Fast Design Optimization Strategy for Radiative Heat Transfer using ANSYS and eArtius Gradient Based Optimization Method – Pt. 2

Vladimir Kudriavtsev, Terry Bluck (Intevac)
Vladimir Sevastyanov
(eArtius, Irvine, CA)

Outline

- Part 1 of this work presented at ANSYS Users Conference 2011 (Santa Clara) and devoted to hybrid genetic and gradient based approach (HMGE)
  
  [http://www.slideshare.net/vvk0/optimization-intevac-aug23-7f](http://www.slideshare.net/vvk0/optimization-intevac-aug23-7f)

- Part 2 of this work is presented here and is devoted to pure gradient based method, which is best used when only limited number of design evaluations is possible (due to CPU time limitations or other reasons)
Current Computational Design Process

8 threads i7 CPU

Computer

240 cores TESLA Graphic Processing Unit GPU (x2)

Human Thinking and Analysis

Ingenious Solutions

fastest component
and grows exponentially faster

slowest component
(meetings, reviews, alignments, cancelations)
Why Optimization by Computer?

Human can not match computer in repetitive tasks and consistency. Assuming computational problem takes 4 hours of CPU time, then in one day (between 8AM to 8 AM) computer is capable of producing 6 design evaluations, with 42 designs completed in just 7 work days.

Coupled with multi-processing capability of i7 workstation this number can easily be multiplied by factors ranging from two to six. Computer will work during the weekend; it will work when user is on vacation, on sick leave or on business trip.

Personal “super computer” cost is now inconsequential for the bottom line.

Software cost sky-rocketed, and its ROI and utilization efficiency is now most important.

Computer needs algorithmic analogy of “human brain” to self-guide solution steps.
New paradigm of **multi-objective computational design** is now being born.

No longer designer needs to approach it through “trial-and-error” simulations, but can rather use “artificial intelligence” of optimization method to automatically seek and to find best combination of input parameters (design). Depending on problem size (CPU time) this process can take from minutes to weeks.

However, now engineer can view tens and hundreds of possible solutions, automatically singling first truly best designs and then evaluate design trade-offs between conflicting objectives (Pareto Frontier).

In many instances, examining dozens and hundreds of computational designs is simply time prohibitive. What to do then?
Intevac c-Si Technology

LEAN SOLAR™
A Revolution in Value for the Solar Industry

Crystalline Silicon

TCO:
Transparent Conducting Oxides such as ITO and ZnO can be deposited using a PVD sputter approach on Lean Solar™ for application on crystalline silicon solar cells such as hetero-junction solar cells.

Metals:
Metallic layers are deposited through PVD sputter processing using Lean Solar™, typically for contact formation and reflector layers on c-Si Solar cells. Metals deposited capability is broad with Lean Solar and can be integrated in stack layers. Capability includes, but is not limited to: Aluminum (Al), Titanium (Ti), Nickel Vanadium (NiV), Copper (Cu) and Molybdenum (Mo).

http://www.intevac.com
Problem Formulation

Minimize thermal variation across single substrate and across a group of substrates during radiant heating stage (TempDiff)

Operate in required process temperature window, \( T_{-dev1} < Top < T_{+dev2} \)

Optimization Formulation

\[ \text{Top} = 400 \text{ deg.C} \]

\[ \begin{align*}
\text{min} \ (\text{TempDiff}) & \quad \text{&} \quad \text{min} \ \text{abs}(T_{\text{max}}-\text{Top}) \quad \text{&} \quad \text{min} \ \text{abs}(T_{\text{min}}-\text{Top}) \\
\end{align*} \]

Constraints to determine design feasibility:

\[ T < T_{\text{max.constr}} \quad \text{&} \quad T > T_{\text{min.constr}}, \text{ where} \]

\[ T_{\text{min.constr}} = \text{Top}-dev1, \quad T_{\text{max.constr}} = \text{Top}+dev2 \]

If dev1 and dev2 are small, then optimization problem is very restrictive.
Gradient method requires path, to enter narrow optimal range (due to nonlinearity) it requires guidance or coincidence. Guidance comes from the previous history (steps taken before, gradients) and coincidence from DOE or random mutations.
MGP Method-- Analogy

DOE #1
DOE #2
DOE #3
DOE #N

Design Space
Initial Design
Initial Tolerance
MGP Design Vector Calculated using
Problem Parameters – Geometry and Temperature

\[
<T_{\text{Tempdiff}}> = T_{\text{max}} - T_{\text{min}}
\]

between 3 substrates

\[ \text{T1, T2 control heat flux from lamps.} \]
Thermal Heating (Radiation) Solution

Interplay between two lamp arrays

Lamp Bank 1

Lamp Bank 2

Substrate Motion Direction

Multi-Step Transient History

Transient Heating Scenario: Row1 of substrates is first heated by Lamp Bank1, then these Substrates moved to Lamp Bank2 and get heated again till desired Top=400 deg.C is reached. Simultaneously, new substrates with T=Tambient populate Row1 and get heated. Thus, Row1 heats from 22 to 250 deg.c and Row 2 from 250 to 400 deg.C. At time t=3.5 sec Row1 T is reset at 22 deg.C; Row2 T is reset at 250 deg.C. At time t=0 sec Row1 T is set at 22 deg.C; Row2 T is set at 250 deg.C.
This Study Consists of Two Parts. In Part 1 (presented in Santa Clara) we focused on hybrid genetic and gradient based method (HMGE). It has lots of positive sides, but generally requires many design points, thus is less suitable for quick improvement studies typical for computational models that require many hours of CPU time.

In Part 2 (presented here) we focus on gradient based approach (MGP) that is generally capable to produce design improvements in just a few design evaluations.

In our study we used modeFrontier as optimization enabling (Scheduler) and statistical data post-processing tool and eArtius multi-objective optimization methods plug-in tool to guide continuous process of selecting better input variables to satisfy multiple design objectives.

This process follows “fire and forget” principle. Gradient based computer thinking combines advantages of precise analytics with human like decision making (selecting roads that lead to improvement, avoiding weak links, pursuing best options, connecting dots).
Fundamental Design Optimization Issues

Study Motivation

The biggest issues of current design optimization algorithms:

- Low computational efficiency
- Low scalability

Reasons:

- Absence of efficient algorithms for estimating gradients
- Curse of Dimensionality Phenomenon
- Searching for optimal solutions in the entire design space while the search space can be reduced
- Approximating the entire Pareto frontier while the user only needs a small part of it

Consequences:

- Artificially reduced task dimensions by arbitrarily excluding design variables
- Overhead in use of global response surfaces and sensitivity analysis
- Have to rely only on use of brute-force methods such as algorithms’ parallelization
**Estimation of Gradients Issues**

Estimation of gradients by the Finite Differences Method (FDM) is resource consuming:

- FDM is performed on each step
- FDM requires $N+1$ model evaluations to estimate a gradient ($N$—the number of design variables)

**Consequences:**
- Task dimension is limited by 5-10 for expensive simulation models
- Development of efficient gradient based techniques with FDM is impossible

Also, gradient based optimization algorithms with FDM cannot be applied to noisy simulation models

eArtius has developed DDRSM method (patent pending) of gradient estimation which overcomes the issues:

- Spends 0-7 model evaluations to estimate gradients
- Equally efficient for any task dimension up to 5,000 design variables
- Not sensitive to noise in optimized models
Example of uniformly distributed points:
- Unit interval—0.01 distance between points—100 points
- 10-dimensional unit hypercube, a lattice with 0.01 between neighboring points—$10^{20}$ sample points (Richard Bellman)

Adding extra dimensions to the design space requires **an exponential increase** in the number of:

- *Sample points* necessary to build an adequate global surrogate model
- *Pareto optimal points* to maintain the same distance between neighboring optimal points in the design space

For Response Surface Methods:
- eArtius DDRSM spends just 0–7 points for local approximations—no global approximations

For Approximation of the Entire Pareto Frontier:
- eArtius performs directed search on Pareto Frontier—no global approximation of the entire Pareto frontier
Current multi-objective optimization algorithms are required to uniformly cover the entire Pareto frontier.

**Curse of dimensionality:** The increase in the number of design variables causes the distance between neighboring points in the design space to be increased exponentially.

Minimize \( f_1 = x_1 \)

Minimize \( f_2 = 1 + x_2^2 - x_1 - 0.1 \cdot \sin(3\pi \cdot x_1) \)

\[ 0 \leq x_1 \leq 1; \quad -2 \leq x_2 \leq 2 \]

1D – 89 Pareto points

Minimize \( f_1 = 3 - (1 + x_3) \cdot \cos(x_1 \cdot \pi / 2) \cdot \cos(x_2 \cdot \pi / 2) \)

Minimize \( f_2 = 3 - (1 + x_3) \cdot \cos(x_1 \cdot \pi / 2) \cdot \sin(x_2 \cdot \pi / 2) \)

\[ 0 \leq x_1 \leq 0.65; \quad 0 \leq x_2 \leq 1; \quad 0.5 \leq x_3 \leq 1 \]

2D – 2225 Pareto points

\( 2225/89 = 25 \) times more!

ND \( \rightarrow \) \( 10^N \) Pareto points
Search in the Entire Design Space

Minimize \( f_1 = x_i \)

Minimize \( f_2 = 1 + x_2^2 - x_i - 0.1 \cdot \sin(6\pi \cdot x_i) \)

0 ≤ \( x_i ≤ 1 \); \( -2 ≤ x_2 ≤ 2 \)

Monte Carlo method:
258 Pareto optimal points (3%) out of 8192 model evaluations

HMGE method:
89 Pareto optimal points (35%) out of 251 model evaluations

Pareto frontier is a straight line \( x_2 = 0 \) in the design space

Why do we need to search in the entire design space?
The search along the line \( x_2 = 0 \) is also possible
Search in the Entire Design Space (continuation)

Minimize \( f_1 = 3 - (1 + x_3) \cdot \cos(x_1 \cdot \pi / 2) \cdot \cos(x_2 \cdot \pi / 2) \)

Minimize \( f_2 = 3 - (1 + x_3) \cdot \cos(x_1 \cdot \pi / 2) \cdot \sin(x_2 \cdot \pi / 2) \)

Minimize \( f_3 = 3 - (1 + x_3) \cdot \cos(x_1 \cdot \pi / 2) \cdot \sin(x_1 \cdot \pi / 2) \)

\( 0 \leq x_i \leq 0.65 \)

\( 0 \leq x_2 \leq 1 \)

\( 0.5 \leq x_3 \leq 1 \)

2225 Pareto optimal points out of 3500 model evaluations

Pareto frontier is located on the flat \( x_3 = 1 \) in the design space

Why do we need to search in the entire design space?
The search on the plane \( x_3 = 1 \) is also possible
Multi-Gradient Pathfinder (MGP) Method

- On the first half-step MGP improves preferable objective ($F_2$)—green arrows

- On the second half-step MGP improves ALL objectives—blue arrows—to maintain a short distance to Pareto frontier

- Then MGP starts the next step from the newly found Pareto optimal point
Directed Optimization on Pareto Frontier

MGP started optimization three times from the same start point \( \{x1=1; x2=1; x3=1\} \), but with different preferable objectives.

**Green trajectory:**
- Min \( f1 \)
- Min \( f2 \)
- Min+ \( f3 \)

**Red trajectory:**
- Min+ \( f1 \);
- Min \( f2 \)
- Min \( f3 \)

**Blue trajectory:**
- Min+ \( f1 \)
- Min \( f2 \)
- Min+ \( f3 \)

Light-green small markers visualize entire Pareto frontier, which is located on the plane \( x3=1 \) in the design space.
Searching the Entire Design Space is Not Productive!

ZDT2 Benchmark Problem: multiple Pareto frontiers

Minimize $F_1 = x_1$

Minimize $F_2 = g \cdot \left[ 1 - \left( \frac{F_1}{g} \right)^2 \right]$

$g = \left[ 1 + \frac{9}{n-1} \sum_{i=2}^{n} x_i \right]$

$0 \leq x_i \leq 1, i = 1,..n$

$n = 30$

MGP—18 global Pareto optimal points out of 38 model evaluations

Pointer—5 optimal points out of 1500 evaluations

NSGA-II & AMGA—FAILED to find a single Pareto optimal point after 1500 evaluations!
Searching the Entire Design Space is Not Productive!

\[ g = 1 + 10 \cdot (n - 1) + (x_2^2 + x_3^2 + \ldots + x_n^2) - 10 \cdot [\cos(4\pi x_2) + \cos(4\pi x_3) + \ldots + \cos(4\pi x_n)], n = 10 \]

\[ h = 1 - \sqrt{F_1 / g} - (F_2 / g) \cdot \sin(10\pi F_3); \quad [X] \in [0;1] \]

Minimize \( F_1 = x_i \)

Minimize \( F_2 = g \cdot h \)

MGP spent 185 evaluations, and found exact solutions

Pointer, NSGA-II, AMGA spent 2000 evaluations each, and failed
OPTIMIZATION RESULTS
MGP – Start from Arbitrary (Bad) design
(TempDiff+, SubMax400+)

Objective 1

TempDiff, deg. C

Objective 2 (secondary consideration)

Improvement direction

Design Table
MGP (TempDiff+, SubMax400+)

Objective 1 (main consideration)

Temp Diff, deg. C

Objective 2 (second consideration)

SubMax400, deg. C

Start from Arbitrary (Bad) Design

Design ID #
MGP – Start from Good Point (Obj1)

Equipment Products Division

Start from Good Design (by Obj1)

Improvement

Obj1 got worse, Obj2 improved

SubMax400, deg. C

Temp Diff, deg. C

Objective 1 (main consideration)

Objective 2 (secondary consideration)
MGP: Start from Small DOE (12 designs)

- **Temp Diff, deg. C**
  - Obj1: 19.7
  - Obj2: 40

- **DOE Points (designs)**
  - Design ID #13, first MGP point (design)
  - Converged improvement

- **Pareto**
  - Obj1: 19.7
  - Obj2: 40

- **Design ID #**
  - 19.7
  - 5.8
  - 0.6
MGP: First Design after DOE (detail of previous slide)

Objective 1 (main consideration) ~with best DOE improvement

Objective 2 (secondary consideration) improvement

DOE Points (designs)

Equipment Products Division
MGP: Start from Several Best Points

### Objective 1

**Temp Diff, deg. C**

<table>
<thead>
<tr>
<th>Obj1</th>
<th>Temp Diff, deg. C</th>
</tr>
</thead>
<tbody>
<tr>
<td>9.97</td>
<td>4.87</td>
</tr>
</tbody>
</table>

**Improvement**

- Best Initial Points
- Improvement
“Sequence Jumping” DOE for MGP

#1, #2,#3 – best points on each step for objective marked “+” (preferred objective)

Multi-Step Fast Start
Multi-Step Fast Start MGP

Step 1 with Initial Tolerance
Steps continue as long as improvement is reached within short number of designs

Quick Search for Good Starting Point: multi-step “short” MGP instead of initial DOE.
Advantage: multi-step approach has solution feedback, DOE does not.
Multi-Step Fast Start MGP (last step)

Step 3: Tolerance Reduced

Results got worse, No need to continue multi-step improvement any more

TempDiff, deg. C

Objective 1 (main consideration)

5.9

6
Computer vs. Human

In head-to-head competitions best “human guided” (case-by-case) studies resulted in system design with ±10-20 deg.C thermal uniformity and took several weeks to accomplish, while FAST MGP method based computer optimization approach allowed to quickly yield design solutions capable of reaching ± 3 deg. C. It took only 8-20 design evaluations for CPU to “independently” accomplish this task.

Such an approach will not allow to uniformly cover entire design space, but will work for engineers who need to find quick improvements for their designs and work with large computational models that take many CPU hours to solve (i.e. hundreds of design evaluations are not an option).

We can conclude that “Optimization Equals Innovation”!
Conclusion: Optimization = Innovation

modeFrontier  ANSYS  eArtius

WorkBench

Equipment Products Division
Acknowledgement

Authors are thankful to ESTECO engineers for developing eArtius MGP modeFrontier plug-in; to Alberto Bassanese (ESTECO) for introducing and helping with modeFrontier; to ANSYS Distributor in Bay Area Ozen Engineering www.ozeninc.com (Kaan Diviringi, Chris Cowan and Metin Ozen) for help and dedicated support with ANSYS Workbench model development and integration with modeFrontier.

Part 1 of this presentation is devoted to thermal optimization when CPU time budget is more flexible to allow many computational design evaluations. It was presented at Santa Clara Aug. 2011 ANSYS Users Conference http://www.slideshare.net/vvk0/optimization-intevac-aug23-7f
SUPPLEMENTS
eArtius – new word in multi-objective optimization capabilities

- Multi-Gradient Explorer (MGE) algorithm uses a conventional approach for optimization practice. It starts from an initial point, and iterates toward Pareto frontier until a Pareto optimal point is found. Then it takes another initial point, iterates again, and so on;

- Multi-Gradient Pathfinder (MGP) algorithm uses Pareto frontier as a search space for multi-objective optimization, and performs in this way directed optimization on Pareto frontier. Directed optimization on Pareto frontier means that a search algorithm steps along Pareto frontier from a given initial Pareto optimal point towards a desired Pareto optimal point;

- Hybrid Multi-Gradient Explorer (HMGE) combines a GA framework with unique eArtius approach to estimate gradients. In this way HMGE combines strengths and avoids weaknesses of two major optimization approaches: gradient-based techniques and Genetic Algorithms (GAs);

- Hybrid Multi-Gradient Pathfinder (HMGp) algorithm is a new multi-objective optimization algorithm which combines elements of MGP (Multi-Gradient Pathfinder) algorithm with elements of genetic algorithms (GA).
Thermal System Optimization Task Formulation

Objectives:

- Minimize $+ \quad$ TempDiff
- Minimize $+ \quad$ SubMax400
- Minimize $\quad$ SubMin400

Constraints:

- Lower $< 35$
- $T_{diff12} < 35$
- Upper $< 45$

Design Variables:

- $Height \in [-0.055, 0.035]$
- $SubMinus \in [-0.015, 0.020]$
- $SubPlus \in [-0.015, 0.030]$
- $T2 \in [735, 940]$
- $T2 \in [735, 940]$

Need to carefully consider

- Minimize+ – preferable objectives
- Minimize – regular objective

278 feasible designs of 317 evaluations

18 Pareto+ designs of 35 Pareto optimal designs
“Fire And Forget” Solution Process - HMGE

Temperature Uniformity

First Wave

Second Wave

Touchdown

Design ID (#)

DOE

Equipment Products Division