Automotive Innovation
https://doi.org/10.1007/s42154-021-00168-8

Performance Limit Evaluation Strategy for Automated Driving Systems

Feng Gao[1,2] · Jianwei Mu[2] · Xiangyu Han[2] · Yiheng Yang[2] · Junwu Zhou[3]

Received: 29 June 2021 / Accepted: 18 November 2021

Abstract
Efficient detection of performance limits is critical to autonomous driving. As autonomous driving is difficult to be realized under complicated scenarios, an improved genetic algorithm-based evolution test is proposed to accelerate the evaluation of performance limits. It conducts crossover operation at all positions and mutation several times to make the high-quality chromosome exist in candidate offspring easily. Then the normal offspring is selected statistically based on the scenario complexity, which is designed to measure the difficulty of realizing autonomous driving through the Analytic Hierarchy Process. The benefits of modified cross/mutation operators on the improvement of scenario complexity are analyzed theoretically. Finally, the effectiveness of improved genetic algorithm-based evolution test is validated after being applied to evaluate the collision avoidance performance of an automatic parallel parking system.

Keywords Autonomous driving · Test and evaluation · Evolution test · Genetic algorithm

Abbreviations

AEB	Automatic emergency baking
AHP	Analytic hierarchy process
APPS	Automatic parallel parking system
APS	Automated parking system
GA	Genetic algorithm
IGA	Improved genetic algorithm
ISO	International standard organization

1 Introduction

To improve traffic safety and efficiency, original equipment manufacturers (OEMs) and governments around the world devote themselves to promoting automated driving systems, such as automatic parking system (APS) and automatic emergency braking (AEB) system [1, 2]. These systems have a stringent requirement on safety and performance, which also benefits their market competitiveness [3]. Therefore, sufficient evaluations are necessary before putting them into the market [4, 5]. One direct way is using the standards delivered by organizations, such as international standard organization (ISO) and government agencies [6, 7]. For example, Christian et al. [8] analyzed their developed AEB under different lateral offsets using the scenario in Euro-NCAP. Since about 2002, ISO has published the specifications for systems from intelligent level 0 to 2 one after another. In 2020, level 2.5 system was first included in the evaluation program executed by iVISTA in China [6]. Such standard test scenarios are limited because only typical conditions are considered. They are far from ensuring coverage, let alone evaluation of performance limits.

Different from traditional onboard systems, such as battery management system [9], functionality and performance of automated driving systems are greatly influenced by traffic environments. The uncontrollability, variety, and indefinability of traffic pose great challenges on evaluation, especially the performance limit. Moreover, automated driving is directly related to safety, so it is critical to find out the boundary of achievable performance to avoid misuse. To achieve this, OEMs have to take a large amount of naturalistic field operational tests. In this way, vehicles equipped with automated driving systems are driven in real traffic for a very long time. The automated driving system is executed under real conditions with random effects. Theoretically, a

Academic editor: Shuo Feng

✉ Feng Gao
gaofeng1@cqu.edu.cn

[1] State Key Laboratory of Vehicle NVH and Safety Technology, Chongqing, China

[2] College of Mechanical and Vehicle Engineering, Chongqing University, Chongqing, China

[3] Passenger Vehicle Technical Center, Dongfeng Liuzhou Motor Co., Ltd, Liuzhou, China

complete evaluation can be achieved with adequate exciting samples [2]. Several such projects have been deployed in the US [10, 11] and Europe [12, 13]. Unfortunately, the probability of critical condition is very low in real traffic, which leads to an inconceivable test cycle and huge costs [10]. Moreover, how to ensure experimental safety in dangerous conditions is still a problem.

Compared with the naturalistic field operational test, the execution can be accelerated by simulation. There are no actual injuries or losses when experiencing dangerous scenarios during simulation [14]. One key problem is the design of the test set, which can be extracted from natural driving or crash databases [15]. But the completeness of the dataset conversely restricts the testing effect [16]. Another way is using some systematic methods, e.g., combinational test. This method can ensure the required coverage and may generate scenarios that do not exist in reality [17, 18].

To accelerate the simulation test, Zhao et al. [19] designed an aggressive driver model to increase the collision possibility, when evaluating the reliability of car following systems. The efficiency is increased by more than 300 times compared with the nominal one. However, the driving behavior is influenced by many factors and has high-order nonlinearity [20]. It is difficult to construct an accurate mechanism model, especially under complex conditions. Duan et al. [21] optimized the test set by raising the proportion of complicated scenarios to increase the fault detection rate. This method was applied to test a traffic jam pilot system by an automated simulation test and evaluation system in Ref. [22].

The aforementioned strategies are open-loop essentially since the test condition is designed irrelevantly to the response. For different systems, the conditions activating its performance limit vary greatly. The evolution test using the genetic algorithm (GA) has been applied to evaluate the performance of automated driving systems, e.g., adaptive cruise control system [23] and APS [24]. The application results show that extreme working conditions can be found with a shorter cycle compared with the random test [23]. But the evolution process is still random because the effectiveness of the generated offspring cannot be evaluated beforehand.

Considering the self-organization, adaption, and learning of GA, the Improved Genetic Algorithm (IGA) is proposed to further accelerate the evaluation process by modifying the crossover and mutation operators. This is because that a complicated test scenario has a higher possibility to find out the performance limit. Based on the scenario complexity designed through the analytic hierarchy process (AHP), the effectiveness of candidate scenarios can be measured without conducting tests. Both crossover and mutation operators of GA are modified by introducing the scenario complexity to generate more effective scenarios. Meanwhile, the advantage of natural evolution is retained. The effectiveness of IGA-based evolution test is analyzed theoretically and validated by the application to evaluate the collision performance of an automatic parallel parking system (APPS).

The rest of this paper is organized as follows: Sect. 2 introduces the IGA based evaluation strategy. Its performance is analyzed theoretically in Sect. 3. Section 4 validates the effectiveness by the evaluation of an APPS as an example and Sect. 5 concludes the paper.

2 IGA-Based Evaluation Strategy

The traditional GA performs crossover/mutation on positions randomly selected with an average distribution. This restricts the possibility of generating better offspring because they cannot be judged without conducting tests [25]. As autonomous driving is more difficult to be realized under more complex conditions, an IGA-based evaluation strategy is proposed as shown in Fig. 1.

The crossover/mutation operators are improved by introducing the scenario complexity to guide the selection of offspring as follows:

(1) Full crossover: The chromosome of parent is crossed at all positions to avoid missing the good offspring. Then each pair of offspring is evaluated by the scenario complexity. The effective offspring is selected to maximize the possibility of bigger complexity.
(2) Multiple mutation: To make the better population appear easily, multiple mutations are conducted. The test scenario of next generation is selected from the candidate populations according to their overall complexity.

2.1 Calculation of Scenario Complexity

From the aforementioned fundamentals of IGA, one of its key components is scenario complexity. It evaluates the effectiveness of the test scenario indirectly and also guides the evolution procedure besides the fitness function. Since the tested system is always a black box, it is hard to establish an analytical description. AHP is adopted to analyze and calculate the scenario complexity as shown in Fig. 2. It generates a more accurate evaluation by comprehensively combining the experience of engineers, technical specifications, and working principles [26, 27].

In Fig. 2, $L_{i,j}$ denotes the j-th influence factor in the i-th layer whose normalized importance degree is $S_{i,j}$, and q_j denotes the discretized value of the bottom factor [28]. They quantitatively characterize the influence of each factor on the system. For continuous or unbounded factors, approaches such as equivalence partitioning and boundary value analysis can be adopted to discretize these factors into multiple but

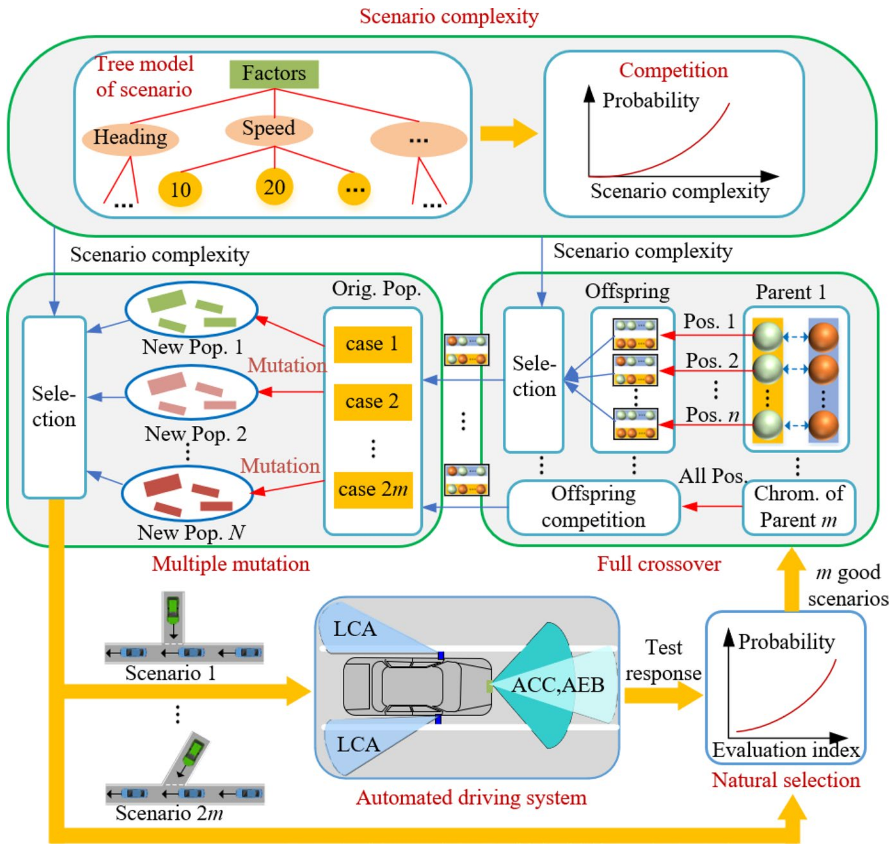

Fig. 1 IGA-based evaluation strategy for automated driving system

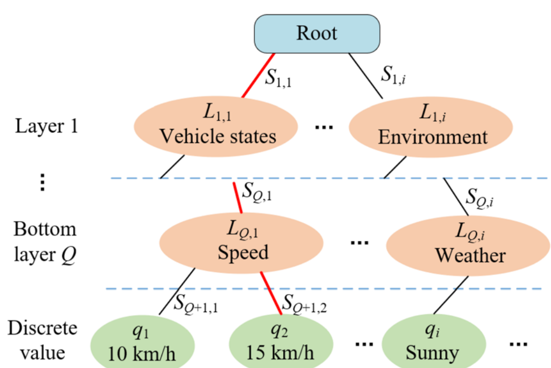

Fig. 2 Hierarchy model for analyzing scenario complexity

limited values [29]. The importance degree of each value relative to the root is calculated by [18]

$$I_n = \prod_{(i,j)\in\Omega} S_{i,j} \tag{1}$$

where I_n is the relative importance degree of q_n and Ω is the set composed of all subscripts from q_n to the root (shown by the red line in Fig. 2 as an example).

A test case, $\boldsymbol{T} = \{v_i, i = 1, \cdots, N_{\mathrm{F}}\}$, is generated by randomly selecting the value v_i in the possible range of each bottom factor. The generated value may be out of the set composed of discrete values. By linear interpolation, the scenario complexity is calculated according to the importance degree with the assumption that $v_i \in [q_j, q_{j+1}]$:

$$D(\boldsymbol{T}) = \sum_{i=1}^{N_{\mathrm{F}}} \left[I_j + \frac{v_i - q_j}{q_{j+1} - q_j}\left(I_{j+1} - I_j\right) \right] \tag{2}$$

where $D(\boldsymbol{T})$ denotes the scenario complexity of $\boldsymbol{T}$.

To better illustrate the calculation process of scenario complexity by AHP, a specific application instance is shown below. For example, as shown in Fig. 2, the influence factors at the bottom layer are "Speed", "Weather", etc. The factor "Speed" is discretized into "10 km/h", "15 km/h", etc. Experts can be invited to evaluate the relative importance of each factor by comparing it with other factors in the same layer. The evaluation score ranges from 0 to 9 (0 is the least important one and 9 is the most important). Then the normalized relative importance $S_{i,j}$ can be obtained by AHP [26, 27]. With the normalized relative importance of each factor, the importance degree of each value relative to the root can be calculated by Eq. (1). When conducting the evolution test, a test scenario is generated by combining all factors at the bottom layer and selecting one value in its range randomly, e.g., $\boldsymbol{T}_i = \{11\ \text{km/h}, \text{Sunny}, \ldots\}$. The scenario complexity of $\boldsymbol{T}_i$ can be obtained from Eq. (2) by linear interpolation according to $S_{Q+1,1}$, $S_{Q+1,2}$, etc.

2.2 Design of IGA-Based Evaluation Process

The overall procedure of IGA is similar to the GA-based evolution test, which includes 6 steps as shown in Fig. 3 [23, 24].

With the definitions in Appendix 1.1, the evolution test process is described as the following.

Step 1: Random generation of initial population $\boldsymbol{X}_G, G = 1$.

Step 2: Test using $\boldsymbol{X}_G$. According to the test results, the pseudo-code for the determination of whether the interactive test process is stopped is as follows:

$k = \arg\min_{1 \le i \le 2m} g_i$
If $g_k \le g_{\text{th}}$
  Output ($\boldsymbol{T}_k$ and g_k)
  Test is stopped because performance limit is detected.
Else if $G \ge G_{\text{th}}$
  If $g^* < g_k$
    Output ($\boldsymbol{T}^*$ and g^*)
  Else
    Output ($\boldsymbol{T}_k$ and g_k)
  End
  Test is stopped because of the limitation of interactive number.
Else
  Go to **Step 3**
End

Step 3: Rearrangement of the individuals $\boldsymbol{T}_i$ in $\boldsymbol{X}_G$ from the smallest to the largest according to g_i. To ensure the global convergence, the following elitist selection strategy is adopted [30]:

If $G = 1$
  $g^* = g_1; \boldsymbol{T}^* = \boldsymbol{T}_1$
Else if $g^* > g_1$
  $g^* = g_1; \boldsymbol{T}^* = \boldsymbol{T}_1$
End

Step 4: Natural selection based on $\boldsymbol{X}_G$ using the linear ranking method to restrain the premature convergence. The $2m$ scenarios for crossover operation are selected with the probability, $\frac{\alpha - \frac{2(\alpha-1)(i-1)}{2m-1}}{2m}$, for each individual $\boldsymbol{T}_i$, where $\alpha \in [1, 2)$ is the selection pressure [31]. The bigger α means it is more likely to choose the individuals with a smaller ranking number.

Step 5. Full crossover (as shown by the "Full crossover" block in Fig. 1), whose pseudo-code is:

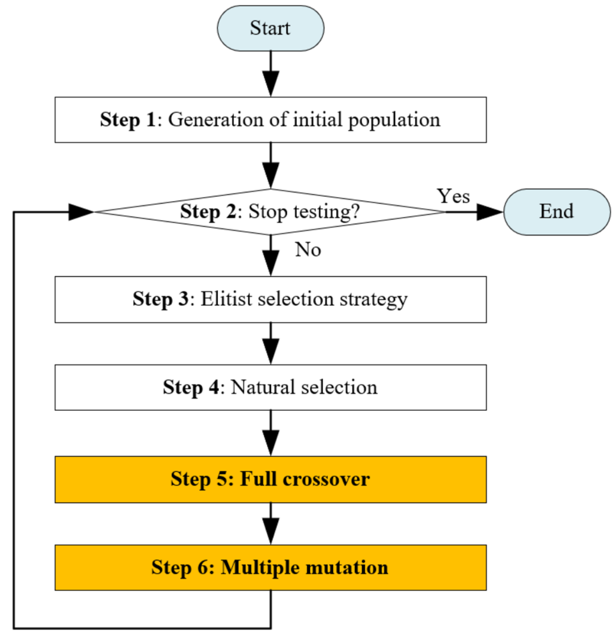

Fig. 3 Evolution test process

For $i = 1:m$
  For $j = 1:L$
    $\boldsymbol{P}^{\text{C}}_{i,j} = \text{Cross}\left(\boldsymbol{P}^{\text{C}}_i, j\right)\left(\boldsymbol{P}^{\text{C}}_i \in \boldsymbol{X}^{\text{S}}\right)$
  End
  Random selection of $\boldsymbol{P}^{\text{C}*}_i$ with the probability:
$$P_{\text{r}}\left(\boldsymbol{P}^{\text{C}*}_i = \boldsymbol{P}^{\text{C}}_{i,j}\right) = \mathrm{e}^{d \times \overline{D}\left(\boldsymbol{P}^{\text{C}}_{i,j}\right)} / \sum_{k=1}^{L} \mathrm{e}^{d \times \overline{D}\left(\boldsymbol{P}^{\text{C}}_{i,k}\right)}$$
  Store $\boldsymbol{P}^{\text{C}*}_i$ in $\boldsymbol{X}^{\text{C}}$
End

where $\text{Cross}\left(\boldsymbol{P}^{\text{C}}_i, j\right)$ denotes the single-point crossover operation on $\boldsymbol{P}^{\text{C}}_i$ at the j-th position [25] and $d \ge 0$ is the influence intensity of scenario complexity on random selection. A larger d means a greater probability of choosing the offspring pair with higher complexity.

Step 6. Multiple mutation (as shown by the "multiple mutation" block in Fig. 1). Its pseudo-code is.

For $i = 1:N$
  $\boldsymbol{X}^{\text{M}}_i = \text{Mute}\left(\boldsymbol{X}^{\text{C}}, p_{\text{M}}\right)$
End
Random selection of $\boldsymbol{X}^{\text{M}*}$ with the probability:
$$P_{\text{r}}\left(\boldsymbol{X}^{\text{M}*} = \boldsymbol{X}^{\text{M}}_i\right) = \mathrm{e}^{d \times \overline{D}\left(\boldsymbol{X}^{\text{M}}_i\right)} / \sum_{k=1}^{L} \mathrm{e}^{d \times \overline{D}\left(\boldsymbol{X}^{\text{M}}_k\right)}$$
$G = G + 1$
$\boldsymbol{X}_G = \boldsymbol{X}^{\text{M}*}$

where $\mathrm{Mute}\left(\boldsymbol{X}^{\mathrm{C}}, p_{\mathrm{M}}\right)$ denotes the canonical mutation operation on $\boldsymbol{X}^{\mathrm{C}}$ with the probability p_{M} [30].

3 Performance Analysis

It has been proved that GA with the elitist selection strategy can achieve global convergence, if the state transition matrix of selection is column-allowable, the state transition matrix of crossover is stochastic, and the state transition matrix of mutation is positive [30]. Compared with GA whose offspring is only determined by the gene of their parents, the generation of IGA's offspring is influenced by both parental gene and scenario complexity. The requirement of global convergence and statistical characteristics of offspring are analyzed theoretically in this section. The definitions of symbols and the details about the proof can be found in Appendix 1.

3.1 Statistical Analysis of Full Crossover Operator

According to **Step 5** of IGA, the probability of selecting $\boldsymbol{P}_{i,j}^{\mathrm{C}}$ as the offspring is

$$P_{\mathrm{r}}\left(\boldsymbol{P}_i^{\mathrm{C}*} = \boldsymbol{P}_{i,j}^{\mathrm{C}}\right) = \mathrm{e}^{d\times\overline{D}\left(\boldsymbol{P}_{i,j}^{\mathrm{C}}\right)} / \sum_{k=1}^{L} \mathrm{e}^{d\times\overline{D}\left(\boldsymbol{P}_{i,k}^{\mathrm{C}}\right)} \tag{3}$$

The same offspring may be generated by the crossover at different positions. If the number of positions where $\boldsymbol{P}_i^{\mathrm{C}*}$ can be generated is n_i, the total probability is

$$P_{\mathrm{r}}\left(C\left(\boldsymbol{P}_i^{\mathrm{C}}\right) = \boldsymbol{P}_i^{\mathrm{C}*}\right) = n_i \mathrm{e}^{d\times\overline{D}\left(\boldsymbol{P}_{i,j}^{\mathrm{C}}\right)} / \sum_{k=1}^{L} \mathrm{e}^{d\times\overline{D}\left(\boldsymbol{P}_{i,k}^{\mathrm{C}}\right)} \tag{4}$$

where $C(\bullet)$ denotes the full crossover operator, i.e., **Step 5** of IGA. From Eq. (4), it is found that the proposed full crossover operator only changes the probability distribution of individual offspring, and the stochastic requirement of the state transition matrix for global convergence is still satisfied.

The proposed operator is compared with the traditional one to further analyze what kind of offspring will be generated by the full crossover operator. When using an equal probability to select the crossover position, i.e. $\boldsymbol{P}_{\mathrm{r}}\left(\boldsymbol{P}_i^{\mathrm{C}*} = \boldsymbol{P}_{i,j}^{\mathrm{C}}\right) = \frac{1}{L}$, the selection probability of offspring for the single-point crossover is [25, 31]:

$$P_{\mathrm{r}}\left(\widehat{C}\left(\boldsymbol{P}_i^{\mathrm{C}}\right) = \boldsymbol{P}_i^{\mathrm{C}*}\right) = \begin{cases} \frac{p_{\mathrm{C}} n_i}{L}, \boldsymbol{P}_i^{\mathrm{C}*} \neq \boldsymbol{P}_i^{\mathrm{C}} \\ 1 - p_{\mathrm{C}} + \frac{p_{\mathrm{C}} n_i}{L}, \boldsymbol{P}_i^{\mathrm{C}*} = \boldsymbol{P}_i^{\mathrm{C}} \end{cases} \tag{5}$$

where $\widehat{C}(\bullet)$ denotes the traditional single-point crossover operator and $p_{\mathrm{C}} \in [0, 1]$ is the crossover probability. From Eqs. (4) and (5), there exist the following two conditions according to whether the offspring is the same as its parent:

(**i**) The offspring is different from its parent, i.e., $\boldsymbol{P}_i^{\mathrm{C}*} \neq \boldsymbol{P}_i^{\mathrm{C}}$.

The inequality, $\boldsymbol{P}_{\mathrm{r}}\left(C\left(\boldsymbol{P}_i^{\mathrm{C}}\right) = \boldsymbol{P}_i^{\mathrm{C}*}\right) > \boldsymbol{P}_{\mathrm{r}}\left(\widehat{C}\left(\boldsymbol{P}_i^{\mathrm{C}}\right) = \boldsymbol{P}_i^{\mathrm{C}*}\right)$, establishes if $e^{d\times\overline{D}\left(P_{i,j}^{C}\right)} > \frac{p_C}{L} \sum_{k=1}^{L} e^{d\times\overline{D}\left(P_{i,k}^{C}\right)}$. This means that compared with the traditional one, the full crossover operator has a greater possibility to select the offspring with higher complexity.

(**ii**) The offspring is the same as its parent, i.e., $\boldsymbol{P}_i^{C*} = \boldsymbol{P}_i^{C}$. If $e^{d\times\overline{D}\left(P_{i,j}^{C}\right)} > \left[(1 - p_C)/n_i + p_C/L\right] \sum_{k=1}^{L} e^{d\times\overline{D}\left(P_{i,k}^{C}\right)}$, then $P_r\left(C\left(\boldsymbol{P}_i^C\right) = \boldsymbol{P}_i^{C*}\right) > P_r\left(\hat{C}\left(\boldsymbol{P}_i^C\right) = \boldsymbol{P}_i^{C*}\right)$. Since $\boldsymbol{P}_i^{C*} = \boldsymbol{P}_i^{C}$, this implies that when the complexity of parent is large enough, the possibility of selecting it as the offspring is higher than the traditional one.

Summarizing the aforementioned discussion, the offspring generated by the full crossover tends to inherit the chromosomes with higher complexity. Furthermore, to analyze the characteristic of offspring quantitatively, the expect of offspring's complexity is studied and summarized by the following Theorem.

Theorem 1 Compared with the single-point crossover with the probability p_C in Ref. [25], the complexity expect of offspring generated by the full crossover satisfies:

(**C1**) When $p_{\mathrm{C}} = 1$, $E\left(\overline{D}\left(C\left(\boldsymbol{P}_i^{\mathrm{C}}\right)\right)\right) \geq E\left(\overline{D}\left(\widehat{C}\left(\boldsymbol{P}_i^{\mathrm{C}}\right)\right)\right)$, where $E(\bullet)$ denotes the expectation of a random signal;

(**C2**) When $0 \leq p_{\mathrm{C}} < 1$, the following inequality establishes:

$$E\left(\overline{D}\left(C\left(\boldsymbol{P}_i^{\mathrm{C}}\right)\right)\right) \geq E\left(\overline{D}\left(\widehat{C}\left(\boldsymbol{P}_i^{\mathrm{C}}\right)\right)\right) \tag{6}$$

if the average complexity of crossover offspring is not smaller than their parent:

$$\frac{1}{L} \sum_{j=1}^{L} \overline{D}\left(\boldsymbol{P}_{i,j}^{\mathrm{C}}\right) \geq \overline{D}\left(\boldsymbol{P}_i^{\mathrm{C}}\right) \tag{7}$$

Proof: The proof is in Appendix

Conclusion **C2** of **Theorem 1** shows that the complexity of offspring generated by the full crossover may be smaller than the traditional one. But this phenomenon hardly appears in practice, because it is found from Eq. (17) in Appendix that this requires the crossover possibility p_{C} to be small enough. This is bad for convergence and easily leads to being premature [32].

Another point that should be noted is that Ω_+^C is required to be nonempty to ensure the inequality in Eq. (12) in Appendix. It is known from Eq. (13) in Appendix that there exists at least one element in Ω_{+}^{C}, otherwise $\sum_{j=1}^{L} D_{i,j}^{C} < p_C \le 1$, which is contradictory to $\sum_{j=1}^{L} D_{i,j}^{C} = 1$.

3.2 Complexity Expect of Mutation Offspring

According to **Step 6** of IGA, the multiple mutation operation can be divided into two steps: (1) The canonical mutation operation repeated for N times [30]; (2) Selection of the offspring from N candidate populations according to the overall complexity. To facilitate the comparative analysis to show the improvement of offspring's complexity, the following lemma is presented to convert the canonical mutation to the similar procedure as the multiple mutation.

Lemma 1: Let $\tilde{M}(\bullet)$ denote the operator, which firstly conducts $N(N \ge 1)$ times canonical mutation and then selects one with the equal probability. Then we have

$$P_{\mathrm{r}}\left(\tilde{M}\left(X^{\mathrm{C}}\right) = X^{\mathrm{M*}}\right) = P_{\mathrm{r}}\left(\hat{M}\left(X^{\mathrm{C}}\right) = X^{\mathrm{M*}}\right) \tag{8}$$

where $\hat{M}(\bullet)$ denotes the canonical mutation operator.

Proof: The proof is in Appendix.

According to Lemma 1, the first step of converted canonical mutation and the proposed multiple mutation is the same. The improvement of offspring's complexity is analyzed only considering the second step and summarized by the following theorem.

Theorem 2: Compared with the canonical mutation operation [30], the offspring complexity expect of the multiple mutation satisfies:

$$E\left(\overline{D}\left(M\left(X^{\mathrm{C}}\right)\right)\right) \ge E\left(\overline{D}\left(\hat{M}\left(X^{\mathrm{C}}\right)\right)\right) \tag{9}$$

Proof: The proof is in Appendix.

Theorem 2 *Ensures that the offspring generated by the multiple mutation operator has a higher overall complexity. This is one of the main objectives of IGA. Besides, the proposed multiple mutation operator only changes the distribution of its offspring population, so the positivity of state transition matrix is not changed.*

4 Application Validation and Analysis

This study focuses on the improvement of GA by modifying the crossover and mutation operators. Since the vehicle could easily fail to park, the collision avoidance performance of APPS is selected to validate the effectiveness of IGA. The test platform integrated with PreScan and Matlab is shown in Fig. 4 [22]. The generated scenarios are automatically constructed by using the API of Prescan. Both the ultrasonic sensors and environments are simulated by Prescan. The tested algorithm and vehicle model are simulated in Matlab/Simulink. The objects and parking space are detected by 12 ultrasonic sensors, the desired parking trajectory composed of arcs is generated using the geometric method, and a preview controller with multiple points is designed to control the steering angle to track the desired trajectory [33].

Considering the following reasons, the simulation scenario is designed as shown in Fig. 5 [33]:

(1) The objects, such as vehicles, can be detected by ultrasonic sensors. If the detected clearance is smaller than the threshold, host vehicle (HV) will be stopped by braking. With this logic, HV will not collide with the surrounding vehicles;
(2) If there exist objects on both sides of the parking space, the parking process is easily stopped by other factors, such as "no available parking space" and "number of moves exceeds the limit". Boundary vehicle (BV) is arranged on one side of the parking space.

The considered influence factors are illustrated by Fig. 5, where F_1 is the distance from BV to curb, F_2 is the heading of BV, F_3 is the distance between BV and HV, F_4 is the heading of HV and F_5 is the speed of HV. The importance degree of each factor was obtained by AHP introduced in Sect. 2.1. Some of them are shown in Table 1 as an example.

4.1 Comparative analysis

The objective function is the minimum distance from HV to curb after parking, and once collision happens the parking process is finished. How to calculate the objective function is illustrated in Fig. 6, where *G, A, B, C,* and *D* denote

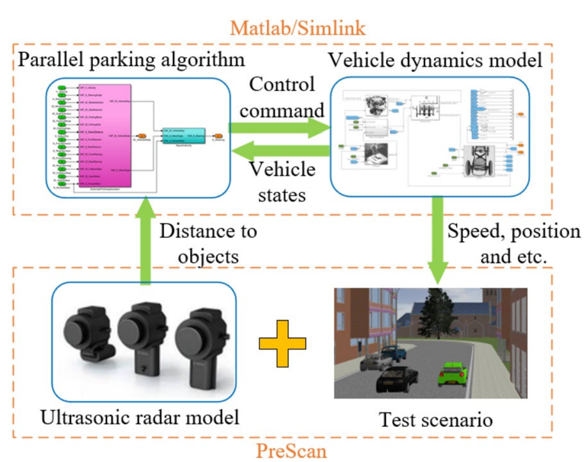

Fig. 4 Test system structure

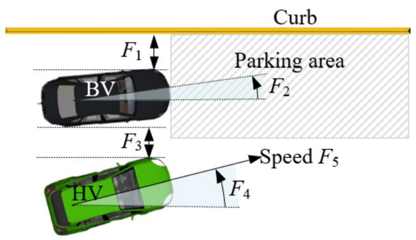

Fig. 5 The diagram of test scenario

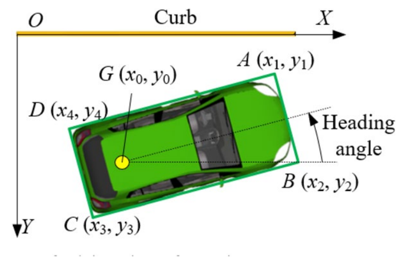

Fig. 6 Diagram of objection function

Table 1 Hierarchal model of influence factors

Layer 1	$S_{1,j}$	Layer 2	$S_{2,j}$	Discrete value	I_n
Distance	0.47	F_1	0.50	0.350 m	0.161
				0.850 m	0.031
				…	…
				1.850 m	0.019
		F_3	0.50	0.400 m	0.120
				0.675 m	0.061
				…	…
				1.500 m	0.008
Heading	0.47	F_2	0.25	−4.000 deg	0.008
				−3.000 deg	0.007
				…	…
				4.000 deg	0.040
		F_4	0.75	−2.000 deg	0.012
				−1.000 deg	0.022
				…	…
				2.000 deg	0.180
Speed	0.07	F_5	1.00	10.000 km/h	0.003
				12.000 km/h	0.010
				…	…
				18.000 km/h	0.021

Table 2 Parameters of evolution tests

Symbol	Description	Value
α	Selection pressure	1.9
m	Number of individuals	5
L	Length of gene	100
p_{M}	Mutation probability	0.009
p_{C}	Crossover probability	1
G_{th}	Maximum number of test interactions	25
g_{th}	Threshold of objective function value	0
N	Number of populations	50
d	Influence intensity of scenario complexity	600

the center of gravity and the four vertices of HV's outline, respectively. The objective function is defined as

$$J = \min_{i=1,2,3,4} y_i \tag{10}$$

where J is the objective function, y_i can be calculated according to the coordinate of G, the body parameters, and the heading angle of HV.

With the parameters defined in Table 2, the comparative results are shown in Fig. 7.

From Fig. 7a, the population complexity of IGA with generation increases much more quickly than that of GA. This implies that the proposed full crossover and multiple mutation operators can inherit the good chromosome of parent to generate the offspring with higher complexity. Accordingly, IGA has a better performance to find the performance limit of the collision avoidance as shown by Fig. 7b. Besides, the convergence speed of IGA is about twice that of GA, and the found minimum distance of IGA is zero. This is much smaller than 0.1 m found by GA. Furthermore, the convergence of object functions of IGA is much better than that of GA. After 15 generations, premature convergence is observed in GA. It is also beneficial to improve the convergence stability by introducing the scenario complexity.

This study aims to solve the problem that autonomous driving is more difficult to be realized under more complex conditions. Accordingly, the index for the evaluation of scenario complex is critical to IGA. The statistical results of scenario complexity calculated by Eq. (2) and the values of the objective function are shown in Fig. 8. It shows that the proposed measurement index of scenario complexity has a statistically obvious correlation with the test objective.

From the theory of GA, it is important to keep a proper mutation probability. It is controlled around 0.01 in general [32]. If the mutation probability is too low, it would be hard to generate better offspring, while a big mutation probability causes the evolution direction to be chaotic. In the proposed multiple mutation operator, the mutation probability is influenced by the scenario complexity. It may deviate from its allowable range. To address this problem, the actual mutation ratio is shown in Fig. 9. It is found that the mutation probability of the multiple mutation operator stays around the predefined probability, and it is similar to that of the canonical mutation operator.

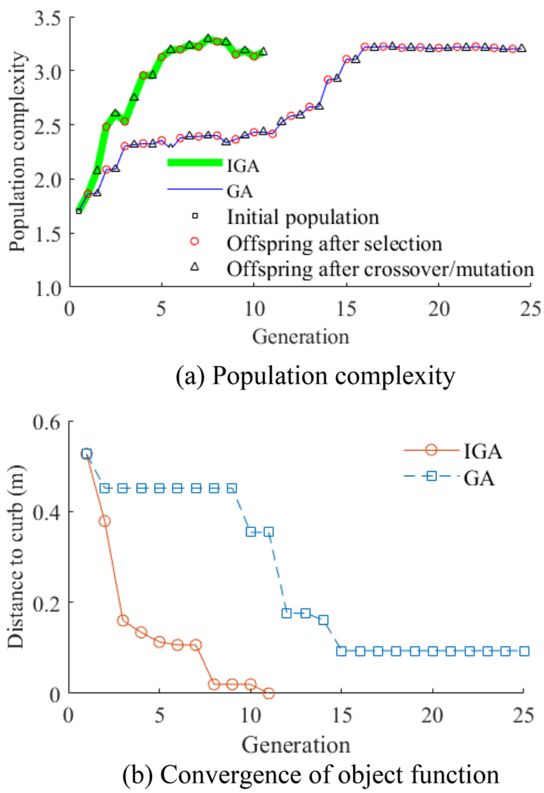

Fig. 7 Comparative evolution test results

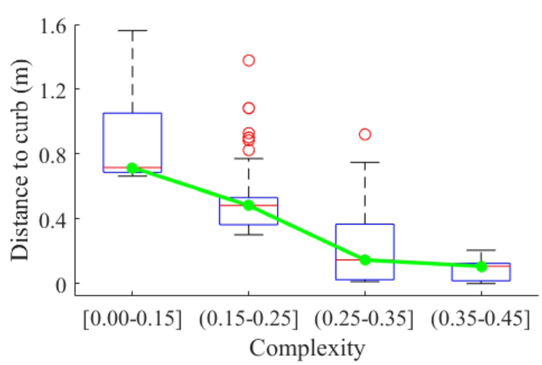

Fig. 8 Relationship between scenario complexity and test effect

4.2 Influence of Algorithm Parameters

Compared with GA, the performance of IGA is influenced by other parameters, i.e., N and d. The former determines the number of generated populations of the canonical mutation operation (see line 1 ~ 3 of **Step 6**) and the latter reflects the influence intensity of scenario complexity on random selection (see line 5 of **Step 5** and line 4 of **Step 6**). This section numerically analyzes the influence of parameters on the performance of test effect and convergence.

Firstly, five groups of tests with $N = 10, 20, 30, 40, 50$ are conducted with each repeated by six times. The average

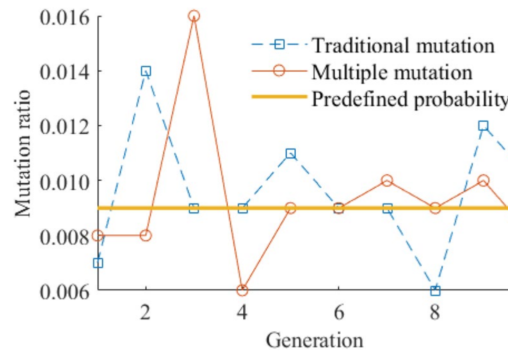

Fig. 9 Comparative results of mutation probability

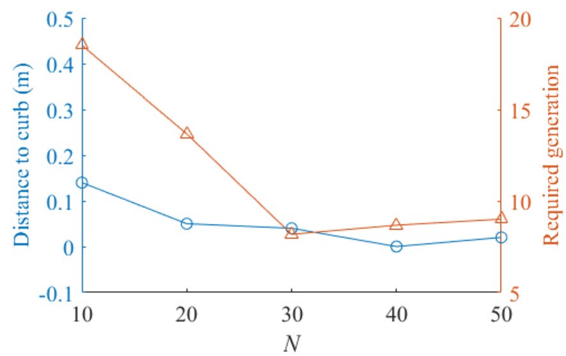

Fig. 10 Influence of mutation times on test performances

results of test effect and convergence are shown in Fig. 10. When N is small, it is more difficult to include the high-quality population in offspring population, which is bad for the test performances. When $N \geq 30$, the test performance tends to be stable, because with the increase of samples the average effect becomes prominent.

Another parameter is analyzed by conducting the tests with $d = 100, 200, 400, 600, 800$. Each type of test is repeated by six times to avoid randomness and average results are shown in Fig. 11. The parameter, d, is positively related to the intervention intensity of complexity. When $d < 400$, the increase of d benefits the test performance. After d reaches 400, the test performance almost keeps unchanged. This is caused by the influence saturation of the complexity intervention.

5 Conclusion

To overcome the challenges of evaluation of performance limit for automated driving systems, an IGA based evolution test strategy is proposed by designing a measurement index of scenario complexity and modifying the original cross/mutation operators to increase test efficiency and

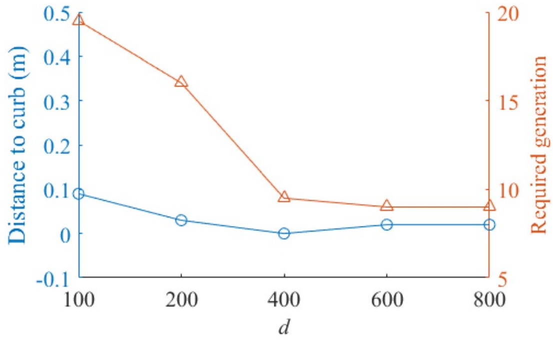

Fig. 11 Influence intensity of scenario complexity

effectiveness. The theoretical analysis and application results on APPS show that:

(1) The proposed complexity index for the test scenario can measure the difficulty of realizing autonomous driving statistically;
(2) The designed full-cross and multiple-mutation operators effectively increase the scenario complexity of offspring;
(3) Compared with GA-based evolution test, IGA-based strategy shows better performance of convergence and test effect.

There remain some issues that are needed to be further studied in the future:

(1) The simulated scenario for the performance limit evaluation of APPS is static. The proposed strategy can be further applied to more complicated intelligent driving systems under dynamic scenarios.
(2) The performance limit may be activated by several different scenarios. The proposed evolution test process is terminated when some possibility is found. How to ensure that all possible scenarios are found needs to be further studied.

Acknowledgments This work is supported by the Open Fund of State Key Laboratory of Vehicle NVH and Safety Technology under Grant NVHSKL-202009 and the Technological Plans of Chongqing under grant cstc2019jcyj-zdxm0022.

Declarations

Conflict of interest On behalf of all the authors, the corresponding author states that there is no conflict of interest.

References

1. Gao, F., Dang, D., He, Y.: Robust coordinated control of nonlinear heterogeneous platoon interacted by uncertain topology. IEEE Trans. Intell. Transp. Syst. (2020). https://doi.org/10.1109/TITS.2020.3045107
2. Dang, D., Gao, F., Hu, Q.: Motion planning for autonomous vehicles considering longitudinal and lateral dynamics coupling. Appl. Sci. **10**(9), 3180 (2020). https://doi.org/10.3390/app10093180
3. Li, L., Huang, W., Liu, Y., Zheng, N., Wang, F.: Intelligence testing for autonomous vehicles: a new approach. IEEE Trans. Intell. Veh. **1**(2), 158–166 (2016)
4. Morton, J., Wheeler, T.A., Kochenderfer, M.J.: Closed-loop policies for operational tests of safety-critical systems. IEEE Trans. Intell. Veh. **3**(3), 317–328 (2018)
5. Tumasov, A.V., Vashurin, A.S., Trusov, Y.P.: The application of hardware-in-the-loop (HIL) simulation for evaluation of active safety of vehicles equipped with electronic stability control (ESC) systems. Proc. Comput. Sci. **150**, 309–315 (2019). https://doi.org/10.1016/j.procs.2019.02.057
6. i-VISTA: i-VISITA intelligent vehicle index. https://www.i-vista.org/download (2020). Accessed 04 March 2021
7. Woo, J.W., Yu, S.B., Lee, S.B.: Design and simulation of a vehicle test bed based on intelligent transport system. Int. J. Automot. Technol. **17**(2), 353–359 (2016)
8. Berger, C., Block, D., Heeren, S.: Simulations on consumer tests: a systematic evaluation approach in an industrial case study. IEEE Intell. Transp. Syst. Mag. **7**(4), 24–36 (2015)
9. Hannan, M.A., Lipu, M.S.H., Hussain, A., Mohamed, A.: A review of lithium-ion battery state of charge estimation and management system in electric vehicle applications: challenges and recommendations. Renew. Sust. Energ. Rev. **78**, 834–854 (2017). https://doi.org/10.1016/j.rser.2017.05.001
10. Zhao, D., Lam, H., Peng, H.: Accelerated evaluation of automated vehicles safety in lane-change scenarios based on importance sampling techniques. IEEE Trans. Intell. Transp. Syst. **18**(3), 595–607 (2017)
11. Perez, M., Angell, L. S., Hankey, J.: Advanced crash avoidance technologies (ACAT) program—final report of the GM-VTTI backing crash countermeasures project. Natural Highway Traffic Safety Admin. (NHTSA), Washington, DC, USA, DOT HS 811 452, (2011)
12. Victor, T., Bärgman, J., Hjälmdahl, M.: Sweden-Michigan naturalistic field operational test (SeMiFOT) phase 1: final report. Chalmers, Goteborg, Sweden, (2010)
13. Burzio, G., Guidotti, L., Perboli, G.: Results and lessons learned of a subjective field operational test on the lane departure warning function. Proc. Soc. Behav. Sci. **48**, 1356–1365 (2012). https://doi.org/10.1016/j.sbspro.2012.06.1111
14. Hou, Y., Zhao, Y., Wagh, A.: Simulation-based testing and evaluation tools for transportation cyber–physical systems. IEEE Trans. Veh. Technol. **65**(3), 1098–1108 (2016)
15. Yang, H., Peng, H.: Development and evaluation of collision warning/collision avoidance algorithms using an errable driver model. Veh. Syst. Dyn. **48**(sup 1), 525–535 (2010)
16. Kang, Y., Yin, H., Berger, C.: Test your self-driving algorithm: an overview of publicly available driving datasets and virtual testing environments. IEEE Trans. Intell. Veh. **4**(2), 171–185 (2019)
17. Gao, F., Duan, J., He, Y., Wang, Z.: A test scenario automatic generation strategy for intelligent driving systems. Math. Probl. Eng. **3737486**, 1–10 (2019). https://doi.org/10.1155/2019/3737486
18. Xia, Q., Duan, J., Gao, F.: Test scenario design for intelligent driving system ensuring coverage and effectiveness. Int. J. Automot. Technol. **19**(4), 751–758 (2018)
19. Zhao, D., Huang, X., Peng, H.: Accelerated evaluation of automated vehicles in car-following maneuvers. IEEE Trans. Intell. Transp. Syst. **19**(3), 733–744 (2018)

20. Wu, P., Gao, F., Li, K.: A vehicle type dependent car-following model based on naturalistic driving study. Electronics **8**(4), 453–468 (2019)
21. Duan, J., Gao, F., He, Y.: Test scenario generation and optimization technology for intelligent driving systems. IEEE Intell. Transp. Syst. Mag. (2020). https://doi.org/10.1109/MITS.2019.2926269
22. Gao, F., Duan, J., Han, Z.: Automatic virtual test technology for intelligent driving systems considering both coverage and efficiency. IEEE Trans. Vehic. Technol. **69**(12), 14365–14376 (2020)
23. Vos, T.E.J., Lindlar, F.F., Wilmes, B.: Evolutionary functional black-box testing in an industrial setting. Softw. Qual. J. **21**(2), 259–288 (2013)
24. Wegener, J., Bühler, O.: Evaluation of different fitness functions for the evolutionary testing of an autonomous parking system. In: Genetic and Evolutionary Computation Conference, Seattle, USA, pp 1400–1412 (2004)
25. Li, J., Li, M.: An analysis on convergence and convergence rate estimate of genetic algorithms in noisy environments. Optik **124**(24), 6780–6785 (2013)
26. Saaty, T.L.: Axiomatic foundation of the analytic hierarchy process. Manage. Sci. **32**(7), 841–855 (1986)
27. Wind, Y., Saaty, T.L.: Marketing applications of the analytic hierarchy process. Manage. Sci. **26**(7), 641–658 (1980)
28. Ishizaka, A., Labib, A.: Review of the main developments in the analytic hierarchy process. Exp. Syst. Appl. **38**(11), 14336–14345 (2011)
29. Barhate, S. S.: Effective test strategy for testing automotive software. In: International Conference on Industrial Instrumentation and Control, Pune, India, 645–649, 28–30 July 2015
30. Rudolph, G.: Convergence analysis of canonical genetic algorithms. IEEE Trans. Neural Netw. **5**(1), 96–101 (1994)
31. Peng, Y., Luo, X., Wei, W.: A new fuzzy adaptive simulated annealing genetic algorithm and its convergence analysis and convergence rate estimation. Int. J. Control Autom. Syst. **12**(3), 670–679 (2014)
32. Bao, X., Xiong, Z., Zhang, N.: Path-oriented test cases generation based adaptive genetic algorithm. PLoS ONE **12**(11), 1–17 (2017)
33. Wang, X.: Study on the key technologies of automatic parking system. Dissertation, Chongqing University (2019).

Feng Gao received the M.S. and Ph.D. degrees from Tsinghua University in 2003 and 2007, respectively. From 2007 to 2013, he was a senior engineer with the Chang'an Auto Global Research and Development Centre, where he has led several projects involving electromagnetic compatibility, durability test of electronic module, ADAS, and engine control. He is currently a Professor with the College of Mechanical and Vehicle Engineering, Chongqing University. His current research interests include robust control and optimization approach with application to automotive systems.

Jianwei Mu received the B. S. degree in Automation from Chongqing University in 2016. He is the author of 2 peer-reviewed papers. His research focuses on the robust designing of autonomous driving system.

Xiangyu Han is an undergraduate student at Chongqing University. He is the co-author of 2 peerreviewed papers. His research interest is the performance boundary analysis of autonomous driving systems.

Yiheng Yang received the B.S. degree in electrical engineering from Nantong University in 2017 and the M.S. in electrical engineering from Chongqing University in 2020. He is the author of 3 peerreviewed papers, and the co-inventor of 3 patents. His research interests are test methods of autonomous driving system.

Junwu Zhou received the B.S. degree from Guangxi University in 2010. Now he serves for Dongfeng Liuzhou Motor Co., Ltd. as the vice director of the Advanced Technology Department of the Passenger Vehicle Technology Center and the secretary general of the Science and Technology Association. His current research interests include new product development and advanced vehicle technologies. He was the recipient of the first prize of the China Innovation Method Competition (Guangxi Division) and the prize for excellence in the National Finals (2020) and the third prize of the Technical Progress Award of Guangxi (2018).

Appendix 1 Definitions and Proofs

Definition of Symbols

g_i: The objective function value corresponding to $\boldsymbol{T}_i$;

g_{th}: The threshold of the objective function value;

G: The generation of population;

G_{th}: The maximum number of interaction test process;

$\boldsymbol{P}_i^{\text{C}}$: The i-th parent individual pair in the natural selected population $\boldsymbol{X}^{\text{S}}$ from $\boldsymbol{X}_G$;

$\boldsymbol{P}_{i,j}^{\text{C}}$: The offspring pair generated by single-point crossover at the j-th position of $\boldsymbol{P}_i^{\text{C}}$;

$\boldsymbol{P}_i^{\text{C}*}$: The selected offspring pair from $\boldsymbol{P}_{i,j}^{\text{C}}, j = 1, \cdots, L$, where L is the number of genes;

$\boldsymbol{T}^*$: The best test case evaluated by g_i, and its objective function value is g^*;

$\boldsymbol{X}_G = \left\{\boldsymbol{T}_i, i = 1, \cdots, 2m\right\}$: The G-th population compose of 2 m individuals;

$\boldsymbol{X}^{\text{C}} = \left\{\boldsymbol{P}_1^{\text{C}*}, \cdots, \boldsymbol{P}_m^{\text{C}*}\right\}$: The offspring population generated by crossover operation;

$\boldsymbol{X}_i^{\text{M}}$: The population generated by conducting the i-th mutation on $\boldsymbol{X}^{\text{C}}$;

$\boldsymbol{X}^{\text{M}*}$: The selected population from $\boldsymbol{X}_i^{\text{M}}, i = 1, \cdots, N$, where N is the mutation times.

$\overline{D}(\mathbf{X}) = \frac{1}{n}\sum_{i=1}^{n} D\left(\boldsymbol{T}_i\right)$: The average complexity of a population $\mathbf{X} = \left\{\boldsymbol{T}_1, \boldsymbol{T}_2, \ldots, \boldsymbol{T}_n\right\}$;

$P_{\text{r}}(\bullet)$: The occurrence probability of an event.

Proof of Theorem 1

From Eqs. (4) and (5):

$$E\left(\overline{D}\left(C\left(\boldsymbol{P}_i^{\text{C}}\right)\right)\right) - E\left(\overline{D}\left(\hat{C}\left(\boldsymbol{P}_i^{\text{C}}\right)\right)\right) = \sum_{j=1}^{L}\left(D_{i,j}^{\text{C}} - \frac{p_{\text{C}}}{L}\right)\overline{D}\left(\boldsymbol{P}_{i,j}^{\text{C}}\right) - (1 - p_{\text{C}})\overline{D}(\boldsymbol{P}_i^{\text{C}}), \tag{11}$$

where $D_{i,j}^{C} = e^{d\times\overline{D}\left(P_{i,j}^{C}\right)} / \sum_{k=1}^{L} e^{d\times\overline{D}\left(P_{i,j}^{C}\right)}$ is the normalized complexity of $\boldsymbol{P}_{i,j}^{C}$. According to whether the full crossover can increase the complexity of offspring, Eq. (11) can be rewritten as

$$\begin{aligned}
&E\left(\overline{D}\left(C\left(\boldsymbol{P}_i^{\text{C}}\right)\right)\right) - E\left(\overline{D}\left(\hat{C}\left(\boldsymbol{P}_i^{\text{C}}\right)\right)\right) \\
&\quad = \sum_{j\in\Omega_+^{\text{C}}}\left(D_{i,j}^{\text{C}} - \frac{p_{\text{C}}}{L}\right)\overline{D}\left(\boldsymbol{P}_{i,j}^{\text{C}}\right) + \sum_{j\in\Omega_-^{\text{C}}}\left(D_{i,j}^{\text{C}} - \frac{p_{\text{C}}}{L}\right)\overline{D}\left(\boldsymbol{P}_{i,j}^{\text{C}}\right) \\
&- (1 - p_{\text{C}})\overline{D}(\boldsymbol{P}_i^{\text{C}}) \geq \min_{j\in\Omega_+^{\text{C}}}\left(\overline{D}\left(\boldsymbol{P}_{i,j}^{\text{C}}\right)\right)\sum_{j\in\Omega_+^{\text{C}}}\left(D_{i,j}^{\text{C}} - \frac{p_{\text{C}}}{L}\right) \\
&+ \max_{j\in\Omega_-^{\text{C}}}\left(\overline{D}\left(\boldsymbol{P}_{i,j}^{\text{C}}\right)\right)\sum_{j\in\Omega_-^{\text{C}}}\left(D_{i,j}^{\text{C}} - \frac{p_{\text{C}}}{L}\right) - (1 - p_{\text{C}})\overline{D}(\boldsymbol{P}_i^{\text{C}})
\end{aligned} \tag{12}$$

where $\Omega_+^{\text{C}} = \left\{j | D_{i,j}^{\text{C}} \geq \frac{p_{\text{C}}}{L}\right\}$ and $\Omega_-^{\text{C}} = \{j | j = 1, \cdots, L\} - \Omega_+^{\text{C}}$.

From Eq. (5) and the definition of $D_{i,j}^{\text{C}}$:

$$\begin{aligned}
&\sum_{j=1}^{L}\frac{p_{\text{C}}}{L} + 1 - p_{\text{C}} = 1 \text{ and } \sum_{j=1}^{L} D_{i,j}^{\text{C}} = 1 \Rightarrow \sum_{j\in\Omega_+^{\text{C}}}\left(D_{i,j}^{\text{C}} - \frac{p_{\text{C}}}{L}\right) \\
&\quad + \sum_{j\in\Omega_-^{\text{C}}}\left(D_{i,j}^{\text{C}} - \frac{p_{\text{C}}}{L}\right) - \left(1 - p_{\text{C}}\right) = 0.
\end{aligned} \tag{13}$$

when $p_{\text{C}} = 1$, the following equation establishes from Eq. (13):

$$\sum_{j\in\Omega_-^{\text{C}}}\left(D_{i,j}^{\text{C}} - \frac{1}{L}\right) = -\sum_{j\in\Omega_+^{\text{C}}}\left(D_{i,j}^{\text{C}} - \frac{1}{L}\right) \tag{14}$$

Then substituting Eq. 14 and $p_{\text{C}} = 1$ to Eq. (12) yields

$$E\left(\overline{D}\left(C\left(\boldsymbol{P}_i^{\text{C}}\right)\right)\right) - E\left(\overline{D}\left(\hat{C}\left(\boldsymbol{P}_i^{\text{C}}\right)\right)\right) \geq \left[\min_{j\in\Omega_+^{\text{C}}}\left(\overline{D}\left(\boldsymbol{P}_{i,j}^{\text{C}}\right)\right) - \max_{j\in\Omega_-^{\text{C}}}\left(\overline{D}\left(\boldsymbol{P}_{i,j}^{\text{C}}\right)\right)\right]\sum_{j\in\Omega_+^{\text{C}}}\left(D_{i,j}^{\text{C}} - \frac{1}{L}\right) \tag{15}$$

According to the definition of Ω_+^{C} and Ω_-^{C}:

$$D_{i,j}^{\text{C}} - \frac{1}{L} \geq 0, \forall j \in \Omega_+^{\text{C}} \text{ and } \max_{j\in\Omega_-^{\text{C}}}\left(\overline{D}\left(\boldsymbol{P}_{i,j}^{\text{C}}\right)\right)\frac{1}{d}\ln\left(\frac{p_{\text{C}}}{L}\sum_{k=1}^{L}\mathrm{e}^{d\times\overline{D}\left(\boldsymbol{P}_{i,k}^{\text{C}}\right)}\right) \leq \min_{j\in\Omega_+^{\text{C}}}\left(\overline{D}\left(\boldsymbol{P}_{i,j}^{\text{C}}\right)\right) \tag{16}$$

Conclusion **C1** is proved by substituting Eq. (16) to Eq. (15).

When $0 \le p_C < 1$, the following equation is deduced by substituting Eq. (13) to Eq. (11):

$$E\left(\overline{D}\left(C\left(\boldsymbol{P}_i^{\mathrm{C}}\right)\right)\right) - E\left(\overline{D}\left(\widehat{C}\left(\boldsymbol{P}_i^{\mathrm{C}}\right)\right)\right) = \sum_{j=1}^{L} D_{i,j}^{\mathrm{C}}\left(\overline{D}\left(\boldsymbol{P}_{i,j}^{\mathrm{C}}\right) - \overline{D}(\boldsymbol{P}_i^{\mathrm{C}})\right) - \frac{p_{\mathrm{C}}}{L}\sum_{j=1}^{L}\left(\overline{D}\left(\boldsymbol{P}_{i,j}^{\mathrm{C}}\right) - \overline{D}(\boldsymbol{P}_i^{\mathrm{C}})\right) \tag{17}$$

Equation (17) is monotonically decreasing with p_{C} as the variable when Eq. (7) establishes, and so **C2** establishes with the conclusion derived from **C1**.

Proof of Lemma 1

Since the N times canonical mutations are independent, the following equations are obtained:

$$\begin{aligned} P_{\mathrm{r}}\left(\tilde{M}\left(\boldsymbol{X}^{\mathrm{C}}\right) = \boldsymbol{X}^{\mathrm{M}*}\right) &= \sum_{i=0}^{N}\left[\frac{i\mathrm{C}_N^i}{N}\mathrm{P}_{\mathrm{r}}^i\left(\widehat{M}\left(\boldsymbol{X}^{\mathrm{C}}\right) = \boldsymbol{X}^{\mathrm{M}*}\right) \times \left(1 - P_{\mathrm{r}}\left(\widehat{M}\left(\boldsymbol{X}^{\mathrm{C}}\right) = \boldsymbol{X}^{\mathrm{M}*}\right)\right)^{N-i}\right] \\ &= P_{\mathrm{r}}\left(\widehat{M}\left(\boldsymbol{X}^{\mathrm{C}}\right) = \boldsymbol{X}^{\mathrm{M}*}\right) \times \sum_{i=1}^{N-1}\left[\mathrm{C}_{N-1}^{i-1}\mathrm{P}_{\mathrm{r}}^{i-1}\left(\widehat{M}\left(\boldsymbol{X}^{\mathrm{C}}\right) = \boldsymbol{X}^{\mathrm{M}*}\right) \times \left(1 - P_{\mathrm{r}}\left(\widehat{M}\left(\boldsymbol{X}^{\mathrm{C}}\right) = \boldsymbol{X}^{\mathrm{M}*}\right)\right)^{N-i}\right] \end{aligned} \tag{18}$$

where $\mathrm{C}_N^i = \frac{N!}{i!(N-i)!}$ denotes the combination number.

According to the binomial expansion

$$\sum_{i=1}^{N-1}\left[\mathrm{C}_{N-1}^{i-1}\mathrm{P}_{\mathrm{r}}^{i-1}\left(\widehat{M}\left(\boldsymbol{X}^{\mathrm{C}}\right) = \boldsymbol{X}^{\mathrm{M}*}\right) \times \left(1 - P_{\mathrm{r}}\left(\widehat{M}\left(\boldsymbol{X}^{\mathrm{C}}\right) = \boldsymbol{X}^{\mathrm{M}*}\right)\right)^{N-i}\right] = \left[P_{\mathrm{r}}\left(\widehat{M}\left(\boldsymbol{X}^{\mathrm{C}}\right) = \boldsymbol{X}^{\mathrm{M}*}\right) + 1 - P_{\mathrm{r}}\left(\widehat{M}\left(\boldsymbol{X}^{\mathrm{C}}\right) = \boldsymbol{X}^{\mathrm{M}*}\right)\right]^{N-1} = 1 \tag{19}$$

Lemma 1: Is proved by substituting Eq. (19) to Eq. (18).

Proof of Theorem 2

According to Step 6 and Lemma 1:

$$E\left(\overline{D}\left(M\left(\boldsymbol{X}^{\mathrm{C}}\right)\right)\right) - E\left(\overline{D}\left(\widehat{M}\left(\boldsymbol{X}^{\mathrm{C}}\right)\right)\right) = \sum_{i=1}^{N}\left(D_i^{\mathrm{M}} - \frac{1}{N}\right) \times \overline{D}(\boldsymbol{X}_i^{\mathrm{M}}) \tag{20}$$

Being similar to the analysis procedure of Eq. (12), Eq. (20) is re-written as

$$\begin{aligned} E\left(\overline{D}\left(M\left(\boldsymbol{X}^{\mathrm{C}}\right)\right)\right) - E\left(\overline{D}\left(\widehat{M}\left(\boldsymbol{X}^{\mathrm{C}}\right)\right)\right) &\ge \min_{i\in\Omega_+^{\mathrm{M}}}\left(\overline{D}\left(\boldsymbol{X}_i^{\mathrm{M}}\right)\right) \\ &\times \sum_{i\in\Omega_+^{\mathrm{M}}}\left(D_i^{\mathrm{M}} - \frac{1}{N}\right) + \max_{i\in\Omega_-^{\mathrm{M}}}\left(\overline{D}\left(\boldsymbol{X}_i^{\mathrm{M}}\right)\right) \times \sum_{i\in\Omega_-^{\mathrm{M}}}\left(D_i^{\mathrm{M}} - \frac{1}{N}\right) \end{aligned} \tag{21}$$

where $\Omega_+^{\mathrm{M}} = \left\{i|D_i^{\mathrm{M}} \ge \frac{1}{N}\right\}, \Omega_-^{\mathrm{M}} = \{i|i = 1, \cdots, N\} - \Omega_+^{\mathrm{M}}$ and

$D_i^M = e^{d\times\overline{D}(X_i^M)} / \sum_{k=1}^{N} e^{d\times\overline{D}(X_k^M)}$ is the normalized complexity. Equation (9) is derived referring to the analysis process from (15) to (16) with the fact that $\sum_{i=1}^{N} D_i^M = 1$.