Constrained Many-objective Optimization: A way forward

Dhish Kumar Saxena, Tapabrata Ray, Kalyanmoy Deb and Ashutosh Tiwari

Abstract— Many objective optimization is a natural extension to multi-objective optimization where the number of objectives are significantly more than five. The performance of current state of the art algorithms (e.g. NSGA-II, SPEA2) is known to deteriorate significantly with increasing number of objectives due to the lack of adequate convergence pressure. It is of no surprise that the performance of NSGA-II on some constrained many-objective optimization problems [7] (e.g., DTLZ5-(5, M/)), M = 10, 20) in an earlier study [19] was far from satisfactory. Till date, research in many-objective optimization has focussed on two major areas (a) dimensionality reduction in the objective space and (b) preference ordering based approaches. This paper introduces a novel evolutionary algorithm powered by epsilon dominance (implemented within the framework of NSGA-II) and controlled infeasibility for improved convergence while the critical set of objectives is identified through a non-linear dimensionality reduction scheme. Since approaching the Pareto-optimal front from within the feasible search space will need to overcome the problems associated with low selection pressure, the mechanism to approach the front from within the infeasible search space is promising as illustrated in this paper. The performance of the proposed algorithm is compared with NSGA-II (original, with crowding distance measure) and NSGA-II (epsilon dominance) on the above set of constrained multiobjective problems to highlight the benefits.

I. INTRODUCTION

Evolutionary Algorithms have so far achieved great success in solving single and multi-objective optimization problems. The state-of-the-art multiobjective optimization algorithms rely on the concept of non-dominance of solutions to drive the population towards the Pareto-optimal front. The deterioration of the performance of evolutionary algorithms (EAs) with increasing number of objectives has already been identified and reported in literature [7], [1], [13], [23]. It is recognized that with an increase in the number of objectives, there is a substantial reduction in the number of fronts, with each front becoming denser. Given this, most solutions are assigned the same rank of non-dominance and Pareto selection weakens since it has to discriminate mostly based on the diversity of solutions. It is also being recognized that for selection to be effective in problems with more than three objectives, a more careful analysis of Pareto-dominance relation is required, in that, the low selection pressure needs to be enhanced. Various researchers have identified that the low selection pressure could be countered by way of inducing a preference ordering over the points in the non-dominated set. As a result, the remedial approaches proposed by them can all be collectively referred as preference ordering based approaches. These approaches primarily include:

- Relaxed form of Pareto-dominance [15], [10]
- Controlling the Dominance Area [17]
- Modification of rank definition [2], [9], [5]
- Utilization of Scalarizing fitness functions [12]
- Indicator based EMO algorithm [28]
- Substitute distance metrics [14] and [23]

These approaches seem promising, and their further investigation will gradually shed light on any limitations, while providing guidelines for the kinds of problems to which, they may be best suited. Preliminary investigations have indicated that loss of diversity of solutions is a common problem associated with scalarization, average ranking and expansion schemes while indicator based approach is known to be computationally expensive.

Another perspective to handling of many-objective problems is based on the notion that given an M-objective optimization problem, not all objectives may be necessary to define the Pareto-optimal front. The issue of dispensability of objectives could be viewed from two perspectives:

I. Non-conflicting objectives (Degenerate problem): that, given an M-objective problem, some of the objectives are non-conflicting to the others, resulting which the P.O.F. is of dimension m, where m < M. In such a case, these m-objectives contributing to the P.O.F. could be labelled as critical, while the rest as redundant.

II. Statistical significance: that, given an M-objective problem, some of the objectives (including the conflicting ones) could be removed without significantly losing the problem information. In such a case, the problem could be treated as a k-objective problem, where k represents only the statistically significant conflicting objectives. Here, these k-objectives could be labelled as critical, while the rest as redundant.

In [19], some of the plausible reasons as to why degenerate problems may exist in practice, are discussed. Guided by this notion, dimensionality reduction procedures [7] (linear) and [20](nonlinear) were introduced in the domain of evolutionary multi-objective optimization (EMO). Another significant contribution in the area of dimensionality reduction has been made in [4], [3]. The latter is a relation based objective reduction method which aims at preserving (most of) the dominance structure and is a pure feature selection method. On the contrary, the former approach
employs principal component analysis (PCA) based reduction. Though PCA is a feature extraction technique, it has been customized for the task of feature selection by exploiting the correlation based notion of conflict between objectives.

To set the context for the ideas being proposed in this paper, the earlier work on dimensionality reduction using principal component based analysis [7], [20] including the formulation of constrained many-objective optimization problems, is summarized in the sections below.

II. CONSTRAINED MANY-OBJECTIVE OPTIMIZATION PROBLEMS

Given the dearth of scalable many-objective test problems on the one hand and the need for a comprehensive test-suite for evaluation of the dimensionality reduction procedures, the existing DTLZ5 problems [8] were generalized in [7] and referred as DTLZ5-(I, M). In the latter, I denotes the dimension of the P.O.F, while M represents the number of objectives in the problem.

\[
\begin{align*}
\min f_1(x) &= r(x_M) \cos(\theta_1) \cdots \cos(\theta_{M-2}) \cos(\theta_{M-1}), \\
\min f_2(x) &= r(x_M) \cos(\theta_1) \cdots \cos(\theta_{M-2}) \sin(\theta_{M-1}), \\
\vdots & \quad \vdots \\
\min f_{M-1}(x) &= r(x_M) \cos(\theta_1) \cdots \sin(\theta_{M-2}), \\
\min f_M(x) &= r(x_M) \sin(\theta_1), \\
\text{where} & \\
r(x_M) &= 1 + g(x_M) = 1 + \sum_{i=1}^{M} x_i (x_i - 0.5)^2, \\
0 & \leq x_i \leq 1, \quad \text{for } i = 1, 2, \ldots, n.
\end{align*}
\]

While the formulation of the objectives remains the same as before (Equation 1), DTLZ5-(I, M) problems differ from the original, at two fronts:

1) in the definition of the parameter \(\theta_i\), given by Equation 2.

\[
\theta_i = \begin{cases} 
\frac{2}{\pi} x_i, & \text{for } i = 1, \ldots, (I-1), \\
\frac{1}{\sqrt{n(x_M)}} (1 + 2g(x_M)x_i), & \text{for } i = I, \ldots, (M - 1) 
\end{cases}
\]

2) in the incorporation of constraints (necessary to retain the dimensionality reduction property of the problem— as discussed in [7]):

\[
\sum_{j=0}^{I-2} f_{M-j} + 2^p f_I^2 \geq 1, \quad \text{for } i = 1, 2 \ldots (M - I + 1)
\]

\[
p_i = \begin{cases} 
M - I, & \text{for } i = 1; \\
(M - I + 2) - i, & \text{for } i = 2 \ldots (M - I + 1)
\end{cases}
\]

The total number of variables involved are \(n = M + k - 1\), where, \(k = |x_M| = 10\) is prescribed. The problem is so designed that the P.O.F. will correspond to (i) a zero value of \(g\) function, in turn implying \(x_i = 0.5\) for \(i = M, \ldots, (M + k - 1)\) (ii) fixed value of \(\pi/4\) for the variables \(x_i\) to \(x_{M-1}\) (iii) independent values for the variables \(x_i\) to \(x_{I-1}\). Hence, by simply setting \(I\) to an integer between two and \(M\), the dimensionality (I) of the Pareto-optimal front can be controlled. It may also be noted that DTLZ5-(M, M) problems are equivalent to and hence would be interchangeably referred to, as DTLZ2(M) [8].

III. PCA BASED DIMENSIONALITY REDUCTION: PREVIOUS CONTRIBUTION

In [7], a principal component analysis based dimensionality reduction procedure within the NSGA-II framework was introduced and referred as PCA-NSGA-II. While it laid the rudiments of dimensionality reduction in an EMO framework, the corresponding results suffered from some anomalies. It was identified [20] that these anomalies emanate from the fundamental limitations of the underlying methodology of PCA itself. Consequently, two nonlinear dimensionality reduction procedures, namely C-PCA-NSGA-II [20] and MVU-PCA-NSGA-II [20] were proposed. While the former of these two was based on Correntropy PCA [27], the latter implemented the Maximum Variance Unfolding [24], [25], [26] principle in a novel way. It should be noted that each of the above procedures were based on the eigen-decomposition and systematic analysis of the resulting eigenvalues/vectors of an underlying matrix— namely the correlation matrix \(R\) in case of PCA-NSGA-II, the correntropy matrix \(V\) in case of C-PCA-NSGA-II and the learnt kernel matrix \(K\) in case of MVU-PCA-NSGA-II. While owing to the context in this paper, the details of the formulations of these matrices are precluded, the dimensionality reduction procedure is explained, below:

Algorithm I: Earlier proposed Dimensionality Reduction Procedure (improved version of [7], [20])

Input:
\(N\): Population Size
\(N_G > 1\): Number of Generations
1) Set: \(t = 0\); initial set of objectives \(U_0 = \{1, 2, \ldots, M\}\).
\textbf{begin}
3) Initialize a random population in \(U_t\), run an EMO, and obtain a population \(P_t\).
4) Perform dimensionality reduction analysis on \(P_t\) to arrive at a reduced set of objectives \(U_{t+1}\), using the predefined TC. Steps of the analysis are as follows:
\(U_{t+1}\) is arrived at from \(P_t\:
\textbf{begin}
5) Compute \(R_{M-M} \) [7], \(V_{M-M} \) [20], \(K_{M-M} \) [20], based on the choice of method
6) Reduction-I: Starting with \(R_{M-M}\), \(V_{M-M}\) or \(K_{M-M}\) - eliminate objectives based on the proposed interpretation scheme for eigen values and eigen vectors (Section III-A.2)
7) Reduction-II: Eliminate further (if possible) the objectives retained after the Reduction-I above, based on the proposed selection scheme (Section III-A.3) for the reduced correlation matrix \(\text{R.C.M.}\)
\textbf{end}
8) if \(U_{t+1} = U_t\) then
\textbf{end}
9) Stop and declare the critical objective set (\(U_c\))
\textbf{else}
10) set \(t = \) \(t + 1\) and go to Step 2.
11) end
12) end
13) \textbf{end}
14) end
A. Components of the earlier proposed Dimensionality Reduction Algorithms

1) Eigenvalue Analysis: This involves the eigen-decomposition of the correlation matrix (R) for linear dimensionality reduction [7]; correntropy matrix (V) or learned kernel matrix (K) for nonlinear dimensionality reduction [20].

2) Interpreting Multiple Principal Components: To make the dimensionality-reduction procedure effective and applicable to various scenarios, we suggest the following detailed procedure, which starts with analyzing the first principal component and then proceed to analyzing the second principal component and so on, till all the significant components are considered. For this purpose, we pre-define TC – a threshold cut (details in [7],[20]) and when the cumulative contribution of the top principal components would exceed TC, we would not analyze any more principal components.

For the first principal component, along with the objective corresponding to the most-positive element, we consider as important, any/all objectives which correspond to a negative component, howsoever small. If in some case, all the elements along PCA-1 are positive, we pick up the objectives corresponding to the first two most positive elements. For subsequent principal components, we first check if the corresponding eigenvalue is greater than 0.1 or not. If not, we only choose the objective corresponding to the highest absolute element in the eigenvector. If yes and also if the cumulative contribution of eigenvalues is less than TC, we consider various cases. If all elements of the eigenvector are positive, we only choose the objective corresponding to the highest element. If all elements of the eigenvector are negative, we choose all objectives. Otherwise, if the value of the highest principal component (p) is less than the absolute value of the most-negative element (n), we check if p ≥ 0.9|n|. If yes, then we choose two objectives corresponding to p and n, otherwise we choose only the objective corresponding to n. Similarly, we also consider the possibility of the absolute value of the most-negative element (n) being less than the highest positive element (p), in which case we further check if |n| ≥ 0.9p. If yes, we choose both objectives corresponding to p and n, otherwise, we only choose the objective corresponding to p.

3) Final Reduction Using the Correlation Matrix: Hopefully, the above procedure identifies most of the redundant objectives dictated by the data set. To consider whether further reduction in the number of objectives is possible, we then return to a reduced correlation matrix (only columns and rows corresponding to non-redundant objectives—adjudged so, by Eigenvalue analysis) and investigate if there still exists a set of objectives having identical positive or negative correlation coefficients with other objectives and having a positive correlation among themselves. This will suggest that any one member from such group would be enough to establish the conflicting relationships with the remaining objectives. For understanding, consider an identically correlated set of S objectives, each being represented by $f_i$, $i = 1 \ldots S$. Further, assume that V principal components were utilized for Eigenvalue analysis each being represented by $v_j$, $j = 1 \ldots V$ and each accounting for a proportion $c_j$, $j = 1 \ldots V$, towards the total variance of the data set. Let the contribution of $f_i$ along $v_j$ be represented by $f_{ij}$. Then compute for each of the identically correlated objectives $f_i$ the value $c_i = \sum_{j=1}^{V} ||(f_{ij} \cdot c_j)||$. Pick from the set \{c_i | i = 1 \ldots S\} of scalars the highest value, say c. Then the objective $f_k$ can be considered as a representative of the set of S objectives. This selection criterion physically implies that an objective which contributes most along the important principal components collectively, deems fit to be the representative. It should also be mentioned, that once NSGA-II is run for sufficiently large number of generations, the correlation matrix stabilizes and correlation patterns turn invariant over number of generations. Hence, the inferences drawn from reduced correlation matrix, can be trusted.

B. Summary of simulation results from the earlier proposed Dimensionality Reduction Algorithms

Table I summarizes the results from the experiments performed with PCA-NSGA-II, C-PCA-NSGA-II (for various choices of a priori chosen kernels with different parameter values) and MVU-PCA-NSGA-II (for varying values of the neighbourhood parameter k). It can be seen that while PCA-NSGA-II failed to accurately solve most of the difficult problems, C-PCA-NSGA-II was relatively more efficient corresponding to specific kernels and specific parameter values. However, given that for an unknown problem it is not known a priori as to which kernel to chose, Maximum Variance Unfolding principle based MVU-PCA-NSGA-II in which kernels specific to the given data set are learned, gains significance.

In the latter, the choice of the neighbourhood parameter k is crucial. It has been well argued in [19], [20], as to why $k = M - 1$ is the safest choice. This claim is based on the fact that for a given data point, k signifies the number of neighbours with which its distances and included angles have to be retained as invariant during the 'unfolding' process. Hence, it is easy to see (Figure 1) that too high a value of k would over-constrain (say, $k = M - 1$) and hence delay the 'unfolding' process, requiring more iterations towards fulfilling the objective. On the other hand, while too low a value of k may help achieve the objective of maximizing the sum total of the pairwise distance fast, it is also risky as it may lead to distortion of local topology. As can be sensed from the case of $k = 1$, even the close neighbours of a given data point may be stretched, eventually leading to 'erroneous unfolding'.

Given the above, it is no surprise that in Table I, the accuracy of the results can be seen to improve with an increase in the value of k. However, what is alarming is the fact that even with $k = M - 1$, the DTLZ5-(5, 10) and the DTLZ5-(5, 20) problems could not be accurately solved. The case of $k = M - 1$ ensures that no distortion to local
isometry occurs during the data ‘unfolding’, hence, there is every reason to expect accurate results. The only plausible reason for the inaccuracies even in this case could be attributed to the following observation. All DTLZ5-\( (I, M) \) problems are by design, such that their P.O.F. corresponds to the parameter \( g = 0.0 \) (Equation 1). It can be seen in the Table II which reports the percentage of solutions with \( g \leq 0.01 \), that only about 0.5% solutions qualify for this condition for the DTLZ5-\( (5, 10) \) and the DTLZ5-\( (5, 20) \) problems. It means that, in the data set to be ‘unfolded’ towards making deductions on the composition of the P.O.F., only about 3 solutions out of 600 (population size) are in close proximity to the P.O.F. Clearly, the data set being investigated contains too meager an information about the P.O.F., at the first step.

In contrast, in case of the other difficult problems like DTLZ5-\( (2, 50) \) and DTLZ5-\( (3, 10) \) which could be accurately solved, about 1.0% solutions achieved a \( g \leq 0.01 \). Though, 1.0% is only marginally better than 0.5% but these problems could perhaps be solved for two reasons, in that (i) the average value of \( g \) parameter (not shown in the above table, though) is noted to be better than the average values in DTLZ5-\( (5, M) \) problem for \( M = 10, 20 \) (ii) a relatively lower-dimensional interactions had to be deciphered, as against the 5-dimensional interactions in the former.

In the wake of the above discussions, it can be concluded that ‘subject to the condition that the data set being investigated contains a minimal degree of convergence information (a subjective issue, though)’, MVU-PCA-NSGA-II with \( k = M - 1 \), is a reliable nonlinear dimensionality reduction procedure that does not require the difficult task of having to chose a particular Kernel \textit{a priori} (as required by C-PCA-NSGA-II). More importantly, the above observations also point to the fact that ‘… to enable any PCA based approach to make accurate inferences on the dimension or composition of the P.O.F, the underlying EMO algorithm (say, NSGA-II) should be able to deliver a reasonably converged set of nondominated solutions with an adequate diversity’.

### IV. Countering Low Selection Pressure: A Way Forward

Backed by evidence in the previous section, this section aims to build upon the premise that for many-objective (\( M \geq 5 \)) problems, NSGA-II in its original form is unable to provide satisfactory level of convergence.

#### A. Epsilon dominance (Eps-dom)

The available literature relating to the causes of poor convergence and the remedial approaches being researched upon are stated in Section I. Of them, two approaches [14] and [23], are pursued in this paper. In these, the relative strengths and weaknesses of various
ranked schemes (crowding, epsilon dominance, subvector fuzzy Pareto dominance, subobjective dominance count etc.) are identified. These studies have revealed that epsilon dominance offers a good balance between convergence and diversity and has a potential to be applied for many-objective optimization problems.

Epsilon dominance ranks a given solution based on the smallest amount that should be subtracted from all objectives of each of the other solutions, such that they dominate the given solution. For a solution \( i \), the value \( \text{Eps-Dom}(i, j) \) denotes the smallest amount to be subtracted from all the objectives of solution \( j \), so that it dominates solution \( i \). Smallest such value among all solutions \( (j \neq i) \) is the distance \( \text{dist}(i) \) assigned to the solution \( i \). A larger Eps-dom measure implies better solution, as it means that a large amount has to be subtracted from the objectives of other solutions to make them dominate the given point.

In this paper, the effect of replacement of crowding distance measure in NSGA-II, by epsilon dominance, on the accuracy of the earlier proposed dimensionality reduction procedures will be demonstrated in a later section. In that, the MVU-PCA dimensionality reduction scheme based on NSGA-II(Eps-dom) will be referred to, as MVU-PCA-NSGA-II(Eps-dom).

B. Infeasibility Driven Evolutionary Algorithm (IDEA)

Another key feature of this paper lies in its pursuing of the notion of Infeasibility Driven Evolutionary Algorithm. Through [21], [22], the authors brought forth the notion that ‘in order to truly maximize/minimize an objective function, the feasible set needs to be optimized’, which in turn calls for exploring the gain in objectives by way of relaxing constraints. True it is, that any compromise made on any constraint, in conventional optimization sense makes the corresponding solution infeasible but to highlight the lack of wisdom in blindly rejecting such solutions (while maintaining the practicability of solutions), is where the motivation to these papers emanated from. This notion of selectively embracing infeasibility has been taken forward in [11] and [16]. In that, the authors have highlighted the benefits of preserving infeasible solutions as they offer a promise to approach the optimal solution from both the feasible and infeasible side of the search space. When the EMO community has started to realize the severity of the curse of the lack of convergence pressure in case of many-objective problems, the notion of Infeasibility Driven Evolutionary Algorithm is anything but short of promise. In that, while the conventional mode of approaching the P.O.F. from within the feasible region of the search space is posing serious challenges to the underlying ranking schemes, the notion of artificially inducing solutions close to the constraint boundary by way of retaining a controlled fraction of infeasible solutions could be a potent remedy to the ills of low convergence pressure. The exploration of this notion and the demonstration of its promise on different versions of a difficult problem, namely DTLZ5-(5, \( M \)) is where a part of the remaining paper is focussed on.

For symbolic visualization of the concept of Infeasibility Driven Evolutionary Algorithms, refer to Figure 2. This figure should be treated as a 2-D representation of a many-objective problem (objective space) with the intent to (a) highlight the curse associated with many-objectives (b) pictorially convey the notion of Infeasibility Driven Evolutionary Algorithms. In that, Figure 2(a) represents the curve associated with many-objective problems in a rather simplified way for visual aid. Solution set A represents the (preconverged) solutions which ideally ought to be converging to the P.O.F. but have failed to do so because of low convergence pressure, given the presence of many objectives. Solution set B represents the set of infeasible solutions (\( \alpha \) proportion of the total population size), which on breeding with the feasible members of the solution set (Figure 2(b)) will lead to creation of new feasible solutions (Figure 2(c)) some of which will be closer to the true P.O.F. It is hoped that this mechanism of pulling the population to the constraint boundary by the use of infeasible solutions is an effective way of countering the low selection pressure on solutions for convergence to the P.O.F, in case of many-objective problems.

A constrained many-objective optimization problem is represented by Equation 4.

\[
\begin{align*}
\text{Minimize} & \quad f_1(x), \ldots, f_M(x) \\
\text{Subject to} & \quad g_j(x) \geq 0, \quad j = 1, \ldots, J 
\end{align*}
\]  

(4)

where \( x = (x_1, \ldots, x_n) \) is the design variable vector bounded by lower and upper bounds. To explore the search space (including the feasible and the infeasible region) under the notion of IDEA, the original \( M \) objective constrained optimization problem is reformulated as an \( M + 1 \) objective unconstrained optimization problem, given by Equation 5.

\[
\begin{align*}
\text{Minimize} & \quad f'_1(x) = f_1(x), \ldots, f'_M(x) = f_M(x) \\
& \quad f'_{M+1}(x) = \text{Violation measure} 
\end{align*}
\]  

(5)

The additional objective represents a measure of constraint violation, which for a solution is the sum of \( J \) ranks corresponding to each constraint. Solutions that do not violate a constraint are assigned a rank of 0.

The steps of the proposed algorithm are outlined in Algorithm 2. In that, simulated binary crossover (SBX) and polynomial mutation operators are used to generate an offspring from a pair of parents, selected using binary tournament as in NSGA-II. Individual solutions in the population are evaluated using the original problem definition (Equation 4) and infeasible solutions are identified. The solutions in the parent and offspring population are divided into a feasible set (\( S_f \)) and an infeasible set (\( S_{in_f} \)). The solutions in the feasible set are ranked using epsilon dominance based on \( M + 1 \) objectives as per the modified problem definition (Equation 5).
The solutions for the next generation are selected from both the sets to maintain infeasible solutions in the population. The process ensures a search through both the feasible and infeasible region of the search space.

Algorithm 2: Infeasibility Driven Evolutionary Algorithm (IDEA)

```
Input:
N: Population Size
N_G > 1: Number of Generations
0 < α < 1: Proportion of infeasible solutions
N_{inf} = α * N
N_f = N - N_{inf}
pop_1 = Initialize()

begin
    for i = 2 to N_G do
        childpop_{i-1} = Evolve(pop_{i-1})
        Evaluate(childpop_{i-1})
        (S_f, S_{inf}) = Split(pop_{i-1} + childpop_{i-1})
        Rank(S_f)
        Rank(S_{inf})
        pop_i = S_{inf}(1, N_{inf}) + S_f(1, N_f)
    end
end
```

A user-defined parameter α is used to maintain a set of infeasible solutions as a fraction of the size of the population. The numbers N_f and N_{inf} denote the number of feasible and infeasible solutions as determined by parameter α. If the infeasible set S_{inf} has more than N_{inf} solutions, then first N_{inf} solutions are selected based on their rank, else all the solutions from S_{inf} are selected. The rest of the solutions are selected from the feasible set S_f, provided there are at least N_f number of feasible solutions. If S_f has fewer solutions, all the feasible solutions are selected and the rest are filled with infeasible solutions from S_{inf}. The solutions are ranked from 1 to N in the order they are selected. Hence, the infeasible solutions selected first will be ranked higher than the feasible solutions selected later. A fixed value of α = 0.05 has been used in this study. It may be noted that the MVU-PCA dimensionality reduction scheme based on IDEA will be referred to, as MVU-PCA-IDEA.

V. SIMULATION RESULTS

Before entering this section, it is important to highlight that the real test of MVU-PCA-NSGA-II(Eps-dom) and MVU-PCA-IDEA lies in their ability/inability to solve problems which could not be solved earlier by even the most reliable procedure, namely MVU-PCA-NSGA-II with \( k = M - 1 \). It could be seen in Table I that DTLZ5-(5, 10) and DTLZ5-(5, 20) were two such problems. Further, it is important to mention that for the results summarized in Table I, population size (N) and the maximum number of generations (G) were taken as 600 and 10000, respectively. The SBX crossover [6] with a probability of 0.9 and index 5 and polynomial mutation with a probability of 0.1 and index of 50, were used. Further, the value TC was pre-set, as 0.95.

For this study, while the rest of the parameters are retained as stated above, the set \{N, G\} is taken as \{100, 5000\}, instead of \{600, 10000\}. This is done to pose a higher level of challenge to the proposed procedures, in that, if they can solve the earlier unsolved problems at a lower computational expense, it would be an added reflection of their efficacy.

Since MVU-PCA-NSGA-II(Eps-dom) and MVU-PCA-IDEA are being tested for \{N, G\} = \{100, 5000\}, it is important to test the response of MVU-PCA-NSGA-II \( (k = M - 1) \) as well. It is observed that the latter still fails to solve the DTLZ5-(5, 10) and DTLZ5-(5, 20) problems. It may be noted (Table II) that for these two problems, with \{N, G\} = \{600, 10000\} only about 0.5% solutions could arrive close to the P.O.F. Now for \{N, G\} = \{100, 5000\}, even though these numbers have improved to 3.0% and 1.0%, respectively (Table III), the problems are still not solved. Given that, earlier MVU-PCA-NSGA-II could accurately solve (Table II) problems with about 1.0% convergence (example, DTLZ5-(2, 50)), the failure could be attributed to the arguments that the check of \( g_{max} < 0.01 \) is not an appropriate measure of the degree to which the population has converged. Perhaps, parameters like \( g_{mean} \) and \( g_{stddev} \) also need to be considered.

On the contrary, both MVU-PCA-NSGA-II(Eps-dom)
and MVU-PCA-IDEA, each with the parameter $k = M - 1$ are found to successfully solve both the problems. Even though the $g_{\text{count}}$ are not drastically high, the reasons for the success of these methods could be found in extremely better values of $g_{\text{mean}}$ and $g_{\text{stddev}}$ on comparative terms (Table III). For instance, in case of DTLZ5-(5,20), while NSGA-II could achieve $g_{\text{mean}} = 2.133$ (against the maximum possible value of 2.5 and the desired value of 0), NSGA-II(Eps-dom) achieved $g_{\text{mean}} = 0.0307$, while IDEA achieved $g_{\text{mean}} = 0.0345$. Clearly, the fact that NSGA-II(Eps-dom) and IDEA could help solve the (earlier unsolved) difficult problems (and in much lower number of function evaluations) only reaffirms the potential of these schemes in handling of many-objective problems.

To demonstrate the relative strengths of NSGA-II, NSGA-II(Eps-dom) and IDEA, the simulation results with MVU-PCA-NSGA-II, MVU-PCA-NSGA-II(Eps-dom) and MVU-PCA-IDEA for an even more difficult problem DTLZ5-(5,30) are presented, for which it is known that the critical objective set is \{f_{26}, f_{27}, f_{28}, f_{29}, f_{30}\}.

<table>
<thead>
<tr>
<th>Table IV</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>DTLZ5-(5, 30): MVU-PCA-NSGA-II, Iteration 1</strong></td>
</tr>
</tbody>
</table>

(a) (i) Eigenvalue Analysis

<table>
<thead>
<tr>
<th>PCA-1 (94.15% variance)</th>
<th>f_{19} f_{26} f_{27} f_{28} f_{29} f_{30}</th>
</tr>
</thead>
<tbody>
<tr>
<td>PCA-2 (93.78% variance)</td>
<td>f_{26}</td>
</tr>
</tbody>
</table>

For MVU-PCA-NSGA-II, it can be seen in Table IV, that the eigen value analysis fails to pick $f_{28}$ as an important objective. As the true critical set cannot be achieved now, the procedure is terminated and no further iterations are performed.

In case of MVU-PCA-NSGA-II(Eps-dom), the eigen value analysis (Table V(a)) picks \{f_{20}, \ldots, f_{30}\} as the important objectives. However, the reduced correlation matrix (Table V(b)) identifies \{f_{20}, \ldots, f_{26}\} as identically correlated for which the selection scheme (Section III-A.3) based on $c_i^T = [0.0159 0.0302 0.0482 0.0709 0.1030 0.1522 0.2145]$ identifies $f_{26}$ as the most important objective in the correlated set. Not shown for brevity though, Iteration 2 is performed with the set of five objectives \{f_{26}, \ldots, f_{30}\}, where it is found further irreducible. This terminates the procedure with the true critical objective set, accurately identified.

Similarly, the Table VI for MVU-PCA-IDEA can be interpreted. The selection scheme for correlated \{f_{19}, \ldots, f_{28}\} picks $f_{26}$ as most important, based on $c_i^T = [0.006 0.016 0.030 0.051 0.080 0.119 0.173 0.240]$. Here too, Iteration 2 performed with \{f_{20}, \ldots, f_{30}\} identifies the set as irreducible. The problem is hence, accurately solved.

**VI. SUMMARY AND CONCLUSIONS**

In this paper, the promise of two algorithms namely NSGA-II(Eps-dom) and IDEA for constrained many-objective optimization problems is demonstrated. Though, this demonstration is based within the dimensionality re-
duction framework, the hallmark of these algorithms is their ability to counter the curse of low selection pressure for convergence to the P.O.F. Amidst the growing interest in the EMO community on handling of many-objective problems and proportionately increasing research effort, the ideas introduced in this paper may significantly contribute towards making the handling of many-objective problems, more realizable. Further studies are currently underway to observe the behavior of the algorithm for more difficult problems. The role of α parameter in case of IDEA is being further investigated.

ACKNOWLEDGEMENTS

The authors appreciate the efforts of Amitay Isaacs and Hemant Kumar Singh from the School of Aerospace, Civil and Mechanical Engineering, University of New South Wales, Australia towards the implementation of various components of the algorithm.

REFERENCES


[9] D. Drechsler, R. Drechsler, and B. Becker. Multi-objective optimization based on relation favo...