Towards an Adaptive One-Shot Approach for Aerodynamic Shape Optimization

Nicolas R. Gauger[1),2)]

1) German Aerospace Center (DLR) Braunschweig
Institute of Aerodynamics and Flow Technology
Numerical Methods Branch ($C^2A^2S^2E$)

2) Humboldt University Berlin
Department of Mathematics

From October 2010 on:
RWTH Aachen University
Dept. of Mathematics / Center for Comp. Eng. Science (CCES)

Towards an Adaptive One-Shot Approach for Aerodynamic Shape Optimization

Related Projects

- EU-Project NODESIM-CFD (Non-Deterministic Simulation for CFD-based Design Methodologies)
- DFG (German Research Foundation) Priority Program SPP 1253 'Optimization with PDEs'
- BMBF-Project DGHPOPT (Discontinuous Galerkin Methods for Robust Aerodynamic Design)

  Title: „Adaptive One-Shot Methods for Robust Design“

  Partners: Trier University, RWTH Aachen University ←HU B, DLR, Airbus, MTU

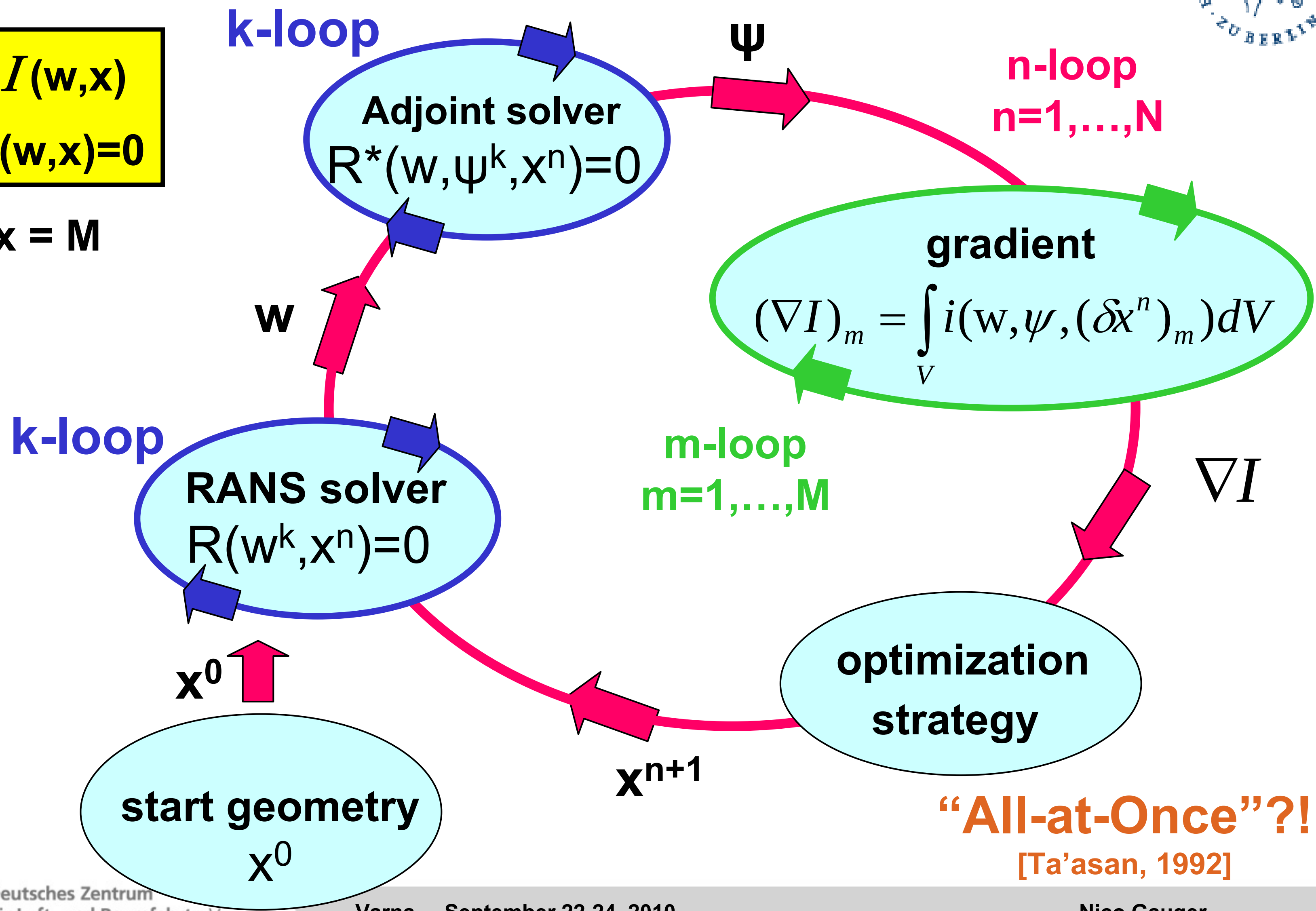
  → Project starts October 2010

Nested approach

min I(w,x)

s.t. R(w,x)=0

dim x = M



"All-at-Once"?!

[Ta'asan, 1992]

Problem Statement

Goal: $\min_u f(y,u)$ **s.t.** $c(y,u)=0,$

where y **and** u **are the state and design variables.**

Given fixed point iteration $y_{k+1}=G(y_k,u)$ **(e.g. pseudo-time stepping) to solve PDE** $c(y,u)=0.$

Assumptions:

- $\dfrac{\partial c}{\partial y}$ **always invertible. IFT** $\Rightarrow$ **given** $u, \exists!\, y$ **s.t.** $c(y,u)=0.$
- $G, f \in C^{2,1}.$
- G **contractive:** $\|G_y(y,u)\| = \|G_y^T(y,u)\| \le \rho < 1$

One-Shot approach

$$L(y,\bar{y},u) = f(y,u) + (G(y,u) - y)^T \bar{y}$$
$$= \underbrace{N(y,\bar{y},u)}_{\text{shifted Lagragian}} - y^T \bar{y}$$

Stationary point:
$$\begin{cases} L_{\bar{y}} = G(y,u) - y = 0 \\ L_y = N_y(y,\bar{y},u)^T - \bar{y} = 0 \\ L_u = N_u(y,\bar{y},u)^T = 0 \end{cases}$$

One-step one-shot (step *k+1*):

$$(OS)\begin{cases} y_{k+1} = G(y_k, u_k) & \text{primal update} \\ \bar{y}_{k+1} = N_y(y_k, \bar{y}_k, u_k)^T & \text{dual update} \\ u_{k+1} = u_k - B_k^{-1} N_u(y_k, \bar{y}_k, u_k)^T & \text{design update} \end{cases}$$

Aims: **Choose *B* such that:**
- **Convergence of *(OS)*.**
- **Bounded retardation.**

Bounded retardation

Jacobian of the extended iteration:

$$J_* = \left.\frac{\partial(y_{k+1}, \bar{y}_{k+1}, u_{k+1})}{\partial(y_k, \bar{y}_k, u_k)}\right|_{(y^*, \bar{y}^*, u^*)} = \begin{pmatrix} G_y & 0 & G_u \\ N_{yy} & G_y^T & N_{yu} \\ -B^{-1}N_{uy} & -B^{-1}G_u^T & I - B^{-1}N_{uu} \end{pmatrix}$$

Whenever we can define B such that

$$\frac{1-\rho(G_y)}{1-\hat{\rho}(J_*)} < const, \quad \text{i.e.} \quad \mathcal{O}(opt) / \mathcal{O}(sim) < const$$

we have bounded retardation.

Necessary condition for contractivity

Eigenvalues of J_* are the zeros of the equation

$$\det((\lambda - 1)B + H(\lambda)) = 0$$

where

$$H(\lambda) = \left(-G_u^T (G_y^T - \lambda I)^{-1}, I\right) \begin{pmatrix} N_{yy} & N_{yu} \\ N_{uy} & N_{uu} \end{pmatrix} \begin{pmatrix} -(G_y - \lambda I)^{-1} G_u \\ I \end{pmatrix}.$$

Necessary (but not sufficient) condition for contractivity:

$$B = B^T \succ 0 \qquad \text{and} \qquad B \succ \frac{1}{2} H(-1).$$

[Griewank, 2006]

Exact penalty function: L^a

Remark:

Deriving (sufficient) conditions on B for J_* to have a spectral radius smaller than 1 has proven difficult.

Instead, we look for descent on the augmented Lagrangian

$$L^a(y,\bar{y},u) := \frac{\alpha}{2}\underbrace{\left\|G(y,u)-y\right\|^2}_{\text{primal residual}} + \frac{\beta}{2}\underbrace{\left\|N_y(y,\bar{y},u)^T-\bar{y}\right\|^2}_{\text{dual residual}} + \underbrace{N-\bar{y}^T y}_{\text{Lagrangian}},$$

where $\alpha > 0$ and $\beta > 0$.

Correspondence condition

The full gradient of L^a is given by

$$\begin{bmatrix} \nabla_y L^a \\ \nabla_{\bar{y}} L^a \\ \nabla_u L^a \end{bmatrix} = -Ms(y,\bar{y},u), \quad \textbf{where} \quad s(y,\bar{y},u) = \begin{bmatrix} G(y,u) - y \\ N_y(y,\bar{y},u)^T - \bar{y} \\ -B^{-1}N_u(y,\bar{y},u)^T \end{bmatrix}$$

and $$M = \begin{bmatrix} \alpha(I - G_y{}^T), & -I - \beta N_{yy}, & 0 \\ -I, & \beta(I - G_y), & 0 \\ -\alpha G_u{}^T, & -\beta N_{yu}{}^T & B \end{bmatrix}.$$

Correspondence condition

Consequence (Correspondence condition):

There is a 1-1 correspondence between the stationary points of L^a and the roots of s if

$$\det[\alpha\beta(I - G_y{}^T)(I - G_y) - I - \beta N_{yy}] \neq 0,$$

for which it is sufficient that

$$\alpha\beta(1-\rho)^2 > 1 + \beta\|N_{yy}\|.$$

[Hamdi, Griewank, 2008]

Descent condition

Theorem (Descent condition):

$s(y,\bar{y},u)$ **is a descent direction for all large positive** B

if and only if

$$\alpha\beta(I-\frac{1}{2}(G_y+G_y{}^T)) > (I+\frac{\beta}{2}N_{yy})(I-\frac{1}{2}(G_y+G_y{}^T))^{-1}(I+\frac{\beta}{2}N_{yy}),$$

which is implied by $\sqrt{\alpha\beta}(1-\rho) > 1+\frac{\beta}{2}\|N_{yy}\|.$

- **Satisfied for** $\beta=\frac{2}{c},\quad \alpha=\frac{2c}{(1-\rho)^2}$ **with** $c=\|N_{yy}\|$.

Theorem: A suitable B is given by:

$$B=\alpha G_u^T G_u+\beta N_{yu}^T N_{yu}+N_{uu}.$$

[Hamdi, Griewank, 2008]

One-step one-shot

Aerodynamic shape design

Descent for $\beta = \frac{2}{c}, \quad \alpha = \frac{2c}{(1-\rho)^2}$ **with** $c = \left\| N_{yy} \right\|$ **.**

(In practice choose $c = 1, \ \Rightarrow \ \beta = 2, \ \alpha >> 1.)$

A suitable B **is given by** $B = \alpha G_u^T G_u + \beta N_{yu}^T N_{yu} + N_{uu}.$

Instead BFGS updates for the Hessian

$$\nabla_u^2 L^a = \underbrace{\alpha G_u^T G_u + \beta N_{yu}^T N_{yu} + N_{uu}}_{B} + \alpha \underbrace{(G - y)}_{\to_* 0}^T G_{uu} + \beta \underbrace{(N_y^T - \bar{y})}_{\to_* 0}^T N_{yuu}.$$

The gradient $\nabla_u L^a = \alpha (G - y)^T G_u + \beta (N_y - \bar{y})^T N_{yu} + N_u$

is evaluated by Algorithmic Differentiation (AD).

[Özkaya, Gauger, 2008]

Automatic Differentiation of Entire Design Chain

design vector (P) $\rightarrow$ defgeo $\xrightarrow{x_{new}}$ difgeo $\xrightarrow{dx}$ meshdefo $\xrightarrow{m}$ flow solver $\rightarrow$ C_D

surface grid ↗ difgeo   grid ↗ meshdefo

- **Adjoint version of entire design chain by ADOL-C / TAPENADE**
- **TAUij (2D Euler) / ELAN (2D/3D RANS) + mesh deformation + parameterization**

$$\frac{dC_D}{dP} = \underbrace{\frac{\partial C_D}{\partial m}}_{\text{TAUij\_AD / ELAN\_AD}} \cdot \underbrace{\frac{\partial m}{\partial (dx)}}_{\text{meshdefo\_AD}} \cdot \frac{\partial (dx)}{\partial x_{new}} \cdot \underbrace{\frac{\partial x_{new}}{\partial P}}_{\text{defgeo\_AD}} \quad \text{and} \quad \frac{\partial (dx)}{\partial x_{new}} = \frac{\partial (x_{new} - x_{old})}{\partial x_{new}} = Id$$

[Özkaya, Gauger, 2008]

[Gauger, Walther, Özkaya, Moldenhauer, 2008]

Treatment of lift constraint by penalty multiplier method

$$\min_u C_D(y,u) \quad s.t. \quad C_L \geq C_{L,target} \quad and \quad y = G(y,u)$$

Penalty function for lift: $h = (C_{L,target} - C_L), \quad h \overset{!}{\leq} 0$

Redefine objective function: $f = C_D + \lambda h$

$$\min_u C_D(y,u) + \lambda h \quad ; \quad h \rightarrow 0$$

Update the penalty parameter in each one-shot step *k*:

$$\lambda_{k+1} = \lambda_k (1 + ch), \qquad c > 0$$

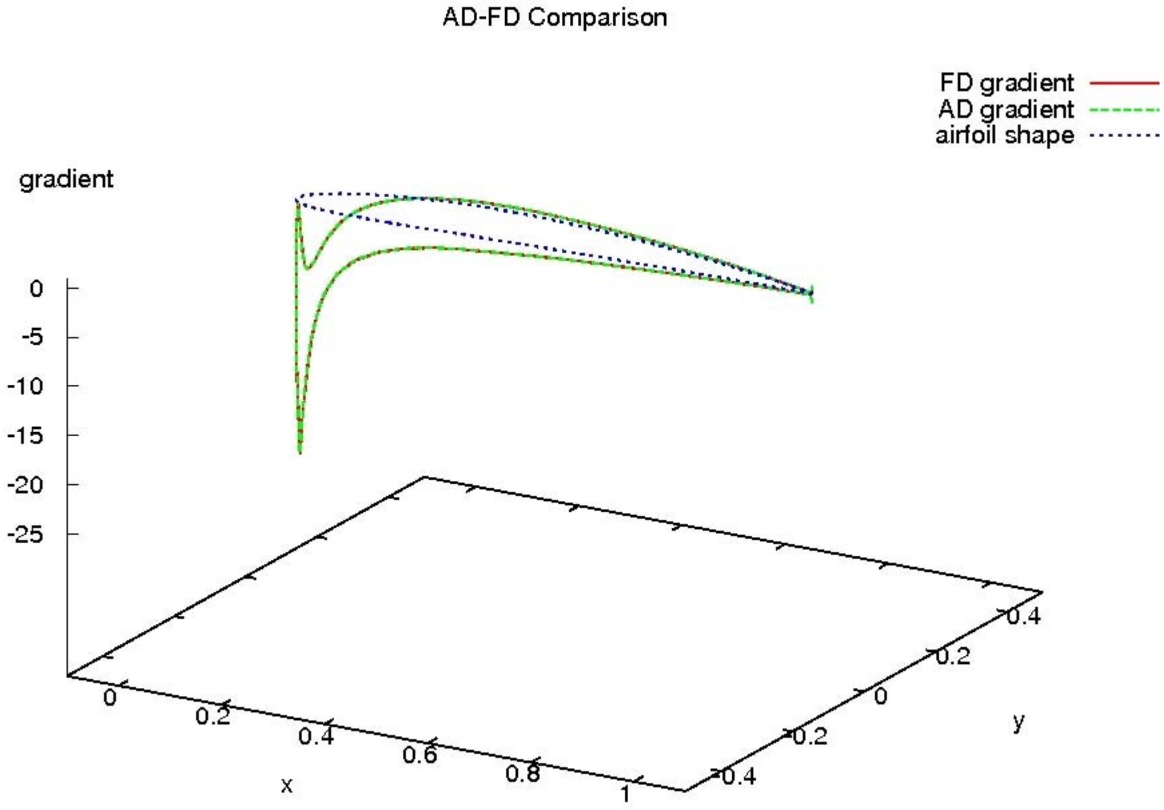
$$h > 0 \quad \Rightarrow \quad \lambda \uparrow, \qquad h < 0 \quad \Rightarrow \quad \lambda \downarrow$$

A good starting value is: $\lambda_0 = \dfrac{\|\nabla_u C_D\|}{\|\nabla_u h\|}$

Extension to Navier-Stokes (ELAN Code)

Flow Solver: ELAN (TU Berlin)
- 3D Navier-Stokes (RANS)
- incompressible with pressure correction
- multiblock
- k-ω (Wilcox) turbulence model (and others)
- Fortran (20.000 lines)

AD Tool: TAPENADE (INRIA)
- source to source
- reverse for **first derivatives**
- tangent on reverse for **second derivatives**



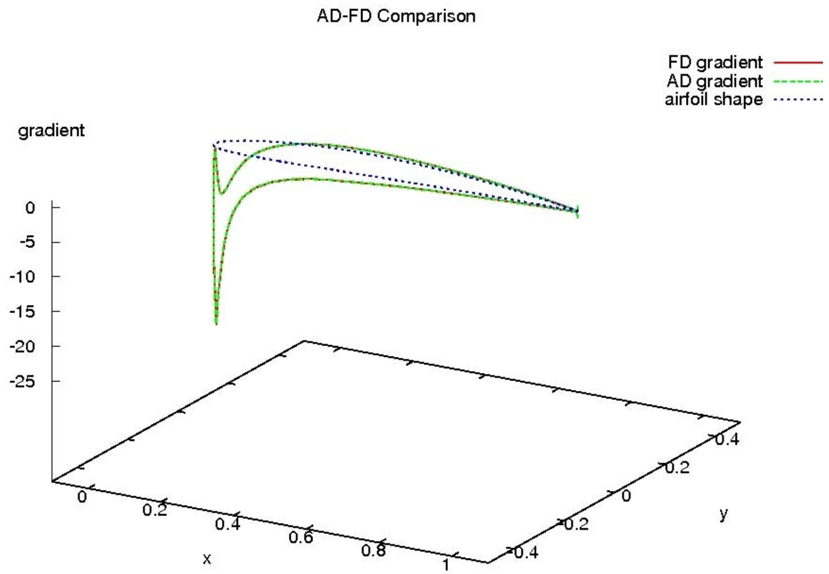
Extension to Navier-Stokes (ELAN Code)

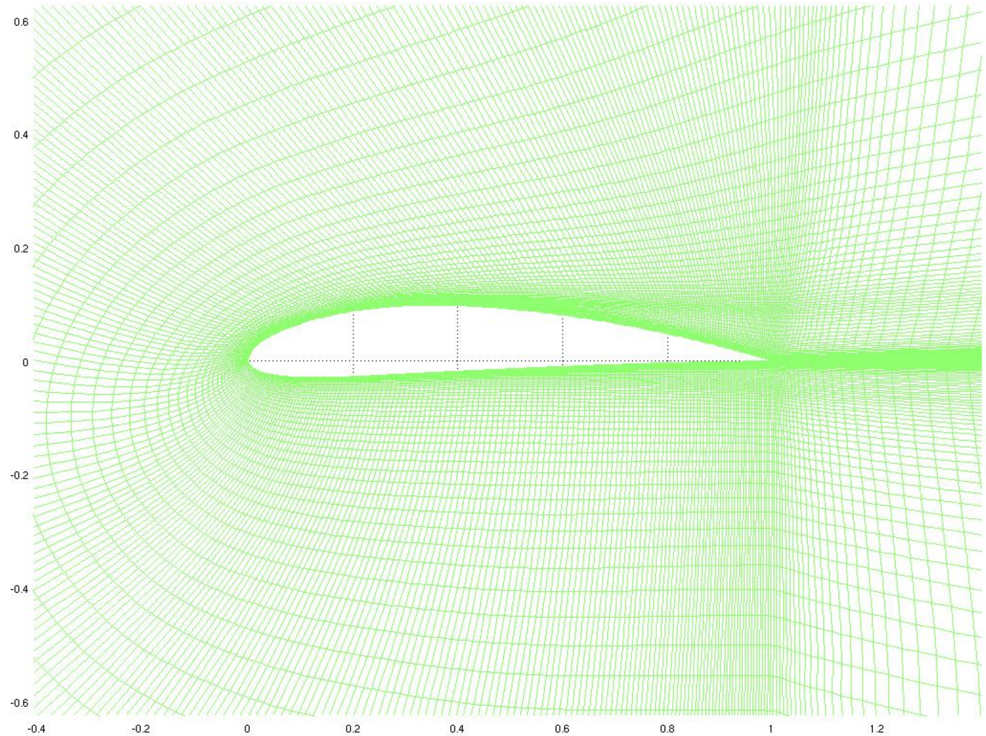
Drag reduction with lift constraint

- NACA 4412
- Re = 1.000.000, α=5.1°
- RANS
- k-ω (Wilcox) turbulence model
- 300 surface mesh points

Approaches for Optimization

- one-shot method
- entire design chain differentiated
- gradient smoothing
- penalty multiplier method





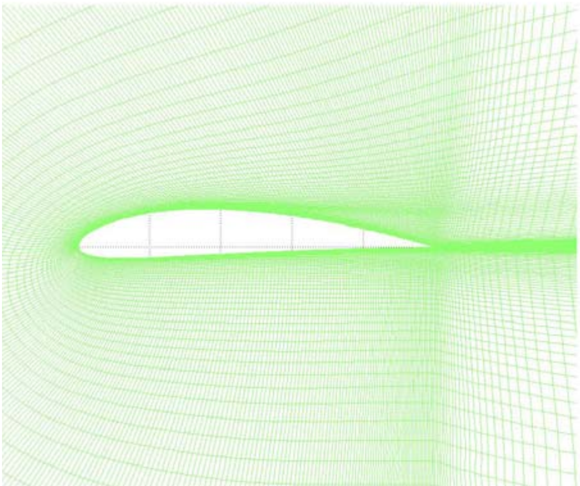
Extension to Navier-Stokes (ELAN Code)

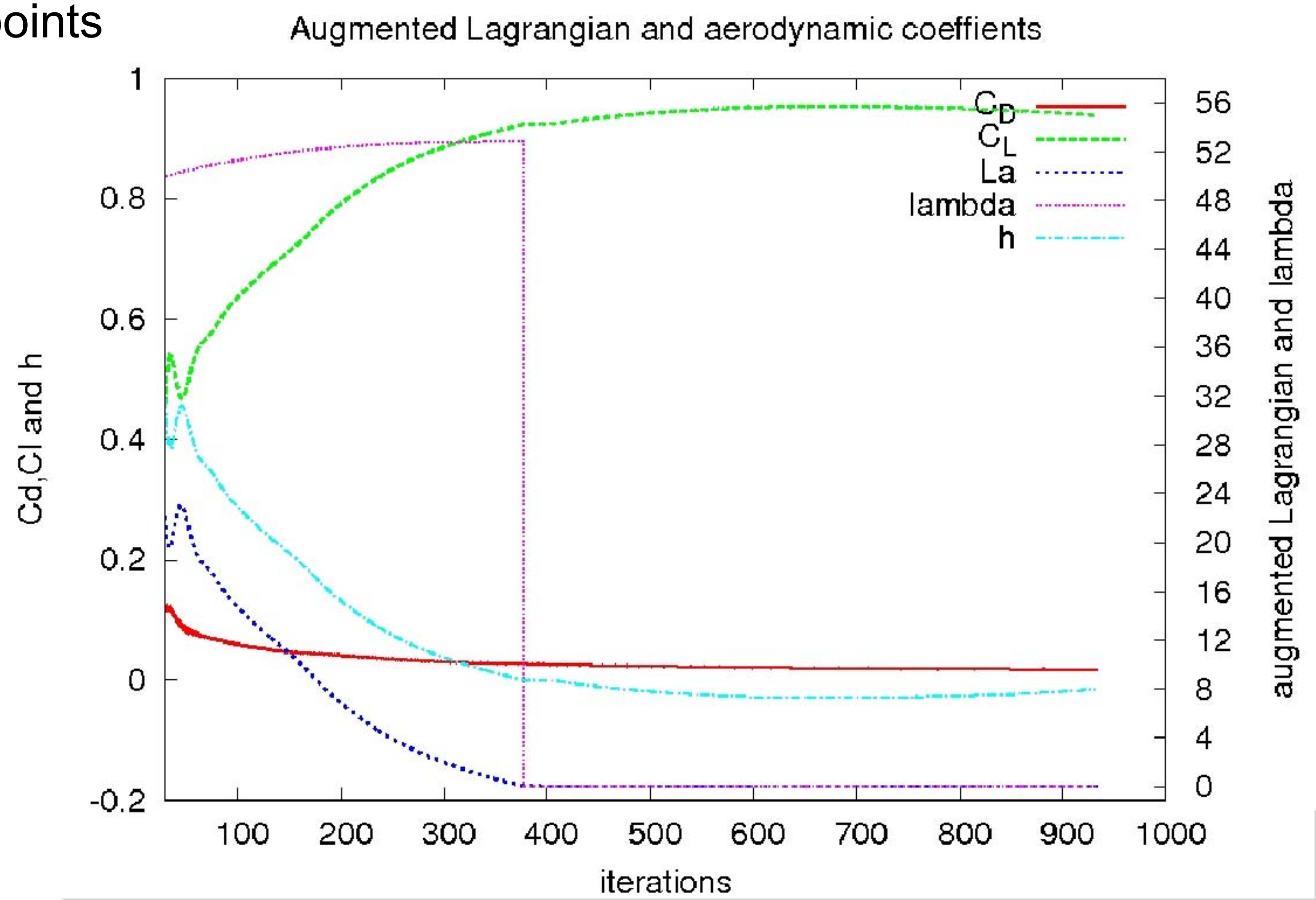
Drag reduction with lift constraint

- NACA 4412
- Re = 1.000.000, α=5.1°
- RANS
- k-ω (Wilcox) turbulence model
- 300 surface mesh points

Approaches for Optimization

- one-shot method
- entire design chain differentiated
- gradient smoothing
- penalty multiplier method

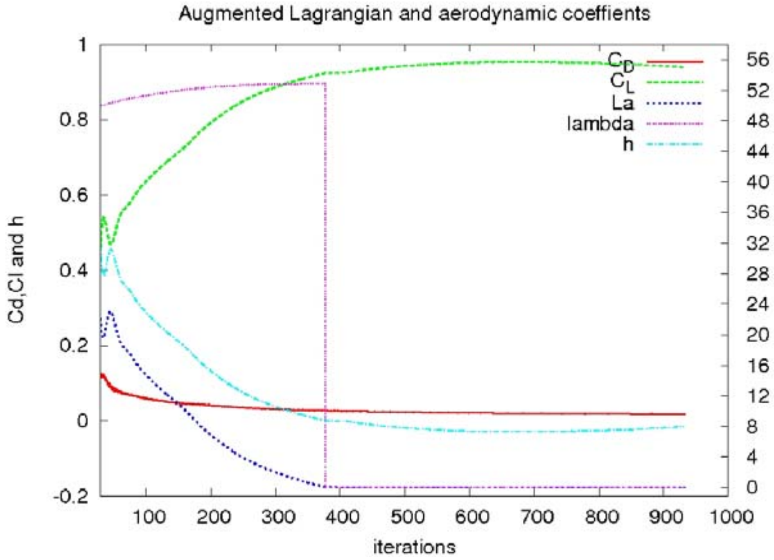




Extension to Navier-Stokes (ELAN Code)

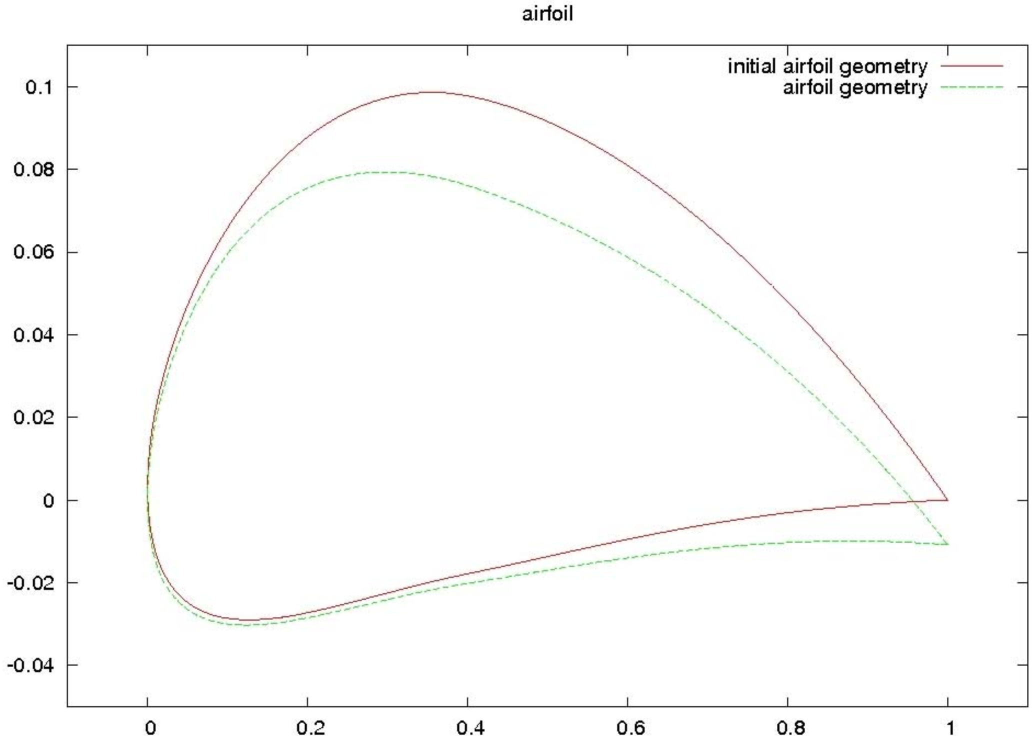
Drag reduction with lift constraint
- NACA 4412
- Re = 1.000.000, α=5.1°
- RANS
- k-ω (Wilcox) turbulence model
- 300 surface mesh points

5% drag reduction



Retardation-Factor = 3

Approaches for Optimization
- one-shot method
- entire design chain differentiated
- gradient smoothing
- penalty multiplier method



[Özkaya, Gauger, 2009]

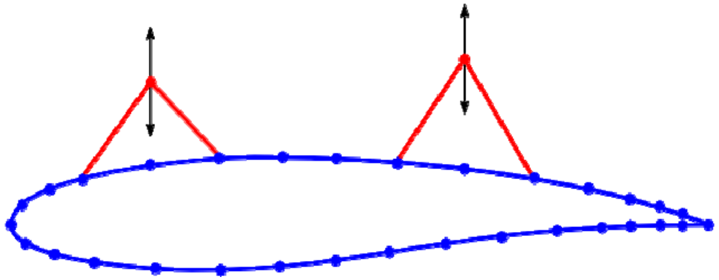
Ingredients for Efficient One-Shot Approach

• **Shape Derivatives:**

$$dF_\ell(\Omega)[V] = \int_\Gamma \langle V, n\rangle \left[\langle \nabla p_\ell n, n\rangle + div_\Gamma\left(p_\ell - (\psi^T w_H)u\right)\right] \; dS$$

[Schmidt, Ilic, Gauger, Schulz, 2008]

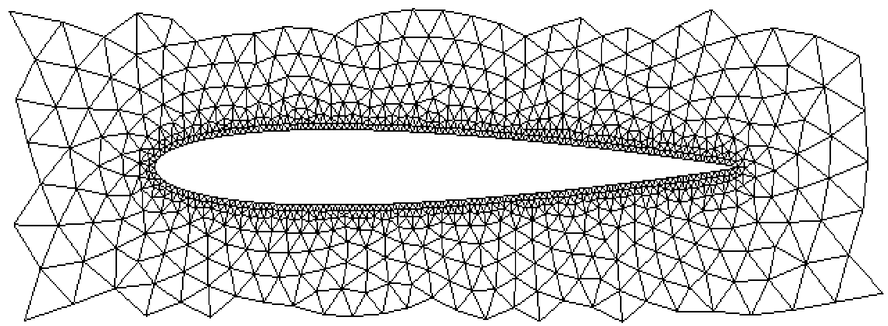
• **Gradient Smoothing / Preconditioning:**



$$\overline{G} - \frac{\partial}{\partial \xi}\varepsilon\frac{\partial}{\partial \xi}\overline{G} = G$$

[Jameson et al., 1994 ff.],…, [Schmidt, Ilic, Gauger, Schulz, 2008]

• **Multilevel-Parameterization:**



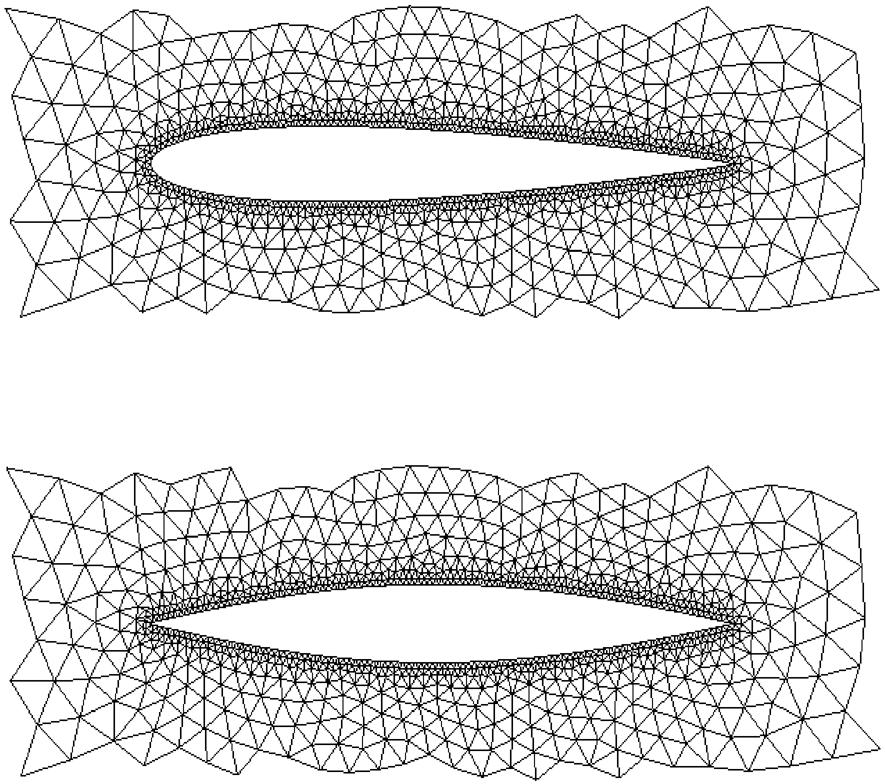
Dual Weighted Residual (DWR)

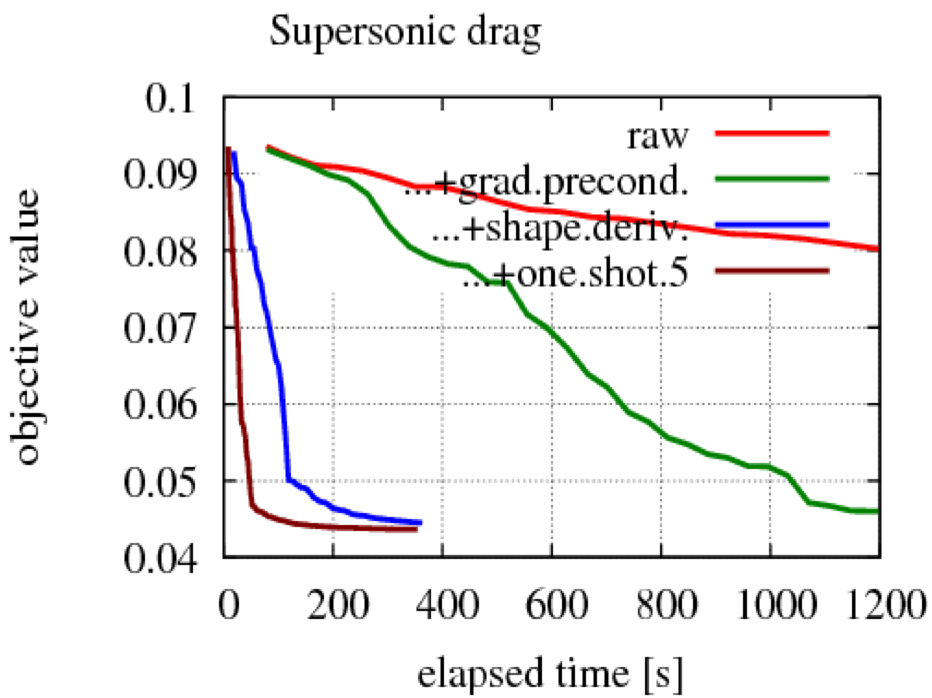
[Rannacher et al., 1996], [Giles 1999], …

[Vollmer, Gauger, 2005], [Kroll, Gauger et al., 2007]

Efficient One-Shot Approach

- NACA0012, M=2.0, α=0°, TAU (Euler)
- Drag Reduction

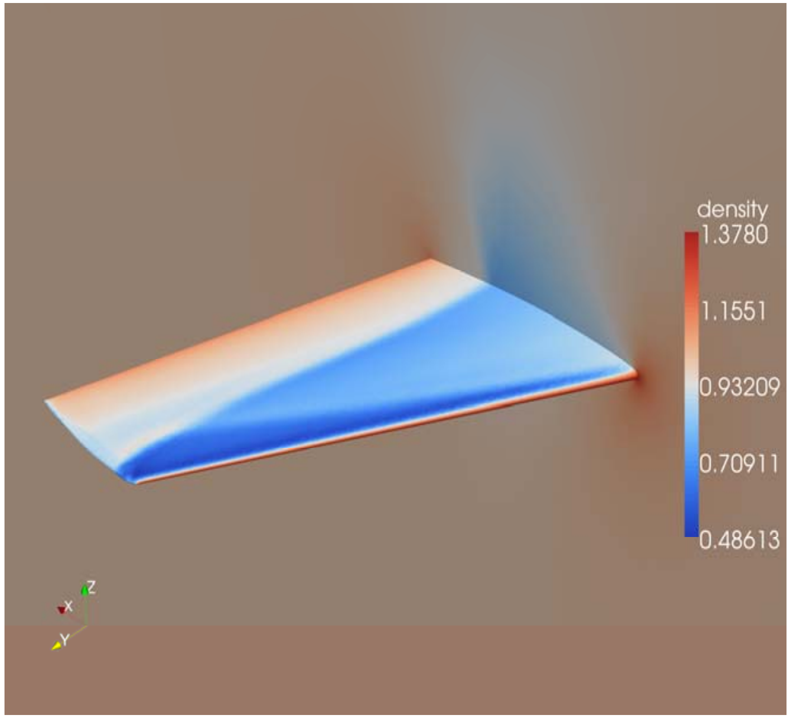




[Schmidt, Ilic, Gauger, Schulz, 2008]

Efficient One-Shot Approach

- ONERA M6, M=0.83, α=3.01°, TAU (Euler)
- Drag reduction by
- Constant lift

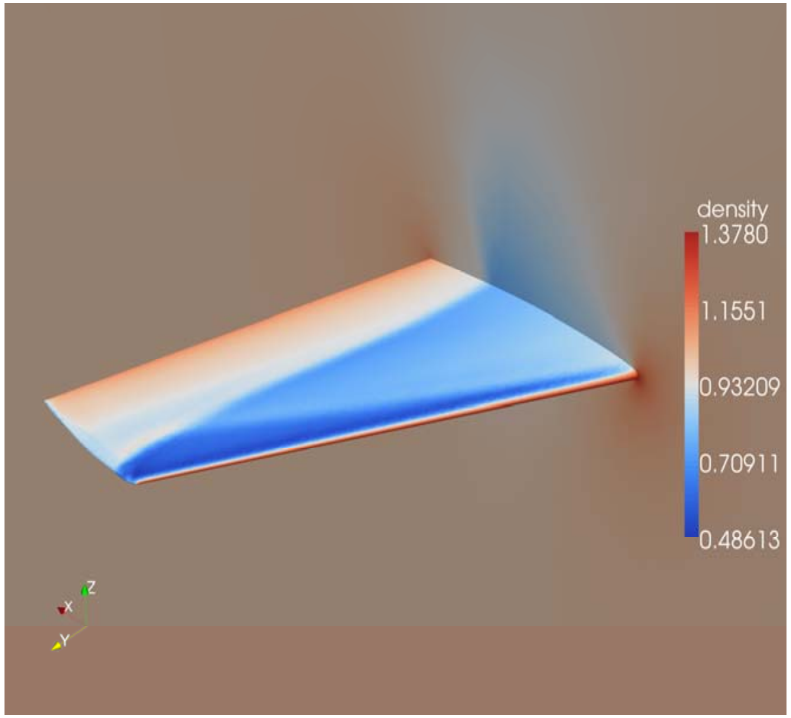


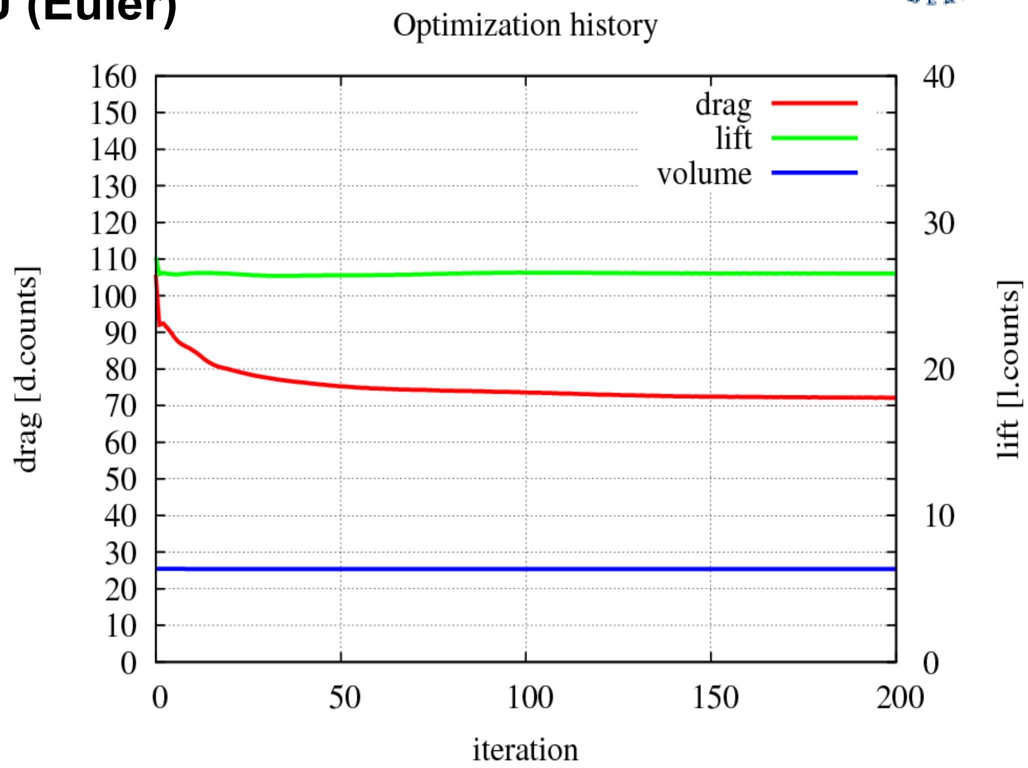
- 867.168 points for volume mesh
- Design variables:
- All 18.285 surface points!
- Approach:
- Shape Derivatives
- Gradient Smoothing / Preconditioning
- One Shot

[Ilic, Gauger, Schmidt, Schulz, 2009]

Efficient One-Shot Approach

- ONERA M6, M=0.83, α=3.01°, TAU (Euler)
- Drag reduction by
- Constant lift





$\mathcal{O}$(opt) / $\mathcal{O}$(sim) = 5 (wall clock time)

= 2 (# iterations)

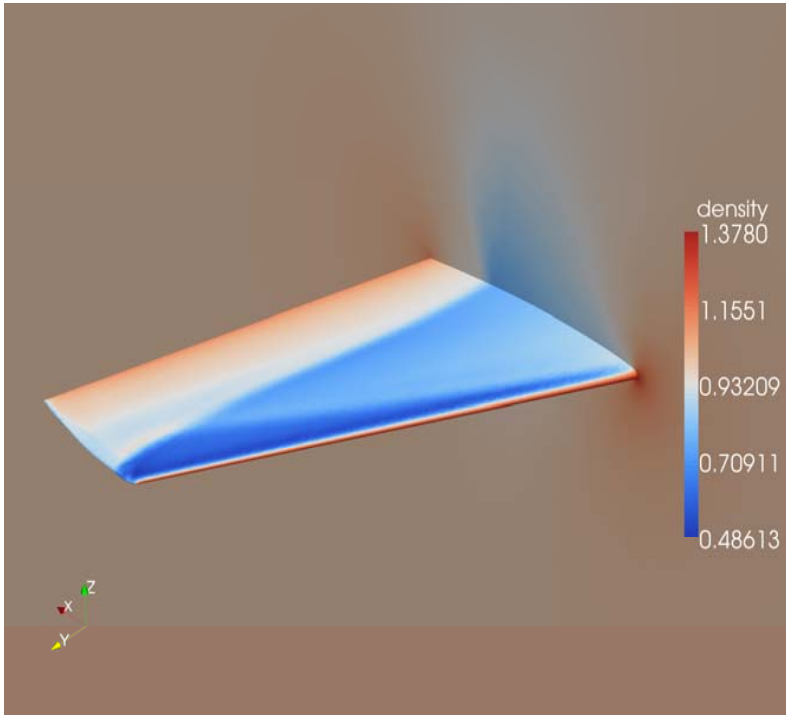
[Ilic, Gauger, Schmidt, Schulz, 2009]

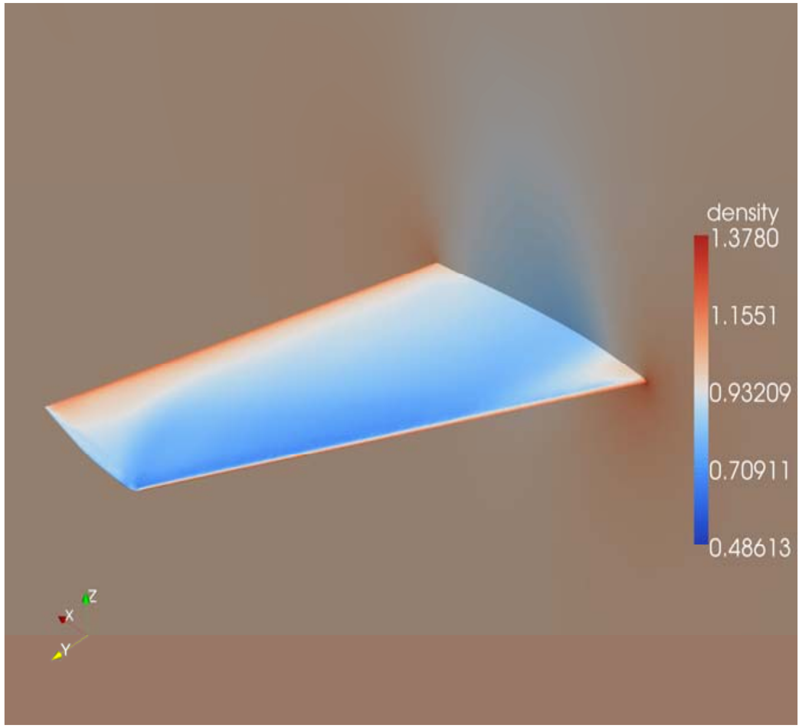
Efficient One-Shot Approach

- Initial geometry
- C_D^{init} = 106 drag counts
- C_L^{init} = 27.6 lift counts

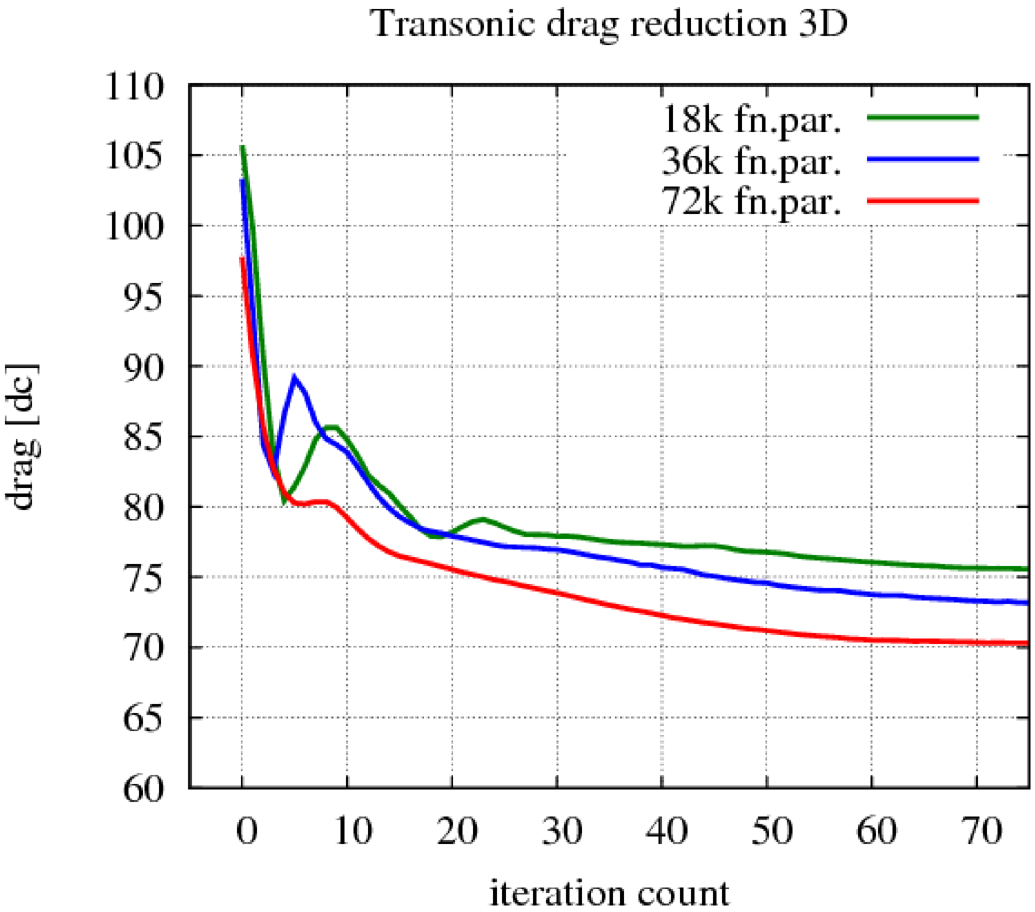
- Optimized geometry
- C_D^{opt} = 72 drag counts
- C_L^{opt} = 26.5 lift counts

32% Drag reduction

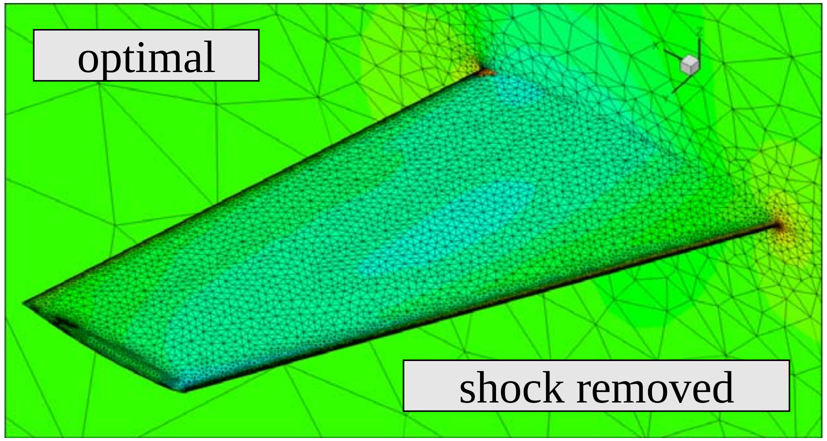




Scaling With Number of Parameters



optimal drag =
induced drag (planform)
+ spurious drag (numerics)

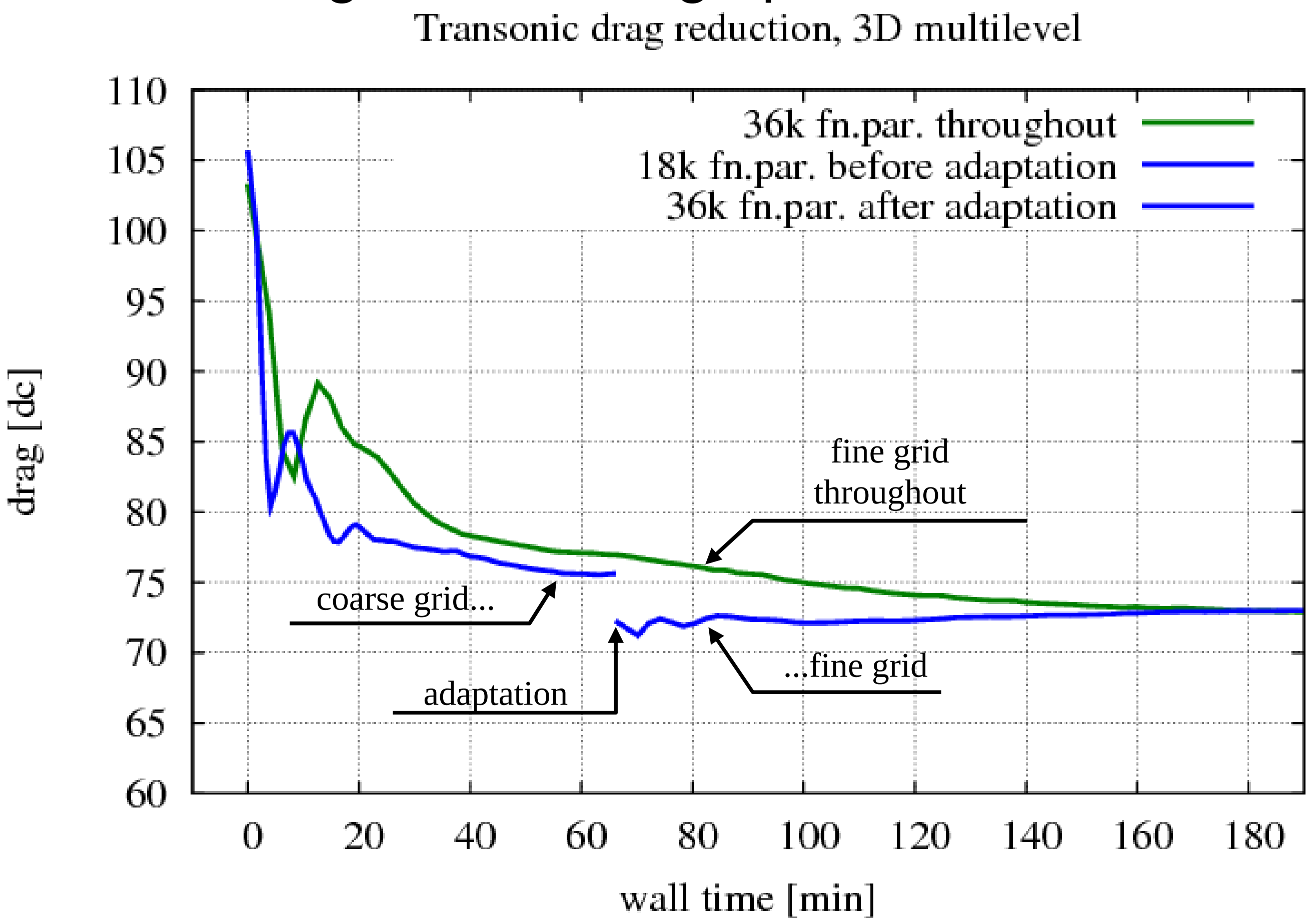


optimization to simulation time ~5

[Ilic, Gauger, Schmidt, Schulz, 2009]

Multilevel Descent

➢2-level: coarse grid **18k** design parameters, fine **36k**



➢ 2-level iteration brings factor **~2** in optimization time

[Ilic, Gauger, Schmidt, Schulz, 2009]

Dual (Adjoint) Weighted Residual Approach (DWR)

Idea stems from ***Rannacher et al.*** **in FEM context.**

Transfer to FVM by ***Venditty & Darmofal*****, but repeated extrapolation between certain mesh levels (coarse** ***H*****, fine** ***h*****) in practice not handy.**

Therefore, instead of $I_h(U_h^H) - I_h(U_h) \approx \left(\psi_h^H\right)^T R_h(U_h^H),$

we consider $I(U_h) - I(U) \approx \psi_h^T R(U_h).$

Now the idea (by ***R. Dwight*****) is, to interpret the discretization error** $R(U_h)$ **as dissipation error. This yields:**

$$I(U_h) - I(U) \approx \psi_h^T \left(k^{(2)} \frac{\partial R}{\partial k^{(2)}} + k^{(4)} \frac{\partial R}{\partial k^{(4)}} \right)$$

[*R. Dwight, JCP, 2008*]

Dual (Adjoint) Weighted Residual Approach (DWR)

The (global) value

$$\psi^T \left(k^{(2)} \frac{\partial R}{\partial k^{(2)}} + k^{(4)} \frac{\partial R}{\partial k^{(4)}} \right),$$

is a scalar product over the computational mesh, i.e. we have a summation over all mesh cells of terms

$$\psi_{i,j}^T \left(k_{i,j}^{(2)} \frac{\partial R_{i,j}}{\partial k^{(2)}} + k_{i,j}^{(4)} \frac{\partial R_{i,j}}{\partial k^{(4)}} \right).$$

Sensor for mesh adaptation (Local dual weighted residual)

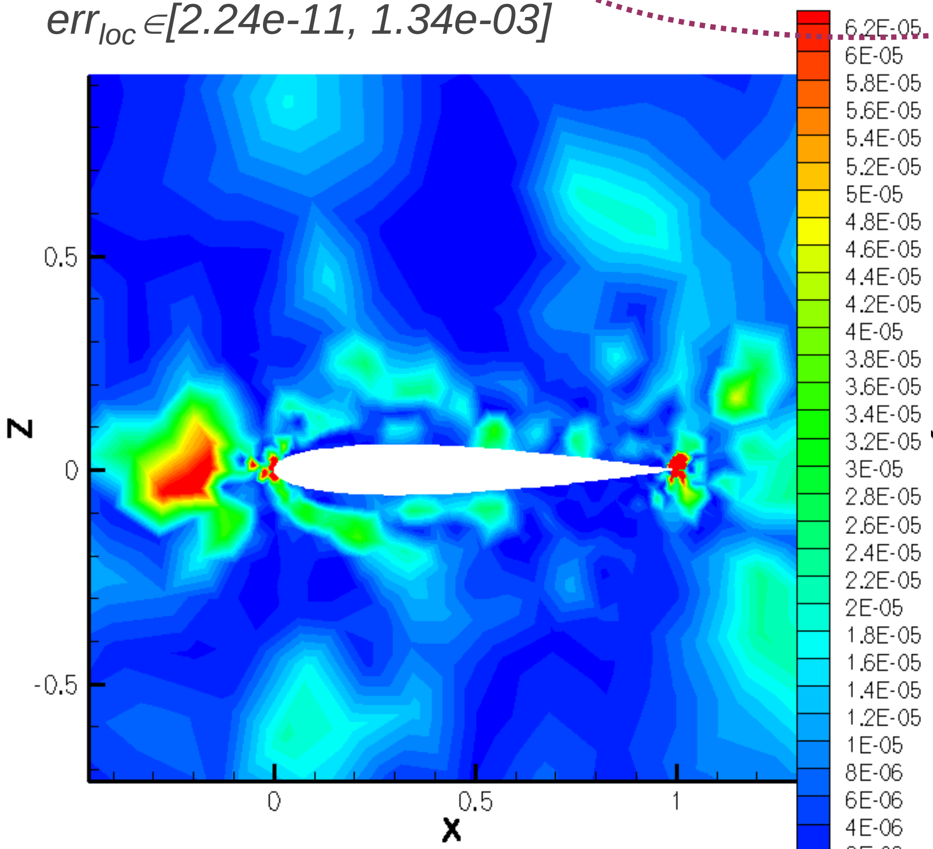
(Local) Adjoint-based Error Estimate / Sensor for Mesh Adaptation by DWR

Sensor

$$abs\left[\psi_{i,j}^{T}\left(k_{i,j}^{(2)}\frac{\partial R_{i,j}}{\partial k^{(2)}}+k_{i,j}^{(4)}\frac{\partial R_{i,j}}{\partial k^{(4)}}\right)\right]$$

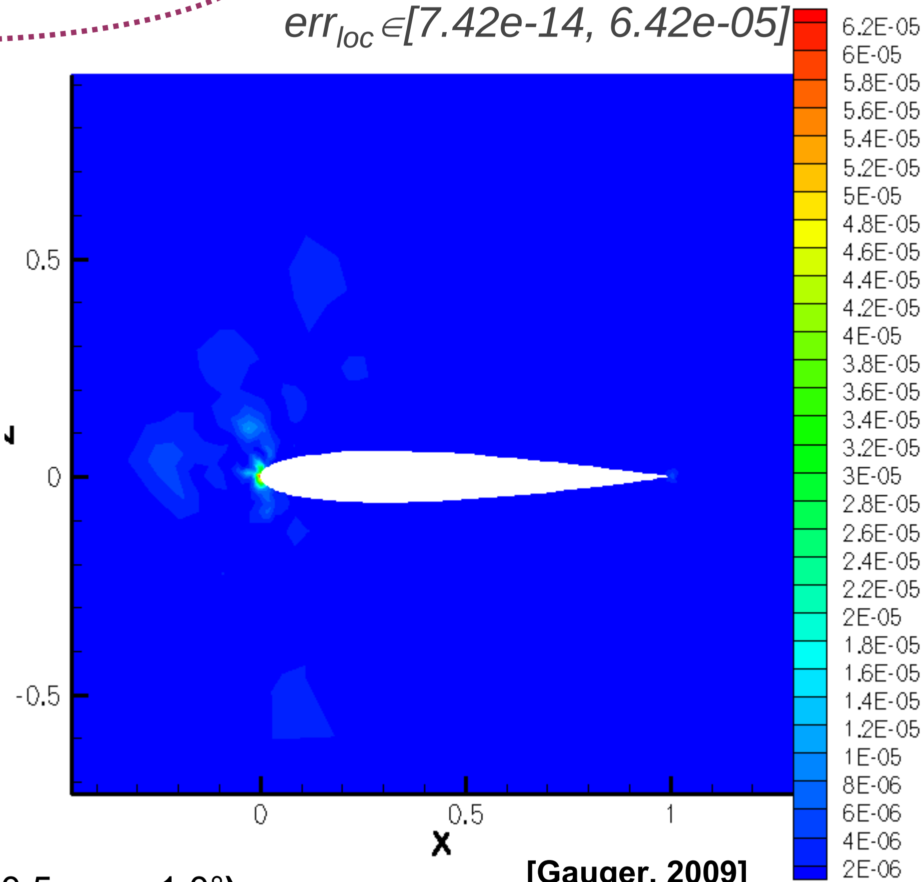
Lift

$err_{loc} \in [2.24e\text{-}11, 1.34e\text{-}03]$



Drag

$err_{loc} \in [7.42e\text{-}14, 6.42e\text{-}05]$



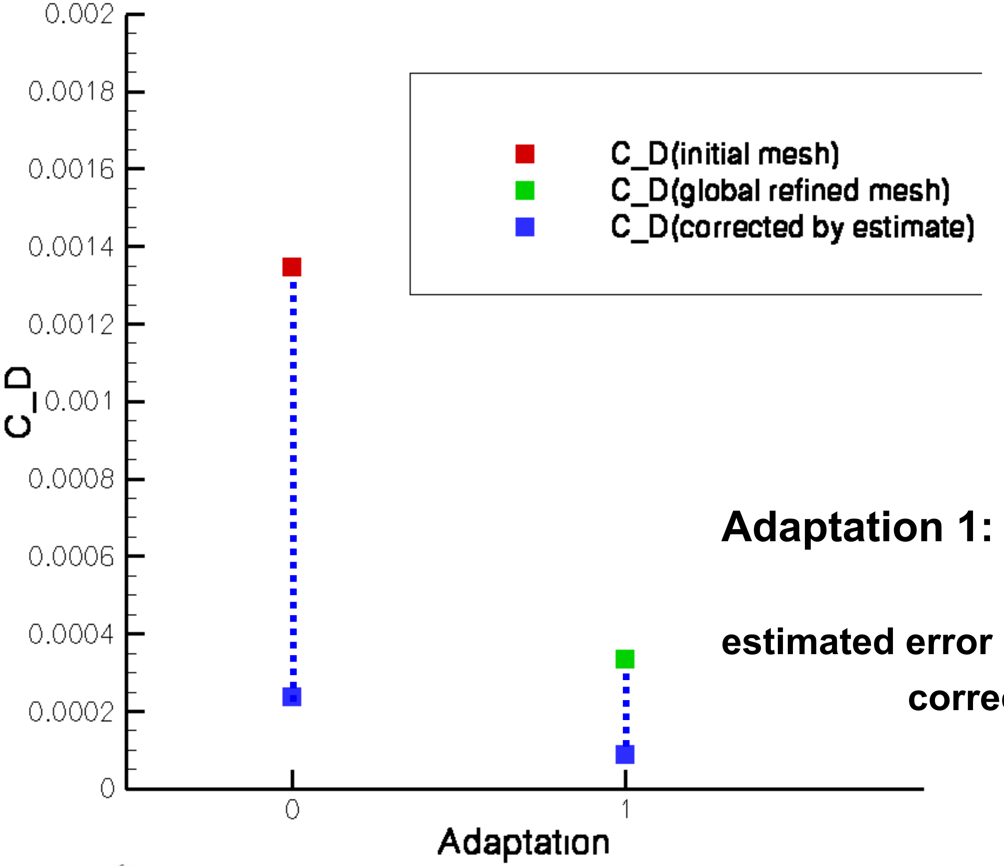
(NACA0012 Ma = 0.5, α = 1.0°)

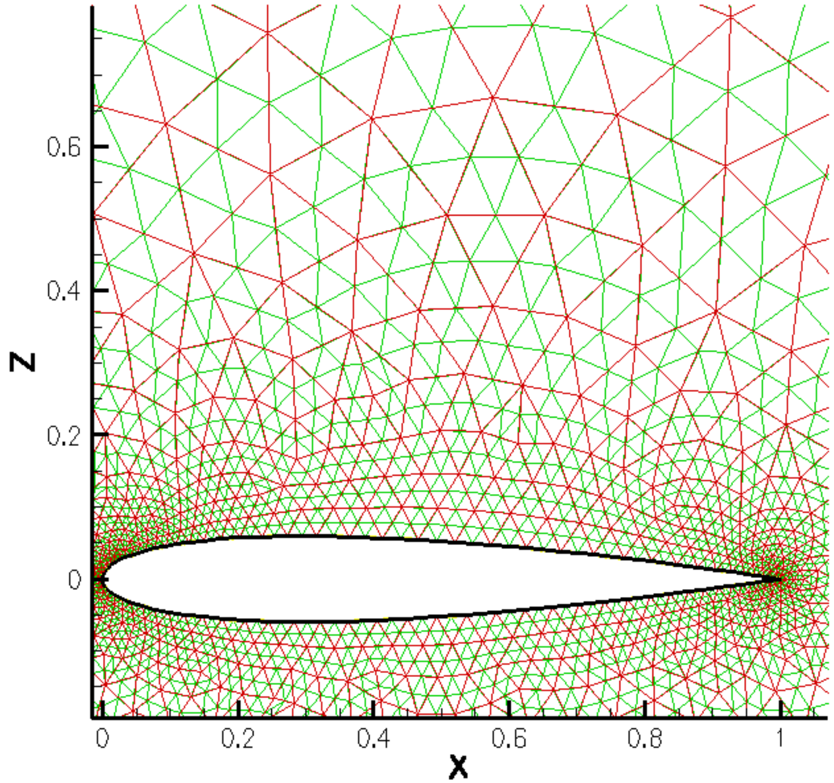
[Gauger, 2009]

Adjoint-based (Global) Error Estimate and Correction

NACA0012 (Euler flow)

Ma = 0.5,   α = 1.0°





Adaptation 1:

C_D(global refined mesh) = 0.00033382

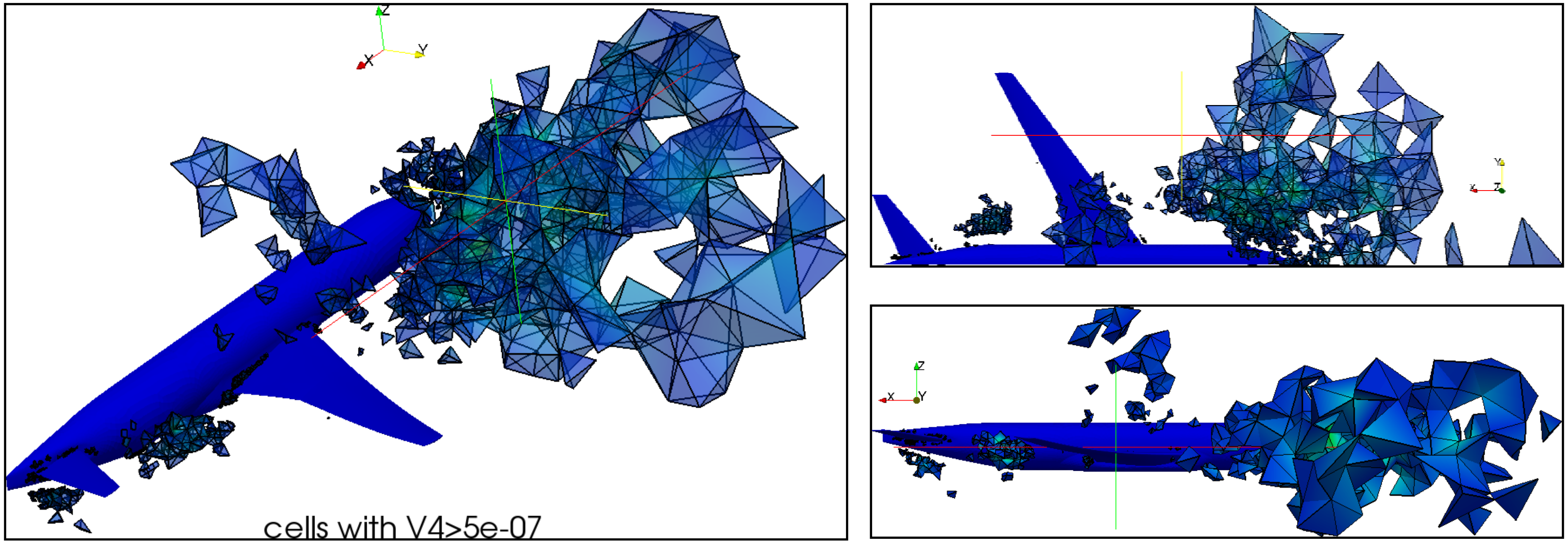
estimated error = Ψ*[(dRdk2 * k2) + (dRdk4 * k4)] = - 0.00024721

corrected C_D for global refined mesh = 0.00008661

[Gauger, 2009]

(Local) Adjoint-based Error Estimate

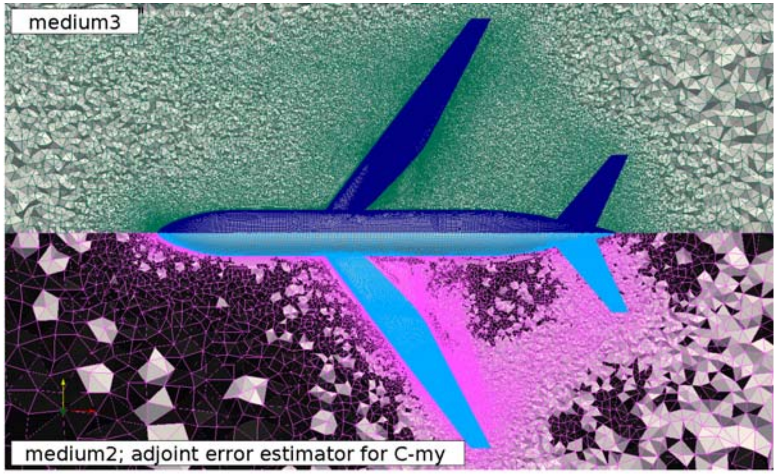
- Adjoint-based error estimates coupled to SOLAR
- Manually performed for the DPW4 meshes
- Error estimate based on the artificial viscosities k2 and k4
- Only field cells with sensor > 5e-07 are plotted in color



[Crippa, 2009]

Adjoint-based Error Estimate and Adaptation

Error estimator and mesh adaptation indicator based on adjoints and explicitly added artificial dissipation



HIRENASD Case (SFB 401)

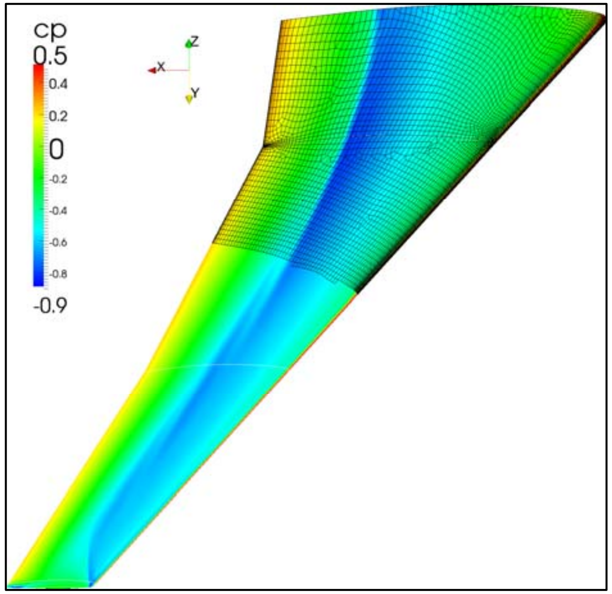
Cd (initial mesh) = 1.181e-02

Adjoint-based Adaptation:

Cd (7.5% of cells refined) = 1.179e-02

Cd (15% of cells refined) = 1.175e-02

Cd (globally refined mesh) = 1.171e-02



(Global) Error estimate for initial mesh = -1.431e-04

⇒ Corrected value for initial mesh = 1.167-02

[Gauger, Orlt, 2010]

Outlook

- Redo one-shot theory in function spaces for coupling with adaptivity
- Combine one-shot with adaptivity in FVM as well as DG context (DLR PADGE Code used in BMBF Project DGHPOPT)
- …

Thanks!