A Self-Optimization Approach for L-SHADE Incorporated with Eigenvector-Based Crossover and Successful-Parent-Selecting Framework on CEC 2015 Benchmark Set

Shu-Mei Guo
Computer Science and Information Engineering
National Cheng-Kung University
Tainan, Taiwan, R.O.C.
Email: guosm@mail.ncku.edu.tw

Chin-Chang Yang
Computer Science and Information Engineering
National Cheng-Kung University
Tainan, Taiwan, R.O.C.
Email: p78981061@mail.ncku.edu.tw

Jason Sheng-Hong Tsai
Electrical Engineering
National Cheng-Kung University
Tainan, Taiwan, R.O.C.
Email: shtsai@mail.ncku.edu.tw

Pang-Han Hsu
Computer Science and Information Engineering
National Cheng-Kung University
Tainan, Taiwan, R.O.C.
Email: p76021093@mail.ncku.edu.tw

***Abstract*—A self-optimization approach and a new success-history based adaptive differential evolution with linear population size reduction (L-SHADE) which is incorporated with an eigenvector-based (EIG) crossover and a successful-parent-selecting (SPS) framework are proposed in this paper. The EIG crossover is a rotationally invariant operator which provides superior performance on numerical optimization problems with highly correlated variables. The SPS framework provides an alternative of the selection of parents to prevent the situation of stagnation. The proposed SPS-L-SHADE-EIG combines the L-SHADE with the EIG and SPS frameworks. To further improve the performance, the parameters of SPS-L-SHADE-EIG are self-optimized in terms of each function under IEEE Congress on Evolutionary Computation (CEC) benchmark set in 2015. The stochastic population search causes the performance of SPS-L-SHADE-EIG noisy, and therefore we deal with the noise by re-evaluating the parameters if the parameters are not updated for more than an unacceptable amount of times. The experiment evaluates the performance of the self-optimized SPS-L-SHADE-EIG in CEC 2015 real-parameter single objective optimization competition.**

Keywords*—*differential evolution; global numerical optimization; noisy optimization

I. INTRODUCTION

Many studies advance differential evolution (DE) to solve global optimization problems, in which the state-of-the-art L-SHADE is a success-history based adaptive differential evolution [1], [2] with linear population size reduction [3] in CEC 2014 real-parameter single objective competition. However, the binomial crossover operator that used by L-SHADE is not rotationally invariant, and thus the performance of L-SHADE is sensitive to the rotation of coordinate systems, which represents highly correlated parameters in real-world optimization problems [4]. To deal with this problem, the eigenvector (EIG) information of the covariance matrix of population can be utilized to provide a rotated coordinate system for crossover operation [4]. On the other hand, L-SHADE may gradually stop generating successful solutions even though the population has not converged to a fixed point. This situation is commonly referred as *stagnation*. To avoid the *stagnation*, a successful-parent-selecting (SPS) framework [5] utilizes recently updated solutions to provide alternative selection of parents.

In the previous research, specifying different parameters for different test functions is not allowed because DE is generally expected to solve a diverse domain of problems. However, in real-world optimization, DE is used to solve a very specific problem. It is important to provide a guideline on how to adjust the parameters of DE for a particular case. In the CEC 2015 competition on learning-based real-parameter single objective optimization, the objective is not to identify the best algorithm for solving each of the benchmark problems. Instead, optimizing the parameters of the proposed optimization algorithm is allowed for solving each problem [6].

DE is a stochastic search algorithm, so that optimizing the parameters of DE is basically an optimization problem with uncertainty of fitness values. To handle the uncertainty, the expected fitness of a solution can be estimated by taking the mean of multiple function evaluations. However, this method dramatically increases the number of function evaluations. In noisy expensive optimization problems, the large number of required function evaluations makes the application impractical.

In this paper, a new DE variant SPS-L-SHADE-EIG is proposed to solve global optimization problems. The proposed SPS-L-SHADE-EIG incorporates L-SHADE with the EIG crossover and the SPS framework. To self-optimize its parameters, an efficient uncertainty handling method is proposed to handle the uncertainty of the performance of the proposed SPS-L-SHADE-EIG. The performance of the self-optimized SPS-L-SHADE-EIG is evaluated on CEC 2015 benchmark set [6] with 10, 30, 50, and 100 dimensions.

This paper is organized as follows. The proposed SPS-L-SHADE-EIG and the proposed uncertainty handling method are presented in Section II. Then, the self-optimization approach and the performance of SPS-L-SHADE-EIG are shown in Section III. Finally, this paper is concluded in Section IV.

II. L-SHADE with Eigenvector-Based Crossover and Successful-Parent-Selecting Framework

This section describes the L-SHADE incorporated with the EIG crossover and the SPS framework, namely SPS-L-SHADE-EIG. The proposed uncertainty handling method is shown in the final subsection.

A. Initialization

The population of NP D-dimensional parameter vectors is defined as:

$$\mathbf{P}_G = \left\{\mathbf{x}_{i,G} \mid \mathbf{x}_{i,G} = (x_{1,i,G}, \ldots, x_{D,i,G})^T, i = 1, \ldots, NP\right\} \tag{1}$$

where G denotes the generation number. The initial population $\mathbf{P}_0$ is randomly initialized within the bounds.

B. Successful-Parent-Selecting Framework

The successful-parent-selecting (SPS) framework has been shown an effective and efficient method to provide more promising solutions and help DE algorithms escaping the situation of stagnation [5]. The basic idea of the SPS framework is to select parents from the most recently updated solutions when stagnation is detected. More specifically, the source of the ith parent $\mathbf{p}_{i,G}$ is given by:

$$\mathbf{p}_{i,G} \in \begin{cases} \mathbf{P}_G, & \text{if } q_{i,G} \le Q \\ \mathbf{A}^{(\mathrm{SPS})}, & \text{otherwise} \end{cases} \tag{2}$$

where $Q > 0$ is a control parameter, $q_{i,G}$ is the number of recently consecutive unsuccessful updates of the ith solution in the Gth generation:

$$q_{i,G+1} = \begin{cases} 0, & \text{if } f(\mathbf{u}_{i,G}) < f(\mathbf{x}_{i,G}) \\ q_{i,G} + 1, & \text{otherwise} \end{cases} \tag{3}$$

with an initial value $q_{i,0} = 0$. $\mathbf{A}^{(\mathrm{SPS})}$ is an archive of a number of NP recently updated solutions. The archive $\mathbf{A}^{(\mathrm{SPS})}$ is initiated to be a copy of the initial population $\mathbf{P}_0$. Then, the successful solutions (i.e., those *trial* vectors $\mathbf{u}_{i,G}$ satisfy $f(\mathbf{u}_{i,G}) < f(\mathbf{x}_{i,G})$) replace the oldest solutions to keep the archive size at NP.

C. Current-to-pbest/1 Mutation Strategy with External Archive

The current-to-*p*best/1 mutation strategy has been shown a powerful method to generate a promising *donor* vector in [1], [2], and [3]:

$$\mathbf{v}_{i,G} = \mathbf{p}_{i,G} + F_{i,G} \cdot (\mathbf{p}_{pbest,G} - \mathbf{p}_{i,G}) + F_{i,G} \cdot (\mathbf{p}_{r_1,G} - \tilde{\mathbf{p}}_{r_2,G}) \tag{4}$$

where $F_{i,G}$ is the ith scaling factor, *p*best is a randomly selected index of the solutions from the top $p \times NP$ best solutions. r_1 and r_2 are randomly selected indexes from [1, NP]. $\tilde{\mathbf{p}}_{r_2,G}$ is selected from the union of the source of the parent $\mathbf{p}_{r_2,G}$ (either $\mathbf{P}_G$ or $\mathbf{A}^{(\mathrm{SPS})}$, depending on the value $q_{i,G}$) and an external archive $\mathbf{A}^{(\mathrm{EXT})}$ which collects a number of $w^{(\mathrm{EXT})} \times NP$ inferior solutions to maintain solution diversity. The external archive $\mathbf{A}^{(\mathrm{EXT})}$ is initiated to an empty set. Then, the parent vector $\mathbf{p}_{i,G}$ is added into the external archive if it is worse than the *trial* vector $\mathbf{u}_{i,G}$. When the size of external archive exceeds the maximum size $w^{(\mathrm{EXT})} \times NP$, a randomly selected solution is removed from the external archive until the size is equal to $w^{(\mathrm{EXT})} \times NP$.

D. Eigenvector-Based Crossover Operator

The binomial crossover operator is denoted as a bivariate function $\mathrm{xover}(\mathbf{p}_{i,G}, \mathbf{v}_{i,G})$ on a parent vector $\mathbf{p}_{i,G} = \{p_{1,i,G}, p_{2,i,G}, \ldots, p_{D,i,G}\}$ and a *donor* vector $\mathbf{v}_{i,G} = \{v_{1,i,G}, v_{2,i,G}, \ldots, v_{D,i,G}\}$. The crossover operation generates a trial vector $\mathbf{u}_{i,G} = \{u_{1,i,G}, u_{2,i,G}, \ldots, u_{D,i,G}\}$ by:

$$u_{j,i,G} = \begin{cases} v_{j,i,G}, & \text{if } rand_{j,i,G}[0,1] \le CR_{i,G} \text{ or } j = j_{rand} \\ p_{j,i,G}, & \text{otherwise} \end{cases} \tag{5}$$

where $CR_{i,G}$ is the crossover rate and j_{rand} is a randomly selected integer between 1 to D to ensure at least one variable in the *donor* vector is copied to the *trial* vector.

Algorithm 1. Success-History Update Algorithm

if $S_G^{(F)}, S_G^{(\mathrm{ER})}, S_G^{(\mathrm{CR})} \neq \varnothing$
    $M_k^{(F)} = \mathrm{mean}_{\mathrm{WL}}(S_G^{(F)})$;
    $M_k^{(ER)} = \mathrm{mean}_{\mathrm{WA}}(S_G^{(ER)})$;
    $M_k^{(CR)} = \mathrm{mean}_{\mathrm{WA}}(S_G^{(CR)})$;
    $k = k + 1$;
    if $k > H$
        $k = 1$;
    end
end

The eigenvector-based (EIG) crossover operator [4] is an improvement of binomial crossover to generate a *trial* vector:

$$\mathbf{u}_{i,G} = \begin{cases} \mathbf{B}_G \cdot \mathrm{xover}(\mathbf{B}_G^*\mathbf{p}_{i,G}, \mathbf{B}_G^*\mathbf{v}_{i,G}), & \text{if } rand_{i,G}[0,1] \le ER_{i,G} \\ \mathrm{xover}(\mathbf{p}_{i,G}, \mathbf{v}_{i,G}), & \text{otherwise} \end{cases} \tag{6}$$

where $rand_{i,G}[0, 1]$ is a random variable between 0 and 1, $ER_{i,G}$ is the eigenvector-based ratio, and $\mathbf{B}_G^*$ denotes the conjugate transpose of $\mathbf{B}_G$ which is a $D \times D$ matrix whose kth column is the eigenvector of the covariance matrix of the population:

$$\mathbf{C}_G = \mathbf{B}_G \mathbf{\Lambda}_G \mathbf{B}_G^{-1} \tag{7}$$

where $\mathbf{\Lambda}_G$ is a diagonal matrix whose elements are eigenvalues of the covariance matrix $\mathbf{C}_G$. In [4], the covariance matrix is directly updated from the covariance of the current population. However, this method generates a covariance matrix which is sensitive to the change of population distribution, and may cause the rotated coordinate system unsteady. To ease the unsteady problem, in this paper, we propose a new update method for the covariance matrix as follows:

$$\mathbf{C}_{G+1} = (1-\alpha_G)\cdot \mathbf{C}_G + \alpha_G \cdot \mathrm{cov}(\mathbf{P}_{G+1}) \tag{8}$$

where $0 < \alpha_G < 1$ is a learning rate, and cov($\mathbf{X}$) computes the covariance matrix of $\mathbf{X}$ by:

$$\mathrm{cov}(\mathbf{X}) = \left[c_{i,j} \;\middle|\; c_{i,j} = \frac{\sum_{k=1}^{NP}\left(x_{i,k} - \bar{x}_i\right)\left(x_{j,k} - \bar{x}_j\right)}{NP - 1} \right] \tag{9}$$

where $c_{i,j}$ is the elements of the covariance matrix in the ith row and the jth column, and $\bar{x}_i = \left(\sum_{k=1}^{NP} x_{i,k}\right) \Big/ NP$ denotes the mean value of the variables in the ith dimension. The learning rate α_G in (8) is initiated to α_{init}, and linearly decreases by:

Algorithm 2. SPS-L-SHADE-EIG

$G = 0$, $NP = NP_{\mathrm{init}}$, $\mathbf{A}^{(\mathrm{EXT})} = \varnothing$;
Initialize population $\mathbf{P}_G = \{\mathbf{x}_{1,G}, \ldots, \mathbf{x}_{NP,G}\}$;
Initialize the archive $\mathbf{A}^{(\mathrm{SPS})} = \{\mathbf{s}_{1,G}, \ldots, \mathbf{s}_{NP,G}\} = \mathbf{P}_G$;
Evaluate function values of $\{\mathbf{x}_{i,G}\}$, and $FES = NP$;
$q_{1,G} = \cdots = q_{NP,G} = 0$;
$\mathbf{C}_G = \mathrm{cov}(\mathbf{P}_G) = \mathbf{B}_G\mathbf{\Lambda}_G\mathbf{B}_G^{-1}$; // see Eqs. (7) and (9)
$\{M_k^{(F)}\} = F_{\mathrm{init}}$, $\{M_k^{(ER)}\} = ER_{\mathrm{init}}$, $\{M_k^{(CR)}\} = CR_{\mathrm{init}}$;
while the termination criteria are not met **do**
    G++; $S_G^{(F)}, S_G^{(ER)}, S_G^{(CR)} = \varnothing$;
    for $i = 1$ **to** NP **do**
        Randomly select r_i from [1, H].
        $F_{i,G} = \mathrm{randc}_{i,G}(M_{r_i}^{(F)}, w^{(F)})$;
        $ER_{i,G} = \mathrm{randn}_{i,G}(M_{r_i}^{(ER)}, w^{(ER)})$;
        $CR_{i,G} = \mathrm{randn}_{i,G}(M_{r_i}^{(CR)}, w^{(CR)})$;
        if $q_{i,G} \le Q$, $\mathbf{p} \equiv \mathbf{x}$;
        else: $\mathbf{p} \equiv \mathbf{s}$;
        end
        $\mathbf{v}_{i,G} = \mathbf{p}_{i,G} + F_{i,G}\cdot(\mathbf{p}_{pbest,G} - \mathbf{p}_{i,G}) + F_{i,G}\cdot(\mathbf{p}_{r_1,G} - \tilde{\mathbf{p}}_{r_2,G})$
        if $rand_{i,G} \le ER_{i,G}$
            $\mathbf{u}_{i,G} = \mathbf{B}_G \cdot \mathrm{xover}(\mathbf{B}_G^*\mathbf{p}_{i,G}, \mathbf{B}_G^*\mathbf{v}_{i,G})$;
        else: $\mathbf{u}_{i,G} = \mathrm{xover}(\mathbf{p}_{i,G}, \mathbf{v}_{i,G})$;
        end
        if we are handling noise **and** $q_{i,G} > Q$
            Re-evaluate the function value of $\mathbf{x}_{i,G}$;
            FES++;
        end
        if $f(\mathbf{u}_{i,G}) < f(\mathbf{x}_{i,G})$
            $\mathbf{x}_{i,G+1} = \mathbf{u}_{i,G}$;
            $S_G^{(F)} = S_G^{(F)} \cup \{F_{i,G}\}$;
            $S_G^{(ER)} = S_G^{(ER)} \cup \{ER_{i,G}\}$;
            $S_G^{(CR)} = S_G^{(CR)} \cup \{CR_{i,G}\}$;
            $\mathbf{A}^{(\mathrm{EXT})} = \mathbf{A}^{(\mathrm{EXT})} \cup \{\mathbf{x}_{i,G}\}$;
            $\mathbf{A}^{(\mathrm{SPS})} = \mathbf{A}^{(\mathrm{SPS})} \cup \{\mathbf{u}_{i,G}\}$;
            $q_{i,G+1} = 0$;
        else
            $\mathbf{x}_{i,G+1} = \mathbf{x}_{i,G}$;
            $q_{i,G+1} = q_{i,G} + 1$;
        end
        FES++;
    end
    Update $\mathbf{C}_{G+1}$ and NP_{G+1} by (8) and (19).
    Update $\{M_k^{(F)}\}$, $\{M_k^{(ER)}\}$ and $\{M_k^{(CR)}\}$ by Algorithm 1.
    Resize the population $\mathbf{P}_{G+1}$ and $\mathbf{A}^{(\mathrm{SPS})}$ with their best NP_{G+1} solutions. Also, resize the archive $\mathbf{A}^{(\mathrm{EXT})}$.
end

$$\alpha_G = \alpha_{\mathrm{init}} \cdot \left(1 - FES/MAXFES\right) \tag{10}$$

TABLE I. OPTIMIZED PARAMETERS OF SPS-L-SHADE-EIG AND THEIR SOLUTION ERROR VALUES ON CEC 2015 BENCHMARK FUNCTIONS FOR $D = 10$ AND $D = 30$

$D = 10$	f_1	f_2	f_3	f_4	f_5	f_6	f_7	f_8	f_9	f_{10}	f_{11}	f_{12}	f_{13}	f_{14}	f_{15}	Default
$NP_{\text{init}} - NP_{\min}$	447	381	155	333	83	74	41	48	74	41	269	159	45	51	1	186
F_{init}	0.4972	0.3116	0.4085	0.2915	0.3045	0.2307	0.1648	0.4923	0.3559	0.3709	0.1738	0.7078	0.5417	0.1492	0.3365	0.5
CR_{init}	0.2892	0.2240	0.1980	0.3603	0.8335	0.6595	0.3418	0.1788	0.7627	0.9553	0.5553	0.7912	0.8090	0.9843	0.8308	0.5
ER_{init}	0.9497	0.8019	0.3172	0.6455	0.6024	0.9822	0.3041	0.9161	0.7073	0.1387	0.3240	0.9078	0.2411	0.5587	0.0670	1
p	0.0956	0.3274	0.2230	0.1907	0.1739	0.3519	0.0447	0.0752	0.0213	0.4105	0.0118	0.0530	0.0399	0.9973	0.2089	0.11
H	422	695	431	61	335	13	956	186	55	730	60	296	405	112	139	6
Q	64	920	151	13	433	634	231	301	428	31	41	9	115	748	41	64
$w^{(\text{EXT})}$	1.9805	2.0466	1.7238	1.5365	2.4490	1.6688	1.4542	1.8254	2.5645	1.4141	1.0355	1.1197	6.0485	2.5510	4.1768	2.6
α_{init}	0.6288	0.2551	0.1720	0.2371	0.8216	0.1777	0.3434	0.7266	0.6565	0.1957	0.6467	0.3753	0.8429	0.3513	0.1143	0.3
$w^{(ER)}$	0.9854	0.7163	0.1717	0.3067	0.7593	0.1636	0.1873	0.1949	0.8203	0.0325	0.8030	0.9077	0.0684	0.6262	0.4346	0.2
$CR_{\min}$	0.6011	0.9567	0.0100	0.5660	0.0867	0.0029	0.1518	0.1026	0.4171	0.9450	0.4862	0.0269	0.0811	0.0087	0.3976	0.05
$CR_{\max} - CR_{\min}$	0.5662	0.2001	0.9248	0.3895	0.2555	0.3763	0.1711	0.1593	0.9230	0.0449	0.8004	0.8083	0.0636	0.0265	0.7673	0.25
$NP_{\min}$	12	6	9	151	9	4	13	4	84	72	80	68	31	6	147	4
$w^{(CR)}$	0.3349	0.2290	0.9933	0.2979	0.0422	0.8292	0.7929	0.3458	0.1223	0.2581	0.3177	0.2641	0.3488	0.7864	0.0037	0.1
$w^{(F)}$	0.1315	0.5178	0.4899	0.0312	0.9708	0.7779	0.4402	0.8679	0.0250	0.8553	0.0422	0.1757	0.5201	0.5410	0.0679	0.1
Solution Error	0.00E+00	2.84E–14	1.14E–12	6.65E–12	3.12E–01	6.82E–13	2.68E–08	2.83E–07	1.00E+02	1.41E+02	2.98E–01	1.00E+02	9.25E–02	1.00E+02	1.00E+02	–
$D = 30$	f_1	f_2	f_3	f_4	f_5	f_6	f_7	f_8	f_9	f_{10}	f_{11}	f_{12}	f_{13}	f_{14}	f_{15}	Default
$NP_{\text{init}} - NP_{\min}$	345	255	77	758	393	359	186	363	267	290	526	301	93	482	41	566
F_{init}	0.4076	0.6028	0.2070	0.4420	0.5752	0.0535	0.7792	0.8954	0.5398	0.2786	0.3284	0.5501	0.0951	0.9684	0.3234	0.5
CR_{init}	0.6209	0.5957	0.9072	0.9855	0.9315	0.3291	0.9532	0.0808	0.5509	0.6071	0.3080	0.3818	0.4193	0.0393	0.2685	0.5
ER_{init}	0.1399	0.8766	0.5580	0.9004	0.4807	0.5878	0.4563	0.8566	0.7752	0.7523	0.7641	0.3944	0.2666	0.2447	0.3126	1
p	0.1340	0.5900	0.9489	0.4366	0.1319	0.0878	0.3195	0.0368	0.1690	0.2006	0.9889	0.2574	0.1689	0.5269	0.2242	0.11
H	90	298	522	21	385	4	918	349	379	865	171	184	19	156	364	6
Q	194	973	473	30	27	15	693	17	41	897	11	3	119	12	662	64
$w^{(\text{EXT})}$	1.9630	2.6171	7.1355	1.6840	2.8030	5.9701	1.2196	4.2297	5.4861	1.7100	1.0189	2.4210	2.1996	6.9982	2.2709	2.6
α_{init}	0.0581	0.1441	0.1117	0.8917	0.3592	0.3142	0.4178	0.5771	0.7888	0.7026	0.9780	0.8440	0.2897	0.5338	0.6709	0.3
$w^{(ER)}$	0.6807	0.2249	0.8222	0.9975	0.5976	0.3504	0.2154	0.0928	0.1086	0.1008	0.2607	0.6085	0.5778	0.9178	0.8731	0.2
$CR_{\min}$	0.3046	0.4223	0.4828	0.9392	0.2826	0.5035	0.7287	0.5542	0.9112	0.0485	0.0114	0.4003	0.4132	0.0809	0.5600	0.05
$CR_{\max} - CR_{\min}$	0.5189	0.7601	0.2096	0.7900	0.1262	0.0099	0.9938	0.2046	0.3206	0.8630	0.0495	0.6821	0.4333	0.1376	0.6806	0.25
$NP_{\min}$	75	221	19	556	169	191	34	39	154	6	5	171	71	5	261	4
$w^{(CR)}$	0.2079	0.1192	0.0213	0.2483	0.4683	0.8719	0.7141	0.2735	0.0458	0.9161	0.4115	0.7242	0.0094	0.5870	0.7065	0.1
$w^{(F)}$	0.3530	0.1542	0.0904	0.0336	0.1027	0.6901	0.7908	0.3104	0.0131	0.8420	0.1191	0.0145	0.4230	0.0946	0.7445	0.1
Solution Error	4.26E–14	8.53E–14	2.00E+01	1.91E–01	2.52E+02	4.16E+01	1.23E+00	3.22E+00	1.00E+02	5.02E+02	3.00E+02	1.06E+02	1.03E–02	3.38E+04	1.00E+02	–

where *FES* is the current number of function evaluations, and *MAXFES* is the maximum number of function evaluations. In this way, the covariance matrix is mainly learned from the current population in the beginning stage. Then, the learning rate of the covariance matrix linearly decreases according to the ratio of the current number of function evaluations to the maximum number of function evaluations.

If the jth variable of a *trial* vector is out of the bounds $\left[x_j^{(\min)}, x_j^{(\max)}\right]$, it is corrected by:

$$u_{j,i,G} = \begin{cases} (x_j^{(\min)} + p_{j,i,G})/2, & \text{if } u_{j,i,G} < x_j^{(\min)} \\ (x_j^{(\max)} + p_{j,i,G})/2, & \text{if } u_{j,i,G} > x_j^{(\max)}. \end{cases} \tag{11}$$

E. Selection

Selection compares each pair of the *target* vector $\mathbf{x}_{i,G}$ and the *trial* vector $\mathbf{u}_{i,G}$ by:

$$\mathbf{x}_{i,G+1} = \begin{cases} \mathbf{u}_{i,G}, & \text{if } f(\mathbf{u}_{i,G}) < f(\mathbf{x}_{i,G}) \\ \mathbf{x}_{i,G}, & \text{if } f(\mathbf{u}_{i,G}) \geq f(\mathbf{x}_{i,G}) \end{cases} \tag{12}$$

where $f(\mathbf{x})$ is the objective function to be minimized.

F. Parameter Adaptation with Success-History

For each generation, the control parameters $F_{i,G}$, $ER_{i,G}$, and $CR_{i,G}$ are generated by:

$$F_{i,G} = \text{randc}_{i,G}(M_{r_i}^{(F)}, w^{(F)}) \tag{13}$$

$$ER_{i,G} = \text{randn}_{i,G}(M_{r_i}^{(ER)}, w^{(ER)}) \tag{14}$$

$$CR_{i,G} = \text{randn}_{i,G}(M_{r_i}^{(CR)}, w^{(CR)}) \tag{15}$$

where $\text{randc}_{i,G}(\mu, \sigma)$ and $\text{randn}_{i,G}(\mu, \sigma)$ are random variables according to Cauchy and normal distributions, respectively, with mean μ and standard deviation σ. r_i is a randomly selected integer within [1, H], where H is the size of the used memory. $M_{r_i}^{(F)}, M_{r_i}^{(ER)}$, and $M_{r_i}^{(CR)}$ are the r_ith historical memories of the

TABLE II. OPTIMIZED PARAMETERS OF SPS-L-SHADE-EIG AND THEIR SOLUTION ERROR VALUES ON CEC 2015 BENCHMARK FUNCTIONS FOR $D = 50$ AND $D = 100$

$D = 50$	f_1	f_2	f_3	f_4	f_5	f_6	f_7	f_8	f_9	f_{10}	f_{11}	f_{12}	f_{13}	f_{14}	f_{15}	Default
$NP_{\text{init}} - NP_{\min}$	478	558	507	861	896	767	218	399	808	579	107	953	969	389	46	946
F_{init}	0.5050	0.7836	0.0726	0.6697	0.2004	0.6280	0.1402	0.1913	0.6965	0.5904	0.2193	0.6753	0.9006	0.5413	0.4691	0.5
CR_{init}	0.2600	0.8215	0.6110	0.7933	0.8269	0.1184	0.2457	0.9433	0.8953	0.8731	0.5892	0.8941	0.3542	0.7982	0.9311	0.5
ER_{init}	0.6033	0.9146	0.9826	0.9669	0.1796	0.8996	0.3367	0.9417	0.4886	0.9376	0.3186	0.4256	0.8719	0.0061	0.2002	1
p	0.2352	0.0724	0.7977	0.9285	0.1559	0.3747	0.8305	0.2056	0.0794	0.3811	0.8741	0.2031	0.0067	0.0873	0.4671	0.11
H	490	531	493	29	44	35	279	32	17	63	970	68	51	842	70	6
Q	176	272	225	34	30	5	495	47	49	38	49	10	932	20	716	64
$w^{(\text{EXT})}$	3.1829	1.0106	1.2716	2.5268	2.3649	3.2268	2.4077	1.6237	3.4823	1.1204	1.4156	1.3381	3.1608	4.9894	3.4008	2.6
α_{init}	0.1344	0.1864	0.1502	0.0178	0.0321	0.7717	0.9528	0.1911	0.5147	0.7665	0.1086	0.0083	0.2883	0.1927	0.2993	0.3
$w^{(ER)}$	0.4667	0.0728	0.2619	0.9183	0.5304	0.0874	0.6076	0.3191	0.7577	0.2523	0.4032	0.9458	0.9727	0.7860	0.7114	0.2
$CR_{\min}$	0.2452	0.3735	0.8540	0.7566	0.0008	0.3688	0.3673	0.1509	0.6930	0.3973	0.0537	0.6937	0.4163	0.7757	0.5730	0.05
$CR_{\max} - CR_{\min}$	0.2418	0.3935	0.7355	0.7444	0.7192	0.8005	0.3580	0.2310	0.7489	0.3715	0.0005	0.9920	0.1039	0.9451	0.6836	0.25
$NP_{\min}$	175	194	34	389	11	4	19	22	41	7	53	15	93	287	385	4
$w^{(CR)}$	0.1461	0.8249	0.9066	0.7701	0.2971	0.1096	0.1549	0.0136	0.0465	0.4223	0.5200	0.0018	0.9693	0.9284	0.7100	0.1
$w^{(F)}$	0.0521	0.0945	0.0882	0.0189	0.5730	0.0463	0.1924	0.6283	0.0533	0.9311	0.2516	0.0774	0.0123	0.4432	0.3853	0.1
Solution Error	5.68E–14	1.14E–13	2.00E+01	5.49E–03	2.30E+03	6.38E+01	6.98E+00	2.01E+01	1.01E+02	3.56E+02	3.00E+02	1.13E+02	2.47E–02	5.27E+04	1.00E+02	–
$D = 100$	f_1	f_2	f_3	f_4	f_5	f_6	f_7	f_8	f_9	f_{10}	f_{11}	f_{12}	f_{13}	f_{14}	f_{15}	Default
$NP_{\text{init}} - NP_{\min}$	244	289	421	1938	1492	1676	452	1766	547	317	1893	1144	1000	685	103	1896
F_{init}	0.6177	0.4541	0.3210	0.8276	0.4057	0.5821	0.5981	0.8913	0.8709	0.9902	0.0229	0.7928	0.6431	0.7055	0.6491	0.5
CR_{init}	0.4542	0.5829	0.9080	0.0045	0.8439	0.9045	0.8028	0.9959	0.4126	0.4740	0.7016	0.2808	0.7738	0.9030	0.7858	0.5
ER_{init}	0.9499	0.5841	0.4612	0.8641	0.9083	0.8432	0.9432	0.0757	0.6982	0.1331	0.5298	0.0082	0.3026	0.6660	0.9485	1
p	0.1261	0.6050	0.6423	0.0450	0.4770	0.0791	0.9447	0.0906	0.0495	0.4982	0.6386	0.4108	0.0956	0.4543	0.5988	0.11
H	26	891	747	527	227	247	739	1863	105	711	657	786	134	1233	1649	6
Q	1475	1007	360	77	46	215	19	628	32	29	1041	72	480	531	361	64
$w^{(\text{EXT})}$	3.0795	4.6265	2.8182	3.2626	3.0129	1.5319	2.9640	1.0144	1.2182	4.4117	2.0672	4.0632	2.7068	4.1751	4.0354	2.6
α_{init}	0.0589	0.0480	0.5875	0.5229	0.8965	0.1539	0.7392	0.9018	0.1829	0.5285	0.0288	0.8383	0.5220	0.3343	0.7215	0.3
$w^{(ER)}$	0.0043	0.6165	0.9838	0.0808	0.1163	0.4033	0.6602	0.9876	0.2127	0.4434	0.9833	0.0402	0.8354	0.1862	0.6825	0.2
$CR_{\min}$	0.1441	0.0876	0.4993	0.7852	0.3581	0.4938	0.3832	0.3708	0.4845	0.9842	0.3406	0.3999	0.1396	0.1872	0.1151	0.05
$CR_{\max} - CR_{\min}$	0.4944	0.5748	0.9387	0.6577	0.4292	0.2376	0.7635	0.8450	0.1101	0.1911	0.9022	0.8771	0.2416	0.2568	0.6618	0.25
$NP_{\min}$	362	172	108	108	26	13	239	78	348	116	50	7	57	463	440	4
$w^{(CR)}$	0.2692	0.0359	0.8088	0.0678	0.2195	0.0100	0.2338	0.1738	0.4188	0.3090	0.1308	0.9457	0.0663	0.0945	0.4178	0.1
$w^{(F)}$	0.0023	0.0384	0.0868	0.0996	0.0001	0.0385	0.0051	0.3038	0.0348	0.0091	0.0327	0.9428	0.1859	0.5857	0.0979	0.1
Solution Error	9.95E–14	1.71E–13	2.00E+01	2.19E+01	4.86E+03	1.25E+03	8.98E+01	9.24E+02	1.05E+02	1.59E+03	3.01E+02	1.11E+02	6.01E–02	1.09E+05	1.00E+02	–

parameters F, ER, and CR, respectively. $w^{(F)}$, $w^{(ER)}$, and $w^{(CR)}$ are real constants between 0 and 1. If the value of $ER_{i,G}$ is outside of the interval [0, 1], it is replaced with the closest limited value (0 or 1). If $F_{i,G} > 1$, the value of $F_{i,G}$ is set to 1, and if $F_{i,G} \leq 0$, the value of $F_{i,G}$ is repeatedly generated by (13) until $F_{i,G}$ is greater than zero. The value of $CR_{i,G}$ is restricted in a predefined range [$CR_{\min}$, $CR_{\max}$]. More specifically, if $CR_{i,G}$ is out of the interval [$CR_{\min}$, $CR_{\max}$], $CR_{i,G}$ is replaced with the value of the closest bounds ($CR_{\min}$ or $CR_{\max}$).

The values of $F_{i,G}$, $ER_{i,G}$, and $CR_{i,G}$ that are used to generate a *trial* vector which successfully survives to the next generation are stored into the sets $S_G^{(F)}, S_G^{(ER)}$, and $S_G^{(CR)}$, respectively. The records of the success-history are updated by Algorithm 1.

In Algorithm 1, the index k determines the position of the success-history to be updated. k is initiated to 1 and incremented when there are new elements in the sets.

The weighted Lehmer mean $\text{mean}_{\text{WL}}(S)$ and the weighted mean $\text{mean}_{\text{WA}}(S)$ are computed by:

$$\text{mean}_{\text{WL}}(S) = \left(\sum_{i=1}^{|S|} w_i \cdot S_i^2\right) \Big/ \left(\sum_{i=1}^{|S|} w_i \cdot S_i\right) \tag{16}$$

$$\text{mean}_{\text{WA}}(S) = \left(\sum_{i=1}^{|S|} w_i \cdot S_i\right) \Big/ |S| \tag{17}$$

$$w_i = \frac{\left|f(\mathbf{u}_{i,G}) - f(\mathbf{x}_{i,G})\right|}{\sum_{k \in I(S_G^{(F)})} \left|f(\mathbf{u}_{k,G}) - f(\mathbf{x}_{k,G})\right|} \tag{18}$$

where $I(S_G^{(F)})$ denotes the set of indexes of stored values $F_{i,G}$. The indexes of stored values $F_{i,G}$, $ER_{i,G}$, and $CR_{i,G}$ are identical because they correspond to the same successful *trial* vectors. Therefore, we have $I(S_G^{(F)}) = I(S_G^{(ER)}) = I(S_G^{(CR)})$.

G. Linear Population Size Reduction

Linear population size reduction [3] has been shown a simple yet effective way to improve DE performance:

$$NP_G = \text{round}\left(NP_{\text{init}} - \frac{FES}{MAXFES} \cdot \left(NP_{\text{init}} - NP_{\min}\right)\right) \tag{19}$$

TABLE III. ERROR VALUES OF SPS-L-SHADE-EIG ON CEC 2015 BENCHMARK FUNCTIONS FOR $D = 10$ AND $D = 30$

	$D = 10$					$D = 30$				
Function	Best	Worst	Median	Mean	Std	Best	Worst	Median	Mean	Std
f_1	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00
f_2	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00
f_3	0.00E+00	2.00E+01	2.00E+01	1.80E+01	6.01E+00	2.00E+01	2.00E+01	2.00E+01	2.00E+01	7.29E–05
f_4	0.00E+00	3.98E+00	9.95E–01	1.17E+00	9.69E–01	1.05E–02	5.23E+01	3.25E+00	1.03E+01	1.41E+01
f_5	3.12E–01	1.45E+02	1.52E+01	2.69E+01	3.43E+01	6.58E+02	3.20E+03	1.49E+03	1.54E+03	5.23E+02
f_6	0.00E+00	1.20E+00	0.00E+00	1.23E–01	2.86E–01	2.68E+01	3.16E+02	7.34E+01	9.91E+01	7.14E+01
f_7	0.00E+00	1.12E–01	1.94E–02	2.78E–02	2.57E–02	6.23E–01	3.95E+00	2.40E+00	2.44E+00	7.44E–01
f_8	6.02E–08	9.72E–03	1.37E–04	8.81E–04	1.92E–03	2.07E+00	1.21E+02	1.40E+01	2.11E+01	2.14E+01
f_9	1.00E+02	1.00E+02	1.00E+02	1.00E+02	3.36E–02	1.02E+02	1.02E+02	1.02E+02	1.02E+02	1.12E–01
f_{10}	2.17E+02	2.17E+02	2.17E+02	2.17E+02	3.78E–03	1.48E+02	6.23E+02	3.68E+02	3.53E+02	1.12E+02
f_{11}	2.61E–02	3.00E+02	7.85E–02	4.72E+01	1.10E+02	3.00E+02	3.01E+02	3.01E+02	3.01E+02	2.16E–01
f_{12}	1.00E+02	1.01E+02	1.00E+02	1.01E+02	1.37E–01	1.02E+02	1.03E+02	1.03E+02	1.03E+02	2.51E–01
f_{13}	3.03E–02	3.05E–02	3.04E–02	3.04E–02	4.67E–05	2.56E–02	2.62E–02	2.59E–02	2.59E–02	2.28E–04
f_{14}	1.00E+02	2.93E+03	1.00E+02	6.56E+02	1.14E+03	3.11E+04	3.32E+04	3.11E+04	3.16E+04	8.32E+02
f_{15}	1.00E+02	1.00E+02	1.00E+02	1.00E+02	0.00E+00	1.00E+02	1.00E+02	1.00E+02	1.00E+02	2.99E–13

TABLE IV. ERROR VALUES OF SPS-L-SHADE-EIG ON CEC 2015 BENCHMARK FUNCTIONS FOR $D = 50$ AND $D = 100$

	$D = 50$					$D = 100$				
Function	Best	Worst	Median	Mean	Std	Best	Worst	Median	Mean	Std
f_1	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00
f_2	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00
f_3	2.00E+01	2.00E+01	2.00E+01	2.00E+01	3.25E–05	2.00E+01	2.02E+01	2.00E+01	2.00E+01	2.25E–02
f_4	9.01E–05	1.29E+01	1.99E+00	2.30E+00	2.25E+00	1.59E+01	7.86E+01	3.68E+01	3.80E+01	1.08E+01
f_5	1.38E+03	3.68E+03	2.69E+03	2.68E+03	5.48E+02	3.98E+03	2.76E+04	1.50E+04	1.57E+04	5.94E+03
f_6	5.18E+01	6.34E+02	2.11E+02	2.39E+02	1.35E+02	9.61E+02	3.04E+03	2.14E+03	2.09E+03	4.90E+02
f_7	6.49E+00	5.25E+01	4.05E+01	2.74E+01	1.77E+01	9.06E+01	9.89E+01	9.33E+01	9.41E+01	2.66E+00
f_8	1.00E+01	2.60E+02	2.62E+01	6.58E+01	7.74E+01	6.87E+02	2.31E+03	1.34E+03	1.34E+03	3.57E+02
f_9	1.03E+02	1.04E+02	1.04E+02	1.04E+02	1.21E–01	1.05E+02	1.06E+02	1.05E+02	1.05E+02	1.17E–01
f_{10}	6.79E+02	1.30E+03	8.09E+02	8.35E+02	1.26E+02	1.59E+03	2.64E+03	1.93E+03	2.00E+03	2.85E+02
f_{11}	3.00E+02	3.00E+02	3.00E+02	3.00E+02	5.20E–02	3.01E+02	1.91E+03	3.01E+02	3.64E+02	3.14E+02
f_{12}	1.03E+02	2.00E+02	1.04E+02	1.10E+02	2.29E+01	1.11E+02	1.13E+02	1.12E+02	1.12E+02	2.95E–01
f_{13}	6.99E–02	8.17E–02	7.57E–02	7.62E–02	3.78E–03	5.96E–02	6.36E–02	6.10E–02	6.14E–02	9.30E–04
f_{14}	4.95E+04	7.43E+04	6.83E+04	6.35E+04	7.40E+03	1.09E+05	1.09E+05	1.09E+05	1.09E+05	1.77E+01
f_{15}	1.00E+02	1.00E+02	1.00E+02	1.00E+02	1.02E–13	1.00E+02	1.00E+02	1.00E+02	1.00E+02	1.43E–13

where NP_G is the population size at the Gth generation, NP_{init} and $NP_{\min}$ are initial and minimum population sizes, respectively. The initial population size must be greater than the minimum population size, specifically, $NP_{\text{init}} \geq NP_{\min}$. The current-to-*p*best/1 mutation strategy requires at least four distinct solutions, so that the minimum population size must be greater than four, specifically, $NP_{\min} \geq 4$.

H. Uncertainty Handling Method

One of the simplest way to optimize the performance of DE is to use DE algorithm to minimize the solution error measure of the DE parameters. However, in real-world optimization problems, the performance of DE might be sensitive to the randomness of mutation, crossover, and parameter/strategy adaptation. The uncertainty is a result of intrinsic noise in the fitness of DE. In most design and implementation, DE algorithm does not re-evaluate a solution because the same solutions have identical function values when the fitness values have no uncertainty. However, if the fitness values are computed with uncertainty, the same solutions may have different function values. In this case, the desired high-quality solution should be robust to the uncertainty. To this end, the candidate solutions are evaluated for multiple times instead of single function evaluation.

The evaluation of DE performance is expensive because it requires a complete run on a specific problem. To efficiently deal with the uncertainty without dramatically increasing the number of function evaluations, the proposed uncertainty handling method re-evaluates a solution only if the solution is not updated for an unacceptable amount of times Q. In this way, a "lucky" solution should be removed from the population when the selection is executed with the re-evaluation.

The overall algorithm of the proposed SPS-L-SHADE-EIG is shown in Algorithm 2. Please note that SPS-L-SHADE-EIG optimizes noise-free optimization problems (in CEC 2015 benchmark set) without the use of the uncertainty handling method to save the number of executed function evaluations.

TABLE V. COMPUTATIONAL COMPLEXITY

	T_0	T_1	$\hat{T}_2$	$(\hat{T}_2 - T_1)/T_0$
$D = 10$	0.1219	0.6453	7.7069	57.9377
$D = 30$	0.1216	1.0444	25.7812	203.5029
$D = 50$	0.1215	1.6736	43.7714	346.4262
$D = 100$	0.1216	4.1918	97.6919	768.9084

The uncertainty handling method is only used for self-optimizing the parameters of the proposed SPS-L-SHADE-EIG.

III. EVALUATING SPS-L-SHADE-EIG ON CEC 2015 BENCHMARK SET

There are 15 parameters in the proposed SPS-L-SHADE-EIG. To find the best set of the parameters for each test function in CEC 2015, the parameters of SPS-L-SHADE-EIG are optimized in terms of the *solution error measure*, which is defined as $f(\mathbf{x}') - f(\mathbf{x}^*)$, where $\mathbf{x}^*$ is the global optimum of the considered test function, and $\mathbf{x}'$ is the best solution achieved by SPS-L-SHADE-EIG with a given set of parameters after $10^4 \times D$ function evaluations. In other words, assuming the set of the parameters of the proposed SPS-L-SHADE-EIG is denoted as $\mathbf{y}$, we define a new objective function $g(\mathbf{y}) = f(\mathbf{x}') - f(\mathbf{x}^*)$ to be minimized by SPS-L-SHADE-EIG with the proposed uncertainty handling method.

The computation of optimizing the parameters of SPS-L-SHADE-EIG is expensive especially when the dimension of the test function increases. To optimize the parameters in limited computational resource, the maximum numbers of the function calls on $g(\mathbf{y})$ are set to 6,000, 2,000, 1,200, and 600 for D = 10, 30, 50, and 100, respectively. In addition, we increase the convergence speed by setting the stagnation tolerance Q to 1, and the initial population size is set to 30. Other parameters are set to the default values as shown in the last column of Table I and Table II. The default values are given by the parameters settings in L-SHADE [3], EIG [4], SPS [5], and our empirical studies on the overall performance of SPS-L-SHADE-EIG in CEC 2015 benchmarks [6]. To ensure satisfactory results, the default parameters are inserted into the initial population, and therefore the optimized parameters should exhibit better performance than the default parameters.

The parameters with minimal solution error in the final population of SPS-L-SHADE-EIG are shown in Table I and Table II on CEC 2015 benchmark functions for D = 10, 30, 50, and 100. For each function, the minimal solution error in the final population is recorded in the row of *Solution Error*. The default parameters are shown in the column of "Default".

The results of SPS-L-SHADE-EIG with the optimized parameters are shown in Tables III and IV. With the rules of CEC 2015 benchmark competition, the search space is set to [–100, 100] for each variable. If the *solution error* is smaller than 10^{-8}, the error is set to 0. The maximum number of function evaluations (*MAXFES*) on each benchmark function is $D \times$ 10,000. The number of runs for each function is 51, and the best, worst, median, mean, and standard deviation of the *solution error* are recorded.

The algorithm complexity of SPS-L-SHADE-EIG is shown in Table V. The simulation is executed on a personal computer with an Intel CPU (3.40GHz) and 8GB RAM, under Matlab 2012b programming environment. T_0 denotes the execution time of the following program:

for i=1:1,000,000

x=0.55+(double)i; x=x+x; x=x/2; x=x*x; x=sqrt(x);

x=log(x); x=exp(x); x=x/(x+2);

end

T_1 denotes the execution time of f_1 for 200,000 evaluations. Let T_2 be the running time of the proposed SPS-L-SHADE-EIG on f_1 for 200,000 evaluations. T_2 is evaluated for five times, and the mean of the five evaluated T_2 is denoted as $\hat{T}_2$. Finally, the algorithm complexity is estimated by $(\hat{T}_2 - T_1)/T_0$.

IV. CONCLUSION

A new DE variant SPS-L-SHADE-EIG is proposed in this paper. The proposed SPS-L-SHADE-EIG is a variant of L-SHADE which is incorporated with the successful-parent-selecting (SPS) framework and a eigenvector-based (EIG) crossover operator. In addition, we generalize the L-SHADE by introducing parameters on some constants in L-SHADE, such as the standard deviation of the Cauchy and Gaussian distribution, the initial values of control parameters (F and CR), and the minimum population size.

The performance of the proposed SPS-L-SHADE-EIG is optimized by itself with the proposed uncertainty handling method for each function in CEC 2015 benchmark set. Optimizing the performance of the proposed SPS-L-SHADE-EIG is a optimization problem with uncertainty of fitness values. To deal with the uncertainty, the proposed uncertainty-handling method re-evaluates a solution if it is not updated for more than an unacceptable amount of times. In this way, the final population should only contain high-quality solutions which are robust to uncertainties. The robust parameters of the proposed SPS-L-SHADE-EIG are obtained from the final population. The optimized parameters, solution error measure, and algorithm complexity are reported in the experimental results.

This study can be extended in a number of directions. Since optimizing the performance of DE is computational expensive, future work may extend the proposed method to expensive noisy optimization, and the effect of the parameters will be studied further.

Acknowledgment

This work was supported by the Ministry of Science and Technology of Taiwan, R.O.C., under Grants MOST 103-2221-E-006-023 and NSC 102-2221-E-006-208-MY3.

References

[1] R. Tanabe and A. Fukunaga, "Success-history based parameter adaptation for differential evolution," in *Proc. IEEE Congr. Evol. Comput.*, Cancún, México, Jun. 2013, pp. 71–78.

[2] R. Tanabe and A. Fukunaga, "Evaluating the performance of SHADE on CEC 2013 benchmark problems," in *Proc. IEEE Congr. Evol. Comput.*, Cancún, México, Jun. 2013, pp. 1952–1959.

[3] R. Tanabe and A. Fukunaga, "Improving the search performance of SHADE using linear population size reduction," in *Proc. IEEE Congr. Evol. Comput.*, Beijing, China, Jul. 2014, pp. 1658–1665.

[4] S.-M. Guo and C.-C. Yang, "Enhancing differential evolution utilizing eigenvector-based crossover operator," *IEEE Trans. Evol. Comput.*, vol. 19, no. 1, pp. 31–49, Feb. 2015.

[5] S.-M. Guo, C.-C. Yang, P.-H. Hsu, and J. S.-H. Tsai, "Improving differential evolution with successful-parent-selecting framework," *IEEE Trans. Evol. Comput.*, in press, DOI: 10.1109/TEVC.2014.2375933.

[6] J. J. Liang, B. Y. Qu, P. N. Suganthan, and Q. Chen, "Problem denifition and evaluation criteria for the CEC 2015 competition on learning-based real-parameter single objective optimization," Zhengzhou Univ., China, and Nanyang Technol. Univ., Singapore, Nov. 2014.