First-Order Frameworks for Managing Models in Engineering Optimization

Natalia M. Alexandrov

Multidisciplinary Optimization Branch

NASA Langley Research Center

Hampton, Virginia

Robert Michael Lewis

College of William & Mary

Williamsburg, Virginia

Slides accessible from http://fmad-www.larc.nasa.gov/mdob/MDOB

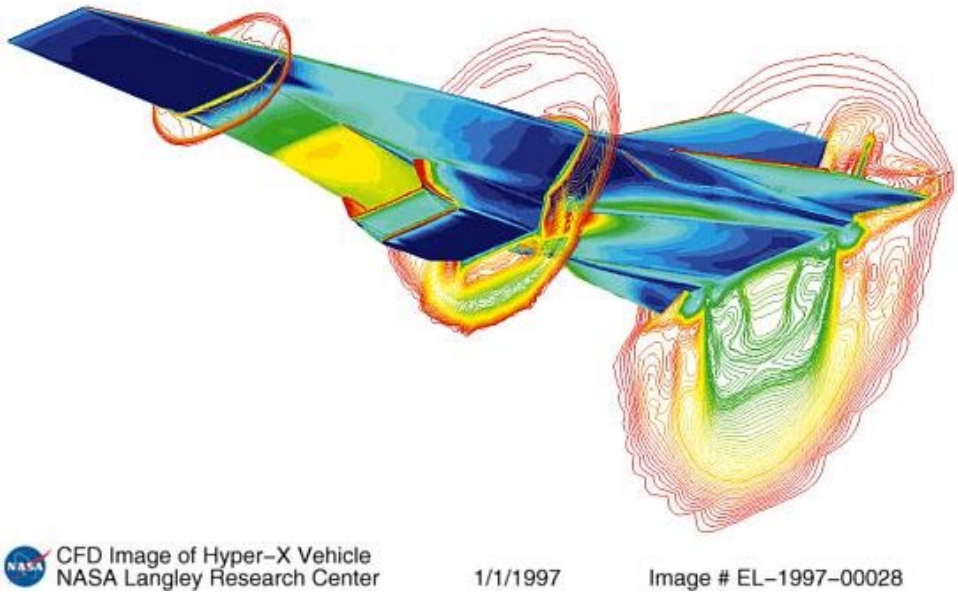
Outline

- **Problem and motivation**
- **First-order AMMO frameworks**
- **Consistency, convergence, performance**
- **Model origins**
- **Computational demonstrations**
- **Applicability**
- **Concluding remarks**

Motivation: optimization of engineering systems governed by differential equations

X-33

Hyper-X



- **Computational models grow in complexity to consume computing power available at any moment**
- **Objectives of an analyst (closest approximation to reality) can be at odds with objectives of a designer (improvement with respect to design criteria)**

Problem

- **The analysis problem:** Given x, solve a disciplinary analysis equation or a system of coupled analysis equations

$$A(x, u(x)) = 0$$

for u that describes the physical behavior of the system.

- **The design problem:** Solve

$$\begin{aligned} \underset{x}{\text{minimize}} \quad & f(x, u(x)) \\ \text{subject to} \quad & h(x, u(x)) = 0 \\ & g(x, u(x)) \geq 0 \\ & x_l \leq x \leq x_u, \end{aligned}$$

where, given x, $u(x)$ is determined from $A(x, u(x)) = 0$.

- The use of high-fidelity models (high-resolution or detailed physics) at every iteration of an optimization procedure can be prohibitively expensive

1st Order Approximation/Model Management Optimization (AMMO)

- **Objective**
  - **Enable design optimization with simulations by minimizing expense of using high-fidelity models in single-discipline optimization and MDO**
- **Idea**
  - **The designer can be satisfied with intermediate results of lower fidelity, provided they lead to improved designs; only the optimal design must be consistent with accurate, high-fidelity analysis**
- **Idea**
  - **For any optimization subproblem, replace the local Taylor series model with a lower-fidelity model (surrogate) that satisfies first-order consistency conditions with respect to a model designated as "truth"**

Related Work

- **Heuristic approximate techniques long in use in engineering optimization (see paper for references)**
- **Systematic first-order model management can be imposed on any gradient-based optimization algorithm**
  - **Alexandrov, 1996; Lewis, 1996 (unconstrained)**
  - **Alexandrov, Dennis, Lewis, Torczon, 1997 (unconstrained)**
  - **Alexandrov, 1998 (constrained and multilevel)**
  - **Alexandrov, Lewis, Gumbert, Green, Newman, 1999 (variable-resolution models with AMMO frameworks based on SQP, augmented Lagrangian and multilevel algorithms)**
  - **Lewis, Nash, 2000 (models based on multigrid approach)**
  - **Alexandrov, Nielsen, Lewis, Anderson, 2000 (variable-fidelity physics models with bound-constrained AMMO)**

Example: First-Order AMMO for Bound Constrained Optimization

Initialize x_c, Δ_c

Do until convergence:

Select model a_c **with** $a_c(x_c) = f(x_c); \nabla a_c(x_c) = \nabla f(x_c)$

Solve approximately for $s_c = x - x_c$**:**

$$\begin{aligned} \underset{s}{\text{minimize}} \quad & a_c(x_c + s) \\ \text{subject to} \quad & l \leq x \leq u \\ & \| s \|_\infty \leq \Delta_c \end{aligned}$$

Compute $\rho_c \equiv \frac{f(x_c) - f(x_c + s_c)}{f(x_c) - a_c(x_c + s_c)}$

Accept s_c **if** $f(x_c) > f(x_c + s_c)$**; otherwise reject**

Update Δ_c

End do

$\rho_c \leq 10^{-5}$	**Reduce** Δ_c
$10^{-5} < \rho_c \leq 0.6$	**Leave** Δ_c **unchanged**
$\rho_c > 0.6$	**Increase** Δ_c

Consistency, Convergence, Performance

- **First-order consistency conditions:**

$$a_c(x_c) = f(x_c)$$

$$\nabla a_c(x_c) = \nabla f(x_c)$$

- a_c **mimics the local behavior of a Taylor-series model around** x_c
- **Easily enforced via "β-correction" due to Chang, Haftka, Giles, Kao, 1993:**
  - **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 at** x_c

Consistency, Convergence, Performance, cont.

- **For convergence, need only a notion of two models, one arbitrarily designated “high fidelity” or “truth”, the other - “low fidelity”**
- **Convergence relies on the consistency conditions and standard assumptions for the convergence analysis of the underlying algorithm (assumption of strict first-order consistency can be relaxed)**
- **Practical efficiency**
  - **Problem/model dependent**
  - **Depends on the ability to transfer computational load onto low-fidelity computation, which...**
  - **Depends on the predictive quality of the low-fidelity models (surrogates)**
  - **In the worst case, AMMO is conventional optimization**

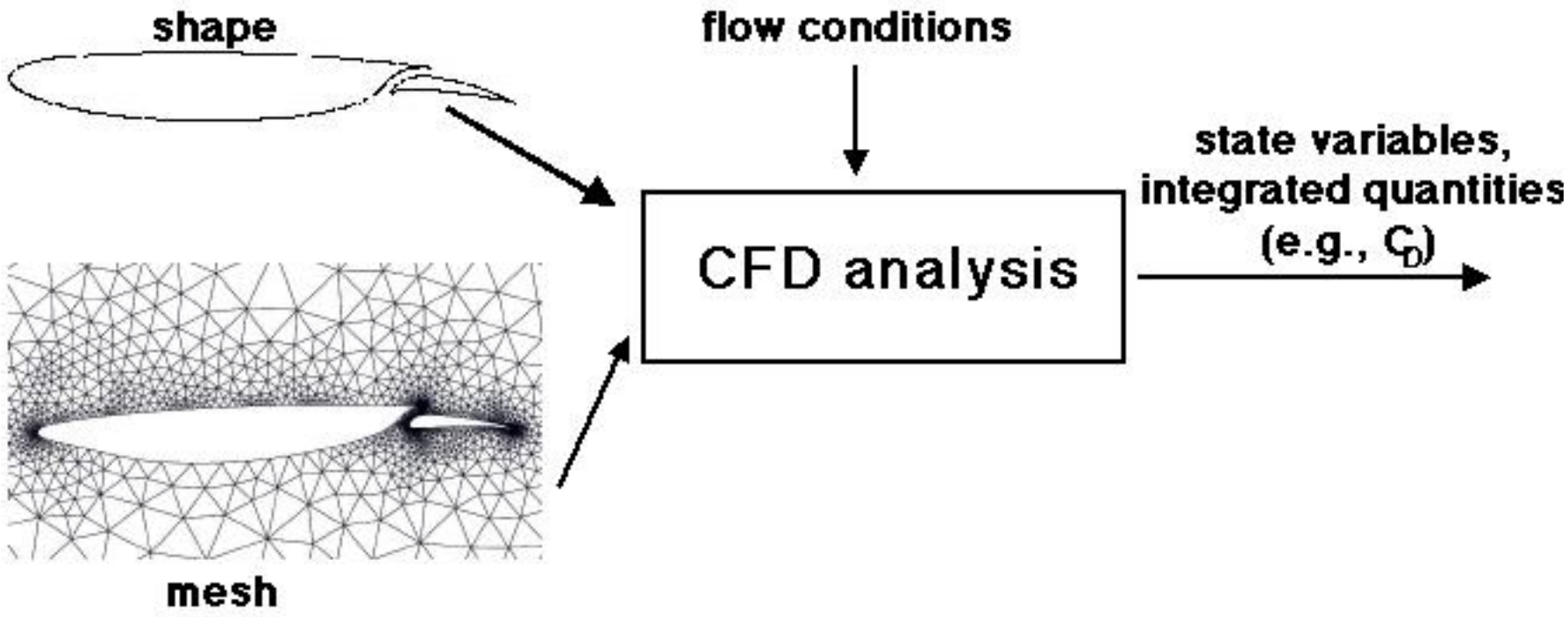
Examples of Variable-Fidelity Models for Use in AMMO

- **Data-fitting models (polynomial RS, splines, kriging)**
  - **Rely directly on hi-fi information; do not require derivatives; simple to construct; difficult to sample; "curse of dimensionality"**
- **Reduced-order models**
  - **Use reduced-order bases (constructed as a span of solutions and possibly derivatives at some points) to represent field variables at other points**
- **Variable-accuracy models**
  - **Converge analyses to a user-specified tolerance**
- **Variable-resolution models**
  - **Executing a single physical model on meshes of varying degree of refinement**
- **Variable-fidelity physics models**
  - **E.g., in aerodynamics, physical models range from inviscid, irrotational, incompressible flow to Navier-Stokes equations for nonlinear viscous flow**

Computational Demonstrations

- **Because of data-fitting model limitations, we have focused on models that are independent of the number of variables**
- **Independence wrt dimension is important: in preliminary design, problems of modest size number O(100) variables**
- **AMMO admits a wide variety of models and algorithms; demonstrations are aimed at accumulating realistic experience to validate the algorithmic performance**
- **Because we cannot predict *a priori* the relative descent characteristics of models, must include cases of favorable and unfavorable relationship between models**
- **Aerodynamic shape optimization is a good test problem: practically important, computationally intensive, comes in a variety of dimensions**

Demonstration Problems: Aerodynamic Optimization



minimize Integrated quantities, such as $-\frac{L}{D}$ $\left(\frac{\text{lift}}{\text{drag}}\right)$ or C_D (drag coefficient)

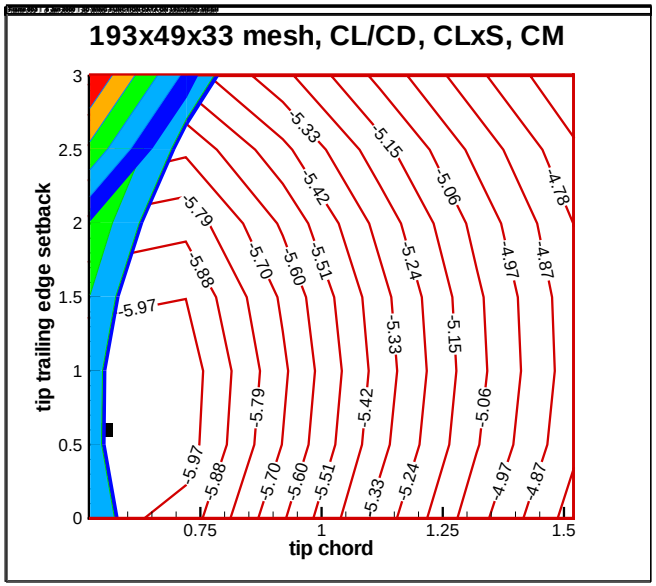
subject to constraints on, e.g., pitching and rolling moment coefficients, etc.

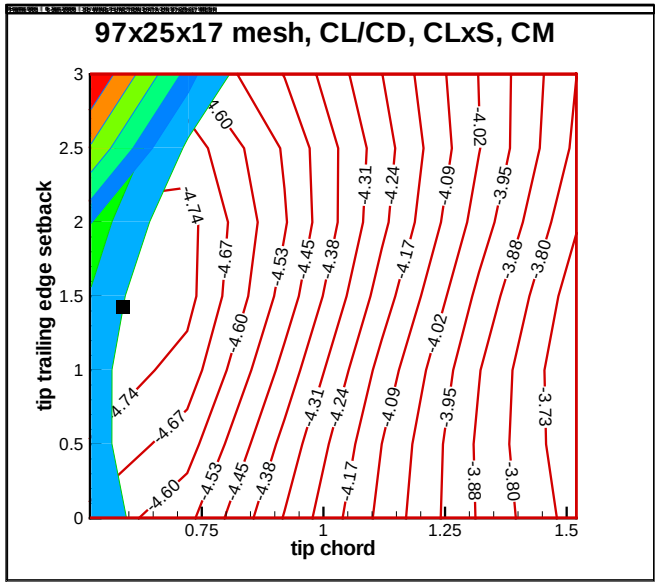
$x_l \leq x \leq x_u$

Managing Variable-Resolution Models: Optimization of a 3D Wing

(AIAA-2000-0841, Alexandrov, Lewis, Gumbert, Green, Newman)

- **Problem: minimize $-\frac{L}{D}$ s.t. nonlinear constraints**
- **Variable-resolution models: Euler equations on a variety of meshes**
- **Algorithms: based on SQP, augmented Lagrangian, a multilevel approach**





- **Threefold improvement in terms of hi-fi function and derivative evaluations across all AMMO algorithms tested**

Managing Variable-Fidelity Physics Models: Multi-Element Airfoil

(AIAA-2000-4886, Alexandrov, Nielsen, Lewis, Anderson)

- **A two-element airfoil designed to operate in a transonic regime — inclusion of viscous effects is very important**
- **Governing equations: time-dependent Reynolds-averaged Navier-Stokes**

$$A\frac{\partial Q}{\partial t} + \oint_{\partial\Omega} \vec{F_i} \cdot \hat{n} dl - \oint_{\partial\Omega} \vec{F_v} \cdot \hat{n} dl = 0,$$

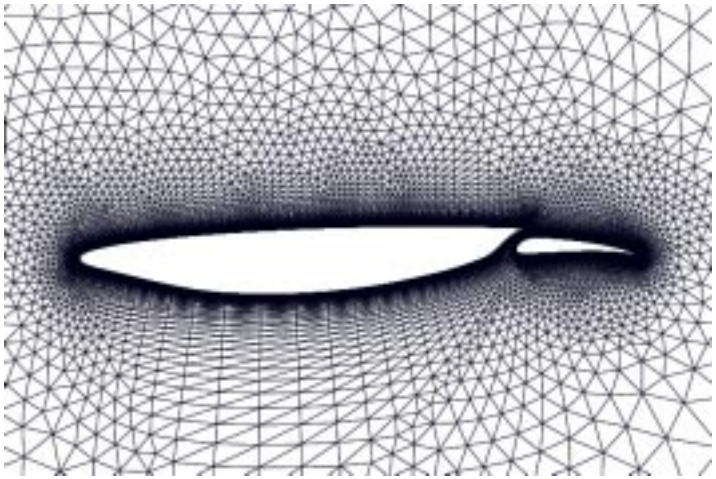
  where $\vec{F_i}$ and $\vec{F_v}$ are the inviscid and viscous fluxes, respectively

- **Flow solver (FUN2D) – unstructured mesh methodology (Anderson, 1994)**
- **Sensitivity derivatives – hand-coded adjoint approach (Anderson, 1997)**
- **Conditions:**
  - $M_\infty = 0.75$
  - $Re = 9 \times 10^6$
  - $\alpha = 1^\circ$ **(global angle of attack)**

Multi-Element Airfoil, cont.

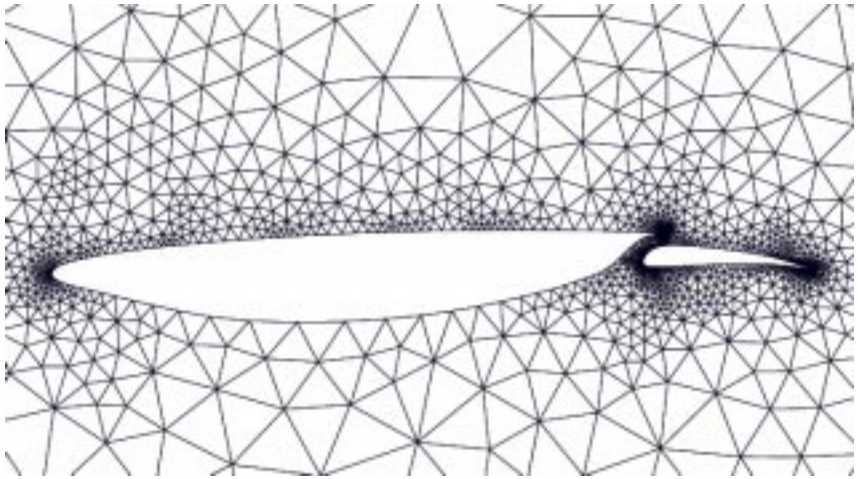
- **Hi-fi model – FUN2D analysis in RANS mode**
- **Lo-fi model – FUN2D analysis in Euler mode**
- **Computing on SGI Origin^{TM} 2000, 4 R1OK processors**

Viscous mesh:
10449 nodes and 20900 triangles



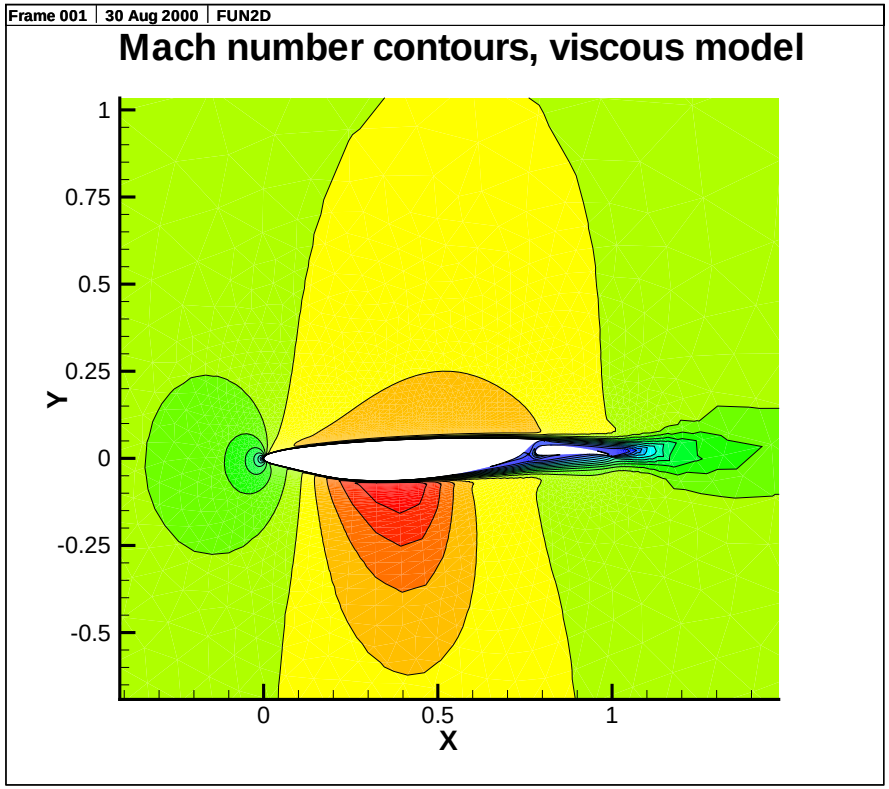
t/analysis $\approx$ 21 min
t/sensitivity $\approx$ 21 or 42 min

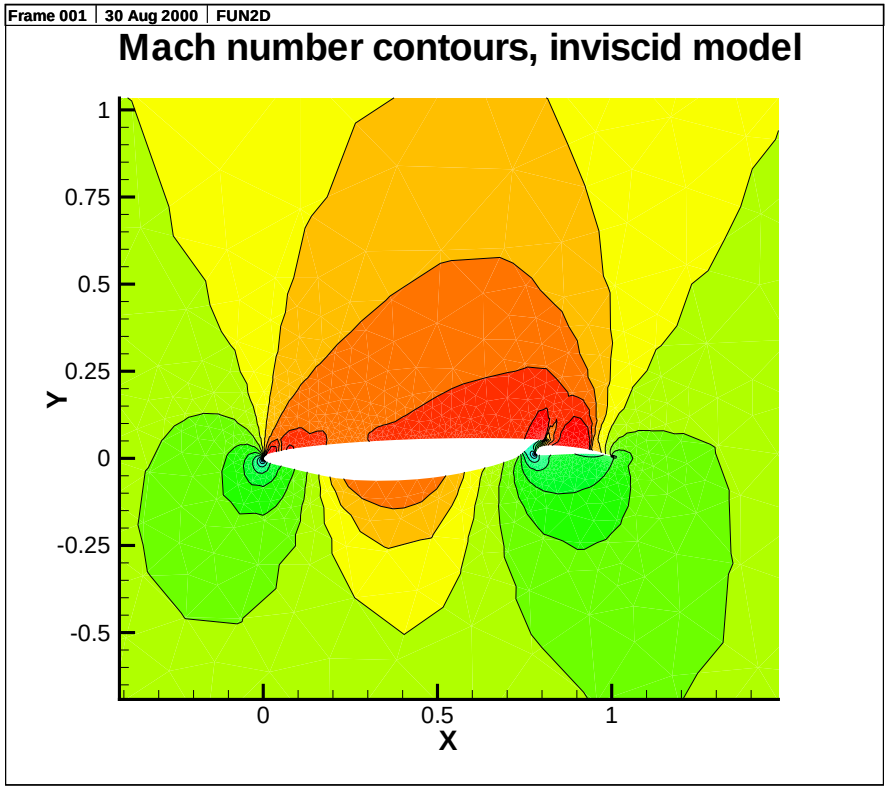
Inviscid mesh:
1947 nodes and 3896 triangles



t/analysis $\approx$ 23 sec
t/sensitivity $\approx$ 100 or 77 sec

Multi-Element Airfoil: Viscous Effects





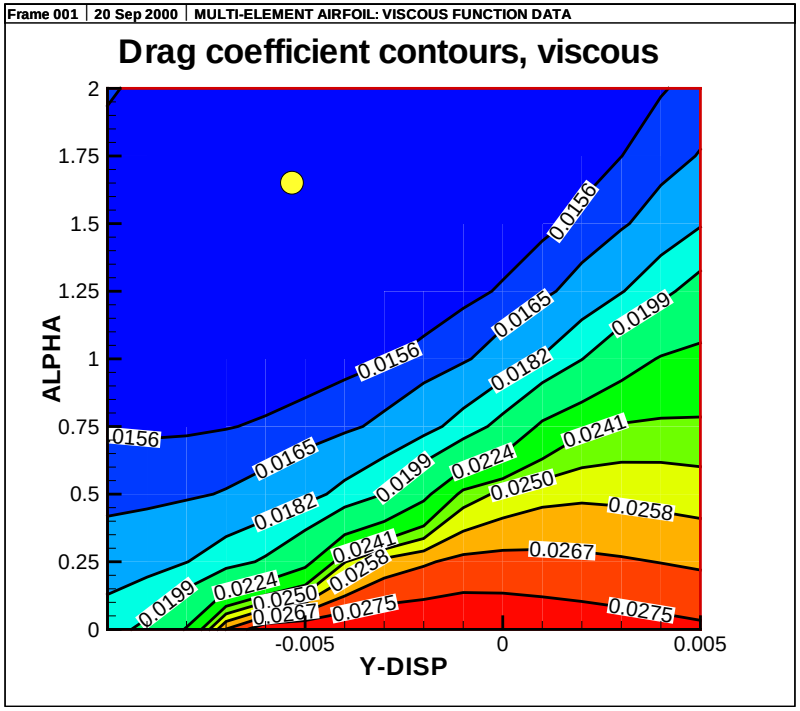
- **Boundary and shear layers are visible in the viscous case.**

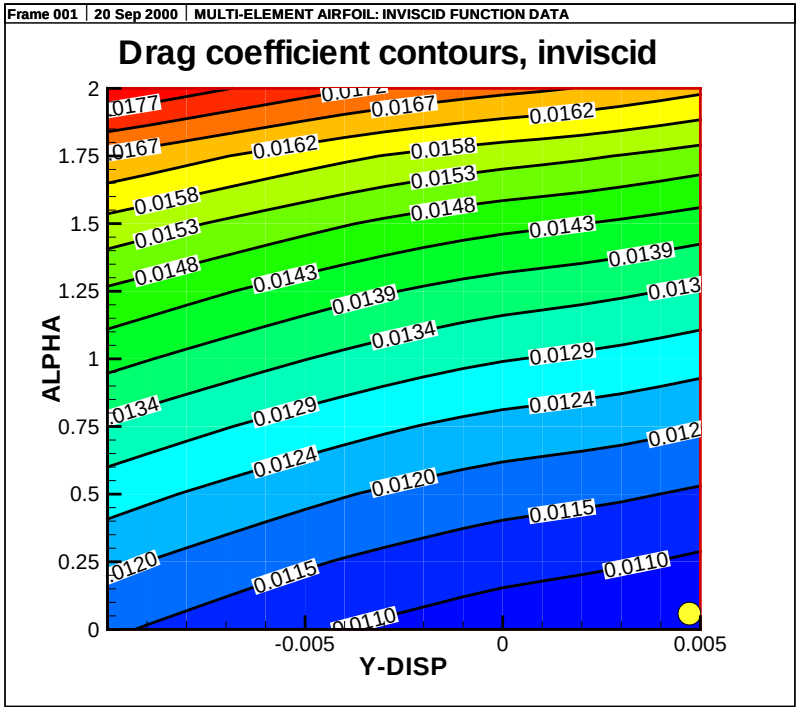
Multi-Element Airfoil: Computational Experiments

- **Objective function:** minimize drag coefficient subject to bounds on variables
- **Case 1:** (for visualization)
  - Variables: angle of attack, y-displacement of the flap
  - Solve problem with hi-fi models alone using a commercial optimization code (PORT, Bell Labs)
  - Solve the problem with AMMO, PORT used for lo-fi subproblems
- **Case 2:**
  - Variables: angle of attack, y-displacement of the flap, geometry description of the airfoil; 84 variables total
  - Same experiment

Multi-Element Airfoil: Models

- **Time/function for inviscid model negligible compared to viscous model**
- **Descent trends are reversed — unusual but a good test**

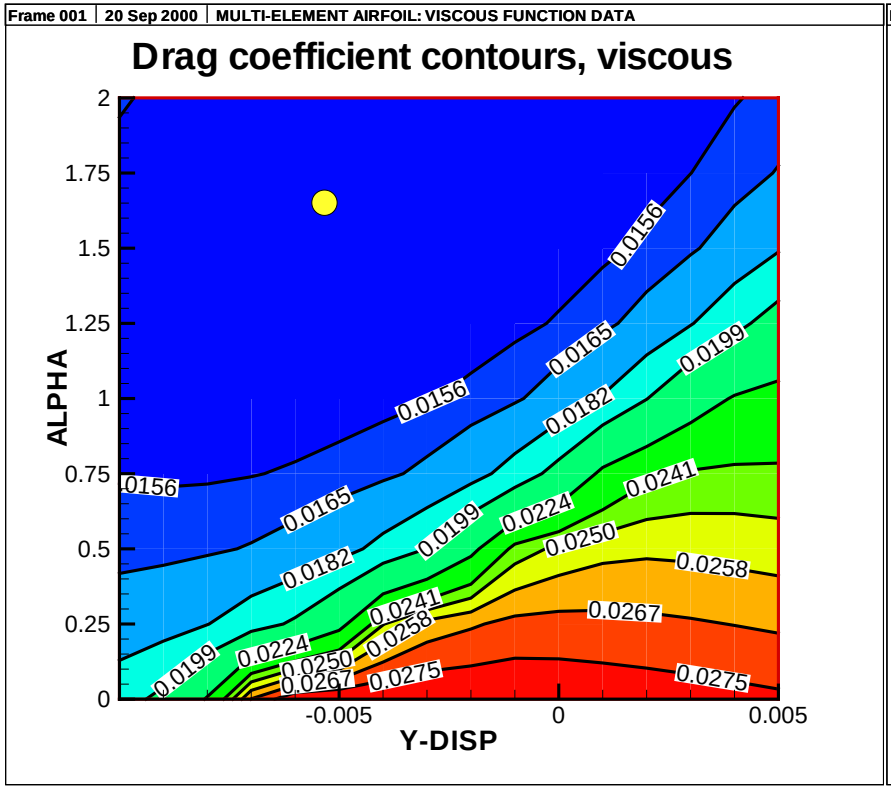


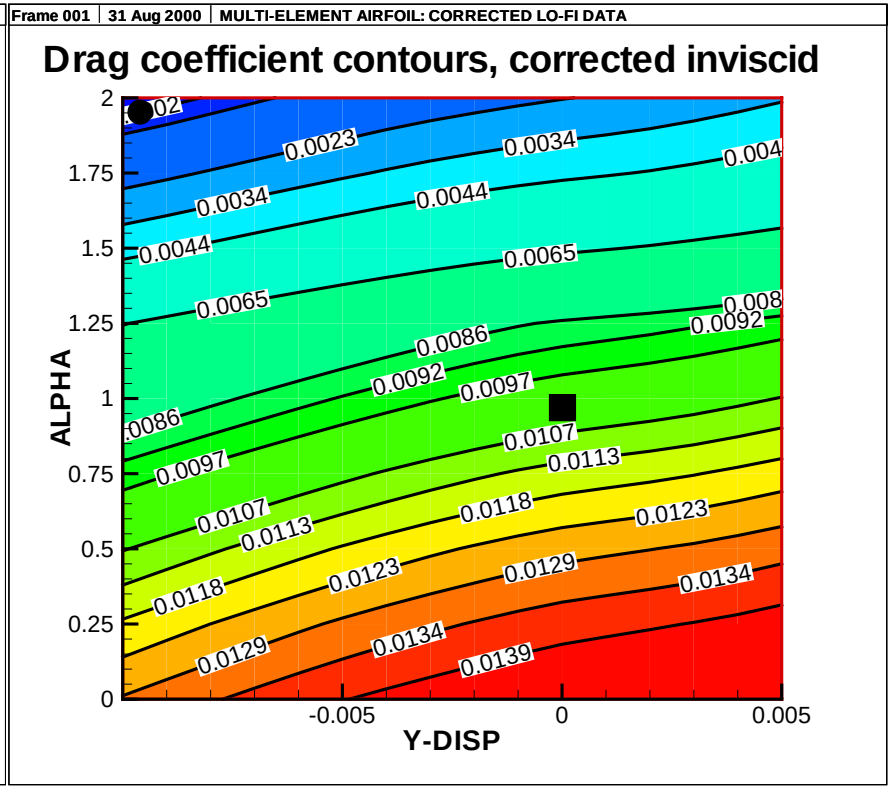


Multi-Element Airfoil: AMMO Iterations with 2 Variables

Iteration 1. Starting point: $\alpha = 1.0$, y-disp $= 0.0$

High-fidelity objective vs. corrected low-fidelity objective





New point: $\alpha = 2.0$, y-disp $= -0.01$

Multi-Element Airfoil: AMMO Iterations with 2 Variables, cont.

- **Similar effect in the next iteration**
- **Solution ($\alpha = 1.6305°$, flap y-displacement $= -0.0048$) located at iteration 2**
- $C_D^{\text{initial}} = 0.0171$ **at ($\alpha = 1°$, flap y-displacement $= 0$)**
- $C_D^{\text{final}} = 0.0148$**, a decrease of approximately 13.45%.**

Multi-Element Airfoil: Performance Summary

Notation: No. functions / No. Gradients

Test	hi-fi eval	lo-fi eval	total t	factor
PORT with hi-fi analyses, 2 var	14/13		$\approx$ 12 hrs	
AMMO, 2 var	3/3	19/9	$\approx$ 2.41hrs	$\approx$ 5
PORT with hi-fi analyses, 84 var	19/19		$\approx$ 35 hrs	
AMMO, 84 var	4/4	23/8	$\approx$ 7.2hrs	$\approx$ 5

Applicability: 1-Order vs 0-Order Model Management (e.g., RSM)

- **Sensitivities**
  - **1-Order:** Need
  - **0-Order:** Do not need, in general
- **Verification of improvement in the true response**
  - **1-Order:** Consistency conditions guarantee that improvement can be ensured by taking small steps, if necessary
  - **0-Order:** No such guarantee; may have to revert to hi-fi optimization
- **Enforcing sufficient decrease conditions**
  - **1-Order:** Having sensitivities, can enforce sufficient decrease
  - **0-Order:** Absence of sensitivity information restricts the class of optimization algorithms used with model management

Applicability: 1-Order vs 0-Order Model Management, cont.

- **Problem size**
  - **1-Order:** Can use models that do not depend on problem size
  - **0-Order:** Difficult to construct approximations for more than, say, 20–40 variables; in general, need a considerable number of samples to capture trends even for relatively simple and small problems
- **Sampling in engineering problems**
  - **1-Order:** Not an issue if we use models not based on data-fitting
  - **0-Order:** May be difficult to guarantee that hi-fi samples correspond to physically meaningful designs. Example: shape optimization problems.
- **Range of models**
  - **1-Order:** The entire range of models may be used
  - **0-Order:** May use variable-accuracy or variable-resolution models as predictors of promising designs (not always); variable-fidelity physics models cannot be used, in general

Concluding Remarks

- **Sensitivity information is becoming increasingly available with analysis codes (automatic differentiation, efficient adjoint techniques) — we believe first-order AMMO will prove very helpful**
- **The results are promising**
  - **Significant savings over conventional optimization in terms of hi-fi evaluations**
  - **Capture descent behavior with the help of corrections, despite sometimes significant dissimilarities between lo-fi and hi-fi models**
  - **First-order information is indispensable when model trends are dissimilar**
- **For large problems, variable-fidelity physics modeling may be recommended**
- **Current work: Incorporation of *a posteriori* error estimates into AMMO (with Patera et al.); other models, AMMO schemes, larger problems**