Managing the Use of Surrogates in General Nonlinear Optimization and MDO

Natalia M. Alexandrov

Multidisciplinary Optimization Branch

NASA Langley Research Center

Hampton, Virginia

Approximation Management Framework

**High-Fidelity,
Expensive
Model**

**Lower-Fidelity,
Cheaper
Model**

**Computing
Improved
Design**

Linking Conceptual Design & Preliminary Design

Approximation Model Management

Conceptual Design

System-level Low-Fidelity Methods

Performance Analysis

Disciplinary High-Fidelity Methods

Meets Requirements?

1-st Order AMF

- **Background:**
  - **Variable-fidelity models (or approximations, or surrogates) used in design with heuristic model selection for a long time.**
  - **Provably convergent 1-st order AMF for unconstrained optimization (Alexandrov et al. '97)—alternate the use of variable-fidelity models *rigorously* to attain an optimum of the high-fidelity problem.**
  - **This work—provably convergent 1-st order AMF for constrained problems.**
  - **Testing done on "toy" problems. Computational experiments of increasingly realistic problems in progress.**
- **Objectives of the talk:**
  - **Outline an AMF for constrained optimization**
  - **State conditions for reaching a solution of the high-fidelity problem**

Underlying framework: MAESTRO class for NLP

(Alexandrov ’93, Alexandrov and Dennis ’94)

- **MAESTRO—a class of multilevel algorithms for large-scale nonlinear programming problems, motivated by MDO problems with arbitrary couplings;**
- **Approach—trust-region methodology (an adaptive move-limit strategy);**
- **Control optimization by varying the size of the trust region;**
- **Not necessary to change models to obtain convergence;**
- **When other models are available (e.g., varieties of structural or aerodynamic models) MAESTRO-based AMF provides the user with guidance on changing the nature of the models.**

MAESTRO Class for Nonlinear Optimization

Decomposition:

$$\text{minimize } f(x) \quad \text{subject to } h(x) = 0$$

(M is the number of fully coupled blocks or disciplines;
x R^n - design variables)

$$\text{minimize } f(x) \quad \text{subject to } h_1(x) = 0,\ h_2(x) = 0,\ \ldots,\ h_M(x) = 0$$

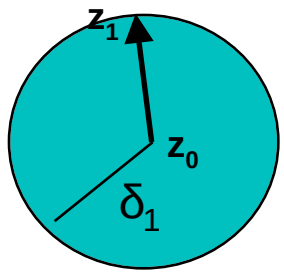
e.g.

Example:

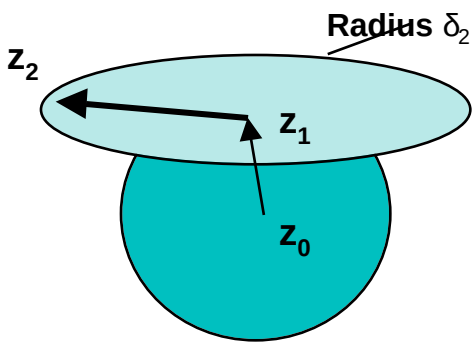
minimize -L / D
subject to AERO
STRUCT
. . .
other disciplines

Computing One Step: (M=2, n=3)

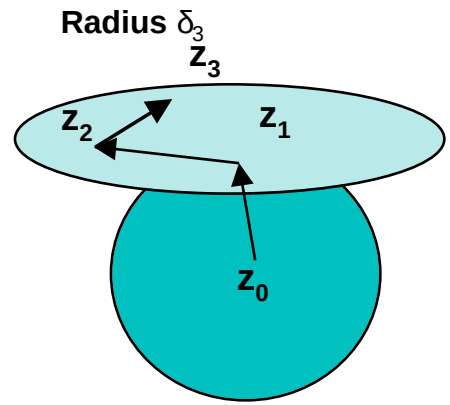
$z_0 = x_c$



min model of $||h_1(y)||^2$ in δ_1



min restricted model of $||h_2(y)||^2$ in δ_2



min restricted model of f(y) in δ_3

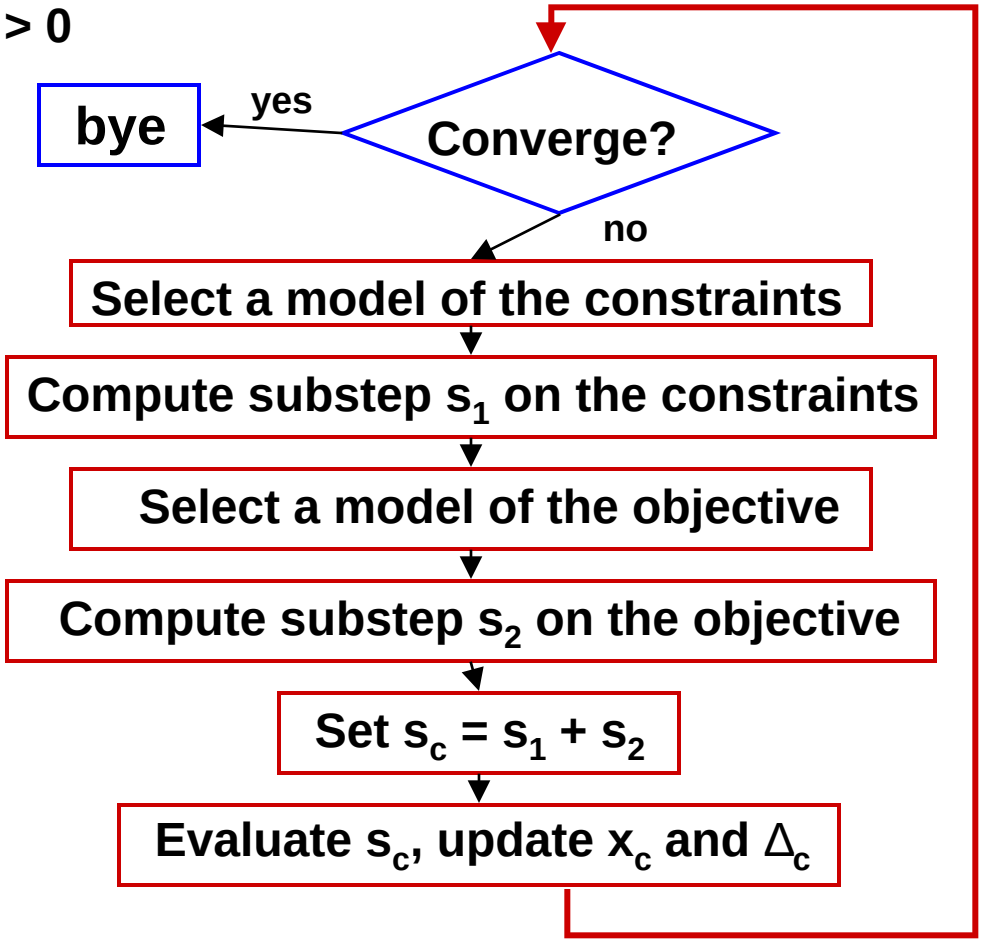
$x_+ = z_3$

Bilevel AMF

$$\text{minimize} \quad f(x)$$
$$\text{subject to } h(x) = 0$$

- AMF for a single block of constraints is the following procedure:

Given x_c $\quad \mathbb{R}^n$, $\Delta_c > 0$



Bilevel AMF: model of constraints and substep

- **Let x_c be the current approximation to a solution and Δ_c be the trust-region radius. Set $z_0 = x_c$.**

- **At x_c, select a model of the constraints a_c^h that satisfies:**

$$\begin{aligned} a_c^h(x_c) &= h(x_c) \\ \nabla a_c^h(x_c) &= \nabla h(x_c). \end{aligned}$$

- **Find a substep s_1 that approximately solves:**

$$\begin{aligned} &\textbf{minimize} && a_c^h(z_0 + s) \\ &\textbf{subject to} && \| s \| \leq \theta \Delta_c, \;\; \theta \in (0.5, 0.6). \end{aligned}$$

- **set $z_1 = z_0 + s_1$.**

Bilevel AMF: model of objective and substep

- **Select a model a_c^f of the objective function that satisfies:**

$$\begin{aligned} a_c^f(x_c + s_1) &= f(x_c + s_1) \\ \nabla a_c^f(x_c + s_1) &= \nabla f(x_c + s_1). \end{aligned}$$

- **Find s_2 that approximately solves:**

$$\begin{aligned} &\textbf{minimize} && a_c^f(z_1 + s) \\ &\textbf{subject to} && \| s \| \leq \sqrt{\Delta_c^2 - \|s_1\|^2}. \end{aligned}$$

- **Set $s_c = s_1 + s_2$.**
- **An extension of MAESTRO—the Gauss-Newton model of the constraints and the quadratic model of the objective replaced by general models that satisfy first-order consistency conditions.**

Bilevel AMF: evaluating the step

- **Merit function:**

$$\mathcal{P}(x;\rho) \equiv f(x)+\rho \parallel h(x) \parallel^2 \quad \textbf{or} \quad \mathcal{L}(x,\lambda;\rho) \equiv f(x)+\lambda^T h(x)+\rho \parallel h(x) \parallel^2$$

- **Penalty parameter ρ is *not* used in computing the step, but only in the merit function.**

- **Penalty parameter is updated in a rigorous manner after the trial step is computed and before it is evaluated.**

- **Define the actual reduction in the merit function as**

$$ared_c \equiv \mathcal{P}(x_c;\rho_c) - \mathcal{P}(x_c+s_c;\rho_c)$$

and the predicted reduction in the merit function as

$$pred_c \equiv [f(x_c) - a_c^f(x_c+s_c)] + \rho_c[\parallel h(x_c) \parallel^2 - \parallel a_c^h(x_c+s_c) \parallel^2]$$

Bilevel AMF: evaluating the step / updating

- **Updating the iterate:**

$$x_+ = \begin{cases} x_c + s_c & \textbf{if } \mathcal{P}(x_c + s_c; \rho_c) < \mathcal{P}(x_c; \rho_c) \\ x_c & \textbf{otherwise.} \end{cases}$$

- **Define**

$$r = \frac{ared_c}{pred_c}$$

and update the radius Δ_c based on the predictive properties of the model:

  - **Reduce Δ_c if r is small (typically $< 10^{-5}$);**
  - **Increase Δ_c if r is large (close to 1);**
  - **Otherwise do not change Δ_c**

Details of computing the trial step

- **To inherit the convergence properties from the MAESTRO class, the substeps s_1 and s_2 computed in AMF have to satisfy:**
  - **A sufficient decrease condition:Each s_k is required to predict at least a fraction of the improvement in model k that is predicted by steepest descent step in the region of interest.**
  - **A boundedness condition:Each substep s_k, $k = 1, \ldots, M$ is required to satisfy**

$$\| s_k \| \leq \mathcal{K} \| h_k(y_{k-1}) \|$$

for some constant $\mathcal{K}$ independent of the iterates.

Both are easily satisfied. The following procedure will suffice. (v_j and p_j used as local variables in both procedures.)

One choice of trial step—constraints (objective analogous)

Given $z_0 = x_c, \Delta_c > 0$**, set** $y_0 = z_0$**,** $\delta_0 < \Delta_c$**,** $v_0 = 0$**.**

For $j = 0, 1, \ldots$**, while remaining within** Δ_c **do** $\{$

Construct $q_j^h(y_j + p) \equiv \| a_c^h(y_j) + \nabla a_c^h(y_j)^T p \|^2$**;**

Find an approximate solution p_j **to**

$$\begin{aligned} &\textbf{minimize} && q_j^h(y_j + p) \\ &\textbf{subject to} && \| p \| \leq \delta_j \\ & && \| y_j + p \| \leq \Delta_c \end{aligned}$$

that satisfies FCD for $\| a_c^h \|^2$ **from** y_j**.**

Compute $r = \dfrac{\| a_c^h(y_j) \|^2 - \| a_c^h(y_j + p_j) \|^2}{\| a_c^h(y_j) \|^2 - q_j^h(y_j + p_j)}$**.**

Evaluate p_j **and update** y_j **and** δ_j**;**

Set $v_{j+1} = v_j + (y_{j+1} - y_j)$**.**

$\}$

Set $s_1 = v_j$**.**

Properties of the trial steps

- **This choice of subproblems for computing s_1 and s_2 results in unconstrained trust-region subproblems.**
- **By results of the unconstrained AMF, the substeps satisfy the sufficient decrease conditions.**
- **Note: in the subproblems, a more stringent rule for updating y_j is required than for updating x_c.**

Multilevel AMF

- **Bilevel AMF—easily extended to multilevel case. Let $h(x)$ be partitioned into M blocks, $h_1(x), \ldots, h_M(x)$.**
- **Let $a_1^h, \ldots, a_M^h$ denote the lower-fidelity models of $h_1(x), \ldots, h_M(x)$, respectively.**
- **Given $x_c \in \Re^n$, $\delta_k > 0, k = 1, \ldots, M+1$, $\rho_k > 0, k = 1, \ldots, M$, the multilevel AMF can be stated as follows:**
  - **Apply the management procedure to consecutive reduced constraint blocks.**
  - **Apply the management procedure to the model of the objective function.**
- **See paper for details.**

Convergence properties

- **Theoretical:**
  - **MAESTRO assumptions: conventional (smoothness and boundedness, full rank for the gradients of constraints and their models), sufficient decrease and boundedness for the substeps.**
  - **AMF assumptions: consistency conditions and uniform boundedness of the Hessian approximations.**
  - **Result: first-order convergence to a critical point of the high-fidelity problem.**
- **Practical:**
  - **Enforce compatibility conditions**
  - **Actual performance will depend on the predictive properties of the model.**

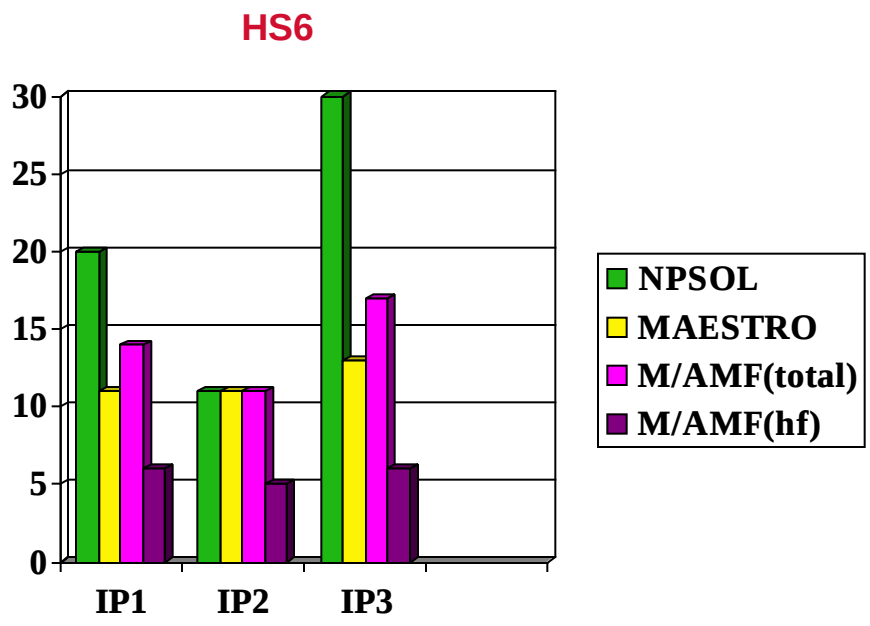
Some computational considerations

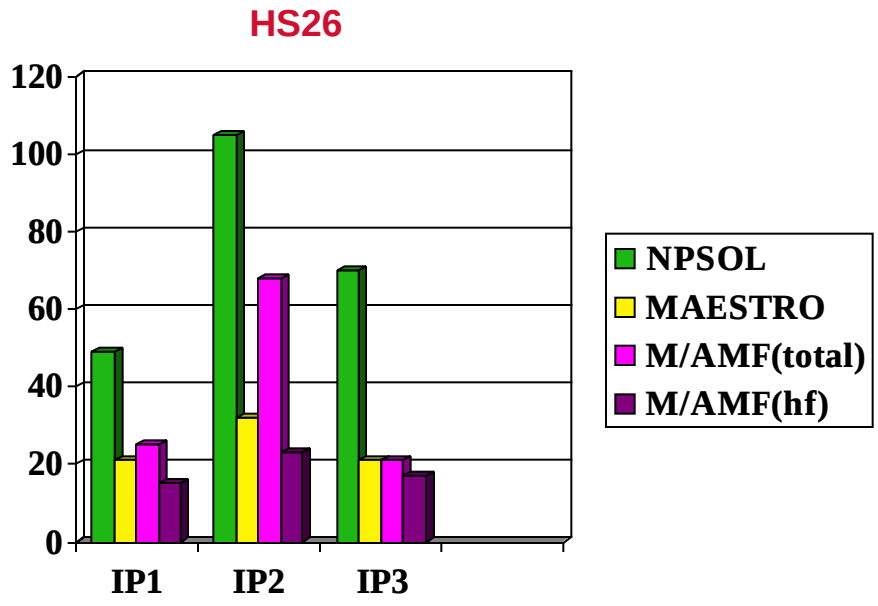
- **Consistency conditions:**
  - **Enforced only at "major" iterates.**
  - **Can be relaxed, but are easily enforced (Chang et al. '93):**
    Given $f_{hi}(x)$ **and** $f_{lo}(x)$**, define** $\beta(x) \equiv \frac{f_{hi}(x)}{f_{lo}(x)}$**.**
    Given x_c**, build** $\beta_c(x) = \beta(x_c) + \nabla\beta(x_c)^T(x - x_c)$**.**
    Then $a_c(x) = \beta_c(x) f_{lo}(x)$ **satisfies the consistency conditions.**
- **Inequality constraints:**
  - **Squared slacks.**
  - **An active set strategy (Alexandrov and El-Alem, '98).**

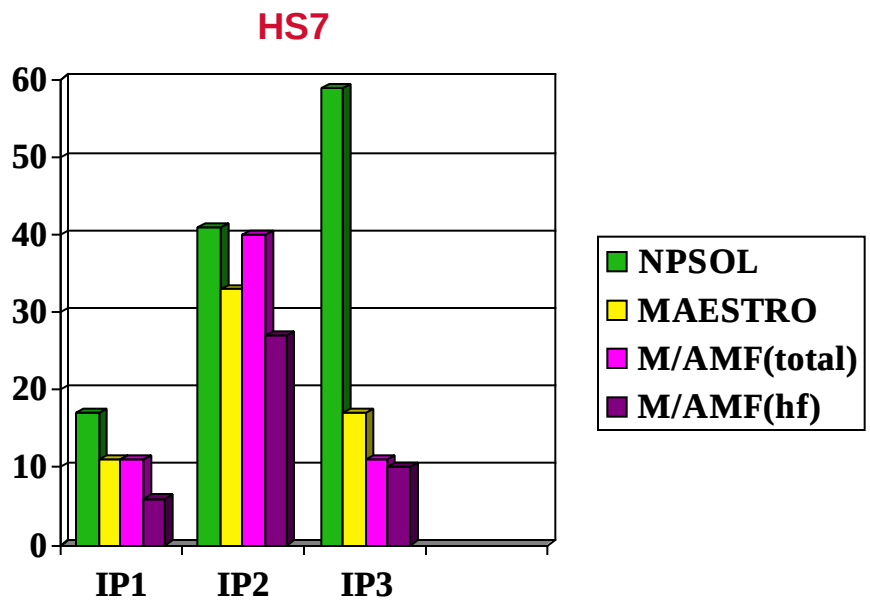
Preliminary numerical results

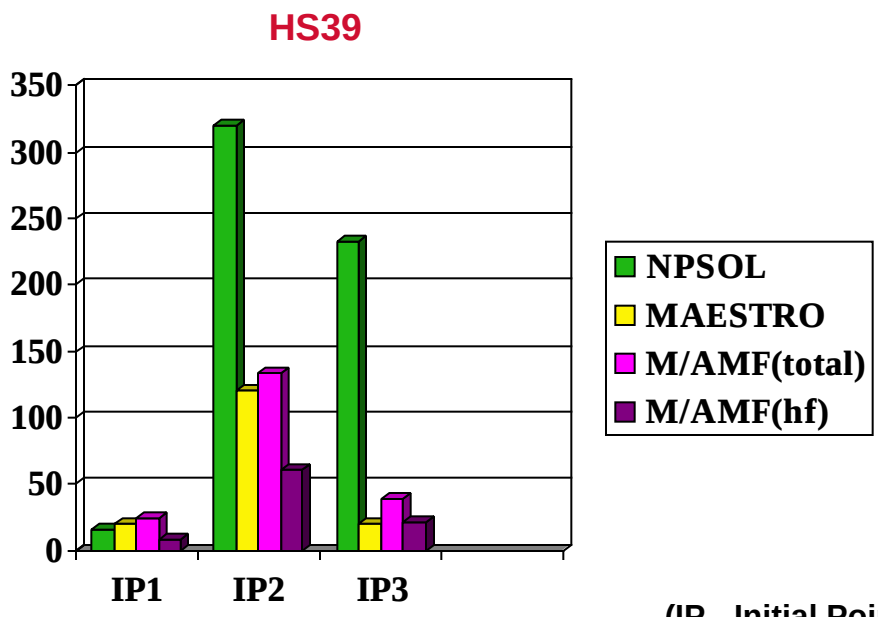
- **The algorithm has undergone initial testing on several simple problems (Hock and Schittkowski and the MDO Test Suite).**
- **The problems were initially solved with**
  - **NPSOL (Gill et al.)—SQP with line searches;**
  - **A research implementation of MAESTRO without approximation management;**
  - **A research implementation of the MAESTRO-based AMF approach. "High-fidelity" models—function evaluations computed to machine precision. "Lower-fidelity" models—evaluations with abbreviated precision and added noise (approx., two to four digits of accuracy after the decimal point). The total number of function evaluations and the number of "higher-fidelity" evaluations is reported.**

Number of Function Evaluations



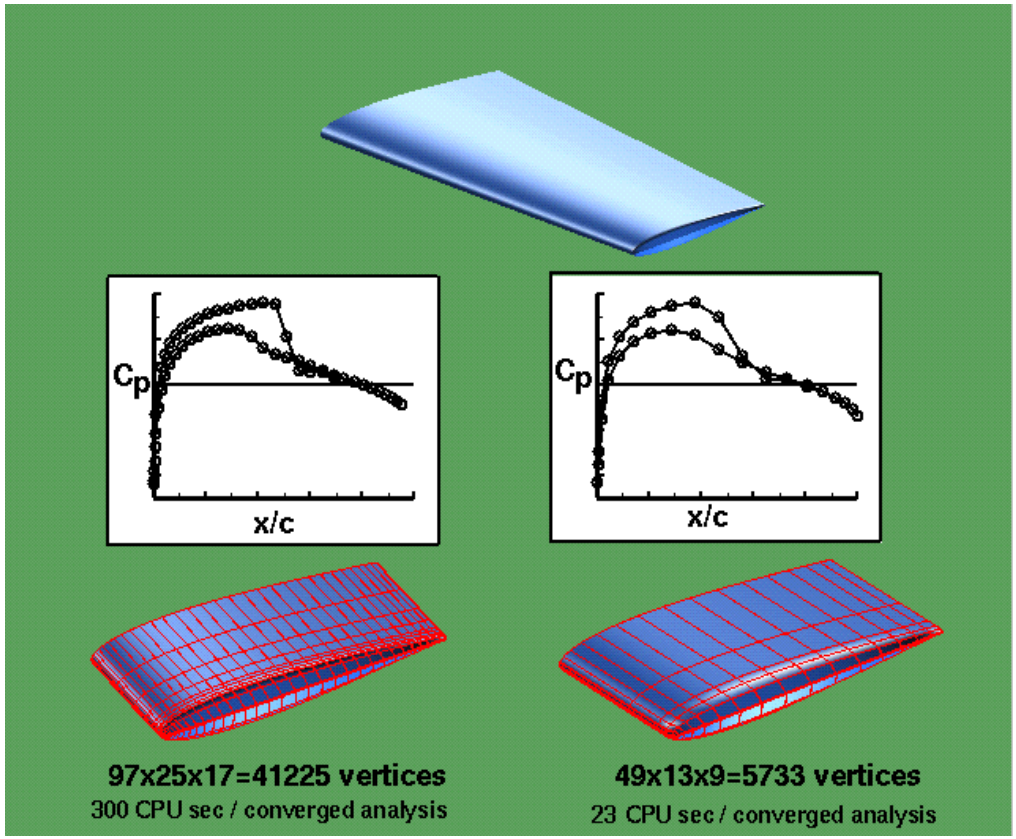






(IP - Initial Point)

Current Demonstration Problem: Aerodynamic Wing Design (with C. Gumbert)



- Euler, 3-D
- Current design variables:
  tip cord and tip trailing setback
- Objective: L/D
- General nonlinear constraints on
  - $C_L S$ (lift coefficient times wing area),
  - C_M (pitching moment coefficient),
  - C_l (rolling moment),
  - C_p (pressure coefficient)

Variable-fidelity models are provided by analysis on a variety of grids:

49 x 13 x 9
97 x 25 x 17
161 x 41 x 25
193 x 49 x 33

CPU time per function evaluation is linear in the number of grid points: order of magnitude increase in execution time between adjacent models

Status of the 1-st order AMF development

- **Introduced a globally convergent MAESTRO-based AMF for constrained optimization;**
- **Empirical verification:**
  - **Demonstrated on “toy” problems.**
  - **Demonstration on increasingly realistic single discipline problems in progress.**
  - **Definition of a multidisciplinary demonstration problem in progress.**
- **Other work includes research on modeling and algorithmic alternatives.**
- **The results give reason for cautious optimism. Considerable amount of experimentation with physics-based models will be required.**
- **The verdict should be out by the next MA&O Symposium**