Why Are Clonal Selection Algorithms Markov?

Khiria E. A. El-Nady
Department of Mathematics, Faculty of Science, Alexandria University, Alexandria, Egypt
pr_dr_kh_elnady@yahoo.com

Usama H. M. Abou El-Enien
Department of Mathematics, Faculty of Science, Alexandria University, Alexandria, Egypt
ossama.hanafy77@gmail.com

Amr A. A. Badr
Department of Computer Science, Faculty of Computers and Information, Cairo University, Giza, Egypt
pr.dr.a.badr@gmail.com

Abstract

Clonal selection algorithms are considered. Two algorithms are designed and executed to obtain purely empirical analysis conclusions in order to turn to purely theoretical analysis results about the behavior of clonal selection algorithms as a finite dimensional Markov and lumped Markov chains, which confirm the conjectures from these experiments and in order to introduce a complete framework toward a new philosophy of machine intelligence. First, we model clonal selection algorithms using a finite dimensional Markov and lumped Markov chains. Second, we carry on a particle analysis (the basic component) and analyze the convergence properties of these algorithms. Third, we produce two unified Markov and lumped Markov approaches for analysis for a complete framework and propose unique chromosomes for a purely successful optimization of these algorithms. Furthermore, for the Markov approach, we obtain purely theoretical analysis for a classification and Stationary distributions of chains. For the lumped Markov approach, we obtain purely theoretical analysis for all possible conditional multivariate normal distributions of transition probability matrices and stationary multivariate normal distributions of chains.
Keywords: Clonal selection algorithms, lumped Markov chains, Classification, central limit theorem, Stationary multivariate normal distribution, Unique chromosomes

1. Introduction

There are works in the literature that attempt to introduce an architecture for the construction of artificial immune systems (e.g., Hofmeyr and Forrest [8]), a formal model for AIS (e.g., Tarakanov and Dasgupta [6]), a physical model for it (e.g., Zak [14]), or even convergence analysis for it (Villalobos-Arias [5]). Nevertheless, these works are usually either too restricted to a particular school or application, or they do not cover enough immune principles and components to provide a general framework in which to engineer an AIS. There has been another recent attempt to provide such a general-purpose framework (in a Ph.D. thesis (de Castro [3])). These works do not develop a unified approach for clonal selection algorithms, analyze their performance and convergence properties to provide a general framework.

These works do not develop a unified stochastic model for clonal selection algorithms, analyze their performance and convergence properties to provide a general framework. Thus it is significant to establish purely theoretical analysis results about the behavior of clonal selection algorithms as a finite dimensional Markov and lumped Markov chains in order to introduce a complete framework.

The main purpose of this paper is to produce a unified Markov and lumped Markov approaches for analysis in order to introduce a complete framework of clonal selection algorithms toward a new philosophy of machine intelligence, and propose unique chromosomes for a purely successful optimization of these algorithms. Through unified Markov approach, we obtain purely empirical analysis conclusions and obtain purely theoretical analysis for a classification and Stationary distributions of chains. Through unified lumped Markov approach, we obtain purely empirical analysis conclusions and obtain purely theoretical analysis for all possible conditional multivariate normal distributions of transition probability matrices and stationary multivariate normal distributions of chains.

The rest of the paper is organized as follows. In Section 2, we give the formulation of the problem, namely: Why is clonal selection algorithms Markov chains?. In Section 3, we state the main result. Then in Section 4, the proof of the main result is given in nine steps. In Section 5, we propose the two algorithms. In Section 6, we give two numerical examples. In Section 7, we give some concluding remarks.
2. Formulation of the problem

All empirical analysis conclusions and theoretical analysis of clonal selection algorithms in literature do not develop a purely unified stochastic model, analyze their performance and convergence properties to provide a general framework. Thus, in this paper, we consider an interesting problem, namely: Why are clonal selection algorithms Markov chains?

Throughout this paper, we consider any objective real valued function of \( n \)-variables

\[
f(x_1, x_2, ..., x_n), \text{ where } a_i \leq x_i \leq b_i \text{ for } i = 1, 2, ..., n
\]

are domains of each variable \( x_i \) and \( a_i, b_i \) are real numbers, given a set of patterns to be recognize (\( S \)), the basic steps of the CLONALG algorithm are as follows:

1. Initialization: create an initial random population of individuals (\( P \)).
2. Antigenic presentation: for each antigenic pattern, do:
   2.1. Affinity evaluation: present it to the population \( P \) and determine its affinity with each element of the population \( P \).
   2.2. Clonal selection and expansion: select \( n_1 \) highest affinity elements of \( P \) and generate clones of these individuals proportionally to their affinity with the antigen: the higher the affinity, the higher the number of copies, and vice-versa.
   2.3. Affinity maturation: mutate all these copies with a rate inversely proportional to their affinity with the input pattern: the higher the affinity, the smaller the mutation rate, and vice-versa. Add these mutated individuals to the population \( P \) and re-select the best individual to be kept as the memory \( m \) of the antigen presented.
   2.4. Metadynamics: replace a number \( n_2 \) of individuals with low affinity by (randomly generated) new ones.
3. Cycle: repeat Step 2 until a certain stopping criterion is met.

Some authors (e.g., Forrest et al.\cite{9}) have argued that a genetic algorithm without crossover is a reasonable model of clonal selection. However, the standard genetic algorithm does not account for the two important immune properties: affinity proportional reproduction and mutation. Other authors (de Castro and Von Zuben \cite{4}) proposed a clonal selection algorithm, named CLONALG, to fulfill these basic processes involved in clonal selection. This algorithm is evolutionary in nature, and it was initially proposed to perform
We restrict an arbitrary uncountable set

\[ S = \{ a_i \leq x_i \leq b_i \text{ for } i = 1, 2, \ldots, n \} \]

to be a subset of \( n \)-space \( \mathbb{R}^n \) as a sample space, restrict an arbitrary countable set \( T \) to be set of all

\[(x_1, x_2, \ldots, x_n) \text{ in } S = \{ a_i \leq x_i \leq b_i \} \]

for which

\[ P(x_1, x_2, \ldots, x_n) > 0 \]
as a sample space.

Let a sample space \( T \) of possible solutions to be coded as strings of \( k \) bits \( \{0, 1\} \). \( T = \{ \text{unique chromosomes} \} \).

We restrict an arbitrary countable set \( U \) to be set of all possible simple random samples with replacement in the \( i^{th} \) trail

\[ D_y^{(i)} = (x_{11}, x_{12}, \ldots, x_{1n}), (x_{21}, x_{22}, \ldots, x_{2n}), \ldots, (x_{m1}, x_{m2}, \ldots, x_{mn}) \]

for \( y = 1, 2, \ldots, h \) (\( h = \text{number of samples} \)) for which

\[ P((x_{11}, x_{12}, \ldots, x_{1n}), (x_{21}, x_{22}, \ldots, x_{2n}), \ldots, (x_{m1}, x_{m2}, \ldots, x_{mn})) > 0 \]
as a sample space, where for each

\[(x_{j1}, x_{j2}, \ldots, x_{jn}) \text{ for } j = 1, 2, \ldots, m(\text{sample size}) \]
in a common sample space \( T \). All possible simple random samples with replacement in the \( i^{th} \) trail can be defined by every possible configuration of an entire population of \( m \) bit strings. There are \( h = 2^{k \times m} \) such samples, where

- \( n \) number of variables
- \( k \) number of bits
- \( m \) is a sample size (\( \equiv \) number of chromosomes)
\( h \) number of all simple random samples with replacement \(( = 2^{k \times m})\)

We restrict an arbitrary countable set \( U \) to be set of all possible repeated dependent two trails without replacement

\[ (D^{(1)}_r \cap D^{(2)}_s) \]

for which

\[ P(D^{(1)}_r \cap D^{(2)}_s) \geq 0 \]

as a sample space, where for each

\[ D^{(i)}_y \text{ for } y = r \text{ or } s, \text{ where } r = 1, 2, \ldots, h \text{ and } s = 1, 2, \ldots, h \]

in a common sample space \( U \). Each conditional probability can be defined and can be obtained by the equation

\[ P(D^{(2)}_s | D^{(1)}_r) = \frac{P(D^{(1)}_r \cap D^{(2)}_s)}{P(D^{(1)}_r)} \geq 0 \text{ such that } P(D_r) \neq 0. \]

There are \( 2(2^{k \times m}) \) such repeated dependent two trials without replacement.

Consider \( q \in Q \) where \( Q \) is the discrete parameter space of the Markov chain process

\[ \{X_q, q = 0, 1, 2, \ldots\} \text{ (parameter homogeneous).} \]

A Markov chains requires a finite collection of states, denoted by

\[ U = \{D_1, D_2, \ldots, D_h\}, \]

an \( h \times h \) probability matrix \( P \) called a transition matrix and a probability vector

\[ p^{(0)} = ( P(D_1), P(D_2), \ldots, P(D_h) ) = (p^{(0)}_1, p^{(0)}_2, \ldots, p^{(0)}_h) \]

called the initial probability vector. We can also interpret \( p^{(0)} \) as stationary distribution, where

\[ h = 2^{k \times m} = \text{number of all simple random samples with replacement} \]

\[ \equiv \text{number of all states} = 2^{k \times m} \]
we Generate all possible combinations of states of unique chromosomes (= \(2^{(k \times m)}\)) and give each state a number.

We apply clonal selection algorithms on each generated state for \(n\)-iterations, where \(n\) is a large number. We Count for each state with specific number the number of times it appeared, Calculate the probability of each state \(P\), where

\[
P = \frac{\text{number of times it appeared}}{n}
\]

and Get the stationary distribution ordered by the state number and its probability \(P\).

For the second approach, let

\[L = \{L_1, L_2, \ldots, L_g\}\]

be a partition of (all possible combinations of states of unique chromosomes = \(U\)). We Get \(L_z\) for each state in combination and Combine states according their equal \(L_z\)'s ( \((\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_n)'s\))

Taking into account that \(L_z\)'s are unique for each set of states.

We restrict an arbitrary countable set \(L\) to be set of all possible outcomes in the \(i^{th}\) trail

\[L_z^{(i)} = (\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_n)\text{ for } z = 1, 2, \ldots, g (g = \text{number of states in } L)\]

for which

\[P(\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_n) > 0\]

as a sample space, where

\[
\bar{x}_k = \frac{x_{1k} + x_{2k} + \ldots + x_{mk}}{m} = \sum_{j=1}^{m} \frac{x_{jk}}{m} \text{ for } k = 1, 2, \ldots, n\text{ and } m = \text{sample size}
\]

and

\[x_{jk} = \text{measurement of the } k^{th} \text{ variable on the } j^{th} \text{ item.}\]

We restrict an arbitrary countable set \(L_1\) to be set of all possible repeated dependent two trails without replacement.
\((L_e(1) \cap L_d(2))\) for which
\[
P(L_e(1) \cap L_d(2)) \geq 0
\]
as a sample space, where for each
\[L_z(i)\] for \(z = e\) or \(d\), where \(e = 1, 2, \ldots, g\) and \(d = 1, 2, \ldots, g\)
in a common sample space \(L\). Each conditional probability can be defined and can be obtained by the equation
\[
P(L_d(2)|L_e(1)) = \frac{P(L_e(1) \cap L_d(2))}{P(L_e(1))} \geq 0 \text{ such that } P(L_e) \neq 0.
\]

Consider \(q \in Q\) where \(Q\) is the discrete parameter space of Markov chain process
\[
\{X_q, q = 0, 1, 2, \ldots \}\text{ (parameter homogeneous)}.
\]
A lumped Markov chains requires a finite collection of states, denoted by
\[
L = \{L_1, L_2, \ldots, L_g\},
\]
an \(g \times g\) probability matrix \(P\) called a transition matrix and a probability vector
\[
p^{(0)} = (P(L_1), P(L_2), \ldots, P(L_g)) = (p_1^{(0)}, p_2^{(0)}, \ldots, p_g^{(0)})
\]
called the initial probability vector. We can also interpret \(p^{(0)}\) as stationary distribution.

Let \(\bar{X}_q\) be an \((n)\)-dimensional random variable defined on a sample space \(L\) during \(q\) (initial value of time parameter), and \(\bar{X}_{q+1}\) be an \((n)\)-dimensional random variable defined on a sample space \(L\) during \(q + 1\). The possible outcomes for \(\bar{X}_q\) are
\[
L_1, L_2, \ldots, L_g,
\]
and the same holds for \(\bar{X}_{q+1}\).

By Theorems for the mean and covariance of the sampling distribution of \(\bar{X}\) and central limit Theorem, we conclude and have the modified forms of the Theorems for an \(n\)-dimensional random variable \(\bar{X}\) defined on a sample space \(L\).
we get sampling multivariate normal distribution of $X$, get expectation of $X$ and get covariance matrix of $X$. We conclude and have that the distribution of $X_q$ is the same as $X$, and the same holds for $X_{q+1}$.

Let $(X_q, X_{q+1})'$ be an $2n$-dimensional random variable defined on a sample space $L_1$ and be distributed as multivariate normal. Then the conditional distribution of $X_{q+1}$, given that $X_q = L_\star = \text{any state of } L$,
is multivariate normal and has

Mean $= \mu_\star$

and

Covariance $= \Sigma_\star$.

We select one state from $U$, get $L_\star$ for the selected state and replicate the selected state $2^{(k \times m)}$ times.

We apply clonal selection algorithms for one transition (iteration) only on the replicated states to produce new states. On the generated new states, we compute $L_z$. On the set of $L_z$'s computed, we compute conditional normal distribution

$N(\mu_\star, \Sigma_\star)$.

By substitution, we Compute non diagonal covariance of the covariance matrix of the distribution of

$(X_q, X_{q+1})'$
in Mean $= \mu_\star$. By a similar argument, we compute non diagonal covariance for all states of $L$, check that non diagonal covariance is the same and check that $\Sigma_\star$ is the same.

3. Main result

In this section, we shall state the main theorem.

**Theorem 3.1.** For any clonal selection algorithm, the following holds:

(1. Markov approach) The sequence of $(m \times n)$-dimensional random matrices
$X_0, X_1, X_2, \ldots$

converges in distribution to the $(m \times n)$-dimensional random matrix $X$ (has unique stationary distribution) if and only if for each $(a_1, a_2, \ldots, a_{(m \times n)})$

$$
\phi_{X_i}(a_1, a_2, \ldots, a_{(m \times n)}) \to \phi_{X}(a_1, a_2, \ldots, a_{(m \times n)}) \text{ as } i \to \infty
$$

(uniqueness of characteristic function), where

$n = \text{number of variables}$

$m = \text{number of sets of measurements on } n \text{ variables } = \text{sample size } = \text{number of chromosomes}$

$\phi_{X}(a_1, a_2, \ldots, a_{(m \times n)})$ is the characteristic function of $X$ of $m \times n$ real variables.

(a) If $P$ is a transition matrix for any clonal selection algorithm (regular chain) and the probability vector $t$ is a fixed point of the matrix $P$, then

$$
\sum_{j=1}^{2^{(k \times m)}} \sum_{i=1}^{2^{(k \times m)}} t_{ij} P_{ij}^n \to (\sum_{j=1}^{2^{(k \times m)}} t_j = 1) \text{ as } n \to \infty.
$$

(b) For any clonal selection algorithm, a real valued function

$$
f(x_1, x_2, \ldots, x_n), \text{ where } a_i \leq x_i \leq b_i \text{ for } i = 1, 2, \ldots, n
$$

is one that contains an infinite number of Markov chains (every chain have different Unique chromosomes for purely successful optimization and have different globally optimum value(s)).

(2. lumped Markov approach) The sequence of $n$-dimensional random vectors

$\mathbf{X}_0, \mathbf{X}_1, \mathbf{X}_2, \ldots$

converges in distribution to the $n$-dimensional random vector $\mathbf{X}$ (has unique stationary multivariate normal distribution) if and only if for each $(a_1, a_2, \ldots, a_n)$

$$
\phi_{\mathbf{X}_i}(a_1, a_2, \ldots, a_n) \to \phi_{\mathbf{X}}(a_1, a_2, \ldots, a_n) \text{ as } i \to \infty
$$

(uniqueness of characteristic function), where

$n = \text{number of variables}$. 
\( \phi_{\mathbf{X}}(a_1, a_2, \ldots, a_n) \) is the characteristic function of \( \mathbf{X} \) of \( n \) real variables.

(a) All possible conditional multivariate normal distributions of transition probability matrix have the following form: (see proof of main theorem)

The conditional distribution of \( \mathbf{X}_{q+1} \), given that \( \mathbf{X}_q = L_\star \), is multivariate normal and has

\[
\text{Mean} = \mu_\star = \mu + \Sigma_{t_1 t_2} \Sigma^{-1}(L_\star - \mu)
\]

and

\[
\text{Covariance} = \Sigma_\star = \Sigma - \Sigma_{t_1 t_2} \Sigma^{-1} \Sigma_{t_2 t_1}.
\]

The distribution of \( \mathbf{X} \) is stationary multivariate normal distribution.

(b) For any clonal selection algorithm, a real valued function

\[
f(x_1, x_2, \ldots, x_n), \text{ where } a_i \leq x_i \leq b_i \text{ for } i = 1, 2, \ldots, n
\]

is one that contains an infinite number of lumped Markov chains (every chain, have different Unique chromosomes for purely successful optimization and have different globally optimum value(s)).

4. **Proof of the main result**

   In this section, we prove the main result in Theorem 3.1. We start with a useful theorem.

**Theorem 4.1.** Let \((S, \beta, P)\) be a probability space and let \( T \) denote the set of all \( x \) in \( S \) for which \( P(x) > 0 \). Then \( T \) is countable.

We shall prove Theorem 3.1 in nine steps.

**Proof of Theorem 3.1.** Step 1. For clonal selection algorithms, we define a probability space

\((S, \beta, P)\).

For clonal selection algorithms, let a real valued function

\[
f(x_1, x_2, \ldots, x_n), \text{ where } a_i \leq x_i \leq b_i \text{ for } i = 1, 2, \ldots, n.
\]
we restrict an arbitrary uncountable set

\[ S = \{ a_1 \leq x_1 \leq b_1 \} \]

to be a subset of the real line

\[ \mathbb{R}, \text{ or of } n\text{-space } \mathbb{R}^n \]

\((S=\{a_i \leq x_i \leq b_i \text{ for } i = 1, 2, \ldots, n\})\)
as a sample space, we shall assume that this set is a Borel set. The Borel subsets of \( S \) themselves form a Boolean \( \sigma \)-algebra \( \beta \). A nonnegative completely additive set function \( P \) defined on \( \beta \) with \( P(S) = 1 \) is called a probability measure. We have a probability space

\((S, \beta, P)\).

Step 2. We prove that \( T \) is a countable subset of \( S \), and define \( n \)-dimensional random variable defined on a sample space \( T \).

**Proof** By Theorem 4.1, we restrict an arbitrary countable set \( T \) to be set of all

\[ x_1 \text{ in } S=\{a_1 \leq x_1 \leq b_1 \} \text{ for which } P(x_1) > 0 \]  
or of all

\[ (x_1, x_2, \ldots, x_n) \text{ in } S=\{a_i \leq x_i \leq b_i \text{ for } i = 1, 2, \ldots, n\} \text{ for which } P(x_1, x_2, \ldots, x_n) > 0 \]
as a sample space. \( T \) denotes a countable subset of \( S \)(whenever we use a set \( T \) of real numbers as a sample space, or, more generally, whenever we use a set \( T \) in \( n \)-space as a sample space, we shall assume that this set is a Borel set).

Let a sample space \( T \) of possible solutions to be coded as strings of \( k \) bits \{0,1\} and let each possible configuration have a fitness

\[ f_i, i = 1, \ldots, 2^k, \]

let \( f^* \) be the globally optimum value. Hence \( T \) is countable.

Let \( T \) be a set in
n-space for some $n \geq 1$

(a Borel set in $n$-space for some $n \geq 1$) and if $\tau$ consists of all subsets of $T$, the probability function $P$ is completely determined on $\tau$. We have a probability space

$$(T, \tau, P).$$

Next, we define an $n$-dimensional random variable defined on a sample space $T$.

Let $X_1$ be a one-dimensional random variable defined on a sample space $T$.

$$X_1(x_1) = x_1 \text{ for each } x_1 \text{ in } T$$

or let $X$ be an $n$-dimensional random variable defined on a sample space $T$.

$$X = (X_1(x_1) = x_1, X_2(x_2) = x_2, ..., X_n(x_n) = x_n)' = (x_1, x_2, ..., x_n)$$

for each

$$(x_1, x_2, ..., x_n) \text{ in } T.$$

We will use notation $x_{jk}$ to indicate the particular value of the $k^{th}$ variable that is observed on the $j^{th}$ item, or trail. That is, $x_{jk} = \text{measurement (commonly called data) of the } k^{th} \text{ variable on the } j^{th} \text{ item}$ consequently, $m$ measurements on $n$ variables can be displayed as a rectangular array. We need to make assumptions about the variables whose observed values constitute the data set.

Step 3. We prove that $U$ is a countable set of all possible simple random samples with replacement in the $i^{th}$ trail, and define an $(m \times n)$-dimensional random variable defined on a sample space $U$.

**Proof.** Let the $jk^{th}$ entry in the data matrix be the random variables $X_{jk}$. Each set of measurements

$$X_j \text{ on } n \text{ variables}$$

is a random vector, and we have the random matrix.
A random sample can now be defined.

\[ \mathbf{X} = \begin{pmatrix} X_{11} & X_{12} & \ldots & X_{1k} & \ldots & X_{1n} \\ X_{21} & X_{22} & \ldots & X_{2k} & \ldots & X_{2n} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ X_{j1} & X_{j2} & \ldots & X_{jk} & \ldots & X_{jn} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ X_{m1} & X_{m2} & \ldots & X_{mk} & \ldots & X_{mn} \end{pmatrix} = \begin{pmatrix} \mathbf{X}'_1 \\ \mathbf{X}'_2 \\ \vdots \\ \mathbf{X}'_j \\ \vdots \\ \mathbf{X}'_m \end{pmatrix} \]

\[ \text{A random sample can now be defined.} \]

\[ \mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_j, \ldots, \mathbf{X}_m \]

form a random sample if their joint probability mass function is given by the product

\[ P(\mathbf{X}_1)P(\mathbf{X}_2) \ldots P(\mathbf{X}_j) \ldots P(\mathbf{X}_m), \]

where

\[ P(\mathbf{X}_j) = P(\mathbf{x}_j) = \sum_{x_{j1}, x_{j2}, \ldots, x_{jk}} \mathbf{x}_{jk} \]

is the probability mass function for the \( j \)th row vector.

\[ \{ \text{Joint probability mass function of } \mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_j, \ldots, \mathbf{X}_m \} \]

\[ = P(\mathbf{X}) = P(\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_j, \ldots, \mathbf{X}_m) \]

\[ = P(\mathbf{X}_1)P(\mathbf{X}_2) \ldots P(\mathbf{X}_j) \ldots P(\mathbf{X}_m). \]

We restrict an arbitrary countable set \( U \) to be set of all possible simple random samples with replacement in the \( i \)th trail

\[ D_y^{(i)} \text{ for } y = 1, 2, \ldots, h \text{ for which} \]

\[ P((x_{11}, x_{12}, \ldots, x_{1n}), (x_{21}, x_{22}, \ldots, x_{2n}), \ldots, (x_{m1}, x_{m2}, \ldots, x_{mn})) > 0 \]

as a sample space, where for each
\((x_{j1}, x_{j2}, \ldots, x_{jn}) \text{ for } j = 1, 2, \ldots, m\)

in a common sample space \(T\).

All possible simple random samples with replacement in the \(i^{th}\) trail can be defined by every possible configuration of an entire population of \(m\) bit strings. There are \(h = 2^{k \times m}\) such samples. Hence \(U\) is countable.

Let \(U\) be a set in \((m \times n)\)-space for some 
\[
(m \times n) \geq (2 \times 1) = 2
\]
(a Borel set in \((m \times n)\)-space for some \((m \times n) \geq 2\)) and if \(\nu\) consists of all subsets of \(U\), the probability function \(P\) is completely determined on \(\nu\). We have a probability space 
\((U, \nu, P)\).

Next, we define an \((m \times n)\)-dimensional random variable defined on a sample space \(U\).

Let \(X\) be an \((m \times n)\)-dimensional random variable defined on a sample space \(U\).

\[
X = (X_1, X_2, \ldots, X_m)
\]

\[
= ( (X_{11}(x_{11}) = x_{11}, X_{12}(x_{12}) = x_{12}, \ldots, X_{1n}(x_{1n}) = x_{1n}),
X_{21}(x_{21}) = x_{21}, X_{22}(x_{22}) = x_{22}, \ldots, X_{2n}(x_{2n}) = x_{2n}),
\ldots, (X_{m1}(x_{m1}) = x_{m1}, X_{m2}(x_{m2}) = x_{m2}, \ldots, X_{mn}(x_{mn}) = x_{mn}) )
\]

\[
= ( (x_{11}, x_{12}, \ldots, x_{1n}), (x_{21}, x_{22}, \ldots, x_{2n}), \ldots, (x_{m1}, x_{m2}, \ldots, x_{mn}) )
\]

for each

\[
(x_{11}, x_{12}, \ldots, x_{1n}), (x_{21}, x_{22}, \ldots, x_{2n}), \ldots, (x_{m1}, x_{m2}, \ldots, x_{mn})
\]

in \(U\), where

\[
X_j = (x_{j1}, x_{j2}, \ldots, x_{jn}) \text{ for } j = 1, 2, \ldots, m
\]

for each

\[(x_{j1}, x_{j2}, \ldots, x_{jn}) \text{ in } T\]
(\(X_j\) are \(n\)-dimensional random variables defined on a common sample space \(T\)).

Step 4. We prove that \(U_1\) is a countable set of all possible repeated dependent two trails without replacement, define an \(2(m \times n)\)-dimensional random variable defined on a sample space \(U_1\), and describe Markov chain process.

**Proof.** We restrict an arbitrary countable set \(U_1\) to be set of all possible repeated dependent two trails without replacement

\[
(D^{(1)}_r \cap D^{(2)}_s) \text{ for which } P(D^{(1)}_r \cap D^{(2)}_s) \geq 0
\]

as a sample space, where for each

\(D^{(i)}_y\) for \(y = r\) or \(s\), where \(r = 1, 2, \ldots, h\) and \(s = 1, 2, \ldots, h\)

in a common sample space \(U\). Each conditional probability can be defined and can be obtained by the equation

\[
P(D^{(2)}_s | D^{(1)}_r) = \frac{P(D^{(1)}_r \cap D^{(2)}_s)}{P(D^{(1)}_r)} \geq 0 \text{ such that } P(D_r) \neq 0.
\]

There are \(2^{(2^k \times m)}\) such repeated dependent two trails without replacement. Hence \(U_1\) is countable.

Let \(U_1\) be a set in

\(2(m \times n)\)-space for some \(2(m \times n) \geq 2(2 \times 1) = 4\)

(a Borel set in \(2(m \times n)\)-space for some \(2(m \times n) \geq 4\)) and if \(\nu_1\) consists of all subsets of \(U_1\), the probability function \(P\) is completely determined on \(\nu_1\). We have a probability space

\((U_1, \nu_1, P)\).

We define an \(2(m \times n)\)-dimensional random variable defined on a sample space \(U_1\).

Consider \(q \in Q\) where \(Q\) is the discrete parameter space of the Markov chain process

\[\{X_q, q = 0, 1, 2, \ldots\}\](parameter homogeneous).

Let
\(X_q\) be an \((m \times n)\)-dimensional random variable defined on a sample space \(U\) during \(q\) (initial value of time parameter), and

\(X_{q+1}\) be an \((m \times n)\)-dimensional random variable defined on a sample space \(U\) during \(q+1\). The possible outcomes for \(X_q\) are

\[D_1, D_2, \ldots, D_h,\]

and the same holds for \(X_{q+1}\).

Let \(X\) be an \(2(m \times n)\)-dimensional random variable defined on a sample space \(U_1\).

\[X = (X_q, X_{q+1}) = (X_q = D_r, X_{q+1} = D_s) = (D_r, D_s) = (X_{b_1}, X_{b_2})\]

\[= ((X_{b_11}, X_{b_12}, \ldots, X_{b_1m}), (X_{b_21}, X_{b_22}, \ldots, X_{b_2m}))\]

for each \((D_r, D_s)\) in \(U_1\), where \(X_v = D_y = (X_{b_1}, X_{b_2}, \ldots, X_{b_m}) = X_b\)

\[= ((X_{b_11}(x_{b_11}) = x_{b_11}, X_{b_12}(x_{b_12}) = x_{b_12}, \ldots, X_{b_1m}(x_{b_1m}) = x_{b_1m}),\]

\[X_{b_21}(x_{b_21}) = x_{b_21}, X_{b_22}(x_{b_22}) = x_{b_22}, \ldots, X_{b_2m}(x_{b_2m}) = x_{b_2m}),\]

\[\ldots, (X_{b_m1}(x_{b_m1}) = x_{b_m1}, X_{b_m2}(x_{b_m2}) = x_{b_m2}, \ldots, X_{b_mm}(x_{b_mm}) = x_{b_mm}))\]

\[= ((x_{b_11}, x_{b_12}, \ldots, x_{b_1m}), (x_{b_21}, x_{b_22}, \ldots, x_{b_2m}), \ldots, (x_{b_m1}, x_{b_m2}, \ldots, x_{b_mm}))\]

for \(v = q\) or \(q+1\), \(y = r\) or \(s\) and \(b = b_1\) or \(b_2\) for each

\[((x_{b_11}, x_{b_12}, \ldots, x_{b_1m}), (x_{b_21}, x_{b_22}, \ldots, x_{b_2m}), \ldots, (x_{b_m1}, x_{b_m2}, \ldots, x_{b_mm}))\]

in \(U\)

\((X_v\) are \((m \times n)\)-dimensional random variables defined on a common sample space \(U\)).

Next, we describe Markov chain process.

A Markov chains requires a finite collection of states, denoted by
$U = \{D_1, D_2, \ldots, D_h\},$

an $h \times h$ probability matrix $\mathbf{P}$ called a transition matrix and a probability vector

$$\mathbf{p}^{(0)} = (P(D_1), P(D_2), \ldots, P(D_h)) = (p_1^{(0)}, p_2^{(0)}, \ldots, p_h^{(0)})$$

called the initial probability vector. We can also interpret $\mathbf{p}^{(0)}$ as stationary distribution.

Step 5. We prove that clonal selection algorithms are regular chains.

**Proof.** From Step 2, We get unique chromosomes ($= 2^k$). From Step 3 and Step 4, we Generate all possible combinations of states of unique chromosomes ($= 2^{(k \times m)}$) and give each state a number.

We apply clonal selection algorithms on each generated state for $n$-iterations, where $n$ is a large number. We Count for each state with specific number the number of times it appeared, Calculate the probability of each state ($P$), where

$$P = \frac{\text{number of times it appeared}}{n}$$

and Get the stationary distribution ordered by the state number and its probability $P$. All states in this chain will be ergodic.

Now we recall a theorem in [12].

**Theorem 4.2.** If $\mathbf{P}$ is a transition matrix for ergodic chain, then:

1. There is a unique probability vector fixed point $t : t = t\mathbf{P}$.
2. All components of $t$ are positive.
3. If $h_j^{(n)}$ is the average number of times the process is in state $j$ in the first $n$ steps, then for any $\epsilon > 0$,

$$P(\left| h_j^{(n)} - t_j \right| > \epsilon) \to 0 \text{ as } n \to \infty$$

no matter what the starting state is.

We conclude and have that all stationary distributions

$$= 2^{(k \times m)}$$

are the same unique probability vector fixed point. If $\mathbf{P}$ is a transition matrix for clonal selection algorithms and the probability vector $t$ is a fixed point of the matrix $\mathbf{P}$, then
\[
\sum_{j=1}^{2^{(k \times m)}} \sum_{i=1}^{2^{(k \times m)}} t_i P_{ij}^n \rightarrow (\sum_{j=1}^{2^{(k \times m)}} t_j = 1) \text{ as } n \to \infty.
\]

Hence clonal selection algorithms are ergodic chains.

We take each generated state and its chain, get the position numbers sequences for each unique state (of \(2^{k \times m}\)) in the chain, starting with index zero in the initial state and check that all sequences of positions are even positions, odd positions, or random positions sequences. Hence clonal selection algorithms are regular chains.

For any clonal selection algorithm, a real valued function
\[
f(x_1, x_2, ..., x_n), \text{ where } a_i \leq x_i \leq b_i \text{ for } i = 1, 2, ..., n
\]
is one that contains an infinite number of Markov chains (every chain have different globally optimum value(s) and have Unique chromosomes for purely successful optimization).

Step 6. We describe the partition \(L\) of all possible combinations of states of unique chromosomes, prove that \(L\) is a countable set and define an \(n\)-dimensional random variable defined on a sample space \(L\).

Let
\[
L = \{L_1, L_2, \ldots, L_g\}
\]

be a partition of all possible combinations of states of unique chromosomes. We get \(L_z\) for each state in combination and Combine states according their equal
\[
L_z \text{'s } (\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_n) \text{'s }
\]
-Taking into account that \(L_z\)'s are unique for each set of states. We prove that \(L\) is a countable set.

We restrict an arbitrary countable set \(L\) to be set of all possible outcomes in the \(i^{th}\) trail
\[
L_z^{(i)} = (\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_n) \text{ for } z = 1, 2, \ldots, g
\]
for which
\[
P(\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_n) > 0
\]
as a sample space, where
\[ \bar{x}_k = \frac{x_{1k} + x_{2k} + \ldots + x_{mk}}{m} = \sum_{j=1}^m \frac{x_{jk}}{m} \text{ for } k = 1, 2, \ldots, n \]

and

\[ x_{jk} = \text{measurement of the } k^{th} \text{ variable on the } j^{th} \text{ item.} \]

Hence \( L \) is a countable.

Let \( L \) be a set in

\[ n \text{-space for some } n \geq 1 \]

and if \( \zeta \) consists of all subsets of \( L \), the probability function \( P \) is completely determined on \( \zeta \). We have a probability space

\[ (L, \zeta, P). \]

Next, we define an \( n \)-dimensional random variable defined on a sample space \( L \).

Let \( \mathbf{X} \) be an \( n \)-dimensional random variable defined on a sample space \( L \).

For each \((\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_n)\) in \( L \),

\[ \mathbf{X} = (X_1(\bar{x}_1) = \bar{x}_1, X_2(\bar{x}_2) = \bar{x}_2, \ldots, X_n(\bar{x}_n) = \bar{x}_n) \]

\[ = (\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_n). \]

Step 7. We prove that \( L_1 \) is set of all possible repeated dependent two trails without replacement, define an \( 2n \)-dimensional random variable defined on a sample space \( L_1 \) and describe lumped Markov chains process.

**Proof.** We restrict an arbitrary countable set \( L_1 \) to be set of all possible repeated dependent two trails without replacement

\( (L_e^{(1)} \cap L_d^{(2)}) \)

for which

\[ P(L_e^{(1)} \cap L_d^{(2)}) \geq 0 \]

as a sample space, where for each

\( L_z^{(i)} \) for \( z = e \) or \( d \), where \( e = 1, 2, \ldots, g \) and \( d = 1, 2, \ldots, g \)

in a common sample space \( L \). Each conditional probability can be defined
and can be obtained by the equation

\[ P(L_d^{(2)} \mid L_e^{(1)}) = \frac{P(L_e^{(1)} \cap L_d^{(2)})}{P(L_e^{(1)})} \geq 0 \text{ such that } P(L_e) \neq 0. \]

Hence \( L_1 \) is countable.

Let \( L_1 \) be a set in

\[ 2n \text{-space for some } 2n \geq 2(1) = 2 \]

(a Borel set in 2n-space for some 2n \( \geq 2 \)) and if \( \zeta_1 \) consists of all subsets of \( L_1 \), the probability function \( P \) is completely determined on \( \zeta_1 \). We have a probability space

\[ (L_1, \zeta_1, P). \]

We define an 2n-dimensional random variable defined on a sample space \( L_1 \).

Consider \( q \in Q \) where \( Q \) is the discrete parameter space of the Markov chain process

\[ \{X_q, q = 0, 1, 2, \ldots \} \text{(parameter homogeneous)}. \]

Let \( \mathbf{X}_q \) be an (n)-dimensional random variable defined on a sample space \( L \) during \( q \) (initial value of time parameter), and

\( \mathbf{X}_{q+1} \) be an (n)-dimensional random variable defined on a sample space \( L \) during \( q + 1 \). The possible outcomes for \( \mathbf{X}_q \) are

\[ L_1, L_2, \ldots, L_g, \]

and the same holds for \( \mathbf{X}_{q+1} \).

Let \( \mathbf{X} \) be an 2n-dimensional random variable defined on a sample space \( L_1 \).

\[ \mathbf{X} = (\mathbf{X}_q, \mathbf{X}_{q+1}) = (\mathbf{X}_q = L_e, \mathbf{X}_{q+1} = L_d) = (L_e, L_d) \]

\[ = (\mathbf{X}_{t_1}, \mathbf{X}_{t_2}) \]

\[ = ((\mathbf{X}_{t_11}, \mathbf{X}_{t_12}, \ldots, \mathbf{X}_{t_1n}),(\mathbf{X}_{t_21}, \mathbf{X}_{t_22}, \ldots, \mathbf{X}_{t_2n})) \]
for each \((L_e, L_d)\) in \(L_1\), where \(\bar{X}_v = L_z = \bar{X}_t\)

\[
(\bar{X}_{t_1}(\bar{x}_{t_1}), \bar{X}_{t_2}(\bar{x}_{t_2}), \ldots, \bar{X}_{t_n}(\bar{x}_{t_n}) = \bar{x}_{t_n})
\]

\[
= (\bar{x}_{t_1}, \bar{x}_{t_2}, \ldots, \bar{x}_{t_n})
\]

for \(v = q\) or \(q + 1\), \(z = e\) or \(d\) and \(t = t_1\) or \(t_2\) for each

\((\bar{x}_{t_1}, \bar{x}_{t_2}, \ldots, \bar{x}_{t_n})\) in \(L\).

\((\bar{X}_v\) are \(n\)-dimensional random variables defined on a common sample space \(L\)).

Next, we describe the lumped Markov chains process.

A Markov chains requires a finite collection of states, denoted by

\[L = \{L_1, L_2, \ldots, L_g\}\]

an \(g \times g\) probability matrix \(P\) called a transition matrix and a probability vector

\[
p^{(0)} = ( P(L_1), P(L_2), \ldots, P(L_g) ) = (p_1^{(0)}, p_2^{(0)}, \ldots, p_g^{(0)})
\]

called the initial probability vector. We can also interpret \(p^{(0)}\) as stationary distribution.

Step 8. By Theorems for the mean and covariance of the sampling distribution of \(\bar{X}\) and central limit Theorem, we conclude and have the following modified forms of the Theorems for an \(n\)-dimensional random variable \(\bar{X}\) defined on a sample space \(L\).

For the mean and covariance Theorems of the sampling distribution of \(\bar{X}\) and central limit Theorem (see [11]).

**The modified form of the mean theorem.** If

\[D_1, D_2, \ldots, D_h\]

represent all possible simple random samples with replacement from a common joint probability mass function

\[P(x_1, x_2, \ldots, x_n)\]

(the parent population, whatever its form, have a mean
\( \mu = (\mu_1, \mu_2, \ldots, \mu_n)' \)

and a finite covariance \( \Sigma_1 \), then an \( n \)-dimensional density for the random vector

\[ \mathbf{X} = (X_1, X_2, \ldots, X_n)' \]

(sampling distribution of \( \mathbf{X} \)) \( r(\mathbf{X}) \)

is centered around the population mean, regardless of sample size \( (E(\mathbf{X}) = \mu) \).

**The modified form of the covariance theorem.** Each covariance

\[ \sigma_{ik}, i, k = 1, 2, \ldots, n \]

of the distribution of \( \mathbf{X} \) decreases with increasing sample size; that is, the distribution of \( \mathbf{X} \) becomes more concentrated around the population mean as the sample size gets larger (a covariance of \( \mathbf{X} \) is denoted by

\[ \text{cov}(\mathbf{X}) = \frac{\Sigma_1}{m}, \text{ where } m \]

(number of chromosomes) is a sample size).

For practical problem, often

\[ m = 100 = \text{chromosomes and bit} = k > 50; \]

thus there may be more than \( 2^{5000} \) possible states in the chain.

The distribution of \( \mathbf{X} \) becomes more symmetrical as the sample size gets larger and is approximately

\[ N_n(\mu, \Sigma = \frac{\Sigma_1}{m}) \]

for large sample size. An \( n \)-dimensional normal density for the random vector \( \mathbf{X} \) has the form

\[ f(\mathbf{x}) = \frac{1}{(2\pi)^{\frac{n}{2}}|\Sigma|^\frac{1}{2}} e^{-\frac{(\mathbf{x} - \mu)'\Sigma^{-1}(\mathbf{x} - \mu)}{2}} \]

where \( -\infty < \bar{x}_i < \infty, i = 1, 2, \ldots, n. \)

**The modified form of central limit Theorem.** Let

\[ X_1, X_2, \ldots, X_m \]
be independent observations from any population with mean \( \mu \) and finite covariance \( \Sigma_1 \). Then \( \bar{X} \) has an approximate
\[ N_n(\mu, \frac{\Sigma_1}{m}) \]
distribution for large sample sizes. Here \( m \) should also be large relative to \( n \). The approximation provided by the central limit theorem applies to discrete, as well as continuous, multivariate populations.

Theorems for the mean and covariance of the sampling distribution of \( \bar{X} \) and central limit Theorem apply to sampling of finite populations if the sampling fraction is 5 percent or smaller (the sample size \( m \) is small relative to the population size \( M \); that is, with fraction small).

Step 9. We prove that clonal selection algorithms are lumped Markov chains.

**Proof.** From Step 2, we get unique chromosomes (= \( 2^k \)). From Step 3 and Step 4, we Generate all possible combinations of states of unique chromosomes(= \( 2^{(k \times m)} \)) and give each state a number.

From Step 6, we get \( L_z \) for each state in combination and Combine states according to their equal \( L_z \)’s - Taking into account that \( L_z \)’s are unique for each set of states. Let each set of states to be a separate new state. From Step 6 and Step 7, we identify new state space of \( L_z \)’s based on set of states of unique \( L_z \)’s, define
\[ L = \{L_1, L_2, \ldots, L_g\} \]
and define a random variable \( \bar{X} \) on state space \( L \).

From Step 8, we get sampling distribution of
\[ \bar{X} \ (N_n(\mu, \frac{\Sigma_1}{m} = \Sigma)) \],
get expectation of sampling distribution
\[ E(\bar{X}) = \mu, \]
where \( \mu \) is the expectation of parent population and get covariance matrix of
\[ \bar{X} \ (= \Sigma = \frac{\Sigma_1}{m}), \]
where \( m \) = sample size and \( \Sigma_1 \) = covariance matrix of parent population. We conclude and have that the distribution of
$\mathbf{X}_q$ is $N_n(\mu, \Sigma)$, and the same holds for $\mathbf{X}_{q+1}$.

Let

$$\mathbf{X} = (\mathbf{X}_q, \mathbf{X}_{q+1})'$$

be distributed as

$$N_{2n}(\mu^*, \Sigma^*)$$

with

$$\mu^* = (\mu = \mu_{t_1}, \mu = \mu_{t_2})',$$

$$\Sigma^* = \begin{pmatrix}
\Sigma_{t_1t_1} = \Sigma & \Sigma_{t_1t_2} \\
\Sigma_{t_2t_1} & \Sigma_{t_2t_2} = \Sigma
\end{pmatrix},$$

and

$$|\Sigma_{t_2t_2}| > 0.$$ 

Then the conditional distribution of $\mathbf{X}_{q+1}$, given that $\mathbf{X}_q = L_*$, is normal and has

Mean = $\mu_* = \mu + \Sigma_{t_1t_2}\Sigma^{-1}(L_* - \mu)$

and

Covariance = $\Sigma_* = \Sigma - \Sigma_{t_1t_2}\Sigma^{-1}\Sigma_{t_2t_1}$. 

We select one state from $U$, get $L_*$ for the selected state and replicate the selected state $2^{k \times m}$ times.

We apply clonal selection algorithms for one transition (iteration) only on the replicated states to produce new states.

On the generated new states, we compute $L_z$. On the set of $L_z$'s computed, we compute conditional normal distribution

$$N(\mu_*, \Sigma_*).$$

We Compute $\Sigma_{t_1t_2}$, where
\[ \mu_* = \mu + \Sigma_{t_1t_2} \Sigma^{-1}(L_* - \mu). \]

By a similar argument, we compute \( \Sigma_{t_1t_2} \) for all states of \( L \), check that \( \Sigma_{t_1t_2} \) is the same and check that \( \Sigma_* \) is the same.

Hence clonal selection algorithms are lumped Markov chains.

We compute all possible conditional normal distributions of transition matrix and can also interpret the distribution of \( \mathbf{X} \) as stationary normal distribution.

On the basis of Steps 1-9, we complete the proof of Theorem 3.1.

5. Proposed algorithms

We prepared programs by using MATLAB 7.5. Afterwards, two algorithms are designed and executed to observe the characteristics and to know the behavior of the following:

5.1. Clonal selection algorithms (as Markov chains) We named the proposed algorithm regular optimization analysis (ROA), the basic steps of the ROA algorithm are as follows:

1. Input number of bits \( k \).
2. Get unique chromosomes \( = 2^k \).
3. Input number of chromosomes \( m \).
4. Get number of states \( = 2^{(k \times m)} \).
5. Generate all possible combinations of states of unique chromosomes \( = 2^{(k \times m)} \).
6. Give each state a number.
7. Pick one state randomly.
8. Apply clonal selection algorithms on the state for \( n \)-iterations, where \( n \) is a large number.
9. Count for each state with specific number the number of times it appeared.
10. Calculate the probability of each state \( (P) \), where

\[ P = \frac{\text{number of times it appeared}}{n}. \]

11. Get the stationary distribution ordered by the state number and its prob-
ability $P$.

12. Taking the randomly chosen state and its chain. Get the position numbers sequences for each unique state (of $2^{k \times m}$) in the chain, starting with index zero in the initial state.

Check if all sequences of positions are either

(A) Even positions or

(B) Odd positions only then the chain is cyclic else if also random positions sequences appear then the chain is regular.

5.2. Clonal selection algorithms (as lumped Markov chains) We named the proposed algorithm Central-Markovian optimization analysis (CMOA), the basic steps of the CMOA algorithm are as follows:\

1. Input number of bits $k$.
2. Get unique chromosomes = $2^k$.
3. Input number of chromosomes $m$.
4. Get number of states = $2^{(k \times m)}$.
5. Get all possible combinations of states of unique chromosomes = $U$ (= also $2^{(k \times m)}$).
6. Get $L_z$ for each state in combination.
7. Combine states according to their equal $L_z$’s - Taking into account that $L_z$’s are unique for each set of states.
8. Let each set of states to be a separate new state.
9. Identify new state space of $L_z$’s based on set of states of unique $L_z$’s.
10. Define $L = \{L_1, L_2, \ldots, L_g\}$.
11. Define a random variable $\mathbf{X}$ on state space $L$.
12. Get sampling distribution of $\mathbf{X}$ ($N_n(\mu, \frac{\Sigma}{m} = \Sigma)$).
13. Get expectation of sampling distribution $E(\mathbf{X}) = \mu$, where $\mu$ is the expectation of parent population.
14. Get covariance matrix of $\mathbf{X}$ ($= \Sigma = \frac{\Sigma_1}{m}$), where $m$ = sample size and $\Sigma_1$ = covariance matrix of parent population.
15. Let $(\mathbf{X}_q, \mathbf{X}_{q+1})$’be distributed as ($N_{2n}(\mu^*, \Sigma^*)$) with

$$
\mu^* = (\mu = \mu_{t_1}, \mu = \mu_{t_2})',
$$

$$
\Sigma^* = \begin{pmatrix}
\Sigma_{t_1 t_1} = \Sigma & \Sigma_{t_1 t_2} = \Sigma \\
\Sigma_{t_2 t_1} = \Sigma & \Sigma_{t_2 t_2} = \Sigma
\end{pmatrix}, \text{ and } |\Sigma_{t_2 t_2}| > 0,
$$

where $q$ is the initial value of time parameter $Q$ of the Markov chain process

$$
\{\mathbf{X}_q, q = 0, 1, 2, \ldots\}$$
that the possible outcomes for $\mathbf{X}_q$ are $L_1, L_2, \ldots, L_g$ and the same holds for $\mathbf{X}_{q+1}$.

16. Get $E(\mathbf{X}_q) = \mu$.
17. Get covariance matrix of $\mathbf{X}_q = \Sigma$.
18. Get $E(\mathbf{X}_{q+1}) = \mu$.
19. Get covariance matrix of $\mathbf{X}_{q+1} = \Sigma$.
20. Group states on $U$ such that

   (1) states are globally optimal (all members of states are identical).
   (2) All or Some members of states have globally optimum values.
   (3) states are not globally optimal and all members of states are identical.
   (4) All or Some members of states have not globally optimum values.

21. Select one state randomly from each group on $U$.
22. Get $L_*$ for the selected states.
23. Replicate each selected state (from each group) $2^{(k \times m)}$ times.
24. Apply clonal selection algorithms for one transition (iteration) only on each replicated selected state to produce new states.
25. On the generated new states, compute $L_z$.
26. On the set of $L_z$'s computed, compute conditional normal distribution $N(\mu_*, \Sigma_*)$.
27. Compute $\Sigma_{t_1 t_2}$, where $\mu_* = \mu + \Sigma_{t_1 t_2} \Sigma^{-1}(L_* - \mu)$.
28. check that $\Sigma_*$ is the same for all four groups.
29. If $\Sigma_{t_1 t_2}$ for the four groups of states are equal then the process is Markov chain.
30. Substitute all possible $L_z$'s to compute all conditional expected values $\mu_*$. As a result this computes all possible conditional normal distributions and The distribution of $\mathbf{X}$ is stationary multivariate normal distribution.

6. Numerical results and discussion

6.1. A numerical example for clonal selection algorithms (as Markov chains)

Consider the following: $f(x) = x \cdot \sin(10\Pi \cdot x) + 1, x \in [1.7, 2]$

$k = 2$ bits, $m = 6$ chromosomes, $n = 70000$ iterations, probability of mutation = 0.9

Unique chromosomes = $\{ 00 = 1.700000, 01 = 1.800000, 10 = 1.900000 \}$
11 = 2.000000 }

Globally optimum value = 1.900000

All possible combinations of states of unique chromosomes =

\{ 0:(00, 00, 00, 00, 00, 00),
1:(00, 00, 00, 00, 00, 01),
\ldots, 4095:(11, 11, 11, 11, 11, 11) \} 

<table>
<thead>
<tr>
<th>Number of iteration</th>
<th>Number of state</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>162</td>
</tr>
<tr>
<td>1</td>
<td>3835</td>
</tr>
<tr>
<td>2</td>
<td>3315</td>
</tr>
<tr>
<td>\vdots</td>
<td>\vdots</td>
</tr>
<tr>
<td>69997</td>
<td>63</td>
</tr>
<tr>
<td>69998</td>
<td>424</td>
</tr>
<tr>
<td>69999</td>
<td>1403</td>
</tr>
</tbody>
</table>

Table(1) $n = 70000$ iterations

<table>
<thead>
<tr>
<th>Ordered number of state</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.000029</td>
</tr>
<tr>
<td>1</td>
<td>0.000014</td>
</tr>
<tr>
<td>2</td>
<td>0.000014</td>
</tr>
<tr>
<td>\vdots</td>
<td>\vdots</td>
</tr>
<tr>
<td>4093</td>
<td>0.000129</td>
</tr>
<tr>
<td>4094</td>
<td>0.000129</td>
</tr>
<tr>
<td>4095</td>
<td>0.004600</td>
</tr>
</tbody>
</table>

Table(2)stationary distribution of Table(1)

Classification of chain (regular)
162 to 3835(neither even nor odd)
Consider the following:

\[ f(x) = x \cdot \sin(10\pi \cdot x) + 1, \quad x \in [1.7, 2] \]

\[ k = 2, \quad m = 6, \quad n = 90000 \text{ iterations, probability of mutation} = 0.9 \]

<table>
<thead>
<tr>
<th>Number of iteration</th>
<th>Number of state</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>380</td>
</tr>
<tr>
<td>1</td>
<td>884</td>
</tr>
<tr>
<td>2</td>
<td>951</td>
</tr>
<tr>
<td>\vdots</td>
<td>\vdots</td>
</tr>
<tr>
<td>89997</td>
<td>2020</td>
</tr>
<tr>
<td>89998</td>
<td>1197</td>
</tr>
<tr>
<td>89999</td>
<td>3213</td>
</tr>
</tbody>
</table>

Table(3)\(n = 90000\) iterations
<table>
<thead>
<tr>
<th>Ordered number of state</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.000122</td>
</tr>
<tr>
<td>1</td>
<td>0.000022</td>
</tr>
<tr>
<td>2</td>
<td>0.000011</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>4093</td>
<td>0.000089</td>
</tr>
<tr>
<td>4094</td>
<td>0.000144</td>
</tr>
<tr>
<td>4095</td>
<td>0.004889</td>
</tr>
</tbody>
</table>

Table 4: Stationary distribution of Table 3

Classification of chain (regular)
380 to 951 (neither even nor odd)
:  
380 to 3226 (odd)
88367
380 to 3142 (odd)
89085
380 to 1922 (even)
89498
380 to 2657 (even)
89650

From Tables 2 and 4, we obtain the same unique stationary distribution and obtain the regular chain.

6.2. A numerical example for clonal selection algorithms (as lumped Markov chains)

Consider the following:

\[ f(x) = x \cdot \sin(10\pi \cdot x) + 1, \quad x \in [1.7, 2] \]

\[ k = 2, \quad m = 6, \quad \text{probability of mutation} = 0.9 \]

Unique chromosomes = \{ 00 = 1.700000, 01 = 1.800000, 10 = 1.900000, \}
11 = 2.000000 

Globally optimum value = 1.900000

We will use the notation $\bar{x}_{value} =$ Value of $\bar{x}$.

All possible combinations of states of unique chromosomes =

\[
\{ (00, 00, 00, 00, 00) \} = 1.700000 = x_{1.700000},
\]

\[
\{ (00, 00, 00, 00, 00, 01), (00, 00, 00, 00, 01, 00), (00, 00, 01, 00, 00, 00), (00, 01, 00, 00, 00, 00), (01, 00, 00, 00, 00, 00) \} = 1.716667 = x_{1.716667},
\]

\ldots, \{ (11, 11, 11, 11, 11, 11) \} = 2.000000 = x_{2.000000} \}

Sampling distribution of $X$ ($N_1(1.850000, 0.000004)$) $\equiv$ stationary distribution.

Covariance = -0.000004 for the four groups of states are equal then the process is lumped Markov chain

Conditional variance = 0.000001 is the same for all four groups

All possible conditional normal distributions of transition matrix =

\[
\{ \bar{x}_{1.700000}:N_1(1.989819,0.000001), \}
\]

\[
\bar{x}_{1.716667}:N_1(1.974284,0.000001),
\]

\ldots, \bar{x}_{2.000000}:N_1(1.710181,0.000001) \}

7. Discussion

In this paper, the main result is the unified theorem of clonal selection algorithms toward a new philosophy of machine intelligence. The unified theorem will help us to propose unique chromosomes for a purely successful optimization of these algorithms. Through unified Markov approach, we obtain purely empirical analysis conclusions and obtain purely theoretical analysis for a classification and Stationary distributions of chains. Through unified lumped Markov approach, we obtain purely empirical analysis conclusions and obtain purely theoretical analysis for all possible conditional multivariate normal dis-
tributions of transition probability matrices and stationary multivariate normal distributions of chains.

References