A Comprehensive Comparison Between Real Population Based Tournament Selection and Virtual Population Based Tournament Selection

Yi Hong, Sam Kwong, Senior Member, IEEE, Qingsheng Ren, and Xiong Wang

Abstract—This paper gives a comprehensive comparison between real population based tournament selection and virtual population based tournament selection both in theory and in experiment. We claim that virtual population based EDA is able to obtain a comparable or even better solution when compared with the one captured by real population based EDA. EDA using virtual population based tournament selection does not store the whole population in the memory. Therefore, less memory is required when compared with the one using real population based tournament selection. Another advantage of EDA using virtual population based tournament selection is higher selection intensity can be achieved, therefore less fitness evaluations are needed to converge.

I. INTRODUCTION

Estimation of distribution algorithm (EDA) [1] [2] is a new class of genetic algorithm (GA) that replaces traditional genetic operators of GA with a probabilistic graphical model to reproduce more promising solutions in a faster and more accurate way. EDA has several advantages when compared with the simple GA. For example, EDA is able to avoid the disruption of important building blocks through explicitly analyzing the relationships among variables. Therefore, it can always achieve a better solution when compared with the simple GA. EDA has been widely applied into many areas such as combinatorial optimization [3], machine learning [4], machine vision [5] and just name a few.

An increasing number of applications have demonstrated the potentials of EDA. However, it has also been recognized that the performance of EDA is sensitive to its population size. To obtain a high quality solution, the population size of EDA should be set very large. Otherwise, its performance can not be guaranteed [6]. But too large a population size results in slow convergence and high memory cost. It is also not an easy task to set an appropriate population size that can draw a good balance between convergent reliability and convergent velocity. In addition, in many application domains such as the wireless network optimization, no enough space is available to simultaneously hold all candidate solutions in the memory. Therefore, we should design a special kind of estimation of distribution algorithm that can be executed with a small memory requirement.

The first attempt to design the compact version of EDA was done in [7]. The authors proposed the Compacted Genetic Algorithm (cGA), which in fact mimics the behavior of the simple GA without storing the population explicitly. This algorithm has the following steps: Firstly, a probability vector is initialized as (0.5, 0.5, …, 0.5). Then the evolving process starts: two individuals are randomly sampled from the probability vector and letting them compete, then the winner is adopted to update the probability vector. The above steps iterate until the finish condition is met. cGA is very convenient to solve problems under the condition lacking of memory and has gained a great success in its applications. For example, cGA has been applied to Feature Subset Selection problem [8], to the TSP problem [9], and to the active interval scheduling in hierarchical sensor networks [10]. However, it has been recognized that the final solution of cGA is always not satisfactory enough due to its low selection pressure and the assumption cGA based on that variables are independent. To increase the selection pressure of cGA, Ahn and Ramakrishna proposed the Elitism-Based Compact Genetic Algorithm [11]. Elitism-Based Compact Genetic Algorithm has a higher selection pressure when compared with cGA, but like the original version of cGA, it simply assumes that all variables are independent.

cEDA is a variant of EDA, where the relationships among variables are considered but the population are directly removed from EDA [12]. It is able to achieve a comparative or even better solution when compared with EDA. The authors gave a generic framework to remove the population from EDA. Another contribution of this work is the concept of selection on virtual population was firstly proposed. EDA using virtual population based tournament selection does not store the whole population in the memory. Therefore, much less memory is required when compared with the one using real population based tournament selection. Another advantage of EDA using virtual population based tournament selection is higher selection intensity can be obtained and therefore less fitness evaluations are needed to converge when compared the real population based tournament selection.

In this paper, we give a comprehensive comparison between real population based tournament selection and virtual population based tournament selection both in theory and in experiment. We claim that virtual population based EDA is as effective as real population based one.

The remainder of this paper is arranged as follows. Section II reviews the tournament selection. In section III, a
theoretical comparison between tournament selection on real population and tournament selection on virtual population is carried out. Numerical results on several benchmark problems are given in section IV to compare the performances of real population based tournament selection and virtual population based tournament selection. Section V concludes this paper.

II. DESCRIPTION OF TOURNAMENT SELECTION

Estimation of distribution algorithm works as follows: a density of the solutions to the problem is initialized and a population of solutions is sampled from the density. Then the evolving process starts: part of promising solutions are selected; the density of the selected promising solutions is estimated and new solutions are sampled from this estimated density function. These steps iterate until the finish condition is met.

Among the above given steps, one important issue of EDA is how to select the promising solutions from the whole population. Several recent papers have been concentrated on the selection operator of EDA. For example, Lima et al. studied the influence of selection operator on the performance of EDA [13]. Santana designed a kind of EDA using selection operator without selected population [14]. Recently, Hong et al. proposed the concept of over-selection and claimed that over-selection based EDA is able to achieve a better solution when compared with the classical EDA [15]. The design of selection operator should favor high-quality solutions such that solutions with higher fitness values also have higher probabilities to be selected. There have been several famous selection schemes used in EDA such as tournament selection, truncation selection and proportional selection. In this paper, we are only interested in tournament selection. Tournament selection can be operated on both real population and virtual population. The purpose of this paper is to give a comprehensive comparison between real population based tournament selection and virtual population based tournament selection both in theory and in experiment.

For real population based selection schemes, there have been several effective methods to analyze their performances [16]-[21]. For example, de la Maze and Tidor analyzed several famous selection schemes according to their scale and translation invariance [16]. Similar work was done by Bäck with respect to their takeover time [17]. The selection intensity was introduced and used to evaluate the proportional selection and the truncation selection by Mühlenbein [18]. More concrete mathematical analysis about different kinds of selection schemes was given by Blickle and Thiele [19]. The authors presented a generic method to analyze the properties of ordinal selection schemes through calculating the change of fitness distribution [20].

In this paper, a comprehensive comparison between virtual population based tournament selection and real population based tournament selection is given. Considered that the behavior of a selection scheme depends only on the fitness values of the solutions in the population, our analysis is based on the concept of fitness distribution [19]. To make this paper self-contained, the following four definitions used in [19] are given as follows:

Definition 1: Let $s(f_i)$ denote the number of individuals whose fitness values equal to $f_i$ in a population.

Definition 2: Let $n$ be the number of unique fitness values and $f_1 < f_2 < \ldots < f_n (n \leq N)$ are ordered fitness values in the population, where $N$ is the number of individuals in the population.

$S(f_i)$ denotes the number of individuals whose fitness values are not higher than $f_i$, i.e.

$$S(f_i) = \begin{cases} 0 & \text{if } i < 0; \\ \sum_{j=1}^{i} s(f_j) & \text{if otherwise}; \\ N & \text{if } i > n; \end{cases} \quad (1)$$

Definition 3: Let $\tilde{S}(f_i)$ denote the mathematical expectation of the number of individuals whose fitness values are not higher than $f_i$ in the selected subpopulation.

Definition 4: Let $S^*(f_i)$ denote the mathematical expectation of the number of individuals whose fitness values equal to $f_i$ in the selected subpopulation.

A. Tournament selection on real population

Real population based tournament selection works as follows: Choose $k$ individuals from the population at random and their fitness values are compared, then the individual with the highest fitness value is selected, where $k$ is commonly called as the tournament size. The above steps repeat for $u$ times and $u$ promising solutions are selected. Often tournament are held only between two individuals, that is the tournament size $k$ equals to 2. To simplify our analysis, the population size $N$ is fixed to $u \cdot k$. The above steps can be concluded as follows:

1) Sample $u \cdot k$ individuals from the density function;
2) Calculate their fitness values;
3) Randomly pick $k$ individuals from the whole population with replacement;
4) The individual with the highest fitness is chosen and stored;
5) Go to 3 and the above steps iterate for $u$ times.

The following two theories are given in [19].

Theorem 1: Given a population, after carrying out real population based tournament selection the mathematical expectation $\tilde{S}(f_i)^R$ of the number of individuals whose fitness values are not higher than $f_i$ equals to

$$\tilde{S}(f_i)^R = u \cdot \left( \frac{S(f_i)}{k \cdot u} \right)^k \quad (2)$$

Theorem 2: Given a population, after carrying out real population based tournament selection the mathematical expectation $S^*(f_i)^R$ of the number of individuals whose fitness values equal to $f_i$ equals to

$$S^*(f_i)^R = u \cdot \left( \frac{S(f_i)}{k \cdot u} \right)^k \cdot \left( \frac{S(f_{i-1})}{k \cdot u} \right)^k \quad (3)$$
The virtual population based selection operators: virtual population based tournament selection and virtual population based truncation selection. After combining the virtual population based selection process and density incremental learning process together, the authors successfully removed the population from EDA. They claimed that virtual population based EDA is able to obtain a comparative or even better solution when compared with real population based EDA. The inspiration of selection on virtual population can be concluded as follows: All solutions are sampled from a density function, but not directly on the population. In this case, the selection process is directly performed on mutation operator in the estimation of distribution algorithm. Therefore, only the selection process is directly performed on virtual population based tournament selection. After combining the virtual population based tournament selection and virtual population based truncation selection, in this paper, we consider the selection process and the density incremental learning process were merged together. Therefore, it is a little complex to analyze the performance of the virtual population based tournament selection [12] as compared with real population based EDA. They claimed that virtual population based EDA is different from both the EDA illustrated in [1] and the cEDA proposed by Hong et al. [12]. Our proposed version of EDA must store the selected subpopulation and the cEDA proposed by Hong et al. [12]. Our proposed version of EDA is given as follows:

1) Set the initial density function \( P_1, i = 1 \);
2) Empty the information holder \( I \);
3) Sample \( k \) individuals from \( P_i \) and let them compete, the winner is used to update \( I \); Remove all \( k \) individuals from the memory. The above steps iterate for \( u \) times.
4) Estimate the density \( P_{i+1} \) from \( I \);
5) If the finish condition is not met, \( i = i + 1 \), go to 2.

In [12], the selection process and a density incremental learning process were merged together. Therefore, it is a little complex to analyze the performance of the virtual population based tournament selection. To more clearly illustrate the virtual population based tournament selection, in this paper, we consider the selection process and the density incremental learning process as two separate processes. Therefore, the selection operator and the density estimation operator are executed sequentially: Firstly, a certain number of promising solutions are selected. These selected solutions make up a subpopulation \( G^u \). Then the density of individuals in \( G^u \) is estimated. The framework of this kind of EDA is given as follows:

1) Set the initial density function \( P_1, i = 1 \);
2) Build \( G^u \) through sampling and simultaneously selecting from \( P_i \);
3) Estimate the density \( P_{i+1} \) of all individuals in \( G^u \);
4) If the finish condition is not met, \( i = i + 1 \), go to 2.

Something worthwhile mentioning is that the above version of EDA is different from both the EDA illustrated in [1] [2] and the cEDA proposed by Hong et al. [12]. Our proposed version of EDA must store the selected subpopulation \( G^u \) in the memory. However, the selection scheme of ours is similar to the virtual population based tournament selection [12] as follows:

1) Sample \( k \) individuals from the density function;
2) Calculate their fitness values;
3) The individual with the highest fitness is chosen and stored;
4) The above steps iterate for \( u \) times.

Their only difference is our method stores the selected individuals, while in [12], only the information about the selected individuals is stored. It is noted that the purpose of this paper is not to propose a novel kind of estimation of distribution algorithm, but to illustrate the virtual population based tournament selection proposed in [12].

For EDA using virtual population based tournament selection, \( u \cdot k \) individuals are sampled and compared in each generation. These \( u \cdot k \) individuals make up a virtual population and \( u \) winners are selected. In addition, \( u \) selected individuals are generated one by one, and the \( i^{th} \) selected individual is the best one among all individuals from the \((i \cdot k + 1 - k)^{th}\) one to \((i \cdot k)^{th}\) one.

**Theorem 3:** Given a population, after performing virtual population based tournament selection the mathematical expectation \( \tilde{S}(f_i)^V \) of the number of individuals whose fitness values are not higher than \( f_i \) is

\[
\tilde{S}(f_i)^V = u \cdot \frac{C_k^u C \tilde{S}(f_i)^{k/u}}{C_k^u}
\]

**Proof:** To randomly pick one individual from the population, the probability that its fitness value is not higher than \( f_i \) equals to \( \frac{C_k^u (k-u) \tilde{S}(f_i)}{C_k^u \tilde{S}(f_i)} \). Then to randomly pick \( k \) individuals from the population, the probability that all of their fitness values are not higher than \( f_i \) equals to \( \frac{C_k^u (k-u) \tilde{S}(f_i)}{C_k^u \tilde{S}(f_i)} \). This is because to select \( k \) individuals from a population of \( k \cdot u \) individuals, the size of different cases is \( C_k^u \). We denote the set of all cases as \( D_0 \). Likewise, to select \( k \) individuals from a population of \( S(f_i) \) individuals, the size of different cases is \( C_k^S(f_i) \). We denote the set of all cases as \( D_1 \). Considered that \( D_1 \) is a subset of \( D_0 \), then randomly pick an element from \( D_0 \), the probability that it belongs to \( D_1 \) equals to \( \frac{C_k^S(f_i)}{C_k^u} \). Considered that all \( u \) selection rounds are independent, the mathematical expectation \( \tilde{S}(f_i)^V \) of the number of individuals whose fitness values are not higher than \( f_i \) after performing \( u \) rounds tournament selection is

\[
\tilde{S}(f_i)^V = u \cdot \frac{C_k^u C \tilde{S}(f_i)^{k/u}}{C_k^u}
\]

**Theorem 4:** Given a population, after performing virtual population based tournament selection, the mathematical expectation \( S^*(f_i)^V \) of the number of individuals whose fitness equal to \( f_i \) is

\[
S^*(f_i)^V = u \cdot \frac{C_k^u \tilde{S}(f_i)^{k/u} - C_k^u C \tilde{S}(f_{i-1})^{k/u}}{C_k^u}
\]

**Proof:** Using the equation (4) and the relation \( S^*(f_i)^V = \tilde{S}(f_i)^V - \tilde{S}(f_{i-1})^V \), formula (5) can be obtained.

III. THEORETIC COMPARISONS

A. General properties

**Theorem 5:** If the number of individuals with the lowest fitness values are smaller than the tournament size \( k \), under
virtual population based tournament selection, all individuals with the lowest fitness must be extinct in the selected subpopulation; but under real population based one, they still have chance to live.

**Proof:** If \( s(f_1) < k \), we get \( S(f_1) < k \), then \( S^*(f_1)^V = 0 \) can be obtained, that means the individuals whose fitness values equal to \( f_1 \) are extinct under virtual population based tournament selection.

If \( s(f_1) < k \), we get \( S(f_1) < k \), then \( S^*(f_1)^R > 0 \) can be obtained, that means the individuals whose fitness values equal to \( f_1 \) still have chance to be selected under real population based tournament selection.

According to the above theorem, if the individual with the lowest fitness value is unique and the tournament size \( k \) equals to 2, the worst individual must be extinct in the selected subpopulation after performing virtual population based tournament selection; but if real population based tournament selection is adopted, it may be selected into the subpopulation. A more generic format of theorem 5 is given in theorem (6).

**Theorem 6:** If \( S(f_i) < k \), under virtual population based tournament selection, all individuals whose fitness values are not higher than \( f_i \) must be extinct in the selected subpopulation; but under real population based one, they still have chance to live.

**Proof:** If \( S(f_i) < k \), the equation \( S^*(f_i)^V = 0 \), \( (j = 1, ... , i) \) can be obtained, that means the individuals whose fitness values are not higher than \( f_i \) are extinct under virtual population based tournament selection.

If \( S(f_i) < k \), the inequality \( S^*(f_i)^R > 0 \), \( (j = 1, ... , i) \) can be obtained, that means the individuals whose fitness values are not higher than \( f_i \) still have chance to live under real population based tournament selection.

**Theorem 7:** If the best individual is unique, it must be selected under the virtual population based tournament selection, but for real population based tournament selection, the best individual may not be chosen.

**Proof:** For virtual population based tournament selection, to randomly pick one individual from the population, the probability that the best individual is not picked equals to \( \frac{k - 1}{k} \). Then to randomly pick \( k \cdot u \) individuals, the best individual is not picked equals to \( \left( \frac{k - 1}{k} \right)^k \). Therefore the probability that the best individual is randomly picked equals to \( 1 - \left( \frac{k - 1}{k} \right)^k \). If and only if the best individual is picked, the best individual can be selected. Considered that \( 1 - \left( \frac{k - 1}{k} \right)^k < 1 \), it means the best individual may not selected under real population based tournament selection.

For virtual population based tournament selection, all individuals must be picked and compared. Therefore the best individual must be selected.

To deeply understand the virtual population based tournament selection, \( S^*(f_i)^V \) and \( S^*(f_i)^R \) are compared.

**Theorem 8:** If tournament size \( k \) equals to 2, there must be a \( f_m \) such that for all individuals whose fitness values are higher than \( f_m \), under virtual population based tournament selection, they have higher probabilities to be chosen when compared with those under real population based tournament selection.

**Proof:**

\[
\nabla = S^*(f_i)^V - S^*(f_i)^R
\]

\[
= u \cdot \left( \left( \frac{C_i}{C_{k \cdot u}} - \frac{C_k}{C_{k \cdot u}} \right) - \frac{S(f_i)}{k \cdot u} - \frac{(S(f_{i-1}) - (S(f_{i-1})^k))}{k \cdot u} \right)
\]

(6)
If $k = 2$, $\nabla$ can be calculated as follows,

$$\nabla = (S(f_1) - S(f_{i-1}))$$

$$\frac{(S(f_1) + S(f_{i-1}) - 1)}{2u} - \frac{S(f_1) + S(f_{i-1})}{2u}$$

$$= \frac{(S(f_1) - S(f_{i-1})) \cdot (S(f_1) + S(f_{i-1}) - 2u)}{2u \cdot (2u - 1)}$$

$$\geq 0$$  

if $S(f_1) + S(f_{i-1}) \geq 2u$  

Considering the definition of $S(f_1)$, we get:

$$\exists m \forall j : j \geq m \rightarrow S(f_1) + S(f_{i-1}) \geq 2u$$

From formula (9), we can draw the conclusion that for all individuals whose fitness values are greater than $f_m$, their probabilities to be selected under virtual population based tournament selection are higher than those under real population based tournament selection.

We give an example to illustrate the above theorems. Providing $f_i = i$ ($i = 1, ..., 100$) and $s(f_i) = 1$ ($i = 1, ..., 100$), the tournament size $k$ is set to 2 and 50 individuals are selected using real population based tournament selection and virtual population based tournament selection, the results are given in Fig.1 and Fig.2. From Fig.1, we can draw the conclusions that although the properties of real population based tournament selection and virtual population based tournament selection are very similar, there are still some differences (See the data listed in Fig. 1). Individuals with high fitness values have higher probabilities to be selected under the virtual population based tournament selection when compared with those under real population based tournament selection (See Fig. 2). Another difference is the best individual must be selected and the worst individual must not be selected under virtual population based tournament selection. But the best individual may not be selected and the worst individual may be chosen when using real population based tournament selection.

B. Production rate

Definition 5: The production rate $R(f_i)$ denotes the ratio of the number of individuals with a certain fitness value $f_i$ after and before selection [19]

$$R(f_i) = \begin{cases} \frac{S^*(f_i)}{S(f_i)} & \text{if } s(f_i) > 0; \\ 0 & \text{if } s(f_i) = 0; \end{cases}$$

Considered that the size of selected subpopulation is smaller than the size of the original population in EDA, the above definition is revised as follows:

$$R(f_i) = \begin{cases} \frac{k \cdot S^*(f_i)}{S(f_i)} & \text{if } s(f_i) > 0; \\ 0 & \text{if } s(f_i) = 0; \end{cases}$$

According to this definition, the production rate of real population based tournament selection is

$$R(f_i)^R = \frac{k \cdot u \cdot \left(\frac{S(f_i)}{\lambda S(f_i)}\right)^k - \left(\frac{S(f_{i-1})}{\lambda S(f_{i-1})}\right)^k}{s(f_i)}$$

and the production rate of virtual population based tournament selection is

$$R(f_i)^V = \frac{k \cdot u \cdot \left(\frac{C^k_{S(f_i)}}{\lambda C^k_{S(f_{i-1})}}\right) - \left(\frac{C^k_{S(f_{i-1})}}{\lambda C^k_{S(f_{i-1})}}\right)^k}{s(f_i)}$$

Theorem 9: If tournament size $k$ equals to 2, there must be a $f_m$ such that for all individuals whose fitness values are higher than $f_m$, under virtual population based tournament selection, they have larger production rates when compared with those under real population based tournament selection.

Proof: From theorem 8, we obtain that for all $f_i > f_m$

$$S^*(f_i)^V > S^*(f_i)^R$$

Therefore, we get $R^*(f_i)^V > R^*(f_i)^R$. ■

C. Selection intensity

Definition 6: The selection intensity $I$ is defined as the expected increase in the average fitness of a population after performing the selection operator [11] [18]. According to this definition, the selection intensity $I^R$ of real population based tournament selection is

$$I^R = \sum_{i=1}^{n} \left[ f_i \cdot \left( \frac{(S^*(f_i))^k}{S(f_i)} - \frac{(S^*(f_{i-1}))^k}{S(f_{i-1})} \right) \right]$$

While for virtual population based tournament selection, its selection intensity $I^V$ equals to

$$I^V = \sum_{i=1}^{n} \left[ f_i \cdot \left( \frac{C^k_{S(f_i)}}{\lambda C^k_{S(f_{i-1})}} - \frac{C^k_{S(f_{i-1})}}{\lambda C^k_{S(f_{i-1})}} \right) \right]$$

In [19], the fitness distribution is considered as the normalized Gaussian distribution $N(0,1)$. In this paper, we assume that fitness values satisfy the uniform distribution, that is $s(f_1) = s(f_2), ..., s(f_n) = \varphi$ and $f_i - f_{i-1} = \lambda$. Based on this assumption, we obtain

$$\varphi = \frac{k \cdot u}{n}$$

In addition

$$S(f_i) = \frac{i \cdot k \cdot u}{n}$$

Then we can verify the following theorem.

Theorem 10: Virtual population based tournament selection has a higher selection intensity when compared with real population based tournament selection.

Proof:

$$\nabla = I^V - I^R$$

$$= \delta \cdot \sum_{i=1}^{n} \left[ (f_i - (i - 1) \lambda) \cdot (4i - 2 - 2n) \right]$$

1This assumption is generic and not prone to high fitness values or to low fitness values.
where
\[ \delta = \frac{1.0}{2 \cdot u \cdot n^2 \cdot (2u - 1)} \]
Considered that
\[ \sum_{i=1}^{n} i^2 = \frac{n \cdot (n + 1) \cdot (2n + 1)}{6} \]  \hspace{1cm} (20)
We obtain
\[ \nabla = \frac{\delta \cdot \lambda \cdot n \cdot (n + 1) \cdot (n - 1)}{3} > 0 \]
Therefore
\[ I^V > I^R \]
The above inequality concludes that virtual population based tournament selection has a higher selection intensity when compared with real population based tournament selection.

D. Loss rate

Like genetic algorithm, estimation of distribution algorithm is also inspired from the principle of survival of the fittest. Both GA and EDA use selection operator to mimic the behavior of survival of the fittest. Under the effect of selection operator, large fitness values are increasing and low ones are losing in fitness distribution with evolving of the population. In this paper, we adopt the concept of lost rate to depict this phenomenon and to evaluate the property of selection operator.

**Definition 7:** The loss rate \( p_l(f_i) \) is defined as the probability that fitness value \( f_i \) is lost in the selected subpopulation. According to this definition, the real population based tournament selection has the loss rate of \( p_l(f_i)^R \) as
\[ p_l(f_i)^R = [1 - \left( \frac{S(f_i)}{k \cdot u} \right)^k]^{u} \]  \hspace{1cm} (21)
While the loss rate of virtual population based tournament selection can be calculated as follows
\[ p_l(f_i)^V = [1 - \left( \frac{C_k S(f_i)}{C_k^u} - \frac{C_k S(f_{i-1})}{C_k^u} \right)^k]^{u} \]  \hspace{1cm} (22)

**Theorem 11:** If tournament size \( k \) equals to 2, there must be a \( f_m \) such that for all \( f_i \) that are higher than \( f_m \), under virtual population based tournament selection, they have lower loss rate when compared with those under real population based tournament selection.

**Proof:** Considered that
\[ 0 \leq \frac{S(f_i)}{k \cdot u} - \frac{S(f_{i-1})}{k \cdot u} \leq 1 \]
and
\[ 0 \leq \left( \frac{S(f_i)}{k \cdot u} \right)^k - \left( \frac{S(f_{i-1})}{k \cdot u} \right)^k \leq 1 \]
In addition, the function \( f(t) \)
\[ f(t) = (1 - t)^n \]  \hspace{1cm} (23)
is a decreasing function in the variable interval \( t \in [0, 1] \). That means \( f(t_1) < f(t_2) \) if and only if \( t_1 > t_2 \), where \( t_1, t_2 \in [0, 1] \).
As theorem 8 has demonstrated that
\[ \exists m \forall i : j > m \rightarrow \frac{C_k S(f_i)}{C_k^u} - \frac{C_k S(f_{i-1})}{C_k^u} > \left( \frac{S(f_i)}{k \cdot u} \right)^k - \left( \frac{S(f_{i-1})}{k \cdot u} \right)^k \]
then we obtain
\[ \exists m \forall i : j > m \rightarrow p_l(f_i)^V < p_l(f_i)^R \]
The above inequality conclude that for all fitness whose values are higher that \( f_m \), they have lower loss rates under virtual population based tournament selection when compared with the ones using real population based tournament selection.

E. Memory cost

**Theorem 12:** The memory cost of virtual population based tournament selection is is denoted as \( M^V \), while the memory cost of real population based tournament selection is denoted as \( M^R \),
\[ \frac{M^V}{M^R} \approx \frac{1.0}{k} < 1 \]  \hspace{1cm} (24)

**Proof:** Real population based tournament selection must store the whole \( u \cdot k \) individuals, while virtual population based tournament selection only needs to store the selected \( u \) individuals. The formula (25) can be obtained.

IV. NUMERICAL EXPERIMENTS

Six benchmark problems are selected to compare the performances of EDA using real population based tournament selection and EDA using virtual population based tournament selection. These benchmarks and their characteristics are described in Table 1. Parameters’ settings to solve these problems using both kinds of EDA are given in Table 2.
For all benchmark problems, EDA using real population based tournament selection and EDA using virtual population based tournament selections are independently executed for 50 runs and their average results are shown in Figure 3. From this figure, we can see virtual population based EDA can achieve the average fitness values of best solutions about 100 for OneMax Problem, about 5250 for Weighted Onemax Problem, about 50 for Quadratic Problem, about 73 for Cubanl−5 Problem, about 77 for Muhl-5 Problem and about 77 for Trap−5 Problem. These results are much better than the ones achieved by real population based tournament selection around 97 for OneMax Problem, around 5150 for Weighted Onemax Problem, around 48 for Quadratic Problem, around 67 for Cuban−5 Problem, around 70 for Muhl-5 Problem and around 77 for Trap−5 Problem. As for convergent velocities of real population based tournament selection and virtual population based tournament selection, they are very similar. It is easy to conclude from the above results that EDA using virtual population based tournament selection is able to obtain a similar or even better solution.
when compared with the one capture by real population based EDA.

V. CONCLUSIONS

This paper has given a comprehensive comparison between real population based tournament selection and virtual population based tournament selection both in theory and in experiment. We found that although their properties are similar, virtual population based tournament selection is a little better than real population based tournament selection. We demonstrated that virtual population based EDA can often achieve a better solution when compared with real population based one. To deeply understand the virtual population based EDA will be of great interest for our future work.

ACKNOWLEDGMENT

This project is supported by the project no. 7002023, City University of Hong Kong. The authors would like to thank the comments and suggestions from the reviewers.

APPENDIX

Benchmark problems:

1. OneMax Problem: \[ \sum_{i=1}^{100} x_i \quad x_i \in \{0, 1\} \]
2. Weighted OneMax Problem: \[ \sum_{i=1}^{100} \frac{x_i}{i} \quad x_i \in \{0, 1\} \]
3. Quadratic Problem: \[ \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{f(x_{2i-1}, x_{2j-1}) - 0.9(u + v) + 1.9uv}{x_{2i-1} \cdot x_{2j-1}} \quad x_{2i-1}, x_{2j-1} \in \{0, 1\} \]
4. Cuban-5 Problem: \[ F_{cuban}(x) = \begin{cases} 3.0 & \text{if } x = (0, 0, 0, 0, 1); \\ 2.0 & \text{if } x = (0, 0, 0, 1, 1); \\ 1.0 & \text{if } x = (0, 0, 1, 1, 1); \\ 0.5 & \text{if } x = (1, 1, 1, 1, 1); \\ 4.0 & \text{if } x = (0, 0, 0, 0, 0); \\ 0.0 & \text{otherwise}; \end{cases} \]
5. Muhl-5 Problem: \[ F_{Muhl}(s) \]
6. Trap-5 Problem: \[ F_{Trap}(s) = \begin{cases} 5 & \text{if } s = 5; \\ 4 - s & \text{if otherwise}; \end{cases} \]

REFERENCES