-based optimization, stability.

I. INTRODUCTION

The term *biogeography* [1]–[6] refers to the study of the geographical distribution of biological organisms over different landscapes. Mathematical models of biogeography describe how species migrate from one island to another, how new species arise, and how species become extinct. In the early 1960s, MacArthur and Wilson together put forward the mathematical models of biogeography, their work culminating with the landmark 1967 publication *The Theory of Island Biogeography* [3]. They showed that the species richness of an area could be predicted in terms of such factors as habitat area, immigration rate, and extinction rate. Since their pioneering work, biogeography has flourished as a major area of research [4]–[6].

In an attempt to use the concepts of island biogeography for the purpose of engineering optimization, Simon [7] recently proposed a metaheuristic algorithm, called biogeography-based optimization (BBO), and demonstrated its efficient performance in optimizing several numerical benchmarks. Central to BBO is the equilibrium theory of island biogeography (proposed by MacArthur and Wilson [3]) that uses a linear time-invariant (LTI) system with zero input to model the immigration and emigration of creatures in an island. BBO uses the states of the biogeography LTI system to determine the mutation rate that has been shown to be important to the algorithm with a

A. Sinha and S. Das are with the Department of Electronics and Telecommunication Engineering, Jadavpur University, Calcutta 700 032, India (e-mail: abhishek_sinha_ju@yahoo.com; swagatamdas19@yahoo.co.in).

B. K. Panigrahi is with the Department of Electrical Engineering, Indian Institute of Technology Delhi, New Delhi 110 016, India (e-mail: bkpanigrahi@ee.iitd.ac.in).

small population size [7]. BBO yields competitive results against many other state-of-the-art metaheuristic algorithms, like particle swarm optimization, ant colony optimization, genetic algorithm, evolutionary strategies, etc., details of which can be found in [8] and the references therein. A careful scrutiny of [7], however, reveals that the biogeography model presented within the framework of BBO lacks sufficient theoretical analysis to explore the important properties like stability, convergence, and characteristics of the equilibrium state, which is very important for understanding the actual biogeography system, as well as improving BBO. Some significant analytical studies on complex systems can be found in [9]–[11]. The author analyzed a simple case of the LTI system model of island biogeography, where the immigration and emigration rates were assumed to be linear with identical maximum. Moreover, the main theoretical analysis was based on a conjecture on the nature of the eigenvalues of the LTI system matrix. The author used the states of the LTI biogeography system to represent the probability distribution of the number of species in an island, which updates with each iteration of BBO. The reasonability of the system state for representing a probability distribution at each iteration was, however, not studied. Finally, the work in [7] did not prove the stability of the LTI migration model of island biogeography system explicitly.

In this correspondence paper, we start by proving that, given the dynamics of an LTI biogeography system, if the initial state is set to be a probability distribution vector (PDV) [12], then the state remains a PDV for all time. We then present a few interesting theorems on the nature and ranges of the eigenvalues of the system matrix in LTI system model by applying the Perron–Frobenius theorem [13]–[16] from linear algebra. Based on the ranges of the eigenvalues so obtained, we show that the general LTI biogeography system is stable in the sense of Lyapunov at the origin. Next, we prove that, given any initial state, the dynamics of the system is convergent and the equilibrium state is equal to the eigenvector that corresponds to the zero eigenvalue of the system matrix. Finally, we prove that the equilibrium PDV is unimodal under two mild restrictions imposed on the system matrix, a feature that contributes to the effectiveness of the properties of the mutation operation in BBO.

II. ISLAND BIOGEOGRAPHY MODEL AND BBO

Mathematical models of biogeography describe the migration, speciation, and extinction of species that migrate between islands. In [7], Simon used the term *island* descriptively rather than literally. That is, an island is any habitat that is geographically isolated from other habitats. In the classic sense of the term, an island is isolated from other habitats by water. However, biogeographical islands can also be habitats that are isolated by stretches of desert, rivers, impassable mountain ranges, predators, or other obstacles. Islands that are well suited as habitats for biological species are said to have a high Island Suitability Index (ISI). Features that correlate with ISI include factors such as rainfall, diversity of vegetation, diversity of topographic features, land area, and temperature. The variables that characterize habitability are called suitability index variables (SIVs). SIVs can be considered the independent variables of the island, and ISI can be considered the dependent variable [3], [7].

The fundamental idea of Simon's BBO stemmed from the phenomena described in the previous paragraph. Suppose that we have some optimization problems and that we also have a certain number of candidate solutions. A good solution is analogous to an island with a

high ISI, and a poor solution is like an island with a low ISI. High-ISI solutions are more likely to share their features with other solutions, and low ISI solutions are more likely to accept shared features from other solutions. A complete description of the algorithm is available in [7]. Instead of reiterating it here, we directly move our attention to the LTI state-space model of the biogeographical migrations used in BBO.

The MacArthur and Wilson's LTI model of island biogeography implicitly describes the variation of the number of species in a given island, which is caused by the immigration and emigration of the species among neighboring islands. Assume that the given island can support at most n species. Let x_i for $i = 1, 2, \ldots, n+1$ denote the probability that the island contains exactly $(i-1)$ species. Then, the probability vector $\vec{x} = [x_1, x_2, \ldots, x_{n+1}]^{\mathrm{T}}$ is modeled as a variable vector that varies with time t according to the following LTI state-space equation [3], [7]:

$$\dot{\vec{x}} = A\vec{x} \qquad (1)$$

where A is given by (2), shown at the bottom of the page. Note that λ_i and μ_i denote the immigration and emigration rates when there are i species in the given island and $i = 0, 1, \ldots, n$. Moreover, here, $\vec{x}(t)$ and A are called *state vector* and *system matrix* in linear systems theory, respectively. From the view of island biogeography, there are two basic constraints on the immigration and emigration rates [3], [7]. First, as the number of species increases, the island gets more crowded; therefore, the immigration rate should decrease, while the emigration rate should increase. Second, if there are n species (i.e., the maximum number of the species) in the given island, then the immigration rate should be zero. On the other hand, if there are no species in the island, then the emigration rate should be zero. To sum up, the two constraints on the immigration and emigration rates are

$$\lambda_0 \geq \lambda_1 \geq \lambda_2 \geq \cdots > \lambda_n = 0 \qquad (3a)$$
$$0 = \mu_0 < \mu_1 \leq \mu_2 \leq \cdots \leq \mu_n. \qquad (3b)$$

In what follows, we shall study certain interesting properties of the LTI biogeographical migration model presented through (1)–(3). Note that the LTI system studied in this correspondence paper is more general than that studied by [7, Th. 1], which only considered the special case that $\alpha_i = (i/n)$ and $\beta_i = ((n-i)/n)$ for $i = 0, 1, 2, \ldots, n$.

III. ANALYSIS OF THE LTI STATE-SPACE MODEL

As discussed in Section II, the state $\vec{x}(t)$ of the LTI system described by (1)–(3) is used to represent the probability distribution of the number of species. Therefore, in order to make the value of the state reasonable, $\vec{x}(t)$ for any $t \geq 0$ should remain a PDV, i.e., $\sum_{i=1}^{n+1} x_i(t) = 1$ and $x_i(t) \geq 0$ for any $i \in \{1, 2, \ldots, n+1\}$. Given an initial state $\vec{x}(0)$ for the LTI system described by (1)–(3), if the initial state $\vec{x}(0)$ is a PDV, then it can be confirmed by the following Lemmas 1 and 2 and Theorem 1 that the state $\vec{x}(t)$ remains a PDV for all $t \geq 0$.

Lemma 1: Given an initial state $\vec{x}(0) = [x_i(0)]_{(n+1)\times 1}$ for the LTI system described by (1)–(3), if $x_i(0) \geq 0$ for every $i \in \{1, 2, \ldots, n+1\}$, then $x_i(t) \geq 0$ for all $i \in \{1, 2, \ldots, n+1\}$ and all $t \geq 0$.

Proof: We have, initially, $x_i(0) \geq 0$ for every $i \in \{1, 2, \ldots, n+1\}$. Let us assume that Lemma 1 is false. Then, there must exist a time $t_1 \geq 0$ and a number $\tau_1 \geq 0$ satisfying

$$x_i(t) \geq 0 \text{ for } i = 1, 2, \ldots, n+1 \text{ and } 0 \leq t \leq t_1$$
$$x_i(t) < 0 \text{ for } i = r \text{ and } t_1 < t \leq \tau_1 \qquad (4)$$

where r is an index and $r \in \{1, 2, \ldots, n+1\}$. $x_r(t)$ for all $t \geq 0$ being a continuous function of time t, we must have

$$x_r(t_1) = 0. \qquad (5)$$

Then, following (1)–(3) and (5), we obtain

$$\dot{x}_r(t_1) = \begin{cases} \mu_1 x_2(t_1) \geq 0, & \text{if } r = 1 \\ \lambda_{r-2} x_{r-1}(t_1) + \mu_k x_{k+1}(t_1) \geq 0, & \text{if } 1 < k < n+1 \\ \lambda_{n-1}.x_n(t_1) \geq 0, & \text{if } k = n+1. \end{cases} \qquad (6)$$

Here, the inequality in each case follows from (4). Since $x_r(t_1) = 0$ and $\dot{x}_r(t_1) \geq 0$ for any choice of r in $\{1, 2, \ldots, n+1\}$, $x_r(t)$ must be an increasing function at $t = t_1$, and hence, there must exist another number $\tau_2 \geq 0$ such that $x_r(t) \geq 0$ for $t_1 < t \leq t_1 + \tau_2$. This, however, contradicts our former assumption that $x_r(t) < 0$ for $t_1 < t \leq \tau_1$. Hence, the earlier assumption must be false, and the lemma holds.

Lemma 2: For the LTI system described by (1)–(3), given an initial state $\vec{x}(0) = [x_i(0)]_{(n+1)\times 1}$, if $z(t) = \sum_{i=1}^{n+1} x_i(t)$, then $z(t) = \sum_{i=1}^{n+1} x_i(0)$ for all $t \geq 0$.

Proof: For all $t \geq 0$, we have

$$\begin{aligned} \frac{dz}{dt} &= \frac{d}{dt}(x_1(t) + x_2(t) + \cdots + x_{n+1}(t)) \\ &= \dot{x}_1(t) + \dot{x}_2(t) + \cdots + \dot{x}_{n+1}(t) \\ &= [1 \; 1 \; \ldots \ldots \; 1].[x_1(t), x_2(t) \; \ldots \ldots, x_{n+1}(t)]^T \\ &= \mathbf{1}_{1\times(n+1)} \cdot \dot{\vec{x}}(t) \\ &\quad [\text{Since } \lambda_n = \mu_0 = 0, \text{ it is easy to see that the sum of} \\ &\qquad \text{elements of the column vector } A.\vec{x}(t) \text{ is zero.}] \\ &= \mathbf{1}_{1\times(n+1)} \cdot A.\vec{x}(t) = 0 \end{aligned}$$

which implies that $z(t)$ is a constant function of time. Thus, $z(t) = z(0) = \sum_{i=1}^{n+1} x_i(0)$ for all $t \geq 0$.

Theorem 1: Given a biogeographical LTI system as per (1)–(3), if the initial state $\vec{x}(0)$ is a PDV, $\vec{x}(t)$ remains a PDV for all $t \geq 0$.

Proof: The proof of Theorem 1 directly follows from Lemmas (1) and (2).

Next, we consider the general linear system model for BBO and derive a necessary property to be followed by the matrix A. This restriction stems from the fact that the state vectors are not any vector from the space $\Re^n$ but with the additional constraint that it is a PDV. Then, we show that the BBO model proposed in [7] is consistent with our theoretical requirement.

Theorem 2: Given that $\vec{x}(t)$ is a PDV for all $t \geq 0$, the sum of the elements of each column of any general *probability transition matrix* A [not necessarily in the form of (2)] is identically equal to zero.

$$A = \begin{bmatrix} -(\lambda_0 + \mu_0) & \mu_1 & 0 & \cdots & & 0 \\ \lambda_0 & -(\lambda_1 + \mu_1) & \mu_2 & \ddots & & 0 \\ \vdots & \ddots & \ddots & \ddots & & \vdots \\ \vdots & & \ddots & \lambda_{n-2} & -(\lambda_{n-1} + \mu_{n-1}) & \mu_n \\ 0 & \cdots & & 0 & \lambda_{n-1} & -(\lambda_n + \mu_n) \end{bmatrix} \qquad (2)$$

Proof: Let us denote the (i, j)th element of A by a_{ij}. Thus, we have to prove that

$$\sum_{i=1}^{n+1} a_{ij} = 0 \quad \forall j \in [1, 2, \dots, n+1].$$

Suppose that the initial state vector is given by

$$\vec{x}(0) = [x_1(0), x_2(0), \dots, x_{n+1}(0)]^{\mathrm{T}}. \tag{7}$$

The solution of (1) in a vector–matrix notation is given by

$$\vec{x}(t) = e^{At}\vec{x}(0). \tag{8}$$

Then, from (1), (7), and (8), it follows that

$$\left.\frac{d\bar{x}}{dt}\right|_{t=0^+} = A\vec{x}(0)$$

$$= \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1j} & \cdots & a_{n+1n+1} \\ a_{21} & a_{22} & \cdots & a_{2j} & \cdots & a_{2n+1} \\ \vdots & \vdots & \cdots & \cdots & \vdots & \vdots \\ a_{n+11} & a_{n+12} & \cdots & a_{n+1j} & \cdots & a_{n+1n+1} \end{bmatrix}$$

$$\times \begin{bmatrix} x_1(0) \\ \vdots \\ x_n(0) \\ x_{n+1}(0) \end{bmatrix}.$$

Thus, we have (comparing each element of both vectors on LHS and RHS term by term)

$$\left.\frac{dx_i(t)}{dt}\right|_{t=0^+} = \sum_{j=1}^{n+1} a_{ij} x_j(0) \quad \forall i = [1, 2, \dots, n+1].$$

Now, summing both sides over all i's, we get

$$\sum_{i=1}^{n+1} \left.\frac{dx_i(t)}{dt}\right|_{t=0^+} = \sum_{i=1}^{n+1}\sum_{j=1}^{n+1} a_{ij}x_j(0)$$

or,

$$\left.\frac{d}{dt}\sum_{i=1}^{n+1} x_i(t)\right|_{t=0^+} = \sum_{j=1}^{n+1} x_j(0)\sum_{i=1}^{n+1} a_{ij}$$

(interchanging the order of summation)   (9)

However, $\vec{x}(t)$ being a PDV, for all time t, we must have $\sum_{i=1}^{n+1} x_i(t) = 1$. Thus, the left side of (9) is identically equal to zero. Now, denoting $\sum_{i=1}^{n+1} a_{ij}$ by S_j, for $j = 1, 2, \dots, n+1$, we have

$$\sum_{j=1}^{n+1} S_j x_j(0) = 0. \tag{10}$$

The aforementioned equation must hold good for *any* selection of the initial state vector $\vec{x}(0)$.

Now, $x_j(0)$'s are linearly dependent with the only linear-constraint equation as $\sum_{j=1}^{n+1} x_j(0) = 1$.

Thus, we may choose only n probability variables independently out of these $n+1$ variables. Suppose now we choose $x_i(0) = 0$ and $i \neq j$; then, we have $x_j(0) = 1$, where j may be repetitively put equal to $1, 2, 3, \dots, n+1$.

Thus, from (10), we get $S_j = 0$, where $j = 1, 2, 3, \dots, n+1$, i.e., $\sum_{i=1}^{n+1} a_{ij} = 0$, where $j \in [1, 2, \dots, n+1]$ (Proved).

Corollary 1: $\phi = 0$ is necessarily an eigenvalue of the matrix A, and hence, A is singular.

Proof: That $\phi = 0$ is an eigenvalue of A may be proved very easily following the theorem we prove before. Here, we will actually show that $\phi = 0$ is an eigenvalue of A^{T}. Since the eigenvalues of A and A^{T} are identical, it is proven that $\phi = 0$ is an eigenvalue of A. Consider the column vector $\vec{x}_1 = [1, 1, \dots, 1]^{\mathrm{T}}$. Then, evidently

$$A^{\mathrm{T}}.\vec{x}_1 = \left[\sum_i a_{i1}, \sum_i a_{i2}, \dots, \sum_i a_{ij}, \dots, \sum_i a_{i(n+1)},\right]^{\mathrm{T}}$$
$$= [0, 0, \dots, 0]^{\mathrm{T}}$$
$$\times \left[\text{Since} \sum_{i=1}^{n+1} a_{ij} = 0 \text{ for every } j = 1, 2, 3, \dots, n+1\right]$$
$$= 0.[1, 1, \dots, 1]^{\mathrm{T}}$$
$$= \phi.\vec{x}_1, \text{ where } \phi = 0.$$

Thus, as per the definition of eigenvalue, $\phi = 0$ is an eigenvalue of A^{T} and, hence, A. Again, we know that the eigenvalues of A matrix satisfy the $(n+1)$th degree algebraic equation (known as the characteristic equation of A): $\det(A - \phi.I) = 0$, where I is the $(n+1)$th order identity matrix.

Since $\phi = 0$ satisfies this equation, we have (putting $\phi = 0$)

$$\det(A) = 0 \tag{11}$$

thus proving that matrix A is singular.

Now, we will prove that the island biogeography system described by (1)–(3) is stable by showing that all the eigenvalues of matrix A are nonpositive and there is only a single eigenvalue (i.e., eigenvalue of multiplicity one) with the value zero. To prove this, we use the famous Perron–Frobenius theorem of linear algebra [13]–[16] for nonnegative and irreducible matrix. Before stating the theorem, we define the following.

Definition 1: A square $n \times n$ matrix $D = [d_{ij}]$ is said to be irreducible if the indices $1, 2, \dots, n$ can be divided into two disjoint nonempty sets $i_1, i_2, \dots, i_\mu$ and $j_1, j_2, \dots, j_v$ (with $\mu + \nu = n$) such that $d_{i_\alpha j_\beta} = 0$, for $\alpha = 1, 2, \dots, n$ and $\beta = 1, 2, \dots, n$. The square matrix D is irreducible if there exist some k's such that $d_{ij}^{(k)} > 0$ for each i and j chosen from $\{1, 2, \dots, n\}$. Here, $d_{ij}^{(k)}$ denotes the (i, j)th element of the matrix D^k for some positive integer k.

Now, the theorem may be stated as follows.

Perron–Frobenius Theorem [13]: If C be a nonnegative and irreducible matrix, then we have the following.

1) *C has a positive eigenvalue q that is equal to its spectral radius $\rho(C)$.*
2) *q is a simple eigenvalue of C, i.e., the multiplicity of q is one.*
3) *There is a positive right eigenvector associated with q.*
4) *One has the eigenvalue estimate*

$$\min_i \sum_j c_{ij} \leq q \leq \max_i \sum_j c_{ij}. \tag{12}$$

However, before we apply this theorem to matrix A, we need to transform it to a suitable form, so that the conditions for applying the Perron–Frobenius theorem (*viz.*, nonnegativity and irreducibility) are met. Here, actually, we obtain the upper bound for the eigenvalues of A^{T}. Since the eigenvalues of matrix A are identically equal to that of matrix $A^{\mathrm{T}} = B$ (for example), in the process, we also obtain the upper bound for the eigenvalues of matrix A.

Let $max_i(\lambda_i + \mu_i) = \gamma$. Thus, $\gamma \geq (\lambda_i + \mu_i)$ for all $i = 1, 2, 3, \dots, n+1$. Let us perform the following linear transformation on B:

$$C = \frac{1}{\gamma}(B + \gamma I) \tag{13}$$

where I is an $(n+1) \times (n+1)$ identity matrix. On the calculation, C takes the form (14), shown at the bottom of the page. Now, it is given that $\lambda_i > 0$ for $i = 0, 1, 2, \ldots, n-1$ and $\mu_i > 0$ for $i = 1, 2, \ldots, n$.

Thus, all the off-diagonal terms are greater than zero. Again, $\gamma \geq (\lambda_i + \mu_i)$; hence, all the diagonal terms are greater than or equal to zero. Now, we prove the following three lemmas on matrix C.

Lemma 3: C is a stochastic matrix and, hence, nonnegative.

Proof: The proof is almost obvious because of the following.

1) All the elements of C are nonnegative and less than one.
2) Sum of each row elements $= 1$ [obvious by the inspection of (14)].

Lemma 4: Matrix C is irreducible.

Proof: Let us denote the (i,j)th element of matrix C raised to power k by $c_{ij}^{(k)}$. From (14), it is clear that, for matrix C

$$c_{i(i+1)}^{(1)} > 0 \quad \text{and} \quad c_{i(i-1)}^{(1)} > 0, \quad \text{for } i = 1, 2, \ldots, n+1 \quad (15)$$

(except, of course, the meaningless c_{10} and $c_{(n+1)(n+2)}$).

Now, by matrix multiplication, we have

$$c_{i(i+2)}^{(2)} = \sum_j c_{ij}^{(1)} c_{j(i+2)}^{(1)} \geq c_{i(i+1)}^{(1)} \cdot c_{(i+1)(i+2)}^{(1)} > 0 \quad (16)$$

[using the nonnegativity of C and condition (15)].

Similarly, we have

$$c_{i(i-2)}^{(2)} = \sum_j c_{ij}^{(1)} c_{j(i-2)}^{(1)} \geq c_{i(i-1)}^{(1)} c_{(i-1)(i-2)}^{(1)} > 0. \quad (17)$$

The aforementioned argument acts as the basis step of our induction.

Now, assuming that $j > i$ and $c_{ij}^{(m)} > 0$, we now prove that $c_{i(j+1)}^{(m+1)} > 0$.

We have

$$c_{i(j+1)}^{(m+1)} = \sum_k c_{ik}^{(m)} \cdot c_{k(j+1)}^{(1)} \geq c_{ij}^{(m)} \cdot c_{j(j+1)}^{(1)} > 0 \quad (18)$$

(using the nonnegativity of C, condition (15), and the induction hypothesis).

Similarly, now, we assume that $j < i$ and $c_{ij}^{(m)} > 0$ and proceed to prove that $c_{i(j-1)}^{(m+1)} > 0$.

We have

$$c_{i(j-1)}^{(m+1)} = \sum_k c_{ik}^{(m)} \cdot c_{k(j-1)}^{(1)} \geq c_{ij}^{(m)} \cdot c_{j(j-1)}^{(1)} > 0 \quad (19)$$

(using the nonnegativity of C, condition (15), and the induction hypothesis).

We can repeat the procedure to prove that, for each (i,j) (satisfying $i \neq j$), there exists some integer k such that $c_{ij}^{(k)} > 0$.

Lastly, for $i = j$, we have

$$c_{ii}^{(2)} = \sum_j c_{ij}^{(1)} c_{ji}^{(1)} \geq c_{i(i+1)}^{(1)} c_{(i+1)i}^{(1)} > 0, \text{ for all } i \in \{1, 2, \ldots, n+1\}.$$

This completes our inductive argument to prove that, for each (i,j), there exists an integer k such that $c_{ij}^{(k)} > 0$. Hence, matrix C is irreducible.

Lemma 5: There is a one-on-one correspondence between the eigenvalues of matrices B and C, and the correspondence is given by

$$\phi_B = \gamma(\phi_C - 1). \quad (20)$$

Proof: We have

$$C = \frac{1}{\gamma}(B + \gamma I)$$

$$\Rightarrow C - \phi_C I = \frac{1}{\gamma}\{B - \gamma(\phi_C - 1)I\}$$

$$det(C - \phi_C I) = \frac{1}{\gamma}.det[B - \gamma(\phi_C - 1)I].$$

Now, the eigenvalues of C are given by the roots of $det(C - \phi_C I) = 0$, which is equivalent to $det(B - \phi_B I) = 0$, where, evidently, $\phi_B = \gamma(\phi_C - 1)$, and hence, the lemma is proved.

Based on Lemmas 3, 4, and 5, we now proceed to prove an important theorem regarding the stability of the biogeography system. We use the eigenvalue criterion from the standard control theory [16], [17] for the determination of system stability.

Criterion 1: Given the LTI system described by $\dot{\vec{x}} = D.\vec{x}$, let $\phi_1, \phi_2, \ldots, \phi_n$ denote the eigenvalues of square $n \times n$ matrix D. The LTI system is stable in the sense of Lyapunov if $Re(\phi_i) \leq 0$ for $i = 1, 2, \ldots, n$, and those ϕ_i for which $Re(\phi_i) = 0$ are simple, where $Re(\phi_i)$ denotes the real part of ϕ_i. Otherwise, the LTI system is unstable.

Theorem 3: The biogeography system described by (1)–(3) is stable.

Proof: From Lemmas 3 and 4, it is clear that C is nonnegative and irreducible. From Perron–Frobenius theorem, if q be the greatest eigenvalue of C, then

$$\min_i \sum_j c_{ij} \leq q \leq \max_i \sum_j c_{ij}.$$

However, the matrix C being stochastic (by Lemma 3), it follows that

$$\min_i \sum_j c_{ij} = \max_i \sum_j c_{ij} = 1.$$

Thus, the greatest eigenvalue is $q = 1$, and according to the Perron–Frobenius theorem, the eigenvalue $q = 1$ is of multiplicity 1. Evidently, $\phi_C = 1$ is a simple eigenvalue of C, i.e., there is one and only one eigenvalue of C of magnitude 1. From Lemma 5, it is easy to see that there is one and only one eigenvalue $\phi_B = 0$ for matrix B. Since $B = A^T$, matrix A also has one and only one eigenvalue of magnitude 0. Again, $q = 1$ being the largest eigenvalue of C, we have all eigenvalues $|\phi_C| \leq 1$. However, from Lemma 5, we also have $\phi_B = \gamma(\phi_C - 1)$. Thus, we conclude that, except only one single eigenvalue at 0, all other eigenvalues of B have negative real parts and the same statement also holds for the eigenvalues of matrix A (as

$$C = \frac{1}{\gamma}. \begin{bmatrix} \gamma - (\lambda_0 + \mu_0) & \lambda_0 & 0 & 0 & \cdots & 0 \\ \mu_1 & \gamma - (\lambda_1 + \mu_1) & \lambda_1 & 0 & \cdots & 0 \\ 0 & \mu_2 & \gamma - (\lambda_2 + \mu_2) & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & \mu_{n-1} & \gamma - (\lambda_{n-1} + \mu_{n-1}) & \lambda_{n-1} \\ 0 & \cdots & \cdots & 0 & \mu_n & \gamma - (\lambda_n + \mu_n) \end{bmatrix} \quad (14)$$
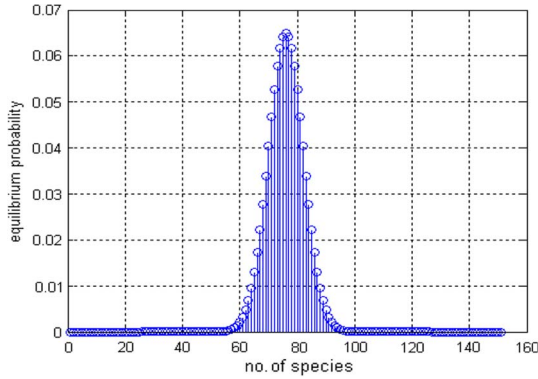
Fig. 1. Equilibrium PDV for island biogegraphy dynamics with number of species $n = 150$ with linear parameters from [7].

$B = A^{\mathrm{T}}$). Since there is only one eigenvalue of A at zero and all other eigenvalues have negative real parts, by Criterion 1, the biogeography system defined by (1)–(3) is stable (Proved).

Observation 1: It is important to remember that all the eigenvalues of C (or A) may not be real. All that we have from the Perron–Frobenius theorem is that $|\phi_C| \leq 1$. Moreover, as $\phi_B = \gamma(\phi_C - 1)$, we have $|\phi_B + \gamma| = \gamma|\phi_C| \leq \gamma$. Thus, the eigenvalues of matrix A will be located at the interior of a circle of radius γ and center at $-\gamma$ in the Argand plane. Fig. 1 shows the locations of the eigenvalues of matrix A.

Theorem 4: Given the LTI biogeography model described by (1)–(3), if the initial state $\vec{x}(0)$ is a PDV, then the equilibrium state of the system is given by the eigenvector $\vec{\nu} = [\nu_i]_{(n+1)\times 1}$, which corresponds to the zero eigenvalue of matrix A and

$$\nu_1 = \frac{1}{1 + \sum\limits_{i=2}^{n+1} \frac{\prod\limits_{k=0}^{i-2} \lambda_k}{\prod\limits_{k=1}^{i-1} \mu_k}} \tag{21}$$

$$\nu_i = \frac{\prod\limits_{k=0}^{i-2} \lambda_k}{\prod\limits_{k=1}^{i-1} \mu_k} \nu_1, \quad \text{for } i = 2, 3, \dots, n+1. \tag{22}$$

Proof: By Definition 1, at equilibrium, we must have $(d\vec{x}(t)/dt) = 0$ at $\vec{x}(t) = \vec{\nu}$.

Thus, if we denote the equilibrium state vector by $\vec{\nu} = [\nu_1, \nu_2, \dots, \nu_{n+1}]^{\mathrm{T}}$, then we must have (23), shown at the bottom of the page. Denoting the jth row of the corresponding augmented matrix of the aforementioned system by R_j, we perform the following elementary row transformations:

$$R_j \to R_1 + R_2 + \cdots + R_j, \quad \text{for } j = 1, 2, 3, \dots, n+1.$$

This results in the following equivalent system:

$$\sim \left\langle \begin{array}{ccccc|c} -\lambda_0 & \mu_1 & 0 & \cdots & 0 & 0 \\ 0 & -\lambda_1 & \mu_2 & \cdots & 0 & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots & 0 \\ \vdots & & \ddots & 0 & -\lambda_{n-1} & \mu_n & 0 \\ 0 & \cdots & 0 & 0 & 0 & 0 \end{array} \right\rangle$$

(remembering $\lambda_n = \mu_0 = 0$)

and finally, solving the resulting equivalent system of equations, we have

$$\nu_2 = \frac{\lambda_0}{\mu_1} \nu_1 \tag{24.1}$$

$$\nu_3 = \frac{\lambda_1}{\mu_2} \nu_2 \tag{24.2}$$

$$\vdots \qquad \vdots$$

$$\nu_{n+1} = \frac{\lambda_{n-1}}{\mu_n} \nu_n \tag{24.n}$$

It is an easy recursion to solve, with the result

$$\nu_i = \frac{\prod\limits_{k=0}^{i-2} \lambda_k}{\prod\limits_{k=1}^{i-1} \mu_k} \nu_1, \quad \text{for } i = 2, 3, \dots, n+1.$$

However, since the initial state $\vec{x}(0)$ is a PDV and, by Theorem 1, the equilibrium state $\vec{\nu}$ will also be a probability vector, we must have

$$\sum_i \nu_i = 1. \tag{25}$$

This gives

$$\nu_1 = \frac{1}{1 + \sum\limits_{i=2}^{n+1} \frac{\prod\limits_{k=0}^{i-2} \lambda_k}{\prod\limits_{k=1}^{i-1} \mu_k}}.$$

Thus, $\nu_i = (\Pi_{k=0}^{i-2}\lambda_k / \Pi_{k=1}^{i-1}\mu_k)\nu_1$ for $i = 2, 3, \dots, n+1$ with ν_1 as evaluated before.

It is easy to observe that, since for the equilibrium state vector $\vec{\nu}$, $A.\vec{\nu} = 0 = 0.\vec{\nu}$; thus, $\vec{\nu}$ is the eigenvector corresponding to the eigenvalue $\phi_A = 0$.

Observation 2: From the linear control theory, it is easy to see that if we have a system with one single-zero eigenvalue, then we have more than one equilibrium point. In fact in this case, we will have a line of equilibrium points, and the direction vector for this line is the eigenvector associated to the eigenvalue zero [16], [17]. Consider, for example, a simple system whose dynamics is expressed as

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{pmatrix} d_{11} & d_{12} \\ d_{21} & d_{22} \end{pmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = D.\vec{x}(t).$$

$$\left. \frac{d\vec{x}(t)}{dt} \right|_{\vec{x}(t)=\vec{\nu}} = A.\vec{\nu} = 0 \Rightarrow \begin{bmatrix} -(\lambda_0 + \mu_0) & \mu_1 & 0 & \cdots & & 0 \\ \lambda_0 & -(\lambda_1 + \mu_1) & \mu_2 & \ddots & & 0 \\ \vdots & \ddots & \ddots & \ddots & & \vdots \\ \vdots & & \ddots & \lambda_{n-2} & -(\lambda_{n-1} + \mu_{n-1}) & \mu_n \\ 0 & \cdots & & 0 & \lambda_{n=1} & -(\lambda_n + \mu_n) \end{bmatrix} \cdot \begin{bmatrix} \nu_1 \\ \nu_2 \\ \vdots \\ \nu_i \\ \vdots \\ \nu_{n+1} \end{bmatrix} = 0 \tag{23}$$

If the two eigenvalues of matrix D are 0 and ϕ_1 and if $\vec{\nu}_1$ and $\vec{\nu}_2$ be the eigenvectors associated with them, respectively, then the general system response may be expressed as

$$\vec{x}(t) = k_1.e^{0.t}.\vec{\nu}_1 + k_2.e^{\phi_1.t}.\vec{\nu}_2. \tag{26}$$

Clearly, if $\phi_1 < 0$, as $t \to \infty$, the system state vector converges to the vector $k_1.\vec{\nu}_1$ along the direction of the eigenvector $\vec{\nu}_1$ corresponding to the zero eigenvalue of D matrix. From Theorems 3 and 4, we find that, since only one eigenvalue of matrix A is 0 and the rest of the eigenvalues are negative, the biogeography system will also have its equilibrium vector along the direction of the eigenvector $\vec{\nu} = [\nu_i]_{(n+1)\times 1}$ given by (21) and (22) for any initial PDV $\vec{x}(0)$.

Observation 3: If in the equation system (23) we assume $\nu_1 = 1$ and, correspondingly, $\nu_i = \prod_{k=0}^{i-2} \lambda_k / \prod_{k=1}^{i-1} \mu_k$ for $i = 2, 3, \ldots, n+1$, then also the vector $\vec{\nu}$ remains an eigenvector of matrix A corresponding to eigenvalue, but in that case, the equilibrium PDV of the system would have been $\vec{\nu}/\Sigma_i \nu_i$, as evident from (21) and (22). This observation is in complete agreement with [4, Th. 1], in which it was proven using the singular value decomposition theory [18]. In what follows, we focus on the nature of the equilibrium PDV. Under two mild constraints, it can be proven that the equilibrium PDV is unimodal. However, before we prove the unimodality of the equilibrium PDV, we prove the following simple lemma. We first give a definition of *modal point* that will be used in proving the unimodality of the equilibrium PDV.

Definition 5: An index "k" is said to be a modal point of matrix A if

$$\lambda_{k-2} \geq \mu_{k-1} \quad \lambda_{k-1} < \mu_k.$$

Lemma 6: For the system given in (2) and (3a) and (3b), there exists one and only one modal point, subject to the constraints $\lambda_0 > \mu_1$ and $\lambda_{n-1} < \mu_n$.

Proof: First, we prove that there must exist one modal point for the system. We prove it by the method of contradiction. On the contrary to the claim, suppose that there is not any modal point for the system. Then, we have

Either $\lambda_{k-2} \geq \mu_{k-1}$, for all k,
Or $\lambda_{k-1} < \mu_k$, for all k

but given that $\lambda_0 > \mu_1$. Thus, only the first option is feasible, and we must have

$$\lambda_{k-2} \geq \mu_{k-1}, \quad \text{for all } k.$$

Now, putting $k = n+1$, we have $\lambda_{n-1} \geq \mu_n$, which contradicts the given constraint. This means that there must exist a modal point for the system. Now, we prove that the modal point for the system is unique. We prove it also by the method of contradiction. On the contrary to the claim, suppose that there are two modal points, namely, k_1 and k_2, with $k_2 > k_1$. Then, as per the definition of the modal points, we must have for $1 \leq i \leq k - 1$

$$\lambda_{k_2-2} \geq \mu_{k_2-1} \quad \lambda_{k_1-1} < \mu_{k_1}. \tag{27}$$

However, since $k_2 > k_1$, we must have [from (3b)]

$$\mu_{k_2-1} \geq \mu_{k_1}. \tag{28}$$

Thus, combining (27) and (28), we have

$$\lambda_{k_2-2} \geq \mu_{k_2-1} \geq \mu_{k_1} > \lambda_{k_1-1}.$$

This implies that $\lambda_{k_2-2} > \lambda_{k_1-1}$, and clearly, it contradicts (3a). Hence, there is one and only one modal point for the given system

described by (2) and (3). Now, we prove the unimodality of the equilibrium PDV in the following theorem.

Theorem 5: Given the system described by (2) and (3), and the equilibrium PDV $\vec{\nu} = [\nu_i]_{(n+1)\times 1}$ as described in Theorem 4, if $\lambda_0 > \mu_1$ and $\lambda_{n-1} < \mu_n$, then $\nu_1 \leq \nu_2 \leq \cdots \leq \nu_k > \nu_{k+1} > \cdots > \nu_{n+1}$, where k is the modal point of $\mathbf{A}$.

Proof: According to the aforementioned lemma, there exists one and only one modal point for $\mathbf{A}$. If k denotes the modal point, then according to (3), $\lambda_{i-1} \geq \lambda_{k-2}$ and $\mu_i \leq \mu_{k-1}$ for $1 \leq i \leq k-1$. Moreover, from (24), we have

$$\frac{\nu_{i+1}}{\nu_i} = \frac{\lambda_{i-1}}{\mu_i} \geq \frac{\lambda_{k-2}}{\mu_{k-1}}, \quad \text{for } 1 \leq i \leq k-1. \tag{29}$$

Since k is the modal point, by definition, we have $\lambda_{k-2} \geq \mu_{k-1}$. Combining the aforementioned results, we obtain $\nu_{i+1} \geq \nu_i$, when $1 \leq i \leq k-1$. When $k \leq i \leq n$, we can prove that $\nu_{i+1} < \nu_i$ in a similar way. Thus, the theorem holds. Fig. 1 shows the equilibrium PDV for $n = 150$.

Observation 4: Theorem 4 implies that, for the LTI biogeography dynamics, the equilibrium state (PDV) is symmetrical, and the modal point is approximately $n/2$ under the conditions $\mu_i = \lambda_{n-i}$ for $i = 0, 1, \ldots, n$. We note that the biogeography system studied in [7] sets $\mu_i = \lambda_{n-i} = i/n$ for $i = 0, 1, \ldots, n$, which satisfies this condition.

Remark 1: The BBO algorithm uses species count probabilities to determine mutation rates. Through Lemma 6 and Theorem 5, we mainly indicated that the equilibrium state has its components farther from the modal point smaller than those near the modal point. Thus, Medium species counts have high probabilities because they are near the equilibrium point. This way, the unimodal characteristic of the equilibrium PDV makes the mutation operation of BBO provide poorer solutions (with low fitness) a chance of improving and at the same time prevents the BBO algorithm from getting trapped in a local optimum.

IV. CONCLUSION

In this correspondence paper, we have first shown that setting the initial state to a PDV ensures that the state remains a PDV for all future time. Next, we have proven that, given the initial state vector is a PDV, the general probability transition matrix A, not necessarily in the form of (2), has sum of the elements in its each column equal to zero. Then, we have calculated the bounds of the eigenvalues of the general LTI system matrix using the concepts of Perron–Frobenius theorem. Based on the bounds of the eigenvalues, the stability of the system was inferred. We found the equilibrium state vector for the system as a vector along the direction of the eigenvector associated with eigenvalue 0. We have also analytically demonstrated the unimodality of the equilibrium PDV and outlined the distribution of the components of the equilibrium state vector to provide a better understanding of the effectiveness of the mutation operation in BBO. Future research may focus on developing suitable Lyapunov energy functions [17] for the stability analysis of the biogeography system. A Markov chain analysis could also be undertaken to better understand the search mechanisms of different variants of the BBO algorithm.

REFERENCES

[1] A. Wallace, *The Geographical Distribution of Animals (Two Volumes)*. Boston, MA: Adamant Media Corporation, 2005.

[2] C. B. Cox and P. D. Moore, *Biogeography: An Ecological and Evolutionary Approach*, 8th ed. Hoboken, NJ: Wiley-Blackwell, May 2010.

[3] R. MacArthur and E. Wilson, *The Theory of Biogeography*. Princeton, NJ: Princeton Univ. Press, 1967.

[4] J. J. Morrone, *Evolutionary Biogeography: An Integrative Approach With Case Studies*. New York: Columbia Univ. Press, Dec. 2008.

[5] M. V. Lomolino, D. F. Sax, and J. H. Brown, Eds., *Foundations of Biogeography*. Chicago, IL: Univ. Chicago Press, Jul. 2004.

[6] J. Wu and J. L. Vankata, "System dynamics model of island biogeography," *Bull. Math. Biol.*, vol. 53, no. 6, pp. 911–940, 1991.

[7] D. Simon, "Biogeography-based optimization," *IEEE Trans. Evol. Comput.*, vol. 12, no. 6, pp. 702–713, Dec. 2008.

[8] A. P. Engelbrecht, *Computational Intelligence*, 2nd ed. Hoboken, NJ: Wiley, 2007.

[9] S. Das, S. Dasgupta, A. Biswas, A. Abraham, and A. Konar, "On stability of chemotactic dynamics in bacterial foraging optimization algorithm," *IEEE Trans. Syst., Man, Cybern. A, Syst., Humans*, vol. 39, no. 3, pp. 670–679, May 2009.

[10] H. de Bruijn and P. M. Herder, "System and actor perspectives on sociotechnical systems," *IEEE Trans. Syst., Man, Cybern. A, Syst., Humans*, vol. 39, no. 5, pp. 981–992, Sep. 2009.

[11] A. Obeidi, D. M. Kilgour, and K. W. Hipel, "Perceptual stability analysis of a graph model system," *IEEE Trans. Syst., Man, Cybern. A, Syst., Humans*, vol. 39, no. 5, pp. 993–1006, Sep. 2009.

[12] A. Graham, *Nonnegative Matrices and Applicable Topics in Linear Algebra*. Hoboken, NJ: Wiley, 1987.

[13] H. Minc, *Nonnegative Matrices*. New York: Wiley, 1988.

[14] O. Perron, "Zur Theorie der Matrices," *Mathematische Annalen*, vol. 64, no. 2, pp. 248–263, 1907.

[15] G. Frobenius, *Ueber Matrizen aus nicht negativen Elementen*. Berlin, Germany: Sitzungsber. Königl. Preuss. Akad. Wiss, 1912, pp. 456–477.

[16] W. L. Brogan, *Modern Control Theory*. Englewood Cliffs, NJ: Prentice-Hall, 1991.

[17] A. Bacciotti and L. Rosier, *Liapunov Functions and Stability in Control Theory*, 2nd ed. New York: Springer-Verlag, Apr. 2005.

[18] G. Strang, *Introduction to Linear Algebra*, 3rd ed. Cambridge, MA: Wellesley-Cambridge, 1998, sec. 6.7.