Adaptive Threshold Parameter Estimation with Recursive Differential Grouping for Problem Decomposition

Yuan Sun\textsuperscript{1,3}  Mohammad Nabi Omidvar\textsuperscript{2}  Michael Kirley\textsuperscript{1}  Xiaodong Li\textsuperscript{3}

\textsuperscript{1}School of Computing and Information Systems, University of Melbourne
\textsuperscript{2}School of Computer Science, University of Birmingham
\textsuperscript{3}School of Science, RMIT University

yuan.sun@unimelb.edu.au
yuan.sun@rmit.edu.au

July 17, 2018
Overview

1. Introduction
2. Background and Related Work
3. Adaptive Threshold Estimation for Recursive Differential Grouping
4. Experimental Results
5. Conclusion
Large-scale (High-dimensional) Continuous Optimization Problems are challenging to solve:

- Search space increases exponentially.
- Problem complexity increases greatly.
- The running time of some evolutionary algorithms increases significantly.
Background: Cooperative Co-evolution

Background: Cooperative Co-evolution

Background: Cooperative Co-evolution

---

Background: Cooperative Co-evolution

Background: Cooperative Co-evolution\textsuperscript{1}

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{fig}
\caption{Diagram of cooperative co-evolution with variables $x_1, x_2, x_3, x_4, x_5, x_6$ and their optimal values $x_1^*, x_2^*, x_3^*, x_4^*, x_5, x_6$.}
\end{figure}

Background: Cooperative Co-evolution

---

Background: Recursive Differential Grouping (RDG)\(^2\)

There exists some interaction between two subsets of decision variables \(X_1\) and \(X_2\) if

\[
\Delta X_1 f(x)|_{X_1=x_1^*,X_2=x_2^1} \neq \Delta X_1 f(x)|_{X_1=x_1^*,X_2=x_2^2},
\]

(1)

where

\[
\Delta X_1 f(x) = f(\cdots, X_1 + \Delta X_1, \cdots) - f(\cdots, X_1, \cdots).
\]

(2)

Background: Recursive Differential Grouping (RDG)²

There exists some interaction between two subsets of decision variables $X_1$ and $X_2$ if

$$\Delta X_1 f(x)|_{X_1=x_1^*, X_2=x_2^1} \neq \Delta X_1 f(x)|_{X_1=x_1^*, X_2=x_2^2},$$  

where

$$\Delta X_1 f(x) = f(\cdots, X_1 + \Delta X_1, \cdots) - f(\cdots, X_1, \cdots).$$

There exists some interaction between two subsets of decision variables $X_1$ and $X_2$ if

$$\Delta X_1 f(x)|_{x_1=x_1^*,x_2=x_2^1} \neq \Delta X_1 f(x)|_{x_1=x_1^*,x_2=x_2^2},$$

(1)

where

$$\Delta X_1 f(x) = f(\cdots, X_1 + \Delta X_1, \cdots) - f(\cdots, X_1, \cdots).$$

(2)
Background: Recursive Differential Grouping (RDG) \(^2\)

There exists some interaction between two subsets of decision variables \(X_1\) and \(X_2\) if

\[
\Delta X_1 f(x)|_{X_1 = X_1^*, X_2 = X_2^1} \neq \Delta X_1 f(x)|_{X_1 = X_1^*, X_2 = X_2^2},
\]

(1)

where

\[
\Delta X_1 f(x) = f(\cdots, X_1 + \Delta X_1, \cdots) - f(\cdots, X_1, \cdots).
\]

(2)

There exists some interaction between two subsets of decision variables $X_1$ and $X_2$ if

$$\Delta X_1 f(x)|_{X_1=X_1^*, X_2=X_2^1} \neq \Delta X_1 f(x)|_{X_1=X_1^*, X_2=X_2^2},$$

(1)

where

$$\Delta X_1 f(x) = f(\cdots, X_1 + \Delta X_1, \cdots) - f(\cdots, X_1, \cdots).$$

(2)
Background: Recursive Differential Grouping (RDG)

There exists some interaction between two subsets of decision variables $X_1$ and $X_2$ if

$$\Delta x_1 f(x)|_{x_1=x_1^*, x_2=x_2^1} \neq \Delta x_1 f(x)|_{x_1=x_1^*, x_2=x_2^2},$$

(1)

where

$$\Delta x_1 f(x) = f(\cdots, x_1 + \Delta x_1, \cdots) - f(\cdots, x_1, \cdots).$$

(2)

---

There exists some interaction between two subsets of decision variables $X_1$ and $X_2$ if
\[ \Delta X_1 f(x)|_{X_1 = X_1^*, X_2 = X_2^1} \neq \Delta X_1 f(x)|_{X_1 = X_1^*, X_2 = X_2^2}, \quad (1) \]
where
\[ \Delta X_1 f(x) = f(\cdots, X_1 + \Delta X_1, \cdots) - f(\cdots, X_1, \cdots). \quad (2) \]

---

There exists some interaction between two subsets of decision variables \( X_1 \) and \( X_2 \) if

\[
\Delta X_1 f(x) \big|_{X_1=X_1^*, X_2=X_2^1} \neq \Delta X_1 f(x) \big|_{X_1=X_1^*, X_2=X_2^2}, \tag{1}
\]

where

\[
\Delta X_1 f(x) = f(\cdots, X_1 + \Delta X_1, \cdots) - f(\cdots, X_1, \cdots). \tag{2}
\]

\[\text{Background: Recursive Differential Grouping (RDG)} \]

There exists some interaction between two subsets of decision variables $X_1$ and $X_2$ if

$$\Delta X_1 f(x)|_{x_1=x_1^*, x_2=x_2^1} \neq \Delta X_1 f(x)|_{x_1=x_1^*, x_2=x_2^2},$$

where

$$\Delta X_1 f(x) = f(\cdots, X_1 + \Delta X_1, \cdots) - f(\cdots, X_1, \cdots).$$
Background: Recursive Differential Grouping (RDG)²

There exists some interaction between two subsets of decision variables $X_1$ and $X_2$ if

$$\Delta X_1 f(x)|_{x_1=x_1^*,x_2=x_2^1} \neq \Delta X_1 f(x)|_{x_1=x_1^*,x_2=x_2^2},$$  \hspace{1cm} (1)

where

$$\Delta X_1 f(x) = f(\cdots, X_1 + \Delta X_1, \cdots) - f(\cdots, X_1, \cdots).$$ \hspace{1cm} (2)

There exists some interaction between two subsets of decision variables $X_1$ and $X_2$ if

$$\Delta X_1 f(x) |_{X_1=\ast_1, X_2=\ast_2} \neq \Delta X_1 f(x) |_{X_1=\ast_1, X_2=\ast_2},$$

(1)

where

$$\Delta X_1 f(x) = f(\cdots, X_1 + \Delta X_1, \cdots) - f(\cdots, X_1, \cdots).$$

(2)

---

In theory, if $\lambda = 0$, $X_1$ and $X_2$ are separable; if $\lambda > 0$, $X_1$ and $X_2$ interact, where $\lambda = |\Delta_1 - \Delta_2|$.
1. In theory, if $\lambda = 0$, $X_1$ and $X_2$ are separable; if $\lambda > 0$, $X_1$ and $X_2$ interact, where $\lambda = |\Delta_1 - \Delta_2|$.

2. In practice, if $\lambda \leq \epsilon$, $X_1$ and $X_2$ are separable; if $\lambda > \epsilon$, $X_1$ and $X_2$ interact.
In theory, if $\lambda = 0$, $X_1$ and $X_2$ are separable; if $\lambda > 0$, $X_1$ and $X_2$ interact, where $\lambda = |\Delta_1 - \Delta_2|$.

In practice,
1. In theory, if $\lambda = 0$, $X_1$ and $X_2$ are separable; if $\lambda > 0$, $X_1$ and $X_2$ interact, where $\lambda = |\Delta_1 - \Delta_2|$.

2. In practice,

![Graph showing $\lambda_{sep}$ and $\lambda_{max}$ for various values of $\lambda$.]
In theory, if $\lambda = 0$, $X_1$ and $X_2$ are separable; if $\lambda > 0$, $X_1$ and $X_2$ interact, where $\lambda = |\Delta_1 - \Delta_2|$. 

In practice,
1. In theory, if $\lambda = 0$, $X_1$ and $X_2$ are separable; if $\lambda > 0$, $X_1$ and $X_2$ interact, where $\lambda = |\Delta_1 - \Delta_2|$.

2. In practice,
1. In theory, if $\lambda = 0$, $X_1$ and $X_2$ are separable; if $\lambda > 0$, $X_1$ and $X_2$ interact, where $\lambda = |\Delta_1 - \Delta_2|$.

2. In practice,
1. In theory, if \( \lambda = 0 \), \( X_1 \) and \( X_2 \) are separable; if \( \lambda > 0 \), \( X_1 \) and \( X_2 \) interact, where \( \lambda = |\Delta_1 - \Delta_2| \).

2. In practice, if \( \lambda \leq \epsilon \), \( X_1 \) and \( X_2 \) are separable; if \( \lambda > \epsilon \), \( X_1 \) and \( X_2 \) interact.
The RDG method estimates a threshold value based on the magnitude of the objective values:

$$\epsilon := \alpha \cdot \min \left\{ |f(x_1)|, \cdots, |f(x_k)| \right\},$$  \hspace{1cm} (3)

where $x_1, \cdots, x_k$ are $k$ randomly generated candidate solutions, and $\alpha$ is the control coefficient \(^3\).

---

The RDG method estimates a threshold value based on the magnitude of the objective values:

\[
\epsilon := \alpha \cdot \min \left\{ |f(x_1)|, \cdots, |f(x_k)| \right\},
\]

(3)

where \(x_1, \cdots, x_k\) are \(k\) randomly generated candidate solutions, and \(\alpha\) is the control coefficient \(^3\).

Limitations:

1. Lack of theoretical foundation.

The RDG method estimates a threshold value based on the magnitude of the objective values:

$$
\epsilon := \alpha \cdot \min \left\{ |f(x_1)|, \ldots, |f(x_k)| \right\},
$$

where $x_1, \ldots, x_k$ are $k$ randomly generated candidate solutions, and $\alpha$ is the control coefficient.

Limitations:

1. Lack of theoretical foundation.
2. Non-trivial to select an appropriate value for $\alpha$.

---

Background: Parameter (Threshold) Setting for RDG

The RDG method estimates a threshold value based on the magnitude of the objective values:

\[ \epsilon := \alpha \cdot \min \{ |f(x_1)|, \ldots, |f(x_k)| \} \tag{3} \]

where \( x_1, \ldots, x_k \) are \( k \) randomly generated candidate solutions, and \( \alpha \) is the control coefficient \(^3\).

Limitations:

1. Lack of theoretical foundation.
2. Non-trivial to select an appropriate value for \( \alpha \).
3. Insufficient to deal with problems with imbalanced components.

The round-off errors involved in the calculation of the non-linearity term \( \lambda = |(f(x_{l,l}) - f(x_{u,l})) - (f(x_{l,m}) - f(x_{u,m}))| \) come from two sources:

\[ \hat{\Delta}_1 = \hat{f}(x_{l,l}) - \hat{f}(x_{u,l}) = (\hat{f}(x_{l,l}) - \hat{f}(x_{u,l})) (1 + \delta_1) \]
\[ \hat{\Delta}_2 = \hat{f}(x_{l,m}) - \hat{f}(x_{u,m}) = (\hat{f}(x_{l,m}) - \hat{f}(x_{u,m}))(1 + \delta_2) \]

where \(|\delta_1|, |\delta_2|, |\delta_3| < \mu_M \).

\[ \hat{\lambda} = |\hat{\Delta}_1 - \hat{\Delta}_2| = |(\hat{f}(x_{l,l}) - \hat{f}(x_{u,l})) (1 + \delta_1)(1 + \delta_3) - (\hat{f}(x_{l,m}) - \hat{f}(x_{u,m}))(1 + \delta_2)(1 + \delta_3)| \]

\(^4\hat{\Delta}\) denotes the floating-point number of \( \Delta \); \( \oplus \) denotes floating-point substraction; \( \mu_M \) is a machine dependent constant (\( \mu_M = 2^{-53} \) in MATLAB).
The round-off errors involved in the calculation of the non-linearity term 
\[ \lambda = \left| (f(x_{l,l}) - f(x_{u,l})) - (f(x_{l,m}) - f(x_{u,m})) \right| \] come from two sources:

**S1:** the arithmetic floating-point subtraction between fitness values \( f(x) \).

---

\( ^4 \hat{\Delta} \) denotes the floating-point number of \( \Delta \); \( \ominus \) denotes floating-point substraction; \( \mu_M \) is a machine dependent constant (\( \mu_M = 2^{-53} \) in MATLAB).
The round-off errors involved in the calculation of the non-linearity term
\[ \lambda = \left| (f(x_l, l) - f(x_u, l)) - (f(x_l, m) - f(x_u, m)) \right| \]
come from two sources:

**S1:** the arithmetic floating-point subtraction between fitness values \( f(x) \).

**S2:** the calculation of the fitness values \( f(x) \).

---

\( \hat{\Delta} \) denotes the floating-point number of \( \Delta \); \( \ominus \) denotes floating-point substraction; \( \mu_M \) is a machine dependent constant (\( \mu_M = 2^{-53} \) in MATLAB).
Adaptive Threshold Estimation: Round-off Errors (S1)

The round-off errors involved in the calculation of the non-linearity term \( \lambda = \left| (f(x_{l,l}) - f(x_{u,l})) - (f(x_{l,m}) - f(x_{u,m})) \right| \) come from two sources:

**S1:** the arithmetic floating-point subtraction between fitness values \( f(x) \).

**S2:** the calculation of the fitness values \( f(x) \).

Round-off Errors (S1):

\[
\hat{\Delta}_1 = \hat{f}(x_{l,l}) \ominus \hat{f}(x_{u,l}) = (\hat{f}(x_{l,l}) - \hat{f}(x_{u,l}))(1 + \delta_1), \quad \text{where } |\delta_1| < \mu_M;^4
\]

---

^4 \( \hat{\Delta} \) denotes the floating-point number of \( \Delta \); \( \ominus \) denotes floating-point substraction; \( \mu_M \) is a machine dependent constant (\( \mu_M = 2^{-53} \) in MATLAB).
Adaptive Threshold Estimation: Round-off Errors (S1)

The round-off errors involved in the calculation of the non-linearity term $\lambda = \left| (f(x_{l,l}) - f(x_{u,l})) - (f(x_{l,m}) - f(x_{u,m})) \right|$ come from two sources:

**S1:** the arithmetic floating-point subtraction between fitness values $f(x)$.

**S2:** the calculation of the fitness values $f(x)$.

**Round-off Errors (S1):**

$$\hat{\Delta}_1 = \hat{f}(x_{l,l}) \ominus \hat{f}(x_{u,l}) = (\hat{f}(x_{l,l}) - \hat{f}(x_{u,l}))(1 + \delta_1), \text{ where } |\delta_1| < \mu_M;$$  \hspace{1cm} (4)

$$\hat{\Delta}_2 = \hat{f}(x_{l,m}) \ominus \hat{f}(x_{u,m}) = (\hat{f}(x_{l,m}) - \hat{f}(x_{u,m}))(1 + \delta_2), \text{ where } |\delta_2| < \mu_M;$$  \hspace{1cm} (5)

---

$\hat{\Delta}$ denotes the floating-point number of $\Delta$; $\ominus$ denotes floating-point substraction; $\mu_M$ is a machine dependent constant ($\mu_M = 2^{-53}$ in MATLAB).
Adaptive Threshold Estimation: Round-off Errors (S1)

The round-off errors involved in the calculation of the non-linearity term

$$\lambda = \left| (f(x_{l,l}) - f(x_{u,l})) - (f(x_{l,m}) - f(x_{u,m})) \right|$$

come from two sources:

S1: the arithmetic floating-point subtraction between fitness values $$f(x)$$.
S2: the calculation of the fitness values $$f(x)$$.

Round-off Errors (S1):

$$\hat{\Delta}_1 = \hat{f}(x_{l,l}) \ominus \hat{f}(x_{u,l}) = (\hat{f}(x_{l,l}) - \hat{f}(x_{u,l}))(1 + \delta_1), \text{ where } |\delta_1| < \mu_M; \quad (4)$$

$$\hat{\Delta}_2 = \hat{f}(x_{l,m}) \ominus \hat{f}(x_{u,m}) = (\hat{f}(x_{l,m}) - \hat{f}(x_{u,m}))(1 + \delta_2), \text{ where } |\delta_2| < \mu_M; \quad (5)$$

$$\hat{\lambda} = |\hat{\Delta}_1 \ominus \hat{\Delta}_2| = |(\hat{\Delta}_1 - \hat{\Delta}_2)(1 + \delta_3)| = |(\hat{f}(x_{l,l}) - \hat{f}(x_{u,l}))(1 + \delta_1)(1 + \delta_3) - (\hat{f}(x_{l,m}) - \hat{f}(x_{u,m}))(1 + \delta_2)(1 + \delta_3)|, \text{ where } |\delta_1|, |\delta_2|, |\delta_3| < \mu_M. \quad (6)$$

$${}^4\hat{\Delta} \text{ denotes the floating-point number of } \Delta; \ominus \text{ denotes floating-point substraction; } \mu_M \text{ is a machine dependent constant (} \mu_M = 2^{-53} \text{ in MATLAB}).$$
Theorem

Given a floating-point number system that satisfies IEEE 754 Standard such that \(|\delta_i| < \mu_M\), and \(k\mu_M < 1\), we have:

\[
\prod_{i=1}^{k} (1 + \delta_i)^{e_i} = 1 + \theta_k, \text{ where } |\theta_k| \leq \frac{k\mu_M}{1 - k\mu_M} := \gamma_k \text{ and } e_i = \pm 1. \tag{7}
\]

\(^a\)Corless R M, Fillion N. A graduate introduction to numerical methods[J]. AMC, 2013, 10: 12, Springer.

Example: \((1 + \delta_1)(1 + \delta_3) = (1 + \theta_2)\), where \(|\theta_2| \leq \gamma_2\).
Theorem

Given a floating-point number system that satisfies IEEE 754 Standard such that $|\delta_i| < \mu_M$, and $k\mu_M < 1$, we have:

$$\prod_{i=1}^{k}(1 + \delta_i)^{e_i} = 1 + \theta_k, \text{ where } |\theta_k| \leq \frac{k\mu_M}{1 - k\mu_M} := \gamma_k \text{ and } e_i = \pm 1.$$  \hspace{1cm} (7)

Example: $(1 + \delta_1)(1 + \delta_3) = (1 + \theta_2)$, where $|\theta_2| \leq \gamma_2$.

Estimating an upper bound for $S1$:

$$\hat{\lambda} = |(\hat{f}(x_{l,l}) - \hat{f}(x_{u,l}))(1 + \theta_2) - (\hat{f}(x_{l,m}) - \hat{f}(x_{u,m}))(1 + \theta'_2)|,$$

where $|\theta_2| \leq \gamma_2$ and $|\theta'_2| \leq \gamma_2$.  \hspace{1cm} (8)

\text{Corless R M, Fillion N. A graduate introduction to numerical methods[J].}
AMC, 2013, 10: 12, Springer.
Adaptive Threshold Estimation: Round-off Errors (S2)

Assumption 1: The number of floating-point operations ($\Phi$) involved in the calculation of a black-box objective function is in the order of $\Theta(n)$, where $n$ is the dimensionality of the objective function$^5$:

$$\Phi \approx n.$$  \hspace{1cm} (9)

---


Adaptive Threshold Estimation: Round-off Errors (S2)

**Assumption 1:** The number of floating-point operations (\(\Phi\)) involved in the calculation of a black-box objective function is in the order of \(\Theta(n)\), where \(n\) is the dimensionality of the objective function:\(^5\)

\[
\Phi \approx n.
\]  

(9)

**Assumption 2:** The round-off error grows with the square root of the number of floating-point operations (\(\Phi\)) involved in a calculation:\(^6\):

\[
k \approx \sqrt{\Phi}.
\]  

(10)

---


Adaptive Threshold Estimation: Round-off Errors (S2)

Assumption 1: The number of floating-point operations (Φ) involved in the calculation of a black-box objective function is in the order of $\Theta(n)$, where $n$ is the dimensionality of the objective function$^5$:

$$\Phi \approx n.$$  \hspace{1cm} (9)

Assumption 2: The round-off error grows with the square root of the number of floating-point operations (Φ) involved in a calculation$^6$:

$$k \approx \sqrt{\Phi}.$$ \hspace{1cm} (10)

Estimating an upper bound for S2:

$$\hat{f}(\mathbf{x}) = (1 + \theta \sqrt{n})f(\mathbf{x}), \text{ where } |\theta \sqrt{n}| \leq \gamma \sqrt{n}.$$ \hspace{1cm} (11)

---


Theorem

*Under Assumption 1 and Assumption 2, an upper bound on the round-off errors associated with the calculation of the non-linearity term $\lambda$ is given by*

$$
|\lambda - \hat{\lambda}| \leq \gamma \sqrt{n+2} \left( |f(x_{l,l})| + |f(x_{u,l})| + |f(x_{l,m})| + |f(x_{u,m})| \right).
$$

(12)
Theorem

Under Assumption 1 and Assumption 2, an upper bound on the round-off errors associated with the calculation of the non-linearity term $\lambda$ is given by

$$|\lambda - \hat{\lambda}| \leq \gamma \sqrt{n+2} (|f(x_{l,l})| + |f(x_{u,l})| + |f(x_{l,m})| + |f(x_{u,m})|). \quad (12)$$

Proof.

Substitute $\hat{f}(x) = (1 + \theta \sqrt{n})f(x)$ into

$$\hat{\lambda} = \left| (\hat{f}(x_{l,l}) - \hat{f}(x_{u,l}))(1 + \theta_2) - (\hat{f}(x_{l,m}) - \hat{f}(x_{u,m}))(1 + \theta_2') \right|. \quad (13)$$
Adaptive Threshold Estimation: An Upper Bound

**Theorem**

Under Assumption 1 and Assumption 2, an upper bound on the round-off errors associated with the calculation of the non-linearity term $\lambda$ is given by

$$|\lambda - \hat{\lambda}| \leq \gamma \sqrt{n+2} \left( |f(x_l,l)| + |f(x_u,l)| + |f(x_l,m)| + |f(x_u,m)| \right).$$

(12)

**Proof.**

Substitute $\hat{f}(x) = (1 + \theta \sqrt{n})f(x)$ into

$$\hat{\lambda} = \left| (\hat{f}(x_l,l) - \hat{f}(x_u,l))(1 + \theta_2) - (\hat{f}(x_l,m) - \hat{f}(x_u,m))(1 + \theta'_2) \right|.$$

(13)

Adaptive Threshold:

$$\epsilon := \gamma \sqrt{n+2} \left( |f(x_l,l)| + |f(x_u,l)| + |f(x_l,m)| + |f(x_u,m)| \right).$$

(14)

Variables are regarded as interacting if $\hat{\lambda} > \epsilon$, and separable if $\hat{\lambda} \leq \epsilon$. 
**Experimental Results: Decomposition Comparison**

**Table:** The decomposition results of the RDG2, RDG (with $\alpha = 10^{-12}$) and DG2 methods when used to decompose the CEC'2013 benchmark problems. “$a$” denotes the decomposition accuracy; “FEs” denotes the function evaluations used.

<table>
<thead>
<tr>
<th>Func</th>
<th>RDG2</th>
<th>RDG ($\alpha = 10^{-12}$)</th>
<th>DG2</th>
</tr>
</thead>
<tbody>
<tr>
<td>$f_7$</td>
<td>100%</td>
<td>9.81e+03</td>
<td>100%</td>
</tr>
<tr>
<td>$f_8$</td>
<td>80.0%</td>
<td>1.91e+04</td>
<td>80.0%</td>
</tr>
<tr>
<td>$f_{10}$</td>
<td>100%</td>
<td>1.93e+04</td>
<td>82.7%</td>
</tr>
<tr>
<td>$f_{11}$</td>
<td>100%</td>
<td>1.93e+04</td>
<td>10.0%</td>
</tr>
</tbody>
</table>
Experimental Results: Decomposition Details ($f_{11}$)
Experimental Results: Decomposition Details ($f_{11}$)

\[ \hat{\lambda}_{\text{int}} \]

\[ \hat{\lambda}_{\text{max}} \]

\[ \hat{\lambda}_{\text{sep}} \]
Experimental Results: Decomposition Details ($f_{11}$)

- $\hat{\lambda}_{int}^{0.25}$
- $\epsilon_{RDG2}$
- $\hat{\lambda}_{sep}^{max}$

Graph showing the values of $\hat{\lambda}_{int}^{0.25}$, $\epsilon_{RDG2}$, and $\hat{\lambda}_{sep}^{max}$ over a range of indices from 1 to 20.
Experimental Results: Decomposition Details ($f_{11}$)

Yuan Sun (University of Melbourne)  
Adaptive Threshold Estimation with RDG  
July 17, 2018 13 / 23
Experimental Results: Decomposition Details ($f_8$)

Yuan Sun (University of Melbourne)

Adaptive Threshold Estimation with RDG

July 17, 2018 14 / 23
Experimental Results: Decomposition Details ($f_8$)

\[ \hat{\lambda}_{int}^{0.25}, \hat{\lambda}_{sep}^{max} \]
Experimental Results: Decomposition Details ($f_8$)
Experimental Results: Decomposition Details ($f_8$)
Experimental Results: Decomposition Details ($f_8$)
Experimental Results: Decomposition Details ($f_8$)
Experimental Results: Optimization Comparison

**Table:** The optimization results of RDG2, RDG and DG2 when embedded into a CC framework to solve CEC’2013 benchmark problems (Wilcoxon rank-sum tests).

<table>
<thead>
<tr>
<th>Func</th>
<th>Stats</th>
<th>RDG2</th>
<th>RDG</th>
<th>DG2</th>
</tr>
</thead>
<tbody>
<tr>
<td>$f_7$</td>
<td>median</td>
<td>3.12e-19</td>
<td>2.93e-20</td>
<td>1.00e+03</td>
</tr>
<tr>
<td></td>
<td>mean</td>
<td>4.04e-16</td>
<td>8.11e-17</td>
<td>1.05e+03</td>
</tr>
<tr>
<td></td>
<td>std</td>
<td>1.48e-15</td>
<td>2.17e-16</td>
<td>2.78e+02</td>
</tr>
<tr>
<td>$f_8$</td>
<td>median</td>
<td>8.15e+06</td>
<td>8.26e+06</td>
<td>3.56e+07</td>
</tr>
<tr>
<td></td>
<td>mean</td>
<td>8.70e+06</td>
<td>8.50e+06</td>
<td>3.84e+07</td>
</tr>
<tr>
<td></td>
<td>std</td>
<td>3.61e+06</td>
<td>2.91e+06</td>
<td>1.08e+07</td>
</tr>
<tr>
<td>$f_{10}$</td>
<td>median</td>
<td>9.05e+07</td>
<td>9.05e+07</td>
<td>9.05e+07</td>
</tr>
<tr>
<td></td>
<td>mean</td>
<td>9.10e+07</td>
<td>9.10e+07</td>
<td>9.13e+07</td>
</tr>
<tr>
<td></td>
<td>std</td>
<td>1.30e+06</td>
<td>1.29e+06</td>
<td>1.50e+06</td>
</tr>
<tr>
<td>$f_{11}$</td>
<td>median</td>
<td>2.81e+03</td>
<td>1.68e+07</td>
<td>1.55e+05</td>
</tr>
<tr>
<td></td>
<td>mean</td>
<td>8.68e+03</td>
<td>1.67e+07</td>
<td>2.47e+05</td>
</tr>
<tr>
<td></td>
<td>std</td>
<td>1.24e+04</td>
<td>1.61e+06</td>
<td>2.36e+05</td>
</tr>
</tbody>
</table>
Figure: The convergence curves of the RDG2, RDG and DG2 methods when embedded into the CC framework to solve the CEC’2013 $f_{11}$. 
Conclusion and Future Work

Conclusion

- Derived an upper bound on the computational round-off errors involved in calculating the non-linearity term for RDG.

Future Work

- Systematically investigate the correlation between the non-linearity term for interacting variables and the weight of the components.
- Generate a more effective decomposition for large-scale problems with overlapping components.
Conclusion

- Derived an upper bound on the computational round-off errors involved in calculating the non-linearity term for RDG.
- Showed that the upper bound was able to be used as the threshold value to identify variable interactions across a wide range of benchmark problems.

Future Work

- Systematically investigate the correlation between the non-linearity term for interacting variables and the weight of the components.
- Generate a more effective decomposition for large-scale problems with overlapping components.
Conclusion and Future Work

1 Conclusion

- Derived an upper bound on the computational round-off errors involved in calculating the non-linearity term for RDG.
- Showed that the upper bound was able to be used as the threshold value to identify variable interactions across a wide range of benchmark problems.

2 Future Work

- Systematically investigate the correlation between the non-linearity term for interacting variables and the weight of the components.
Conclusion

- Derived an upper bound on the computational round-off errors involved in calculating the non-linearity term for RDG.
- Showed that the upper bound was able to be used as the threshold value to identify variable interactions across a wide range of benchmark problems.

Future Work

- Systematically investigate the correlation between the non-linearity term for interacting variables and the weight of the components.
- Generate a more effective decomposition for large-scale problems with overlapping components.
Thank You! & Questions?
Interaction Structure

\[ x_1, x_2, x_3, x_4, x_5 \]

\[ \Delta_1 \neq \Delta_2 \]
Back-up: Decomposition Process of RDG

Interaction Structure

Decomposition Process

\[ x_2 \]
\[ x_1 \quad x_3 \]
\[ x_4 \quad x_5 \]

\[ \Delta_1 \neq \Delta_2 \]

Yuan Sun (University of Melbourne)
Adaptive Threshold Estimation with RDG
July 17, 2018
Back-up: Decomposition Process of RDG

Interaction Structure

- $X_2$
- $X_1$
- $X_3$

Decomposition Process

- $X_1$
- $X_2, X_3, X_4, X_5$

Yuan Sun (University of Melbourne)
Adaptive Threshold Estimation with RDG
July 17, 2018 20 / 23
Back-up: Decomposition Process of RDG

Interaction Structure

- $x_2$
- $x_1$
- $x_3$
- $x_4$
- $x_5$

Decomposition Process

- $\Delta_1 \neq \Delta_2$
- $x_1 \rightarrow x_2, x_3, x_4, x_5$
Back-up: Decomposition Process of RDG

Interaction Structure

Decomposition Process

\[ \Delta_1 \neq \Delta_2 \]

\[ x_1 \]

\[ x_2, x_3, x_4, x_5 \]

\[ x_2, x_3 \]

\[ x_4, x_5 \]

Yuan Sun (University of Melbourne)
Adaptive Threshold Estimation with RDG
July 17, 2018 20 / 23
Back-up: Decomposition Process of RDG

Interaction Structure

Decomposition Process

\[ \Delta_1 \neq \Delta_2 \]

\[ \Delta_1 = \Delta_2 \]
Back-up: Decomposition Process of RDG

Interaction Structure

Decomposition Process

$\Delta_1 \neq \Delta_2$

$\Delta_1 = \Delta_2$

$\Delta_1 \neq \Delta_2$
Back-up: Decomposition Process of RDG

Interaction Structure

Decomposition Process

\[ \Delta_1 \neq \Delta_2 \]

\[ \Delta_1 = \Delta_2 \]

\[ \Delta_1 \neq \Delta_2 \]

\[ \Delta_1 = \Delta_2 \]
Back-up: Decomposition Process of RDG

Interaction Structure

Decomposition Process

\[ \Delta_1 \neq \Delta_2 \]

\[ \Delta_1 = \Delta_2 \]

Yuan Sun (University of Melbourne)
Back-up: Decomposition Process of RDG

Interaction Structure

Decomposition Process

\[ \Delta_1 \neq \Delta_2 \]

\[ x_1, x_2 \rightarrow x_3, x_4, x_5 \]
Back-up: Decomposition Process of RDG

Interaction Structure

Decomposition Process

$x_1, x_2 \Rightarrow \Delta_1 \neq \Delta_2 \Rightarrow x_3, x_4, x_5$

Yuan Sun (University of Melbourne)
Back-up: Decomposition Process of RDG

Interaction Structure

Decomposition Process

\[ \Delta_1 \neq \Delta_2 \]

\[ \Delta_1 = \Delta_2 \]
Back-up: Decomposition Process of RDG

Interaction Structure

Decomposition Process

\[ \Delta_1 \neq \Delta_2 \]

\[ \Delta_1 = \Delta_2 \]
Back-up: Decomposition Process of RDG

Interaction Structure

Decomposition Process

\[ x_1, x_2 \]
\[ \Delta_1 \neq \Delta_2 \]
\[ x_3 \]

\[ x_1, x_2 \]
\[ \Delta_1 = \Delta_2 \]
\[ x_3 \]
\[ x_4, x_5 \]

\[ x_1, x_2, x_3 \]
\[ \Delta_1 = \Delta_2 \]
\[ x_4, x_5 \]
Back-up: Decomposition Process of RDG

Interaction Structure

Decomposition Process

X₁, X₂, X₃

X₂

X₁
X₃

X₄
X₅

Δ₁ ≠ Δ₂
Back-up: Decomposition Process of RDG

Interaction Structure

Decomposition Process

$x_1$, $x_2$, $x_3$

$x_1$

$x_2$

$x_3$

$x_4$

$x_5$

$\Delta_1 \neq \Delta_2$

Yuan Sun (University of Melbourne)
Back-up: Decomposition Process of RDG

Interaction Structure

Decomposition Process

\[ \Delta_1 \neq \Delta_2 \]
Interaction Structure

Decomposition Process

\[ \Delta_1 \neq \Delta_2 \]
Back-up: Decomposition Process of RDG

Interaction Structure

Decomposition Process

\[ \Delta_1 \neq \Delta_2 \]
Back-up: Time Complexity of RDG

Time Complexity: $O(n \log(n))$

1. Fully separable problem: $3n \in \Theta(n)$.
2. Fully non-separable problem: $6n \in \Theta(n)$.
3. Partially separable problem: $6n \log_2(n) \in \Theta(n \log(n))$.
4. Overlapping problem $6n \log_2(n) \in \Theta(n \log(n))$. 