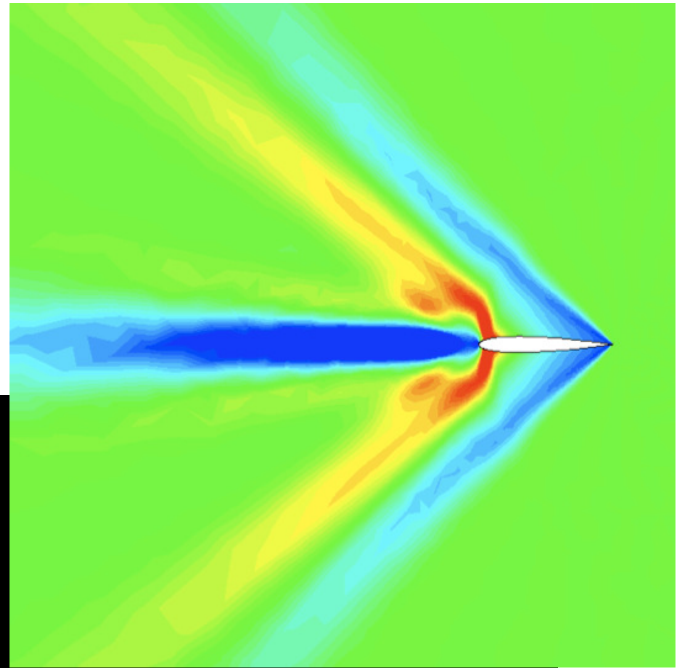
Applications of an Adjoint Capability



Richard Dwight, Jouke de Baar, Dirk-Jan Boon,
Hester Bijl, Fulvio Scarano
Aerodynamics Group TU Delft

Adjoint applications

1. Adjoint solution fields

2. Applications
   - Error estimation and goal-oriented adaptation
   - Response surfaces in high-dimensional spaces
     - Kriging
     - Global optimization
     - Uncertainty quantification
   - 4D-var data assimilation
   - Frequency-domain solvers

Adjoint solution fields

- Consider the PDE-constrained minimization:

$$\min_{\alpha \in \Omega} I(\mathbf{w}, \mathbf{x}, \alpha)$$

- subject to

$$R(\mathbf{w}, \mathbf{x}, \alpha) = 0$$

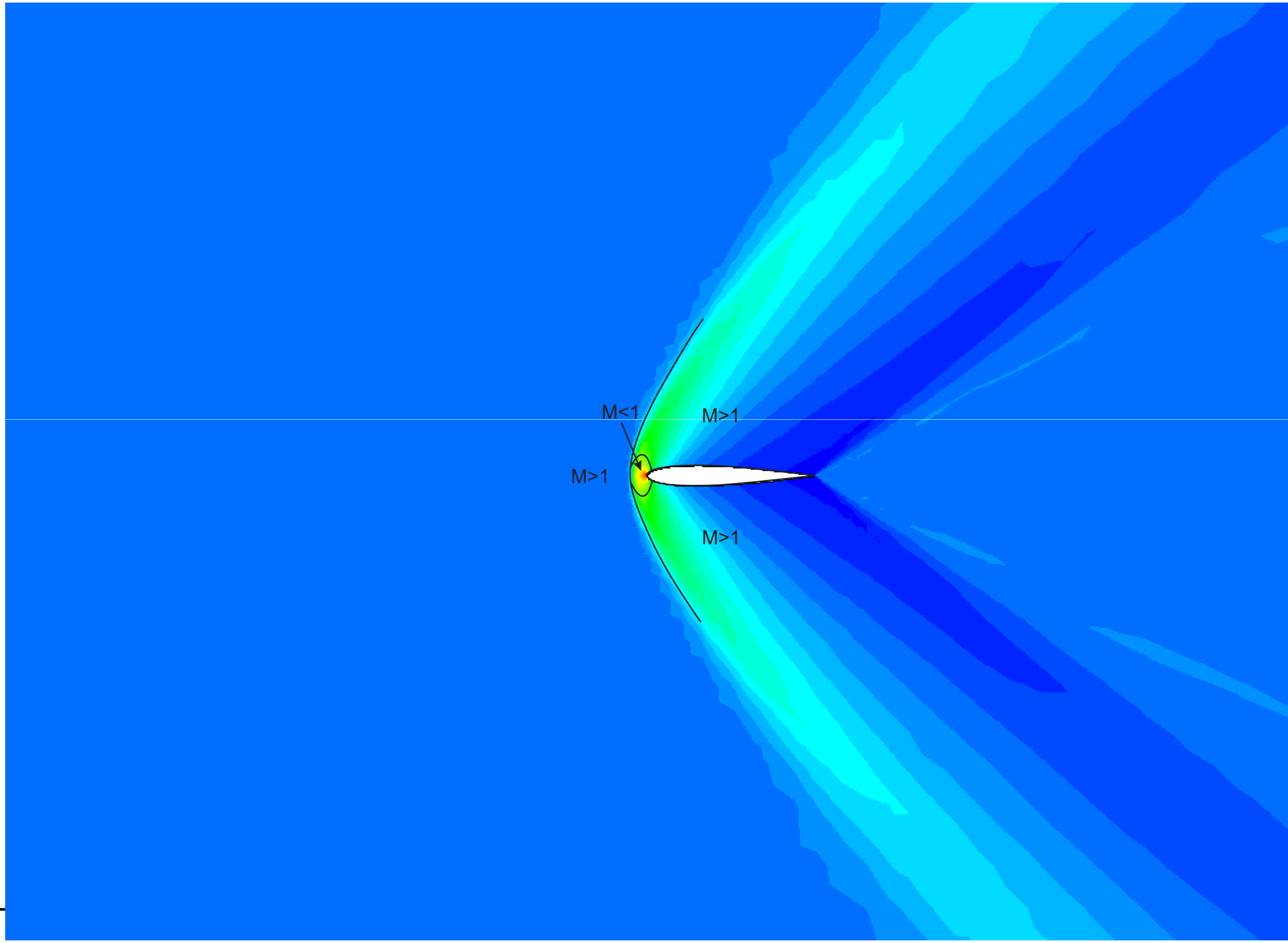
with W the flow variables, X the mesh and D the design variables.

- **Goal:** determine the derivatives of I with respect to *D.*
- We define the Lagrangian which is identical to I and its derivatives with respect to the design variables *D*:

$$L = I + \Lambda^T R$$

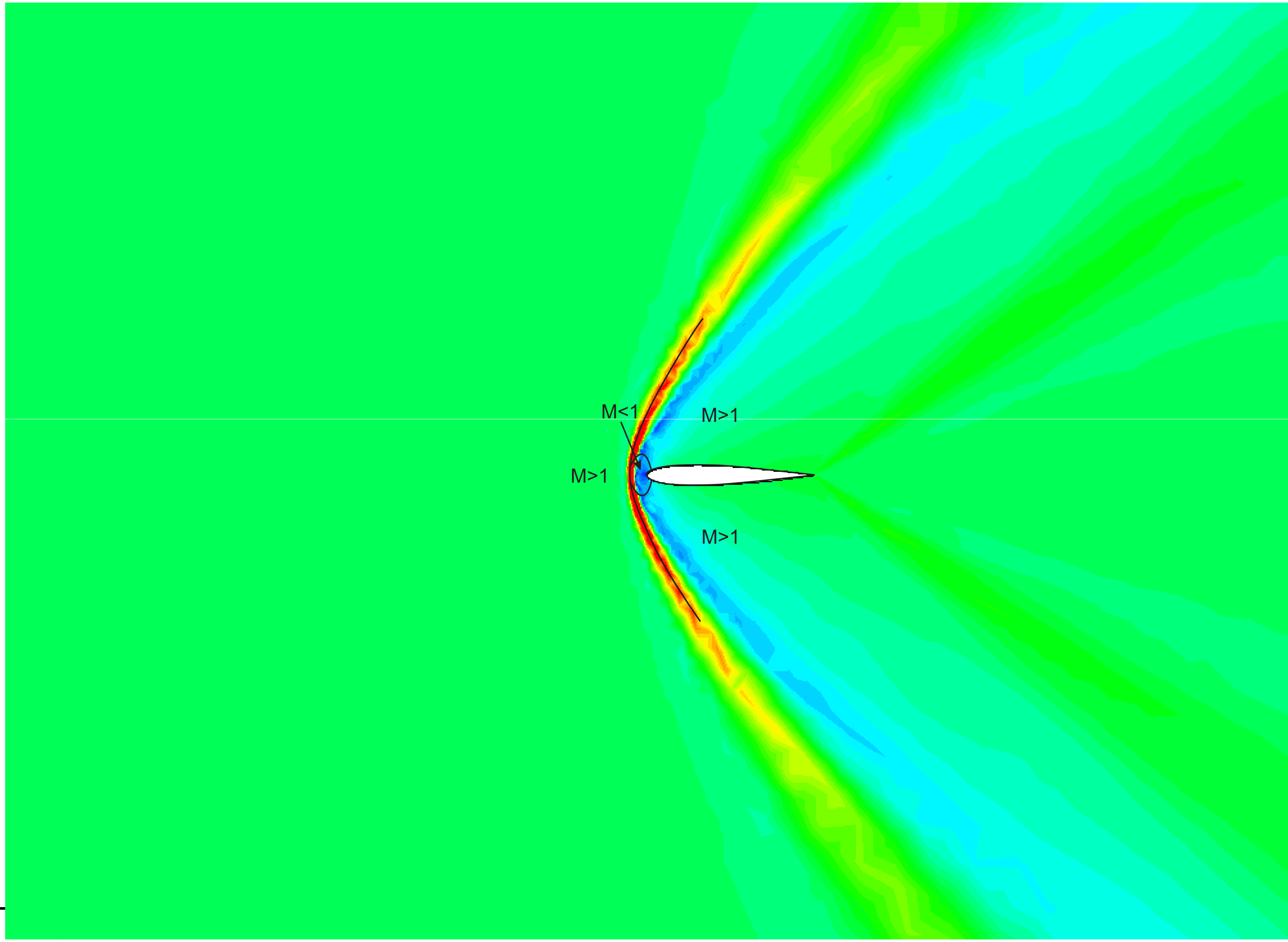
Supersonic NACA0012, Mach=1.5

Information transport in hyperbolic problems



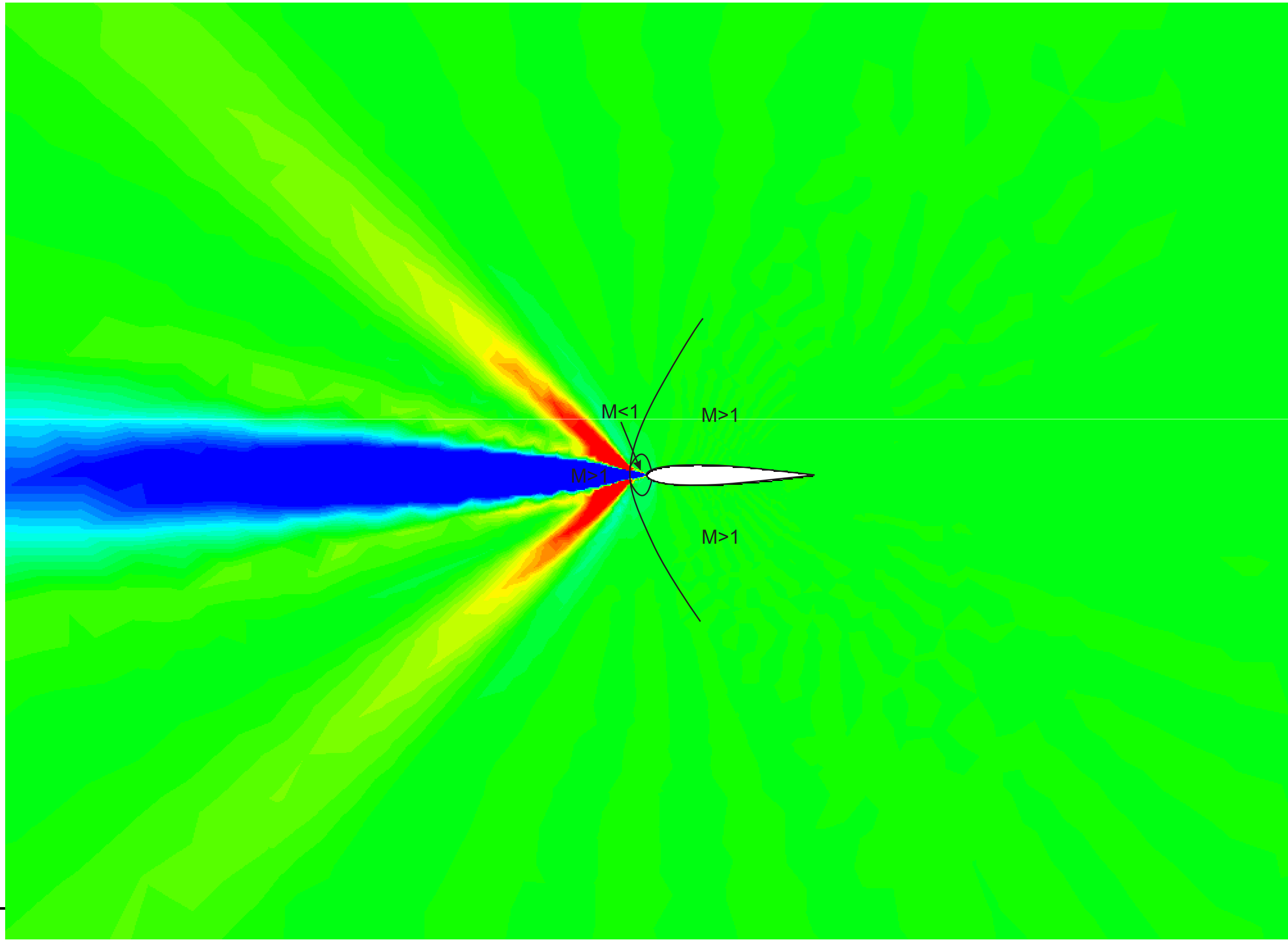
Supersonic NACA0012 Primal

Design parameter = Mach number



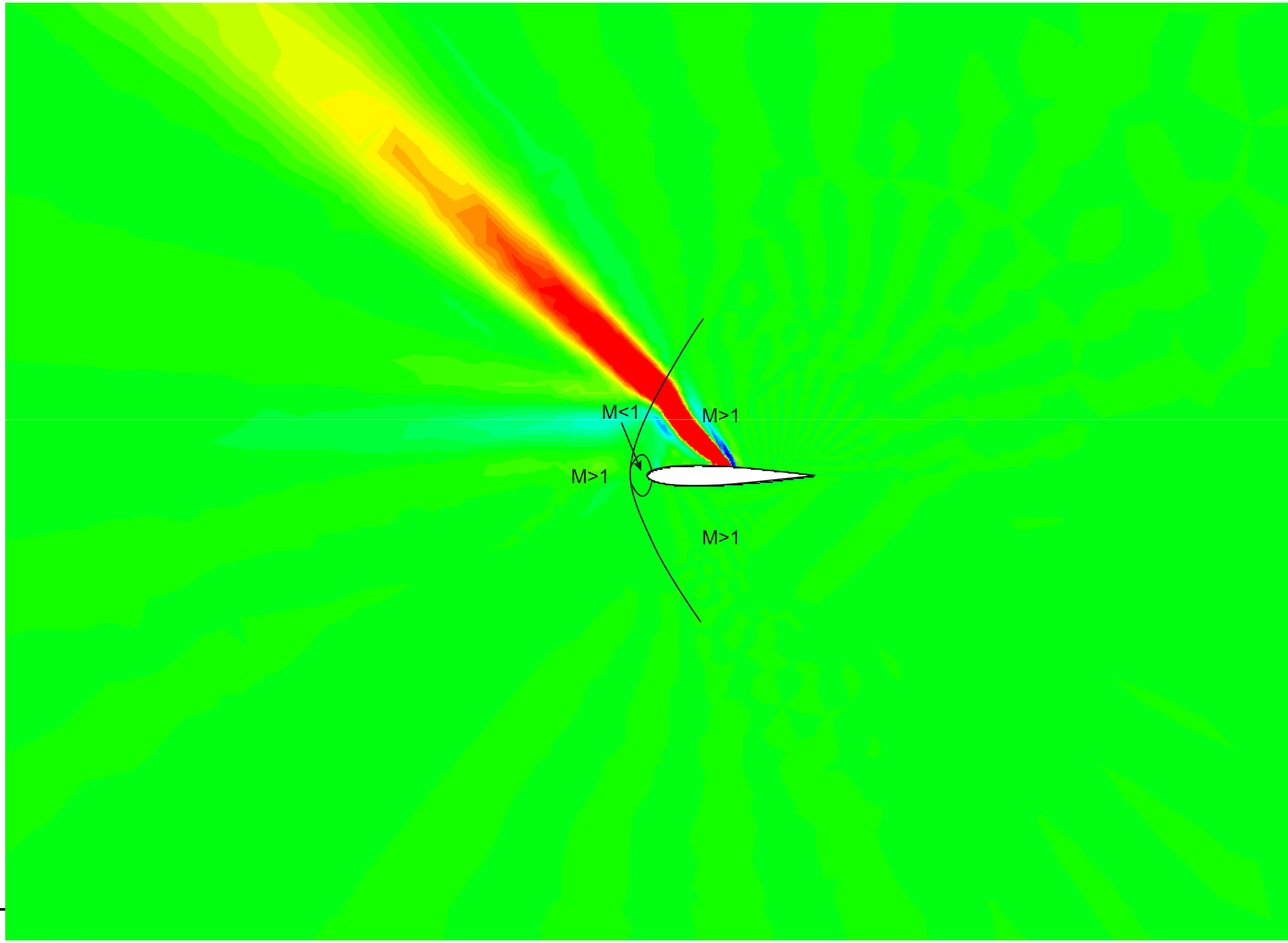
Supersonic NACA0012 Adjoint

Cost function = Stagnation pressure



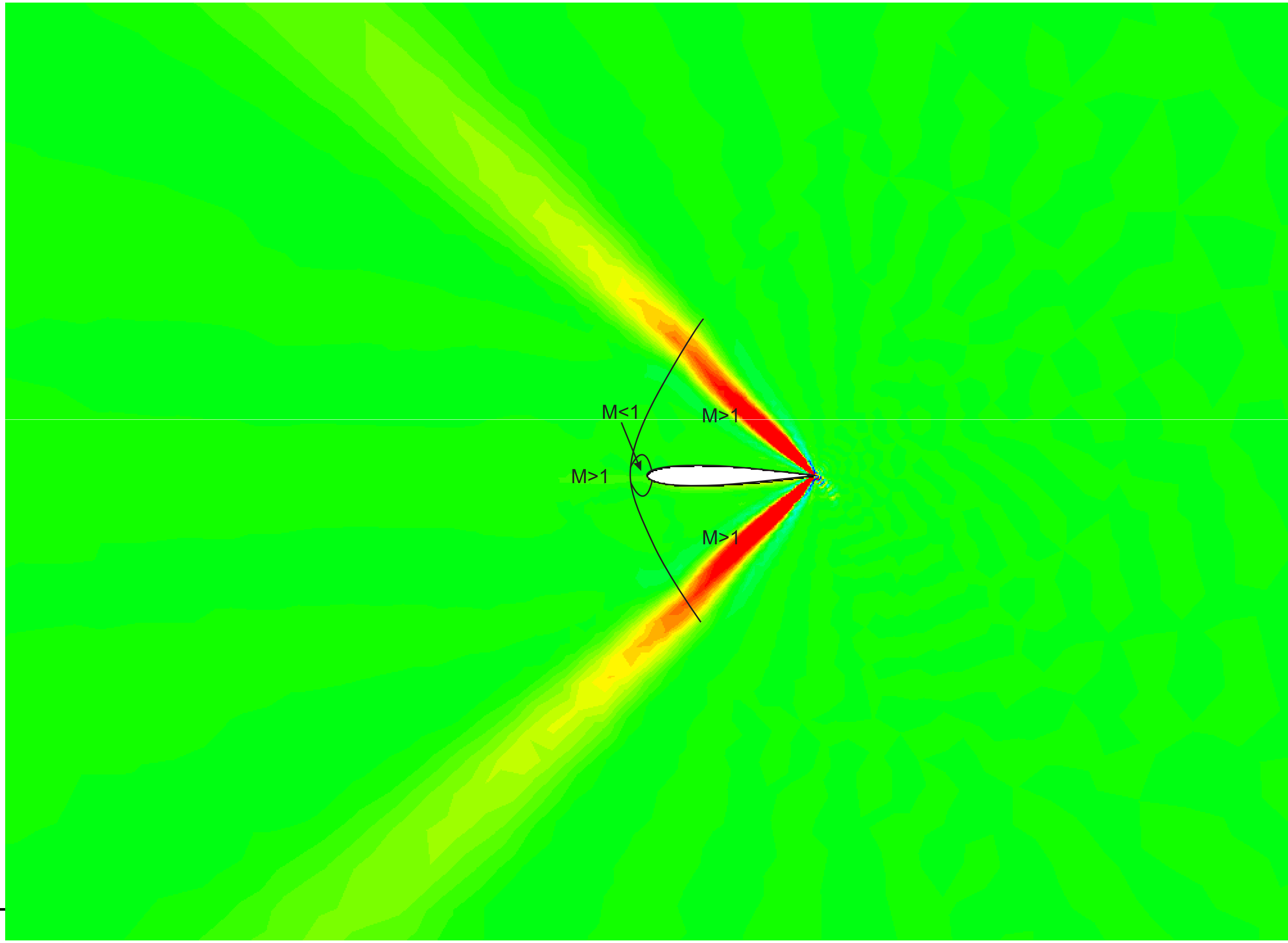
Supersonic NACA0012 Adjoint

Cost function = Point pressure



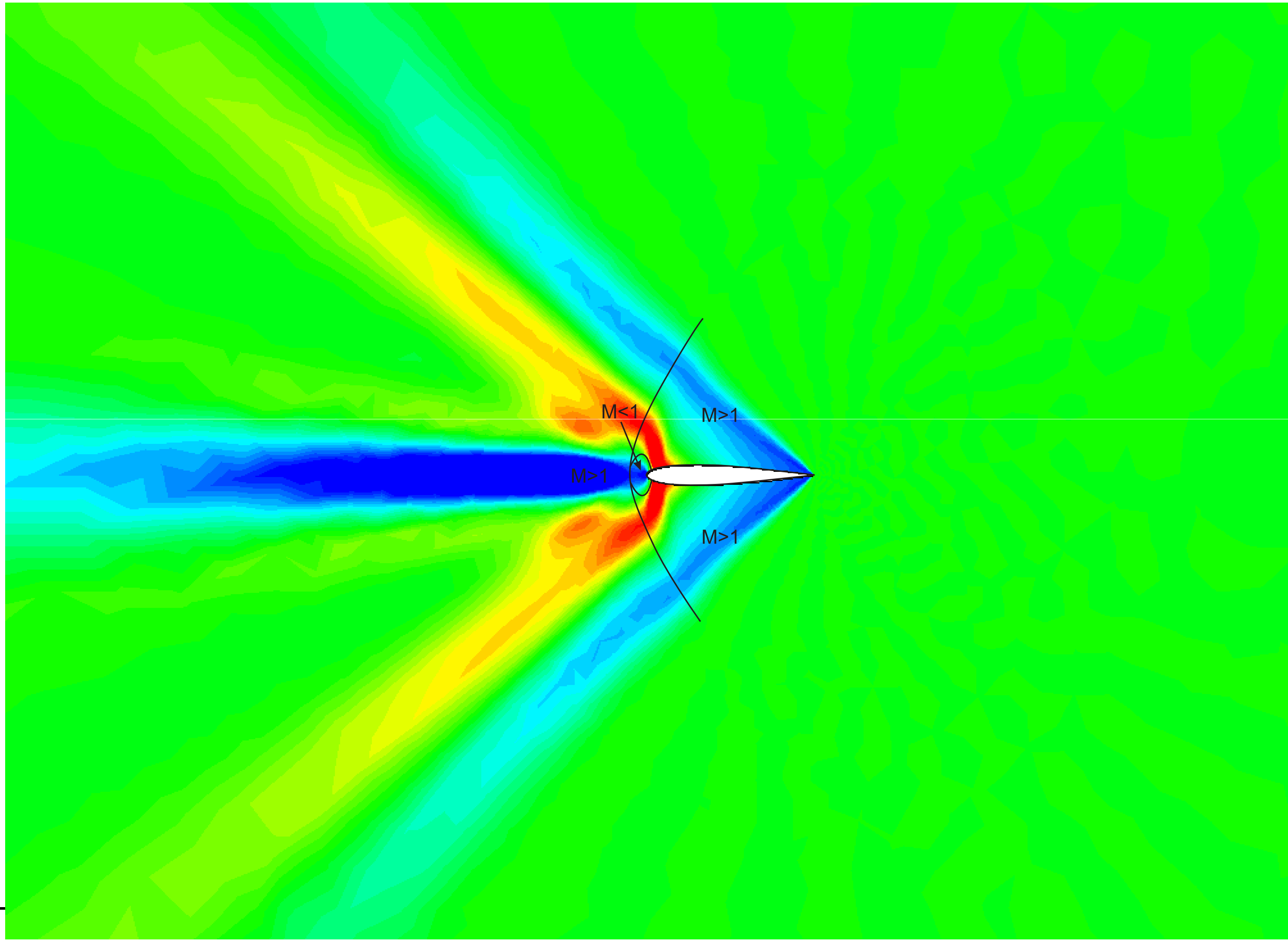
Supersonic NACA0012 Adjoint

Cost function = Trailing edge pressure

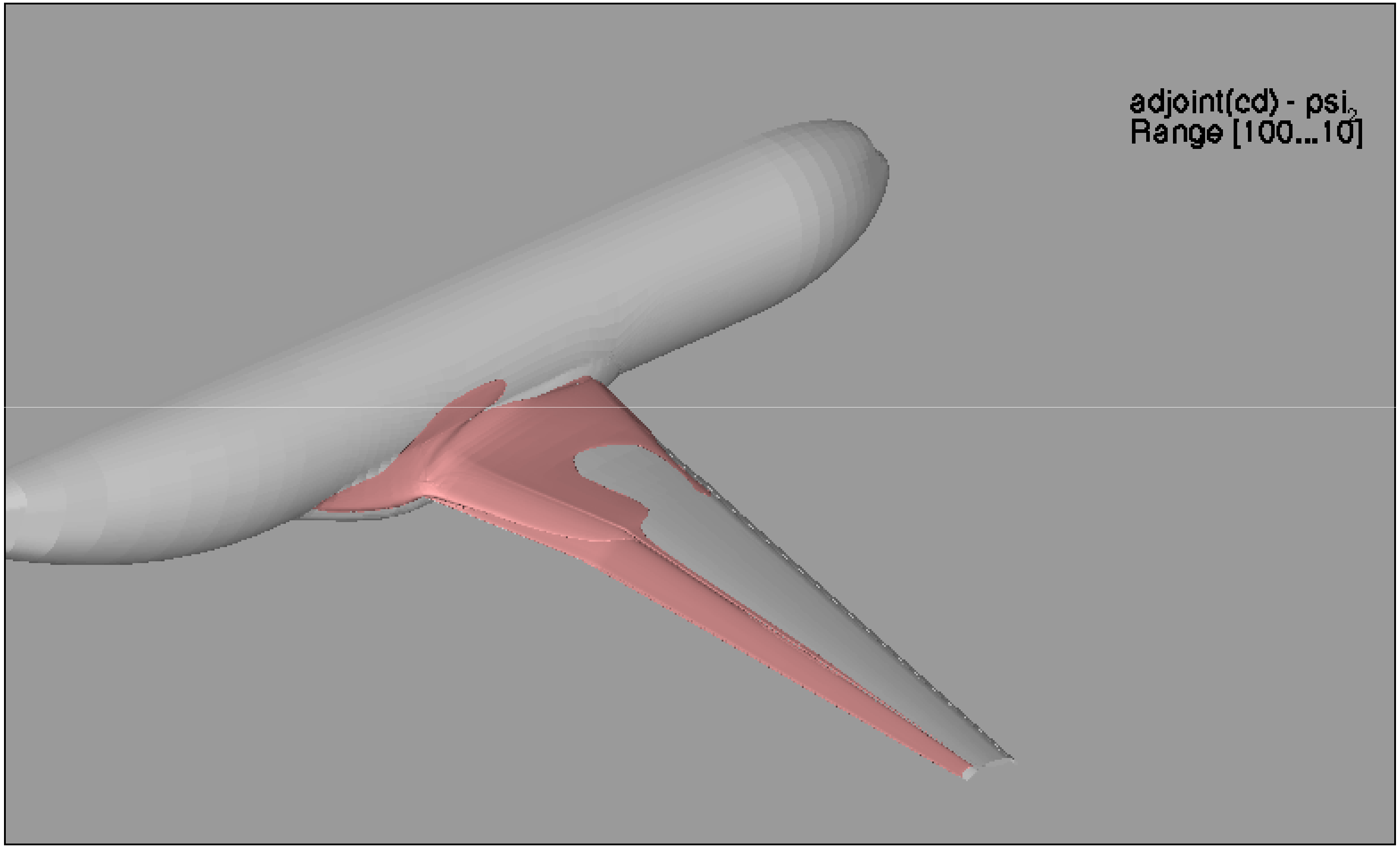


Supersonic NACA0012 Adjoint

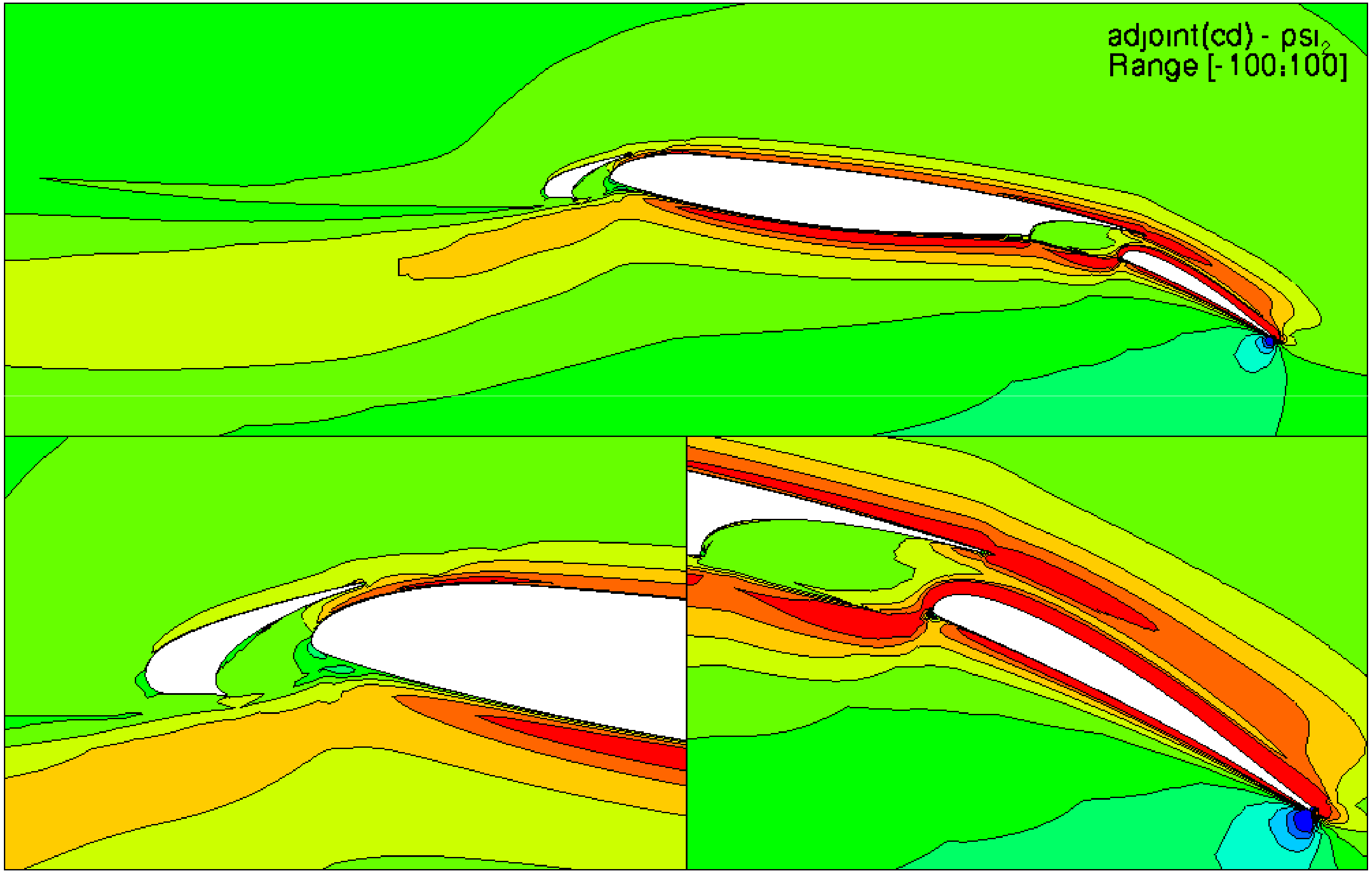
Cost function = Drag



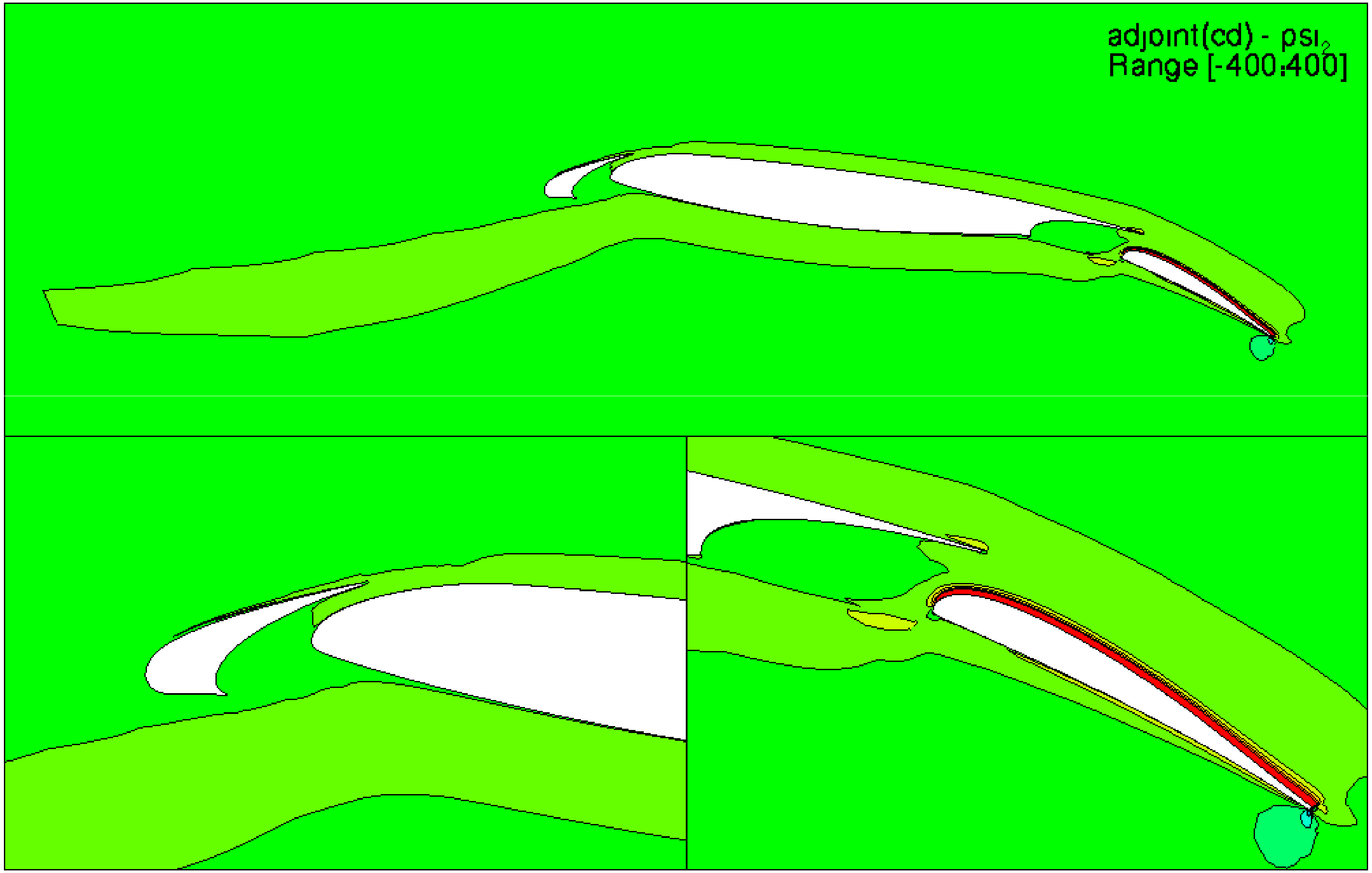
3d High-lift Gradient Evaluation
Adjoint Field



3d High-lift Gradient Evaluation
Adjoint Field



3d High-lift Gradient Evaluation
Adjoint Field

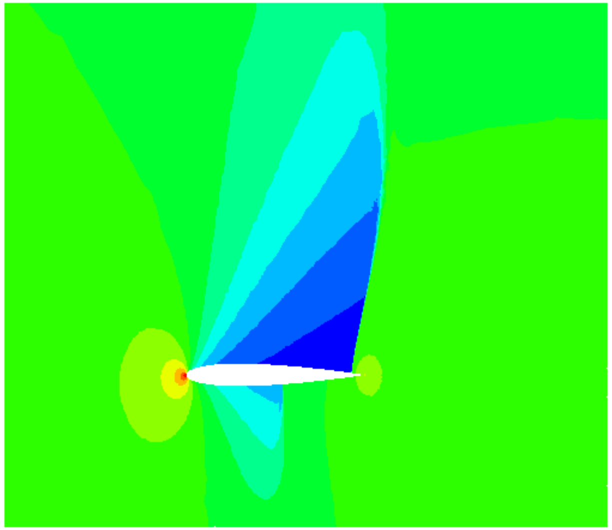


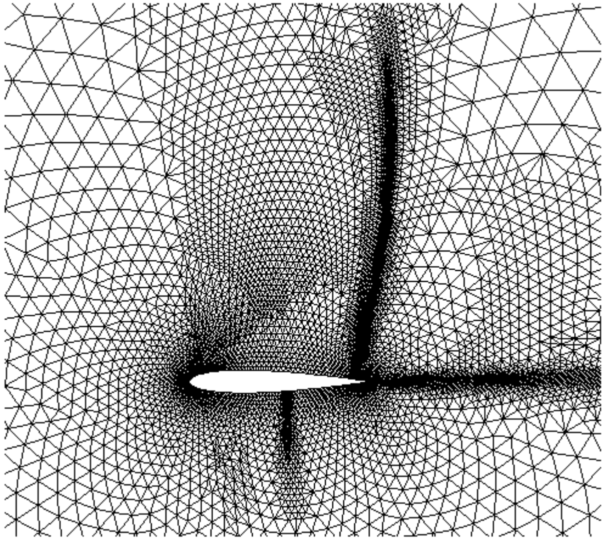
Overview of applications – *as of today*

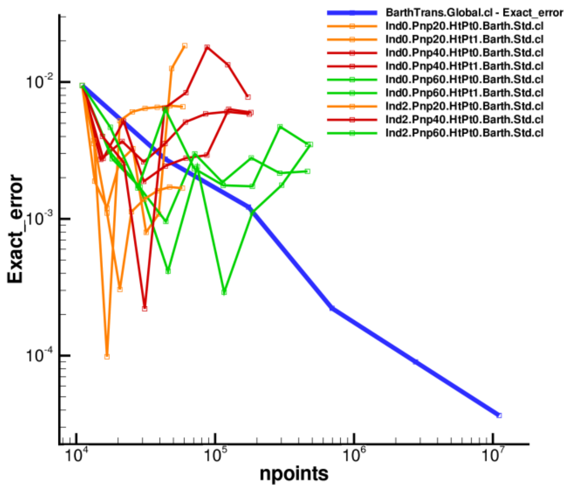
- Shape optimization (Pironneau 1973, Jameson 1988, Eleshaky 1997)
- Efficient global optimization (EGO) (Ely 2004, Ghattas 2009)
- Error estimation
  - *A posteriori* (Giles 1997, Becker and Rannacher 1998)
  - Error transport equation (Roache 1998, Cesnik 2002)
- Goal-oriented mesh adaptation (Darmofal 2004, Barth 2005, Dwight 2007)
  - Frequency-domain methods
  - Linear frequency-domain solver (Campobasso 2007)
  - Schur's-complement eigenvalue solver (Badcock 2010)
- Uncertainty quantification
  - First-order moment methods
  - Parameter reduction (Loeven 2009)
  - Surrogate modelling (Mavriplis 2010, Dwight 2010)
- Inverse problems + data assimilation
  - Deterministic (4D-var, Least-squares) (Dwight 2011)
  - Stochastic (McMC, Hessian) (Fox 2009, Ghattas 2010)
- Petrov-Galerkin Reduced-order modelling (ROMs) (Farhat 2010)

Error estimation + Goal-oriented adaptation

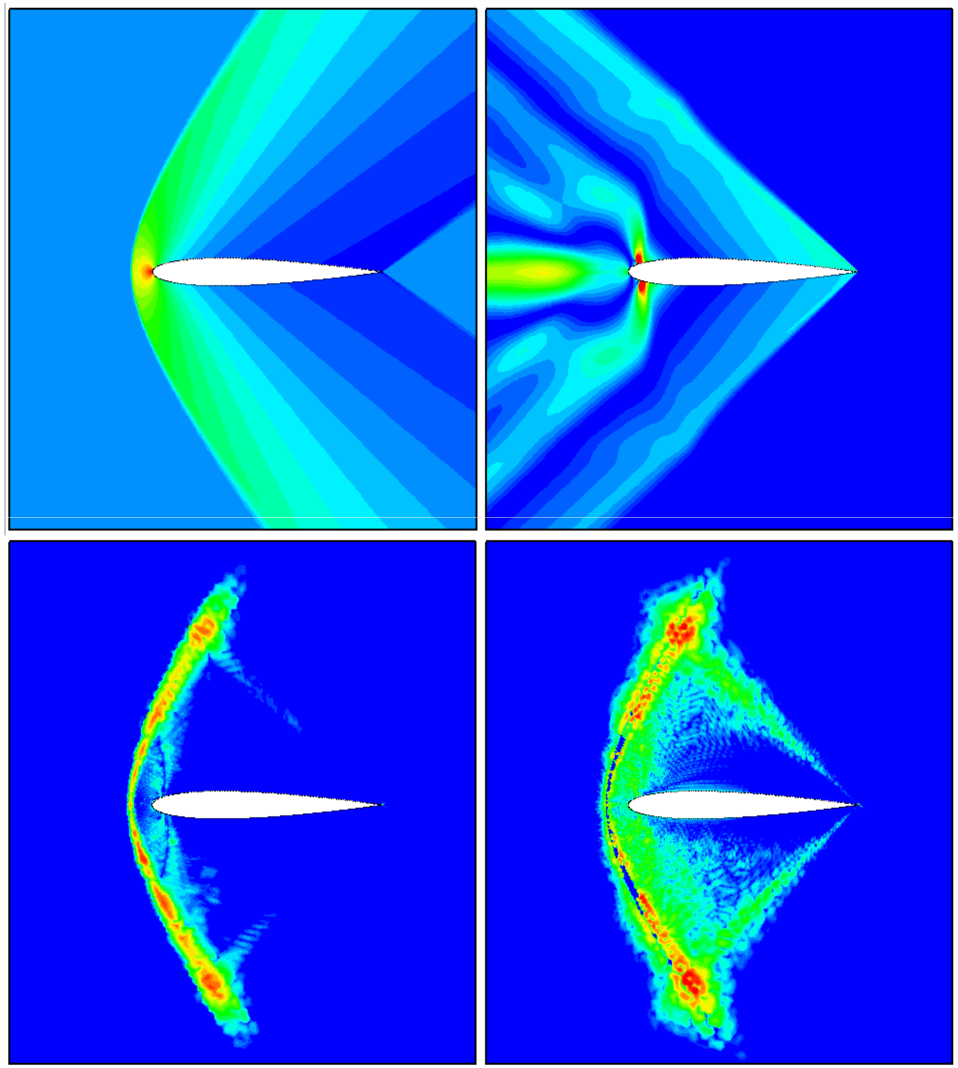
Problem: Local gradient adaptation *doesn't work* for hyperbolic problems (e.g. Euler equations):







Goal-oriented adaptation: Idea



A Posteriori Error Estimation

Error representation for linear problems

- Consider the linear problem:

$$\mathcal{L}u = f,$$

- and let the *functional of interest* be of the form:

$$J = (g, u),$$

- where the inner product may be defined by e.g.:

$$(u, v) = \int_{\Omega} u \cdot v \, \mathrm{d}\Omega.$$

A Posteriori Error Estimation

Error representation for linear problems

- Also define the adjoint operator *L** of *L* as:

$$(g, u) = (\mathfrak{L}^* v, u) = (v, \mathfrak{L} u) = (v, f),$$

- giving the adjoint problem:

$$\mathfrak{L}^* v = g,$$

- Now for a linear problem the error satisfies the same eqn as the solution:

$$\mathfrak{L} \epsilon_h = r_h,$$

- where:

$$\epsilon_h := (u - u_h) \qquad r_h := f - \mathfrak{L} u_h$$

TUDelft

A Posteriori Error Estimation

Error representation for linear problems

- But the error in J can also be repesented in terms of the solution of the adjoint problem:

$$\begin{aligned} J(u) - J(u_h) &= (g, u) - (g, u_h) \\ &= (g, u - u_h) \\ &= (\mathfrak{L}^* v, u - u_h) \\ &= (v, \mathfrak{L}(u - u_h)) \\ &= (v, f - \mathfrak{L} u_h) \\ &= (v, r_h), \end{aligned}$$

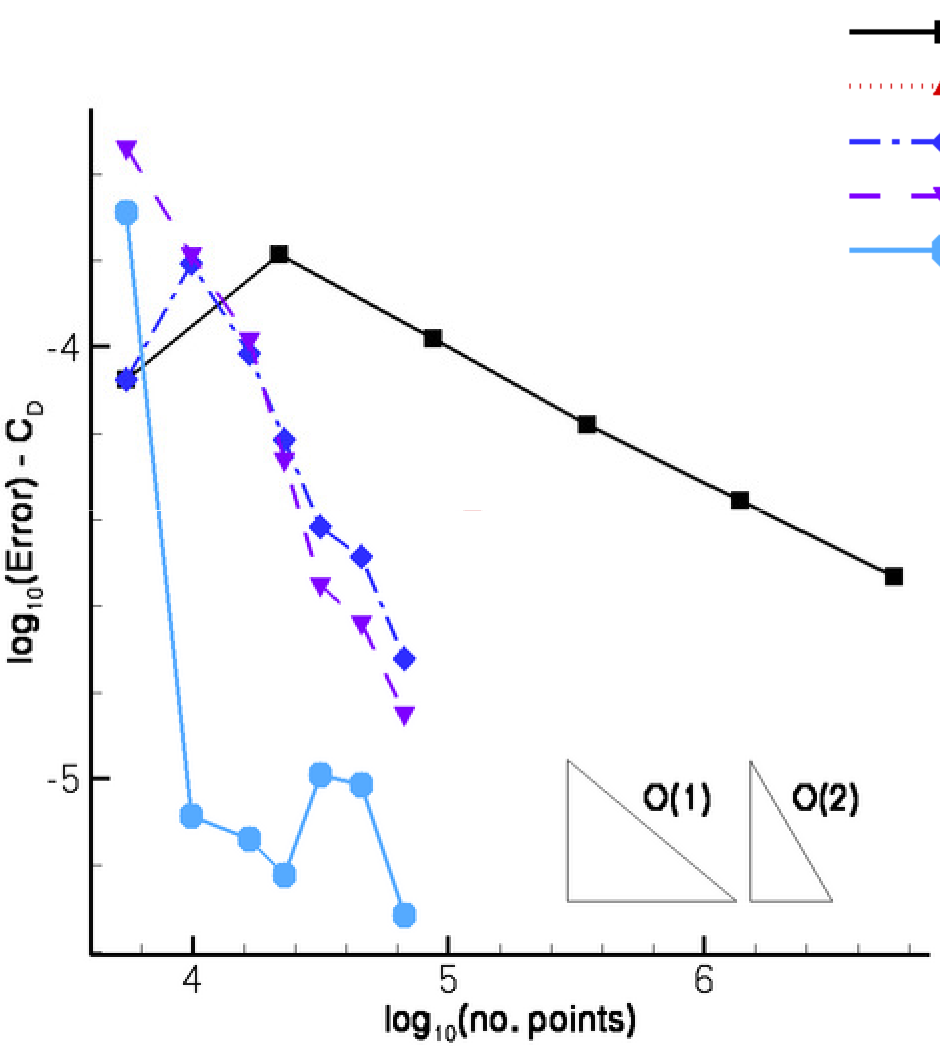
- So two formulations for the error are:

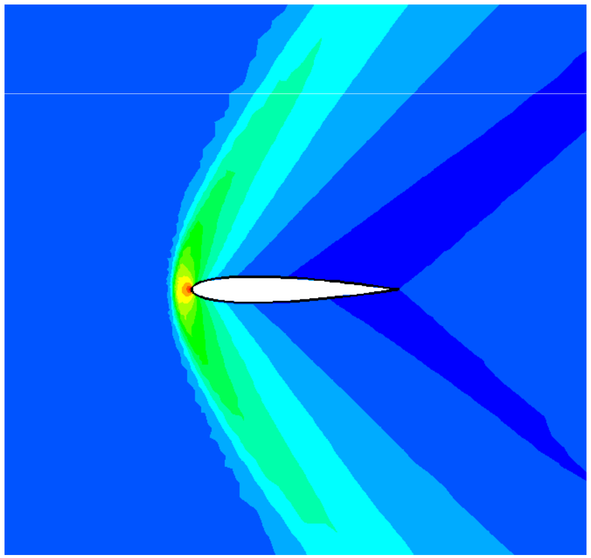
$$J(u) - J(u_h) = (g, \epsilon_h) = (v, r_h),$$

- Compare with the dual equivalence relation:

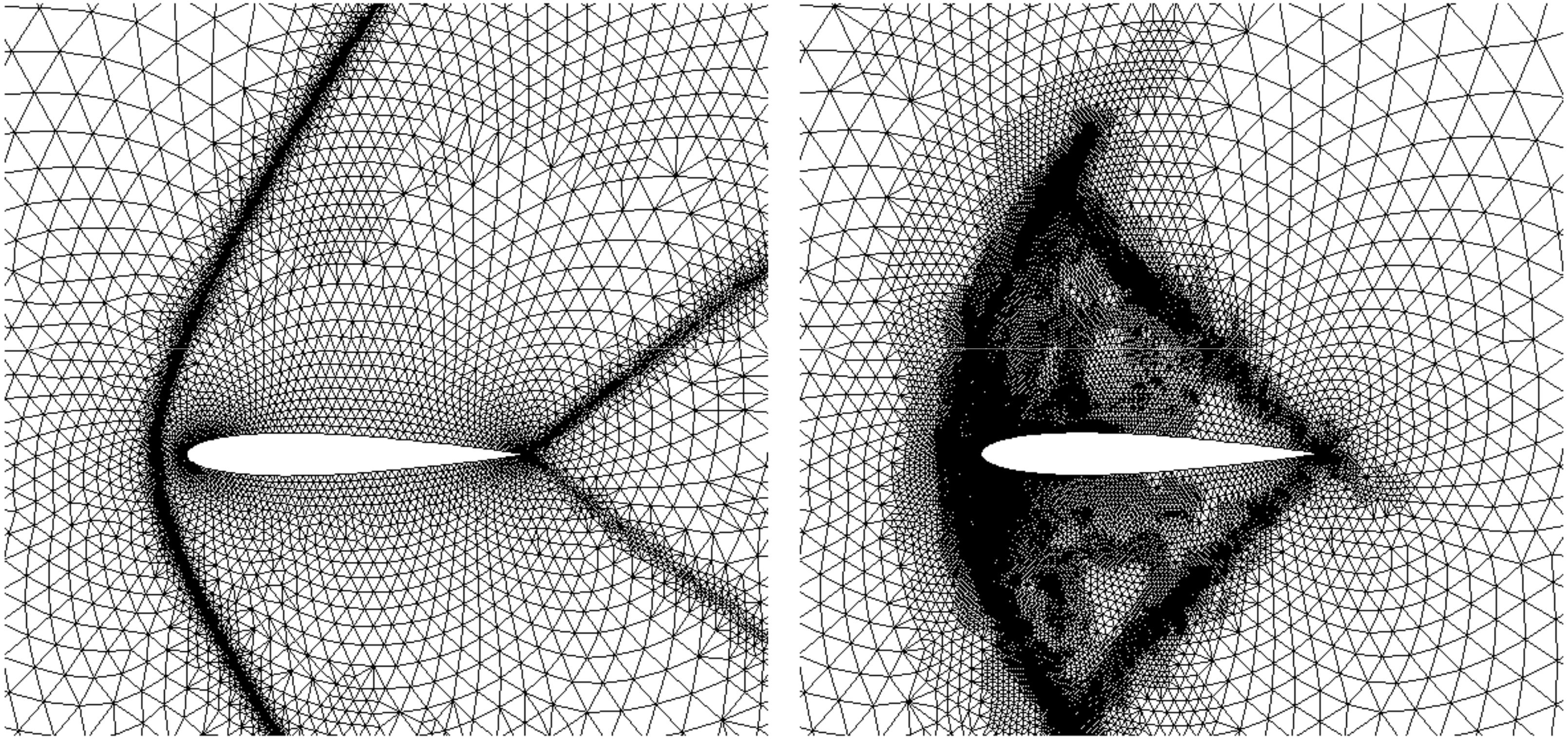
$$(g, u) = (\mathfrak{L}^* v, u) = (v, \mathfrak{L} u) = (v, f),$$

Convergence of error: Drag

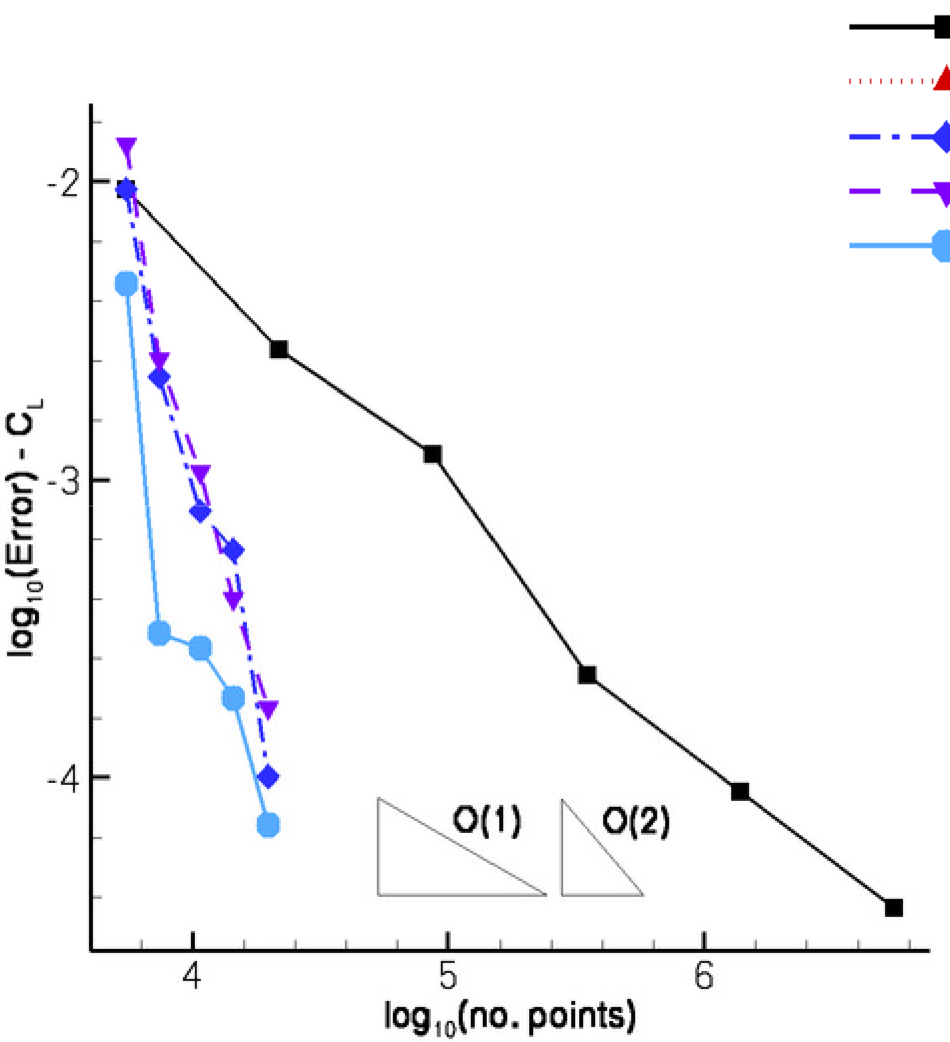


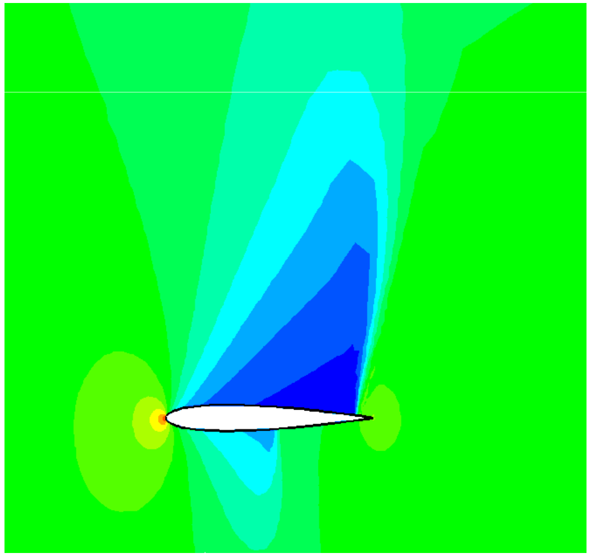


Location of error: Drag

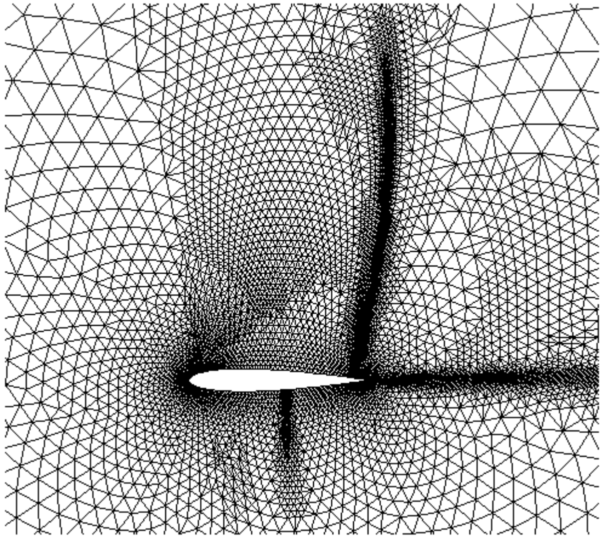


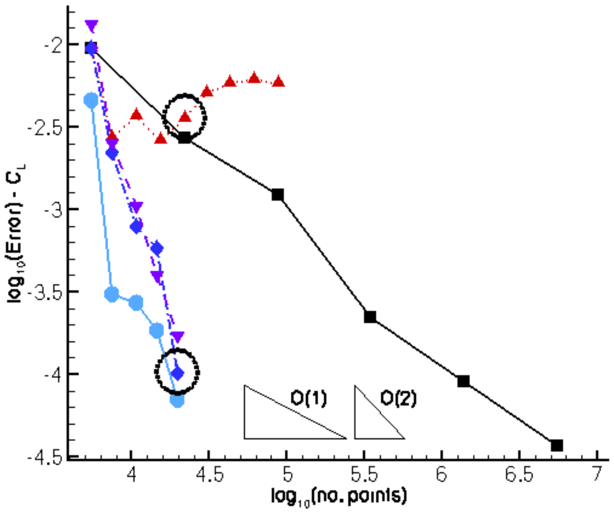
Convergence of error: Lift

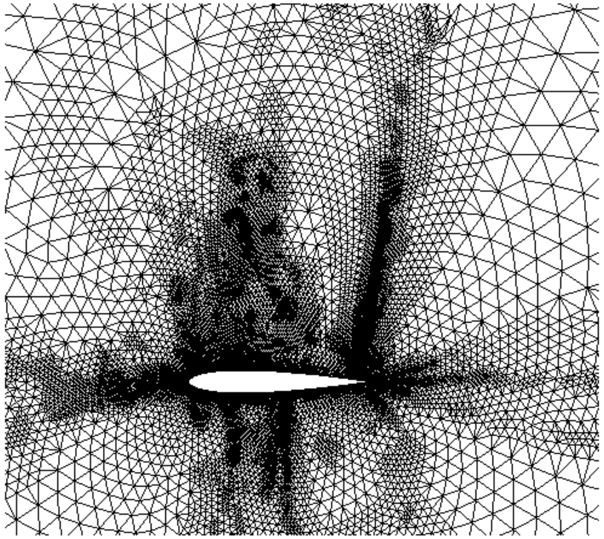


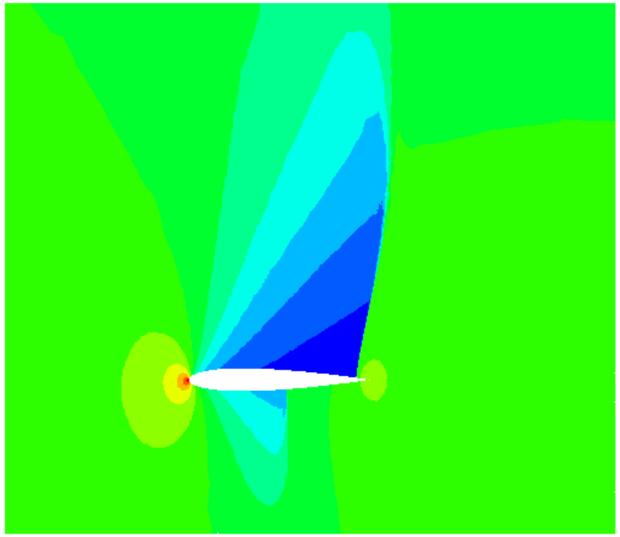


Location of error: Lift

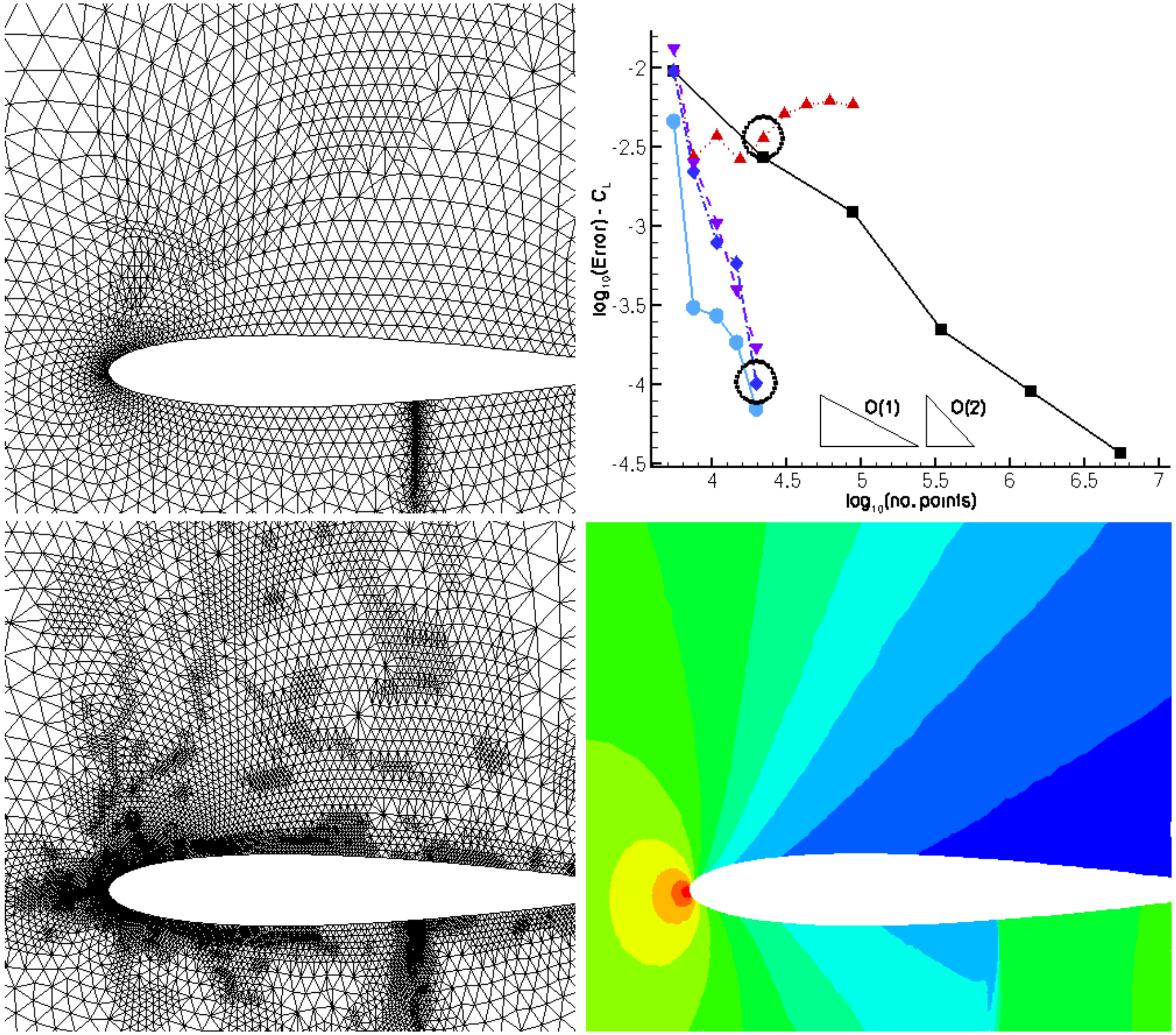




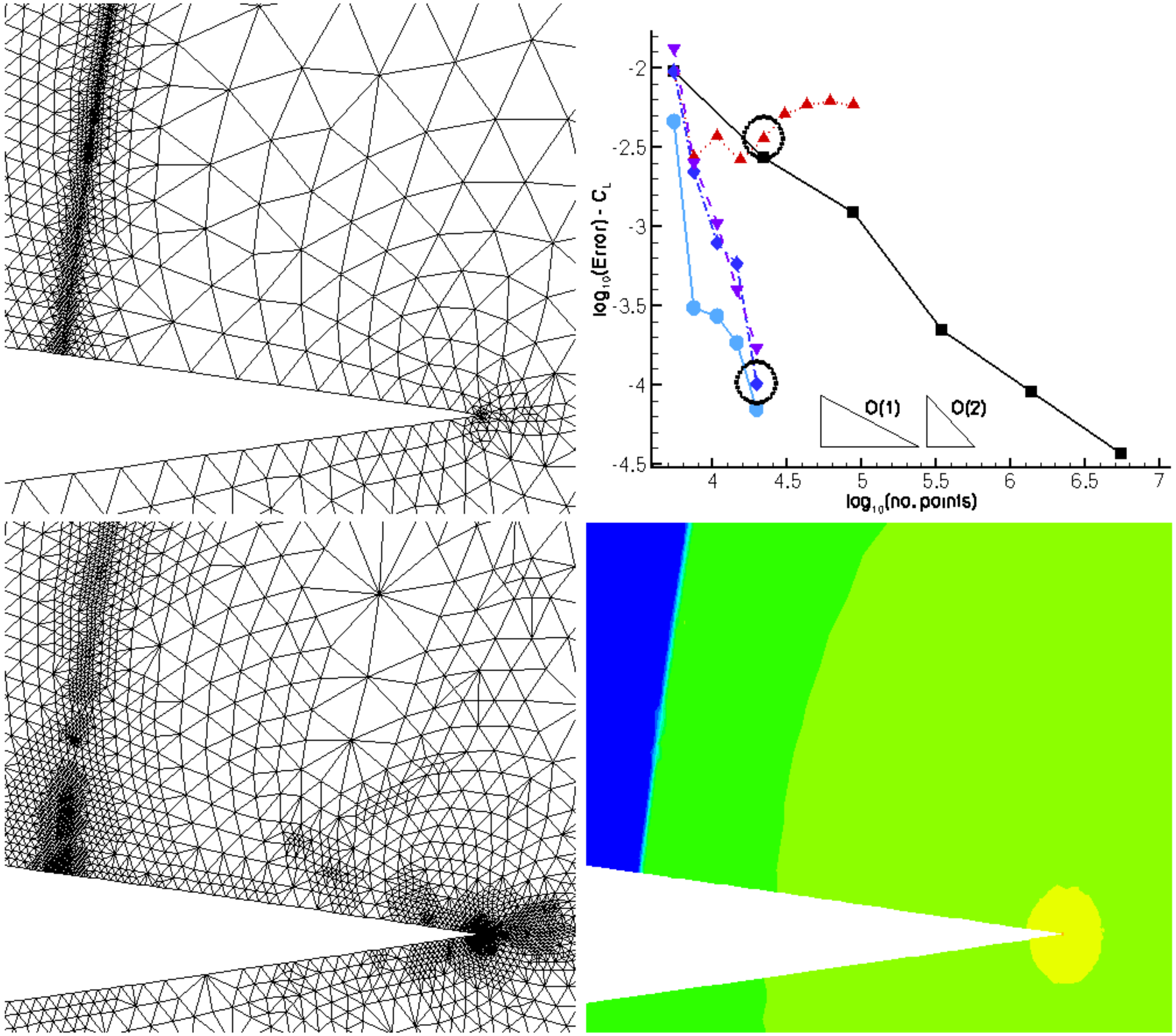




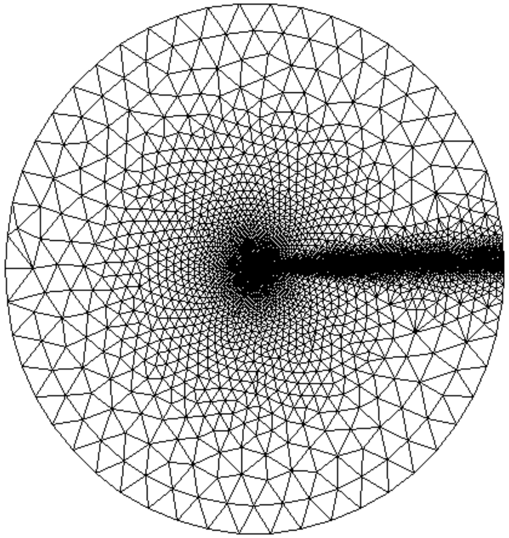
Location of error: Lift

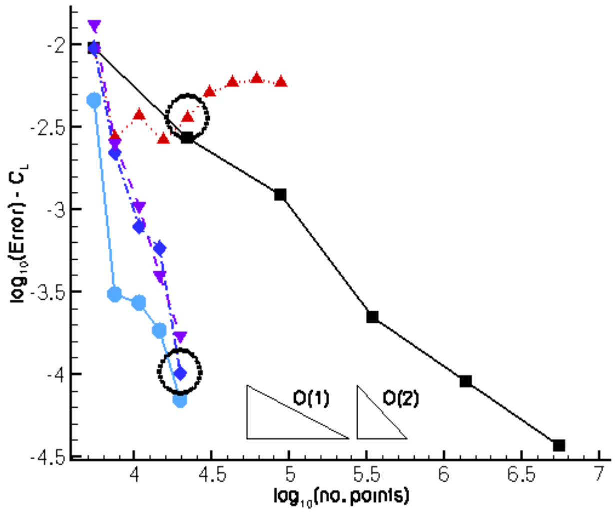


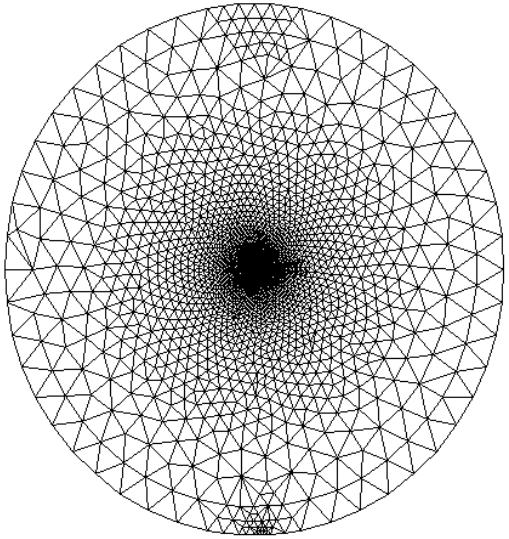
Location of error: Lift

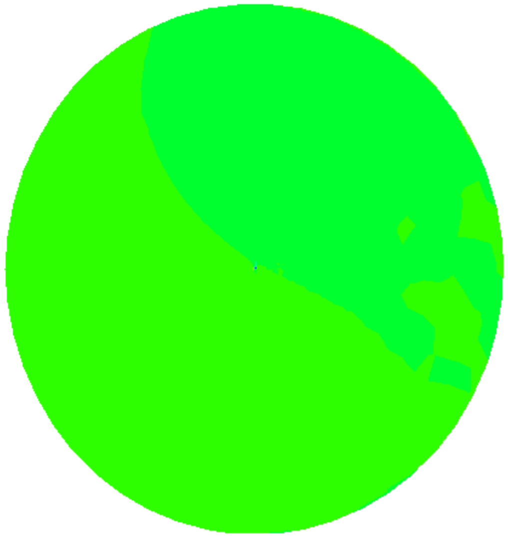


Location of error: Lift

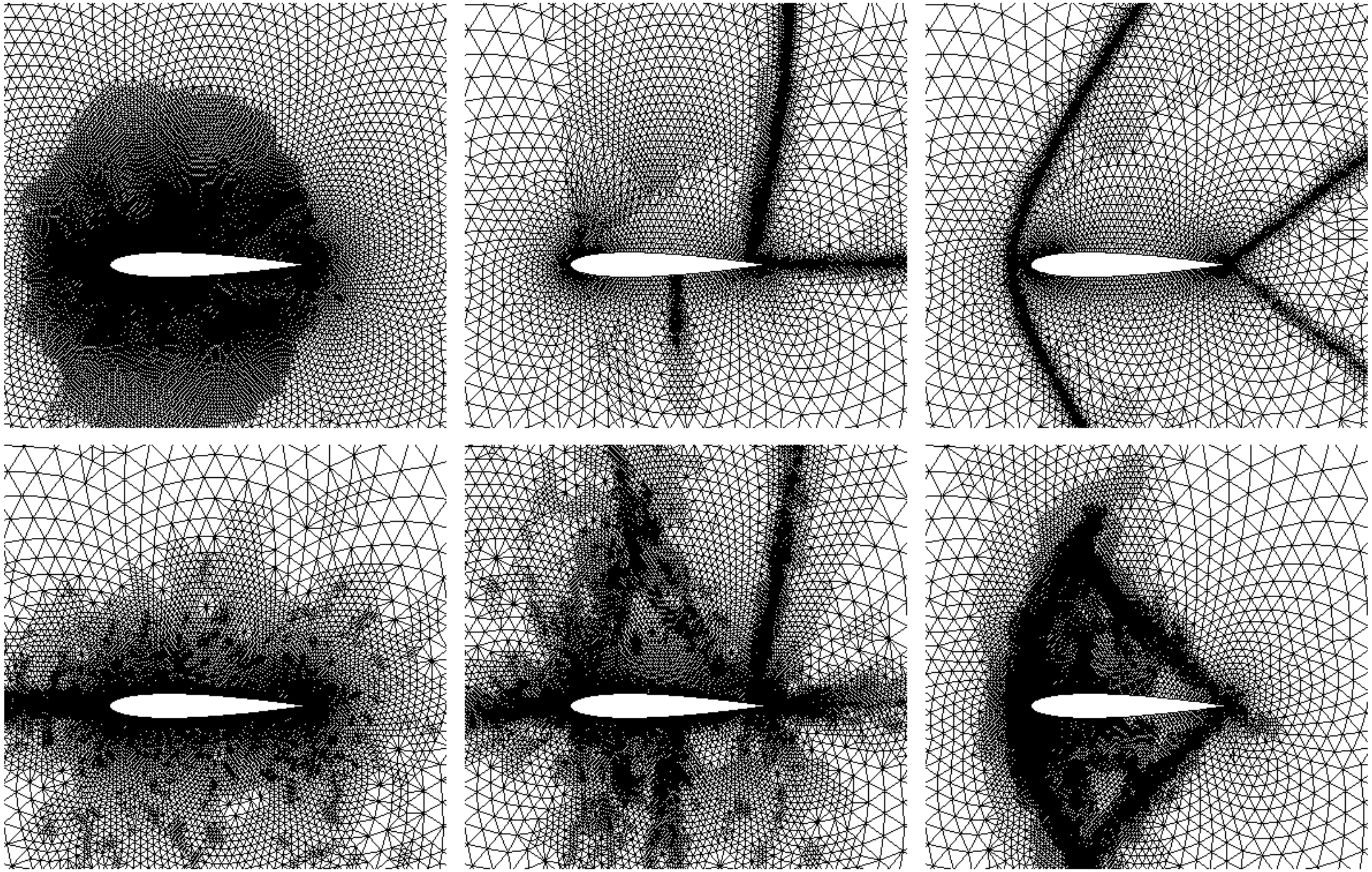




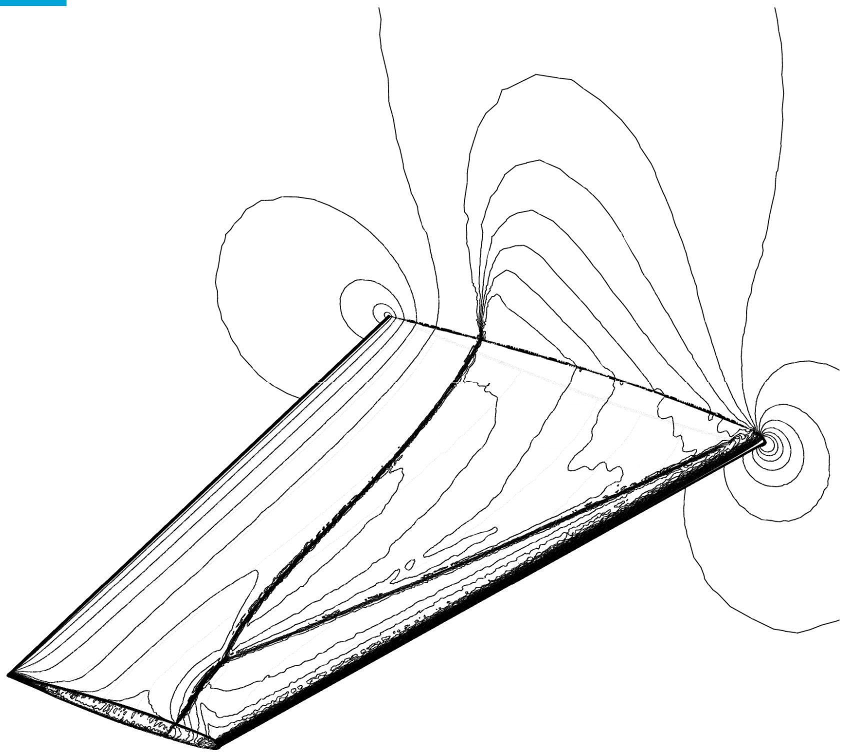


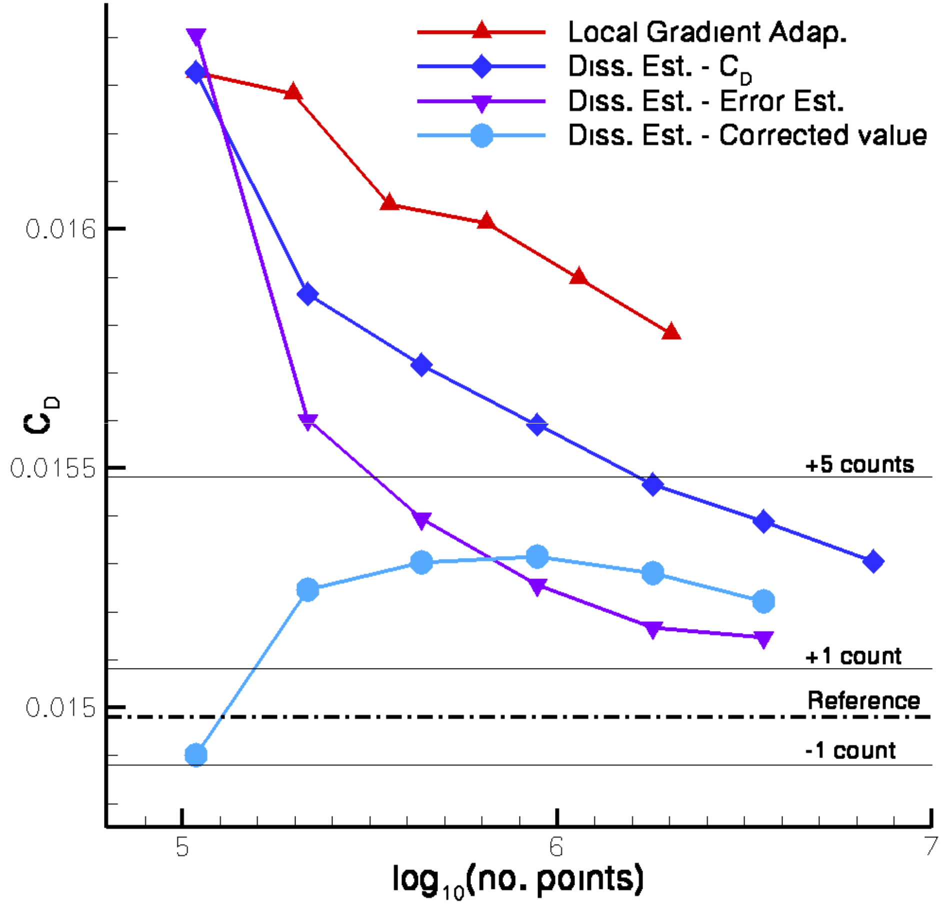


Convergence of all cases: Lift

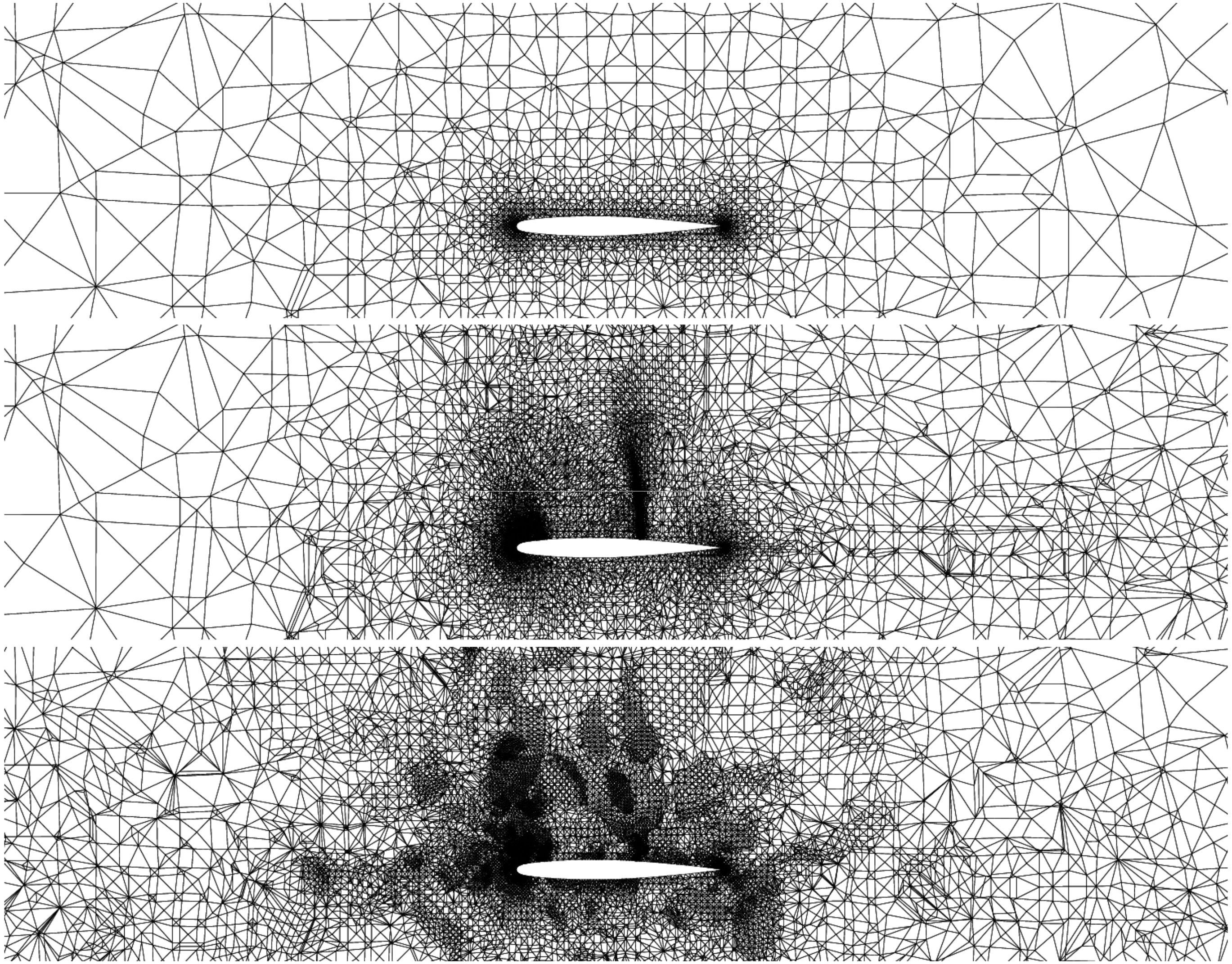


Convergence of error: ONERA M6

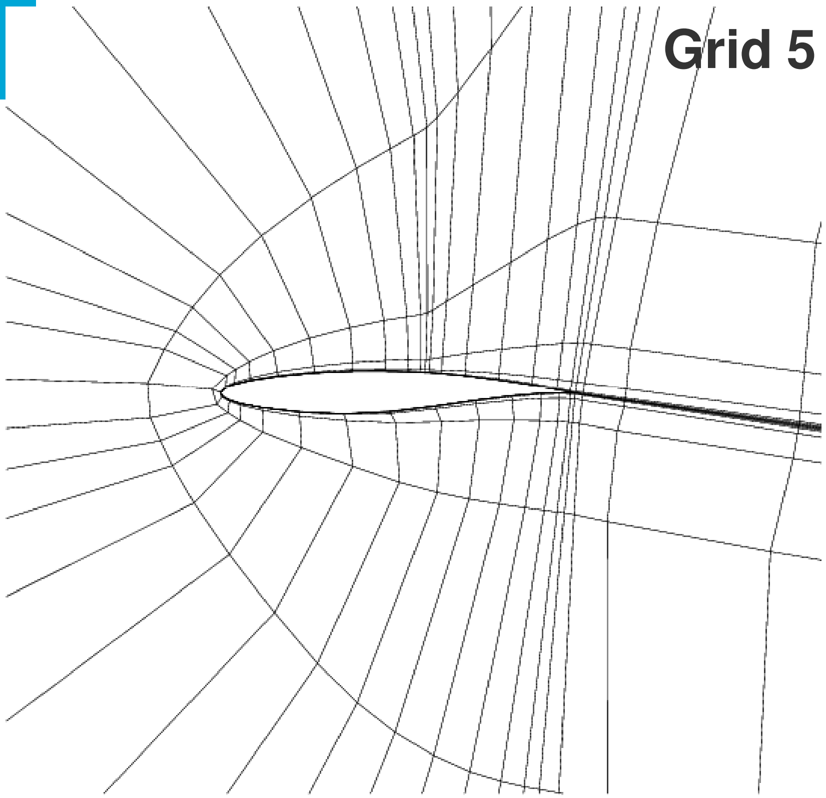


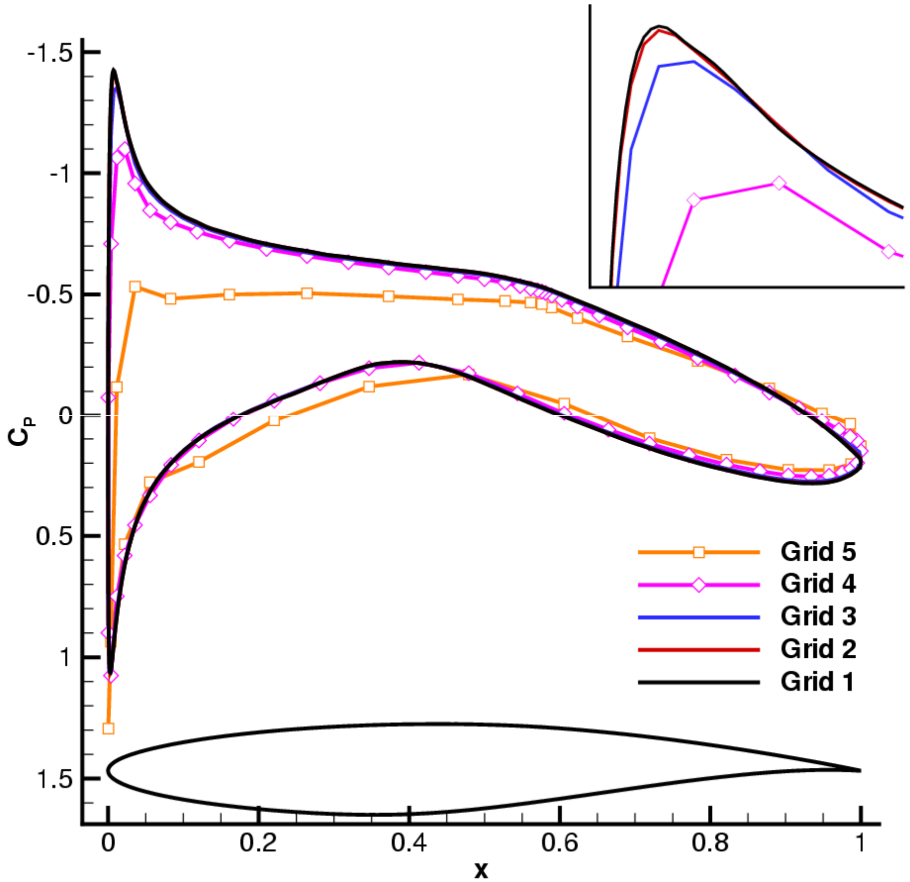


Convergence of error: ONERA M6

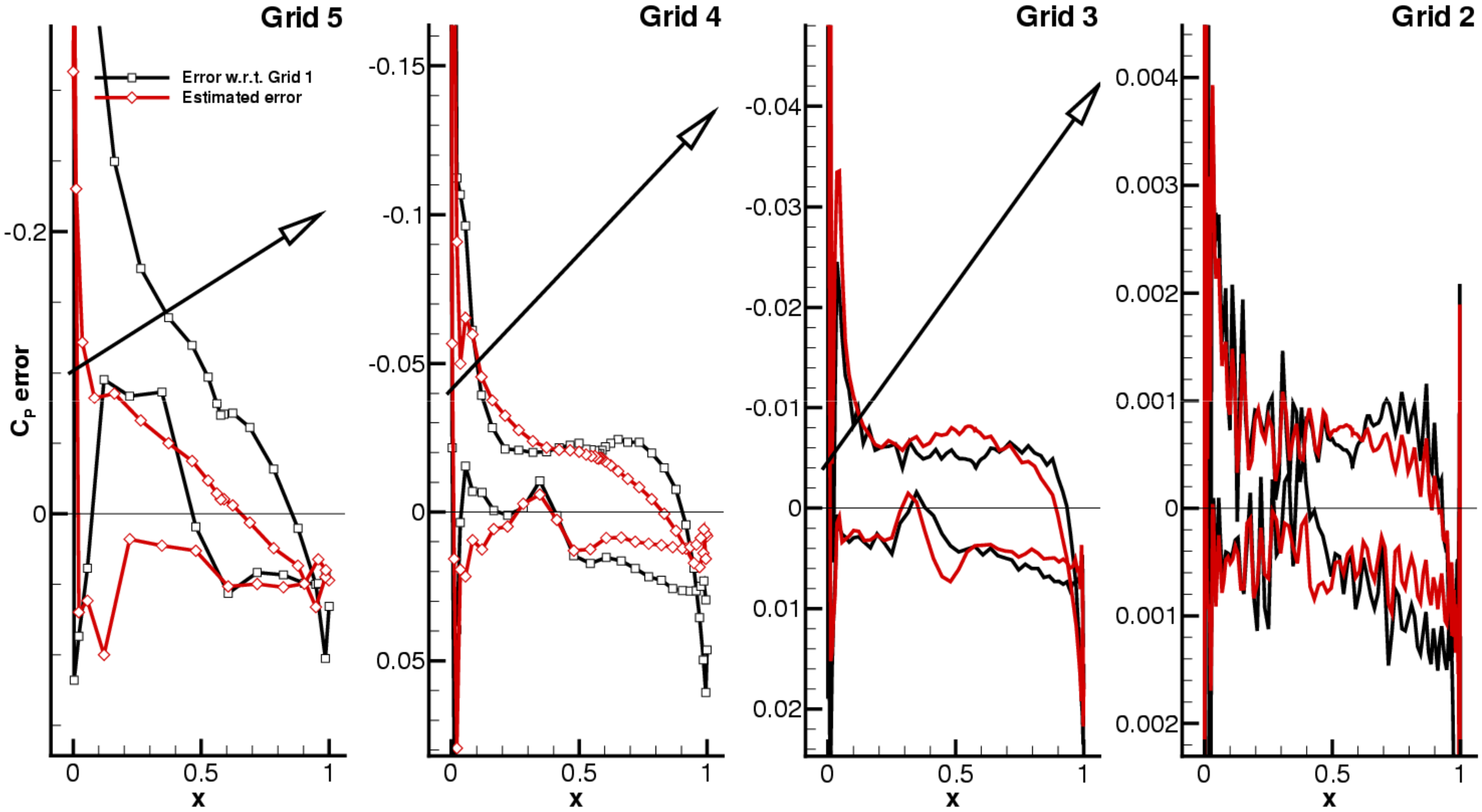


Error Transport Equation

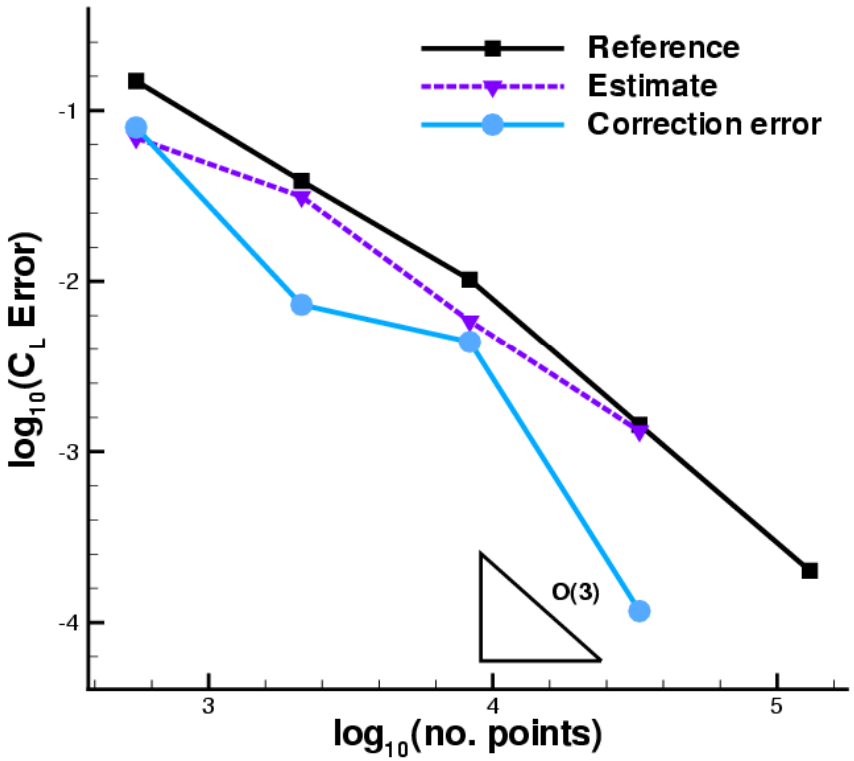


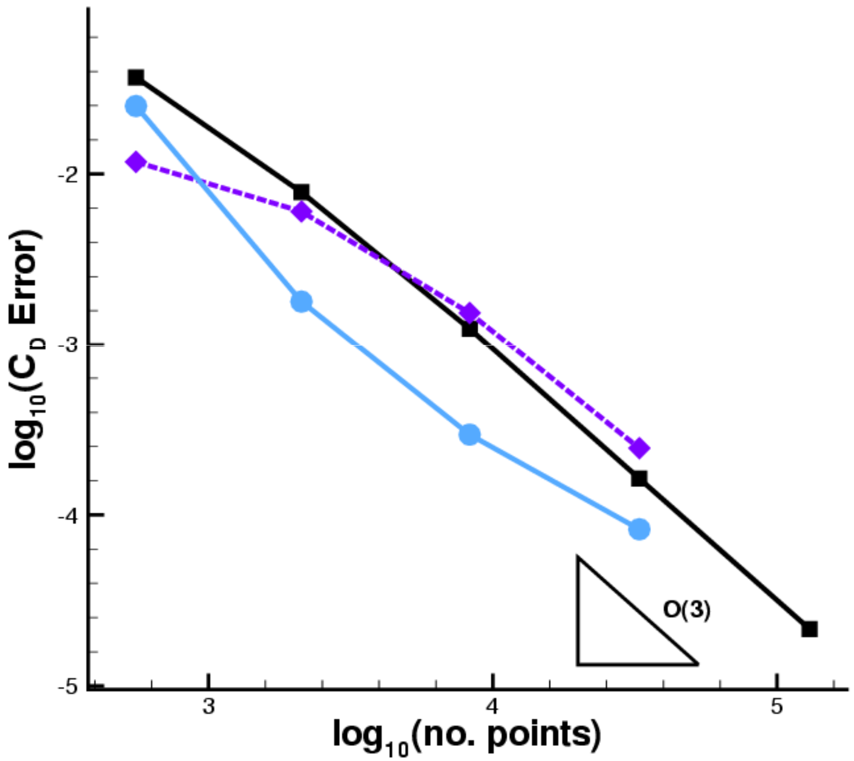


Error Transport Equation: RAE2822

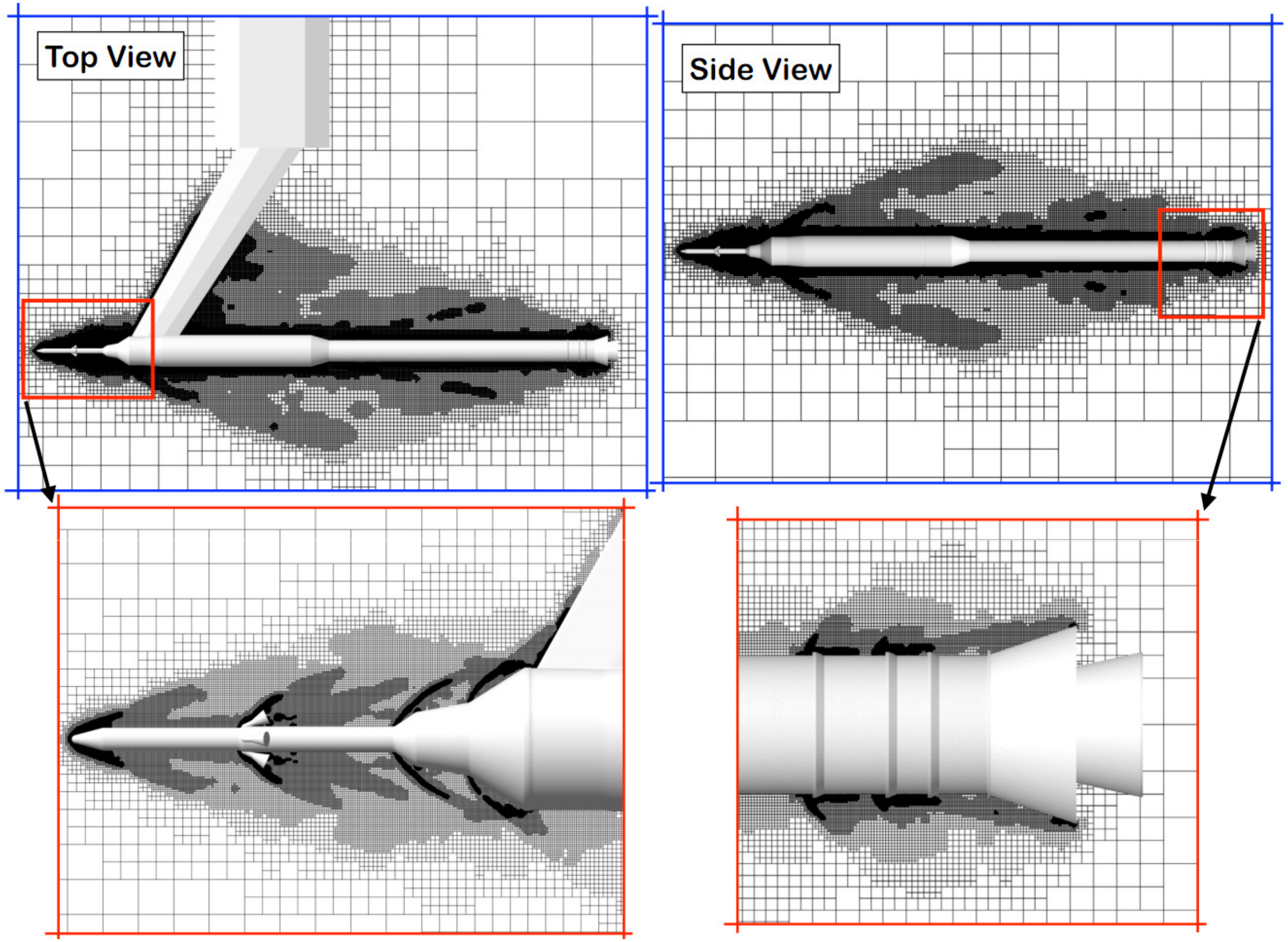


Error estimate: Lift and Drag





Current best (NASA-Cart3d)

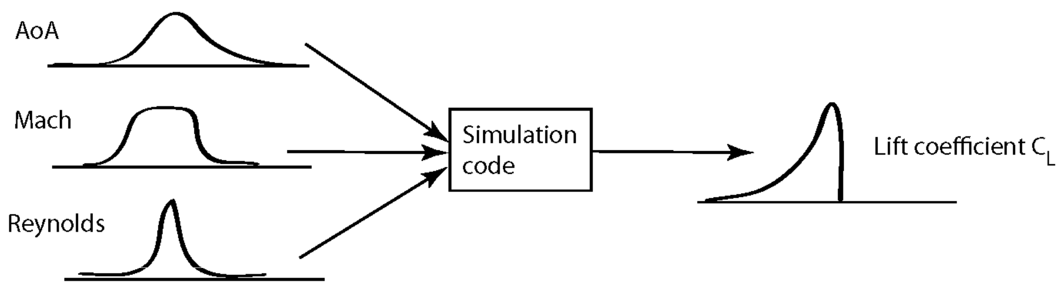


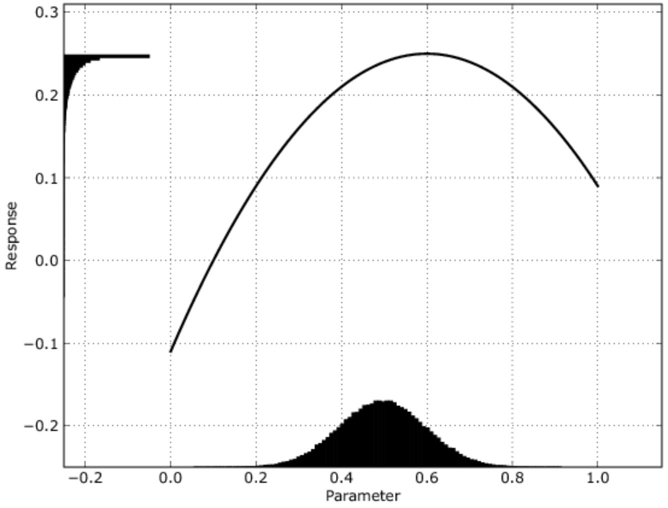
With permission, from: M. Nemec, M. Aftosmis, *"Adjoint-Based Adaptive Mesh Refinement for Complex Geometries"*, AIAA-2008-0725, 2008.

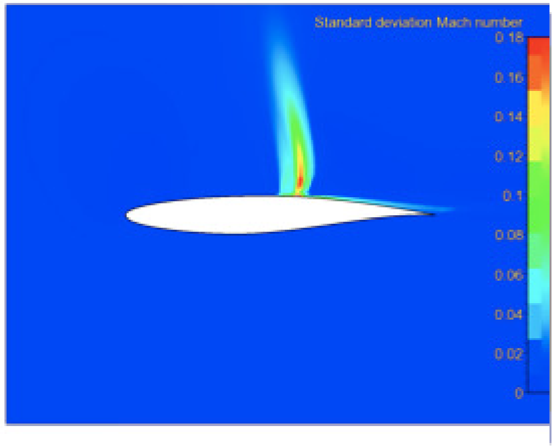
Response surfaces with imprecise gradients

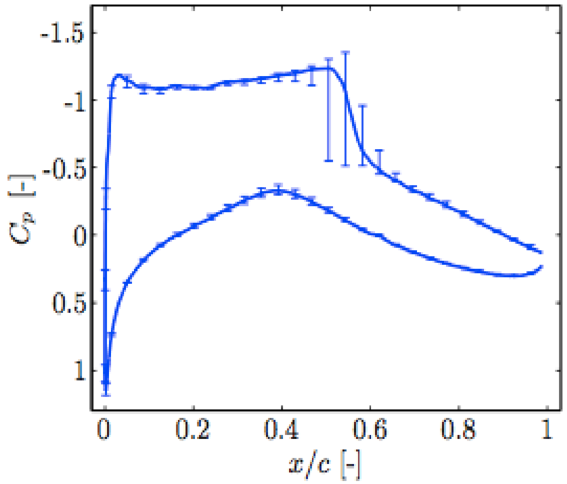
Uncertainty quantification

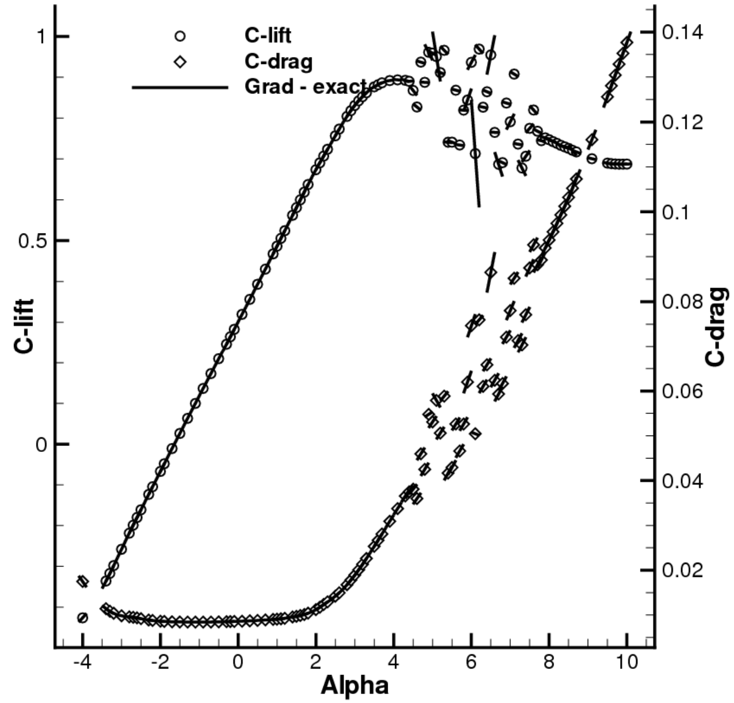
Let $X : \Omega \to \mathbb{R}$ a RV, with known pdf $\rho_X(x)$.
Let $Y = f(X)$. What is the pdf of Y: $\rho_Y(y)$?

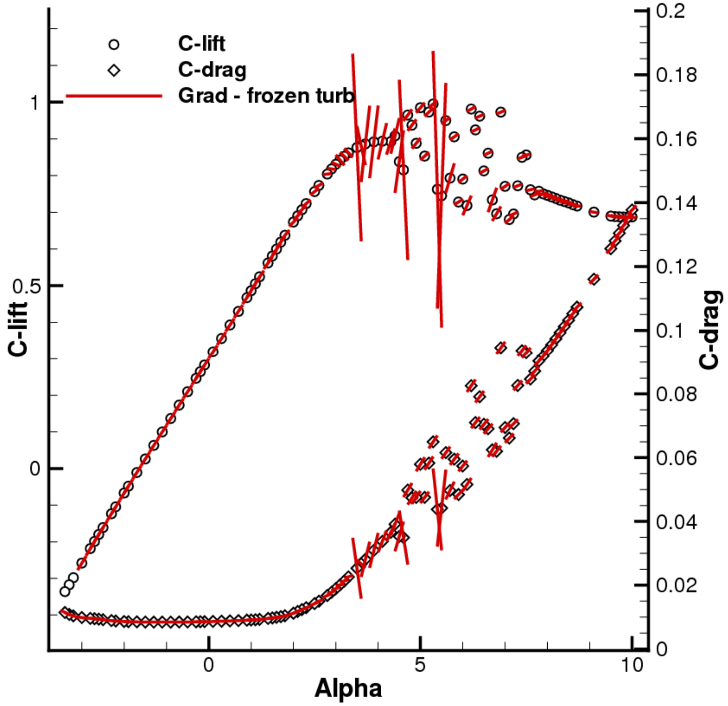


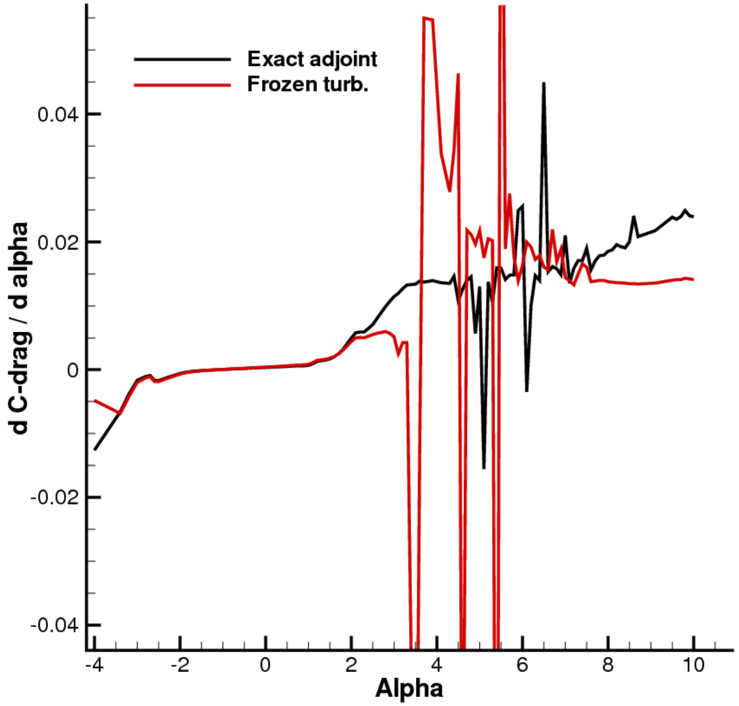


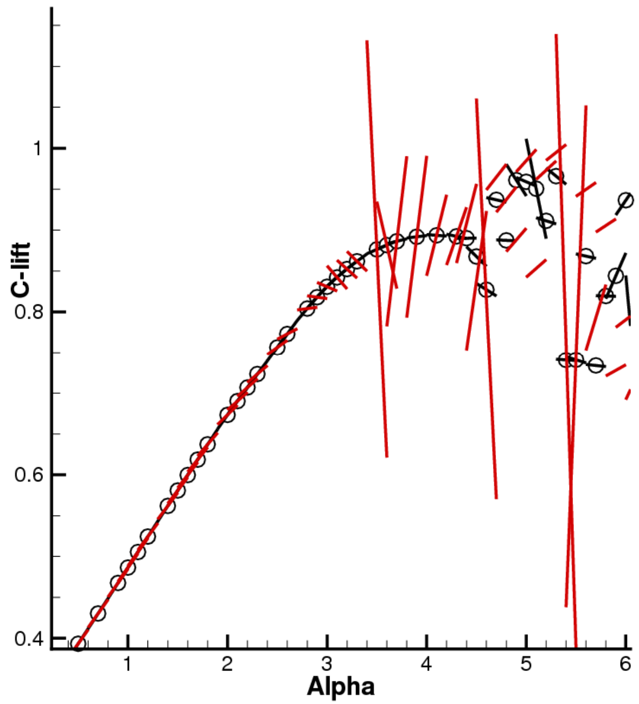




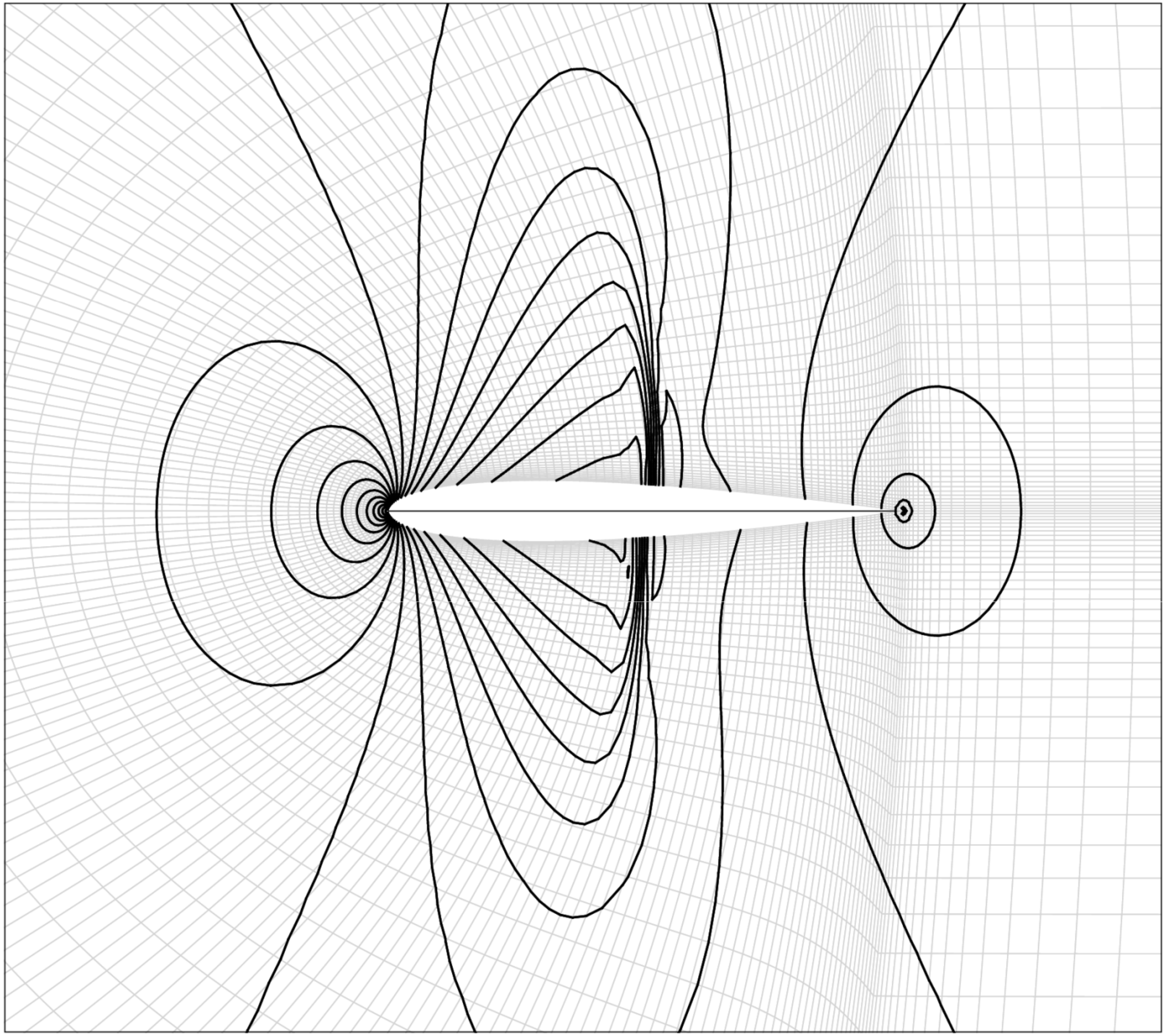

C-lift
C-drag
Grad - exact
Alpha
-4
-2
0
2
4
6
8
10
0
0.5
1
0.02
0.04
0.06
0.08
0.1
0.12
0.14


C-lift
C-drag
Grad - frozen turb
Alpha
0
5
10
0
0.5
1
0
0.02
0.04
0.06
0.08
0.1
0.12
0.14
0.16
0.18
0.2

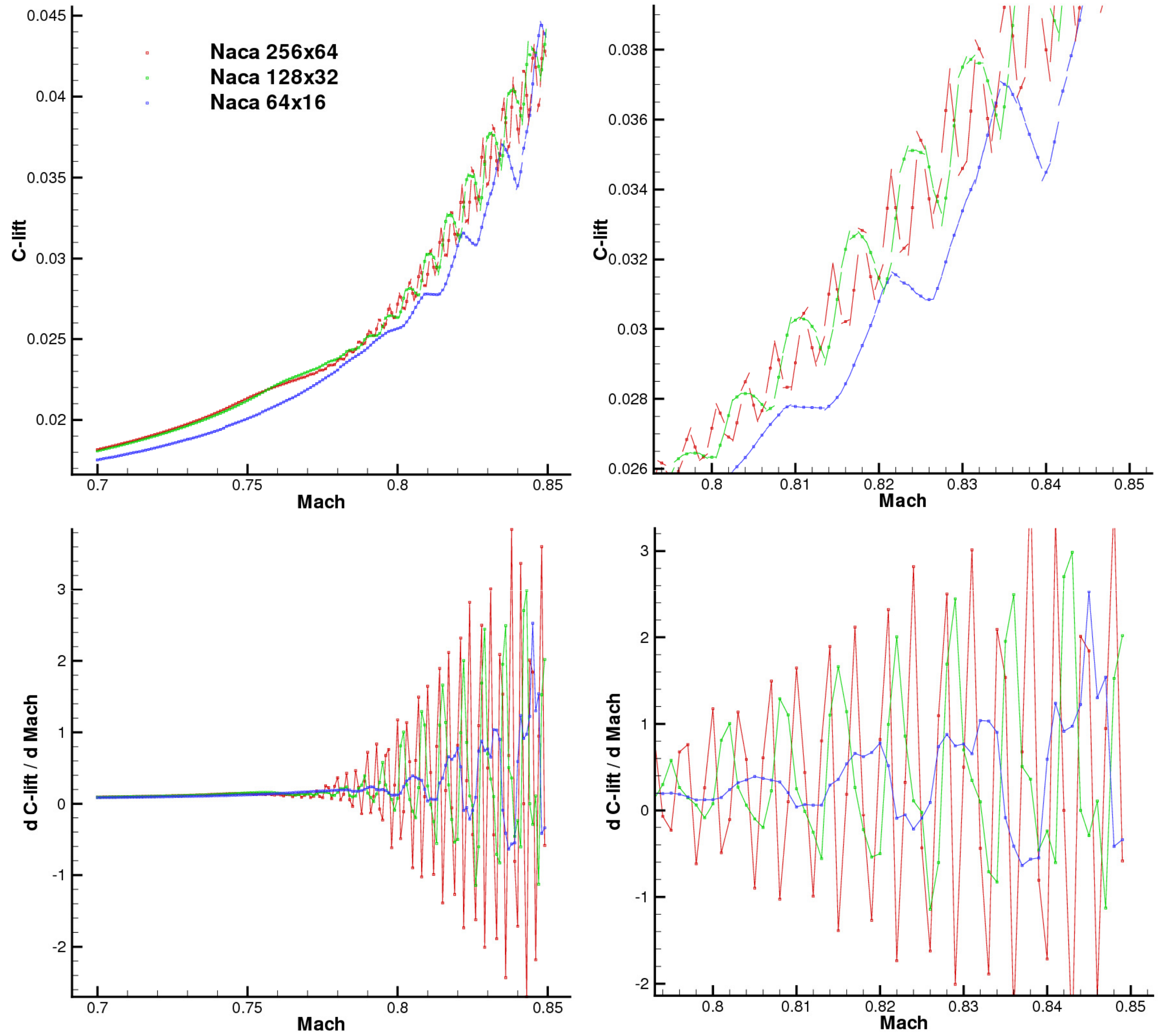

Exact adjoint
Frozen turb.
d C-drag / d alpha
Alpha
-4
-2
0
2
4
6
8
10
-0.04
-0.02
0
0.02
0.04


C-lift
Alpha
1
2
3
4
5
6
0.4
0.6
0.8
1

Shocks



Shocks



Aside: Conditional probability

Bayes' theorem

$$\rho(X|d) \propto \rho(d|X) \cdot \rho_0(X)$$

Combining a:

1. *Prior* belief on parameters.
2. *Likelihood* of observations.

Bayes' theorem gives a *posterior* (updated belief) on parameters.

Uncertainty *Reduction* with Bayes

$$\rho(X|d) \propto \rho(d|X) \cdot \rho_0(X)$$

Answer!

Prior: best guess at input uncertainties

Statistical model: relates the simulation to the data – e.g.:

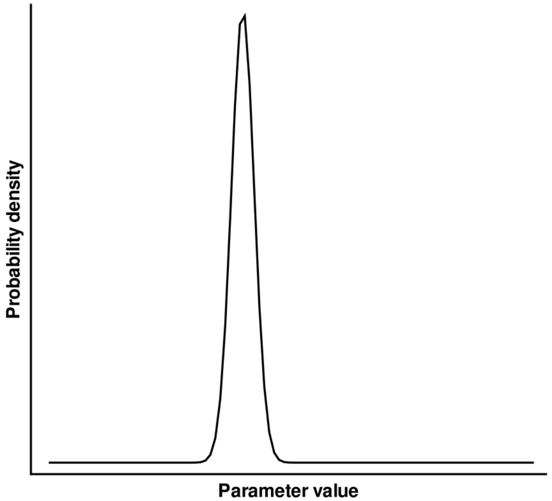
$$d = m(X) + \epsilon$$

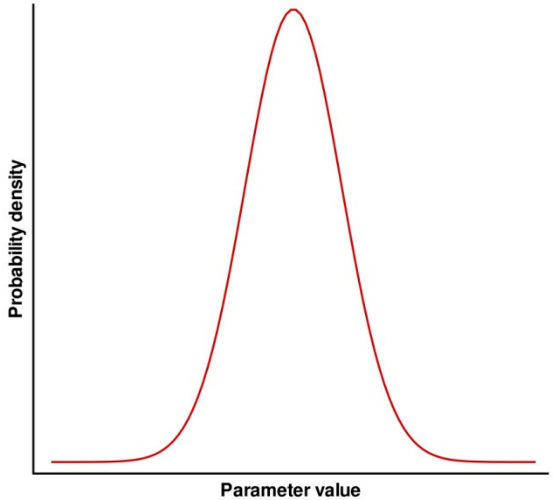
$$\epsilon \sim \mathcal{N}(0, \sigma_d^2)$$

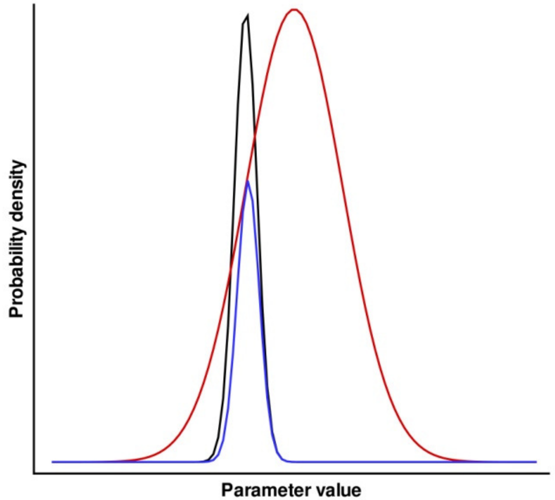
TU Delft

Uncertainty *Reduction* with Bayes

Likelihood	**Prior**		**Posterior**
$\rho(d\|X)$	$\rho_0(X)$	$\propto$	$\rho(X\|d)$

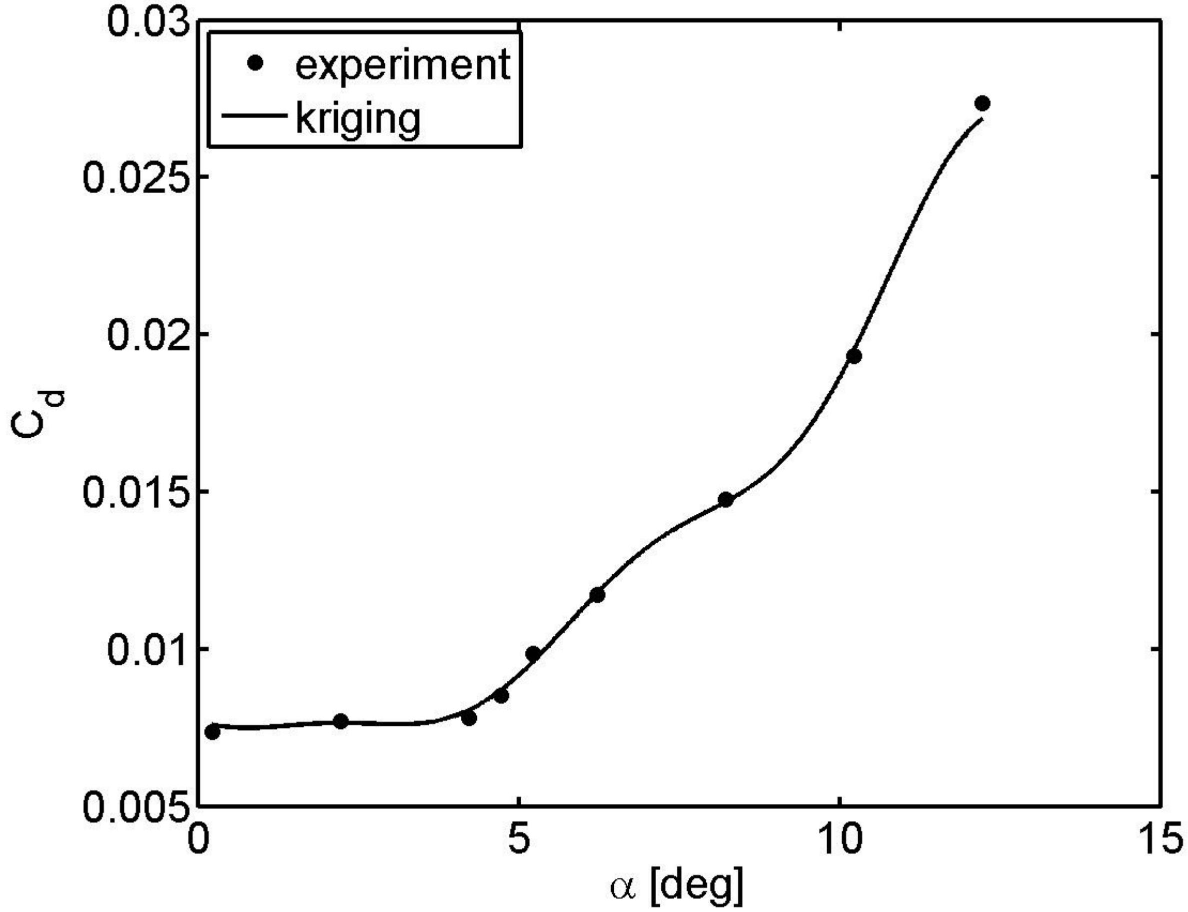






Kriging interpolation

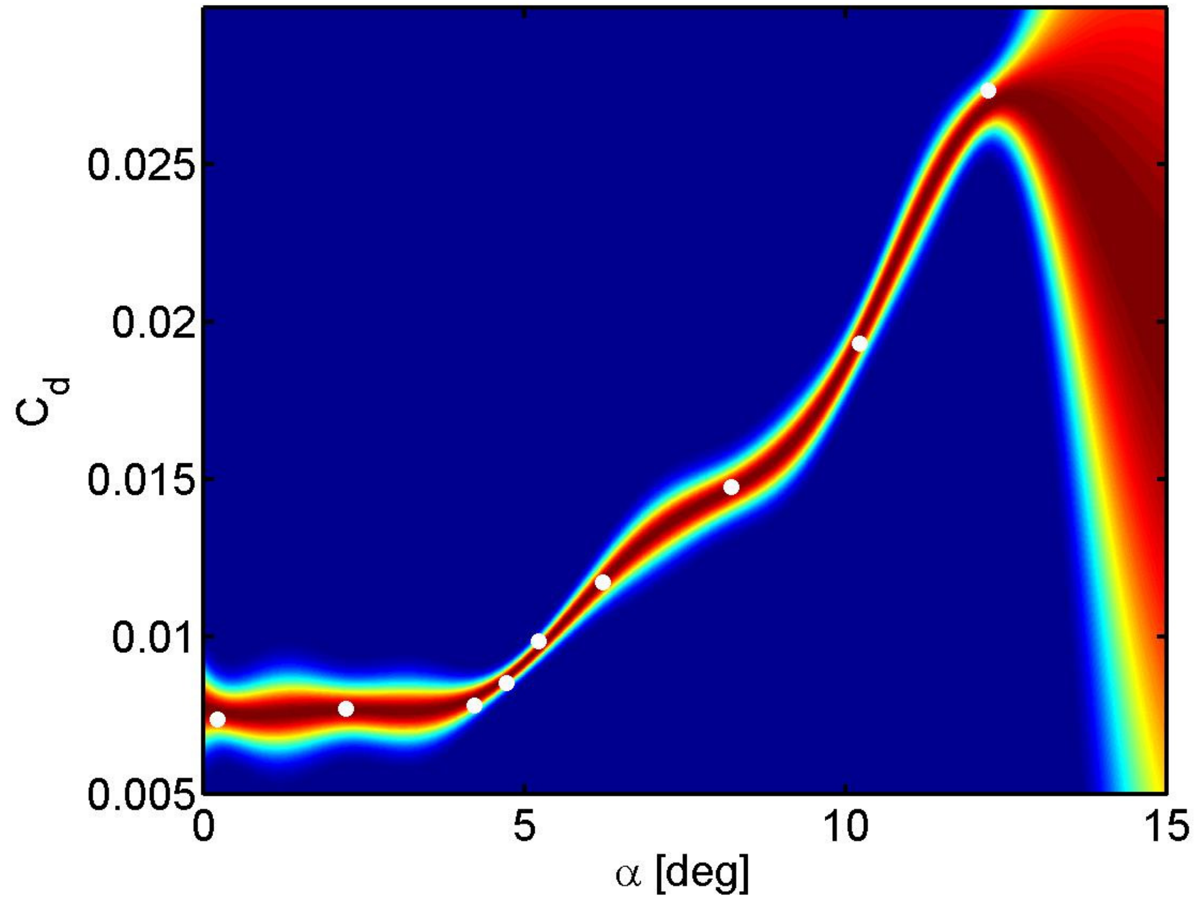
Data from pedagogical windtunnel tests



Results from group 19 (2011)

Kriging interpolation

Interpolant with uncertainty estimate



Results from group 19 (2011)

Kriging – From a Bayesian Perspective

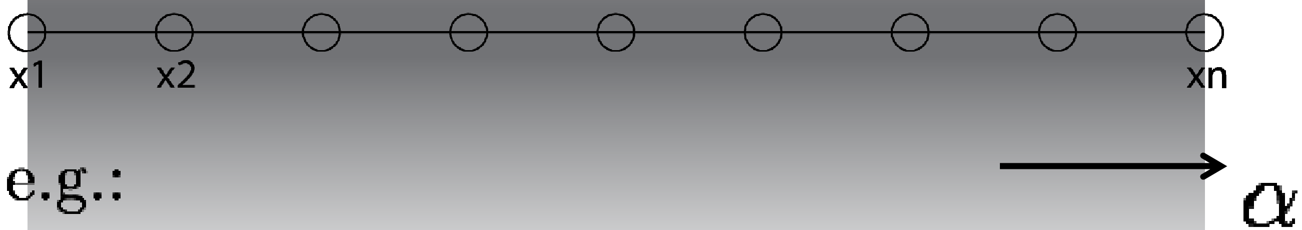
Multivariate normal distributions

$$X \sim \mathcal{N}(\mu, \Sigma)$$

μ – Vector mean (for each point location)

Σ – Covariance matrix

Independent $x_i \implies \Sigma = \sigma^2 I$



Dependent x_i, e.g.:

$$r(|\alpha_i - \alpha_j|) = r(h_{ij}) = \exp\left\{-\theta h_{ij}^2\right\}$$

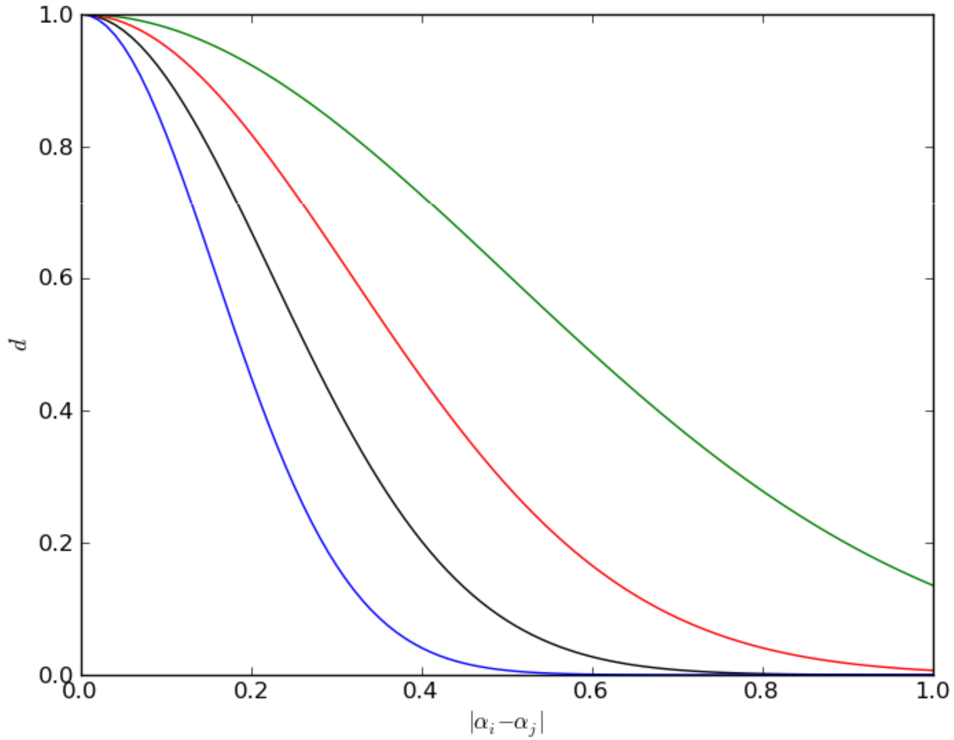
$$\Sigma_{ij} = \sigma^2 r(h_{ij})$$

Kriging – From a Bayesian Perspective

Multivariate normal distributions

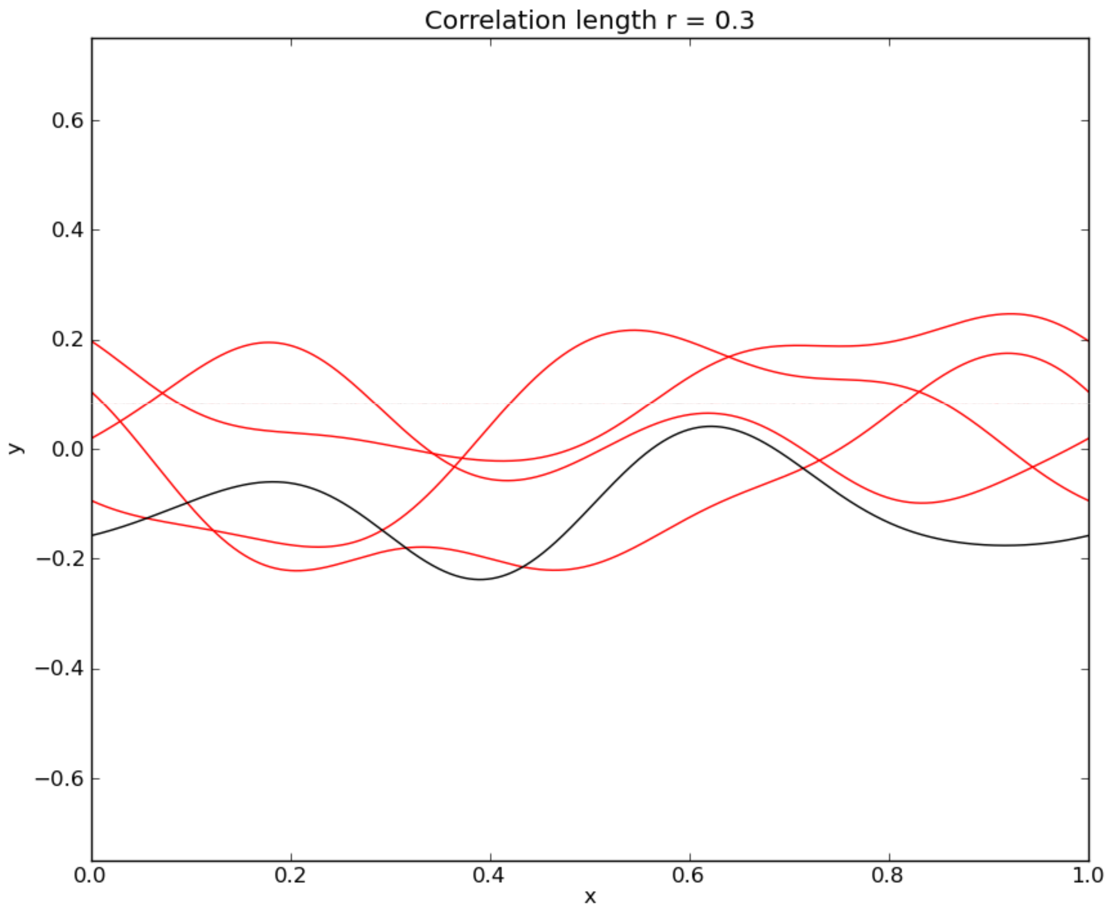
$$r(|\alpha_i - \alpha_j|) = r(h_{ij}) = \exp\left\{-\theta h_{ij}^2\right\}$$

$$\Sigma_{ij} = \sigma^2 r(h_{ij})$$



Kriging – A Bayesian Perspective

Multivariate normal distributions



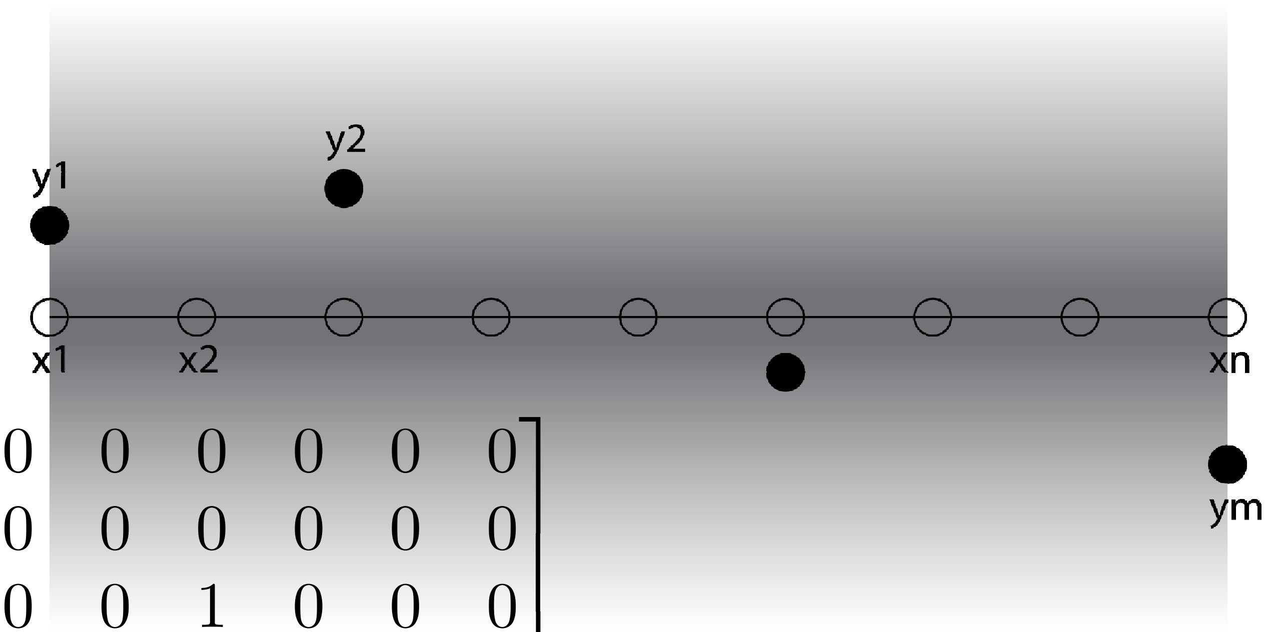
Kriging – From a Bayesian Perspective

Incorporating observations

$$Y \sim \mathcal{N}(Hx, R)$$

H – observation operator

$$R = \sigma_d^2 I$$



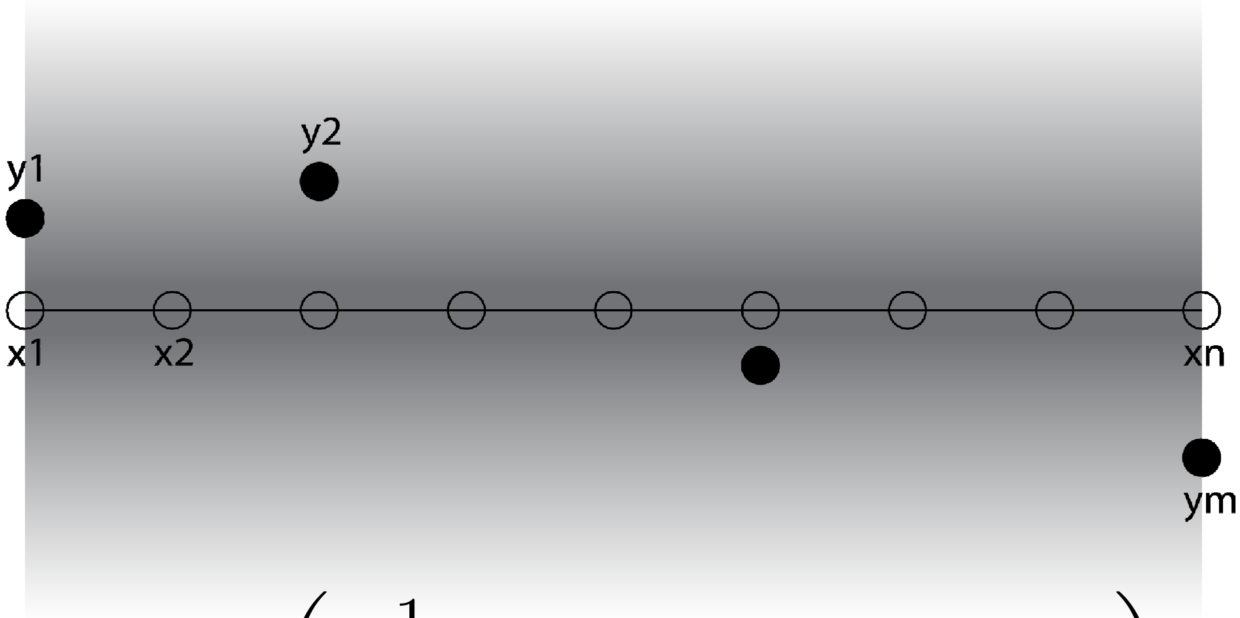
$$H = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

Kriging – From a Bayesian Perspective

Incorporating observations

Bayes!

$$\rho(X|Y) \propto \rho(Y|X)\rho_0(X)$$

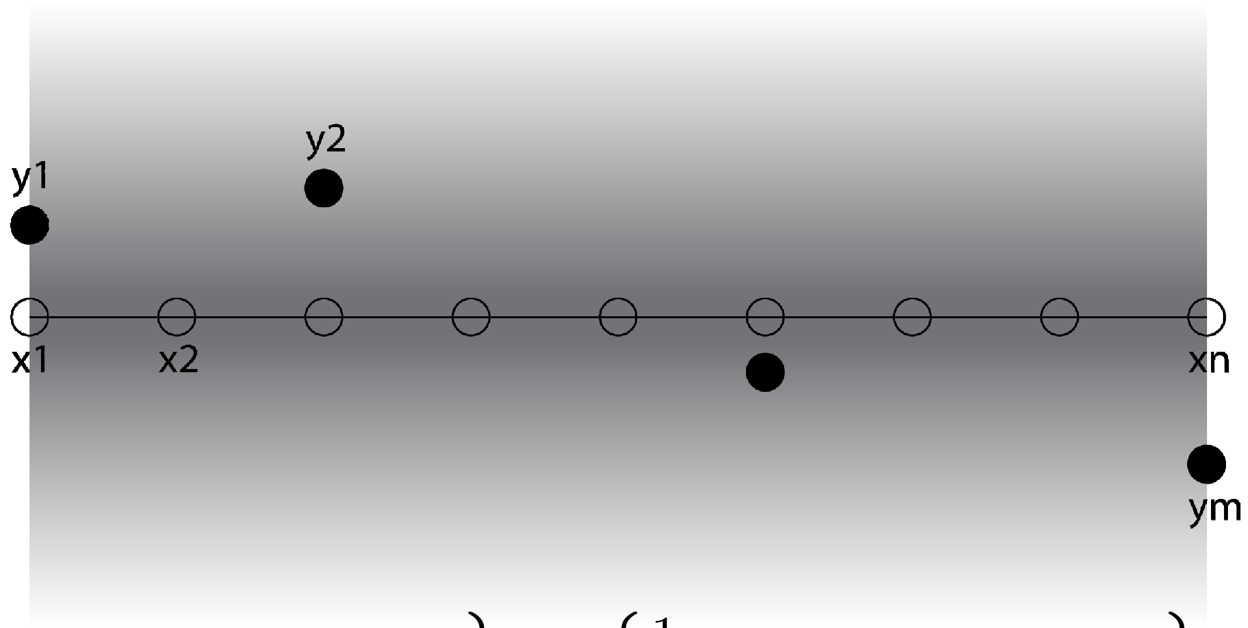


Prior: $$\rho_0(X) \propto \exp\left(-\frac{1}{2}(X-\mu)^T\Sigma^{-1}(X-\mu)\right)$$

Kriging – From a Bayesian Perspective

Incorporating observations

Likelihood: $Y|X \sim \mathcal{N}(y - Hx, R)$



$$\rho(X|Y) \propto \exp\left\{\frac{1}{2}(y - Hx)^T R^{-1}(y - Hx)\right\} \exp\left\{\frac{1}{2}(x-\mu)^T \Sigma^{-1}(x-\mu)\right\}$$

Kriging – From a Bayesian Perspective

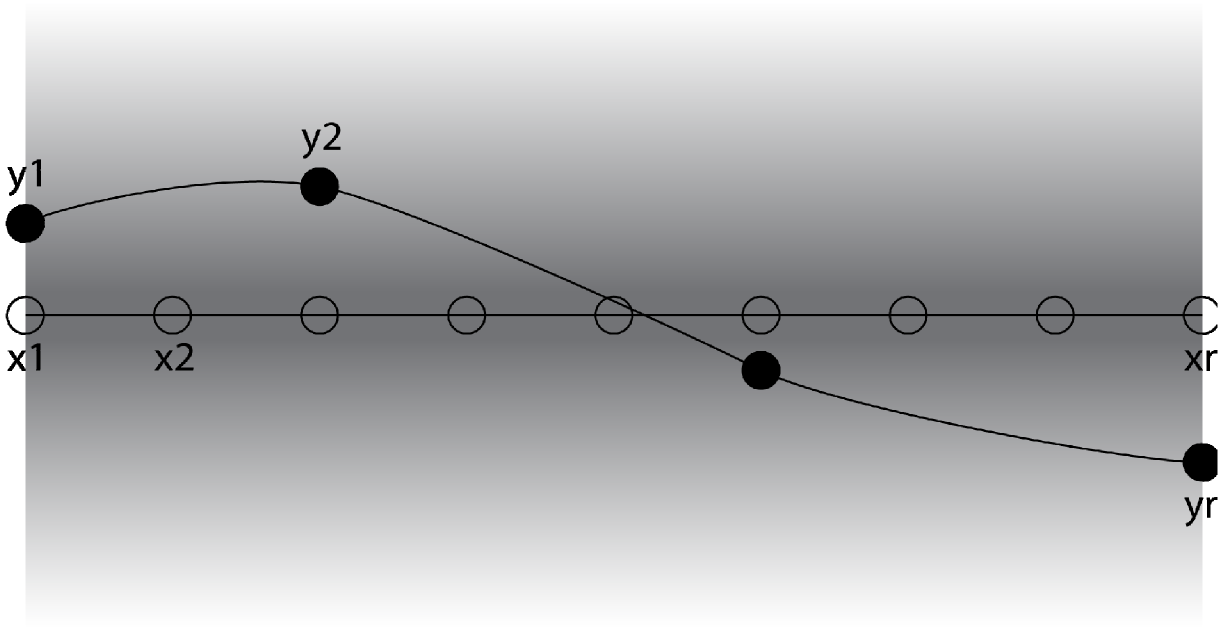
Incorporating observations

$$X|Y \sim \mathcal{N}(\cdot,\cdot)$$

$$\mathbb{E}(X|Y) = \mu + K(y - H\mu)$$

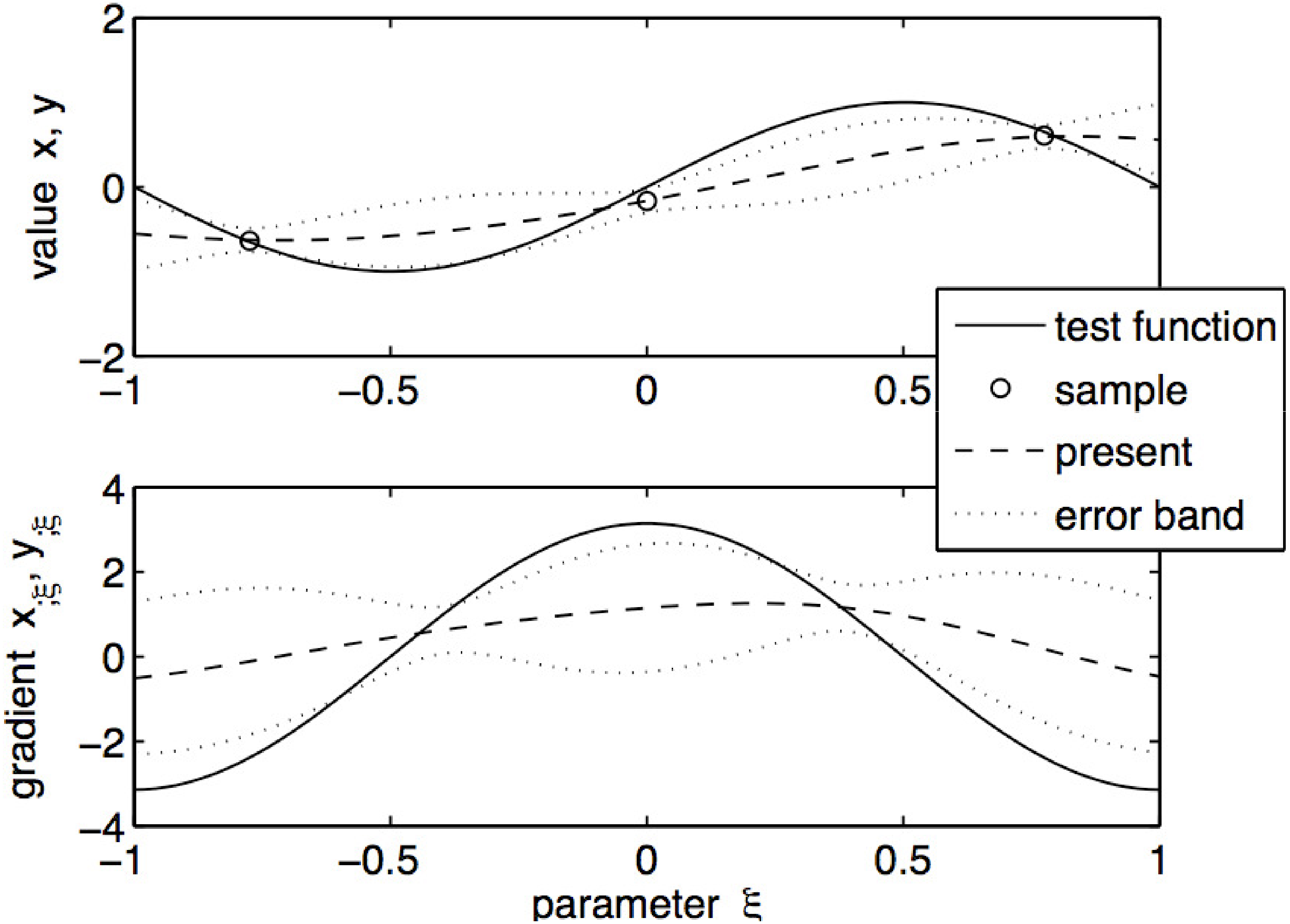
$$\mathrm{var}(X|Y) = (I - KH)\Sigma$$

$$K := \Sigma H^T (R + H\Sigma H^T)^{-1}$$

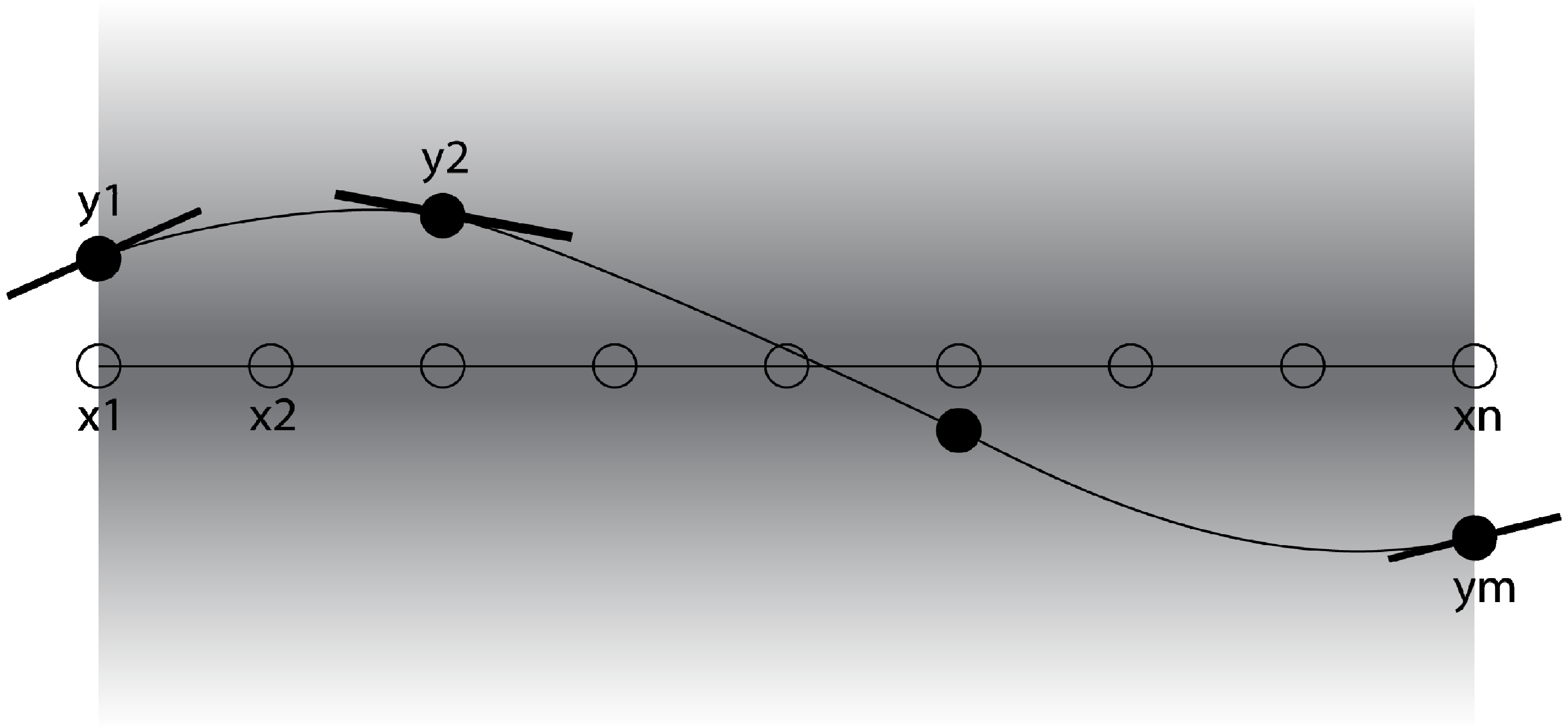


Posterior distribution = Kriging surface

Mean +- 1 sd



Incorporating gradients



Gradient information

Prior on the gradients

- Augment unknown value vector with unknown gradients:

$$X_c = [X, X']$$

- Prior is still a multivariate Gaussian (i.e. discrete random field)

$$X \sim \mathcal{N}(\mu, \Sigma)$$

$$X' \sim \mathcal{N}(0, \Sigma_g)$$

- So:

$$X_c \sim \mathcal{N}(\mu_c, \Sigma_c)$$

Gradient information

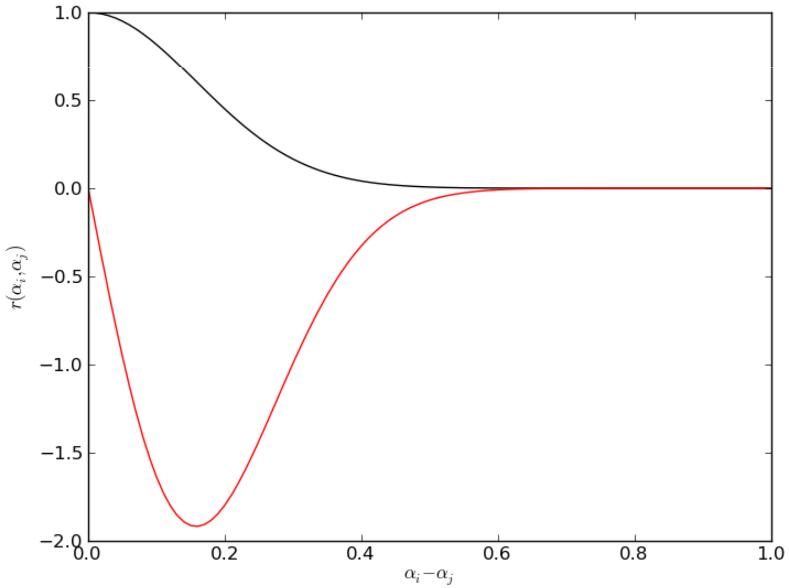
Defining the covariance matrix

- Value and gradient are not correlated at a point

$\Rightarrow$ can not apply Gaussian

$$\Sigma_c = \begin{pmatrix} P_{00} & P_{01} \\ P_{10} & P_{11} \end{pmatrix}$$

$$\begin{aligned}
\boldsymbol{P}_{00}^{ij} &= E\left[f(\xi^i)f(\xi^j)\right] \approx r(h_{ij}), \\
\boldsymbol{P}_{10}^{ij} &= \frac{\partial}{\partial \xi^i} E\left[f(\xi^i)f(\xi^j)\right] \approx -\frac{\partial}{\partial h} r(h_{ij}), \\
\boldsymbol{P}_{01}^{ij} &= \frac{\partial}{\partial \xi^j} E\left[f(\xi^i)f(\xi^j)\right] \approx \frac{\partial}{\partial h} r(h_{ij}), \\
\boldsymbol{P}_{11}^{ij} &= \frac{\partial^2}{\partial \xi^i \partial \xi^j} E\left[f(\xi^i)f(\xi^j)\right] \approx -\frac{\partial^2}{\partial h^2} r(h_{ij}).
\end{aligned}$$



Apply Bayes to this prior

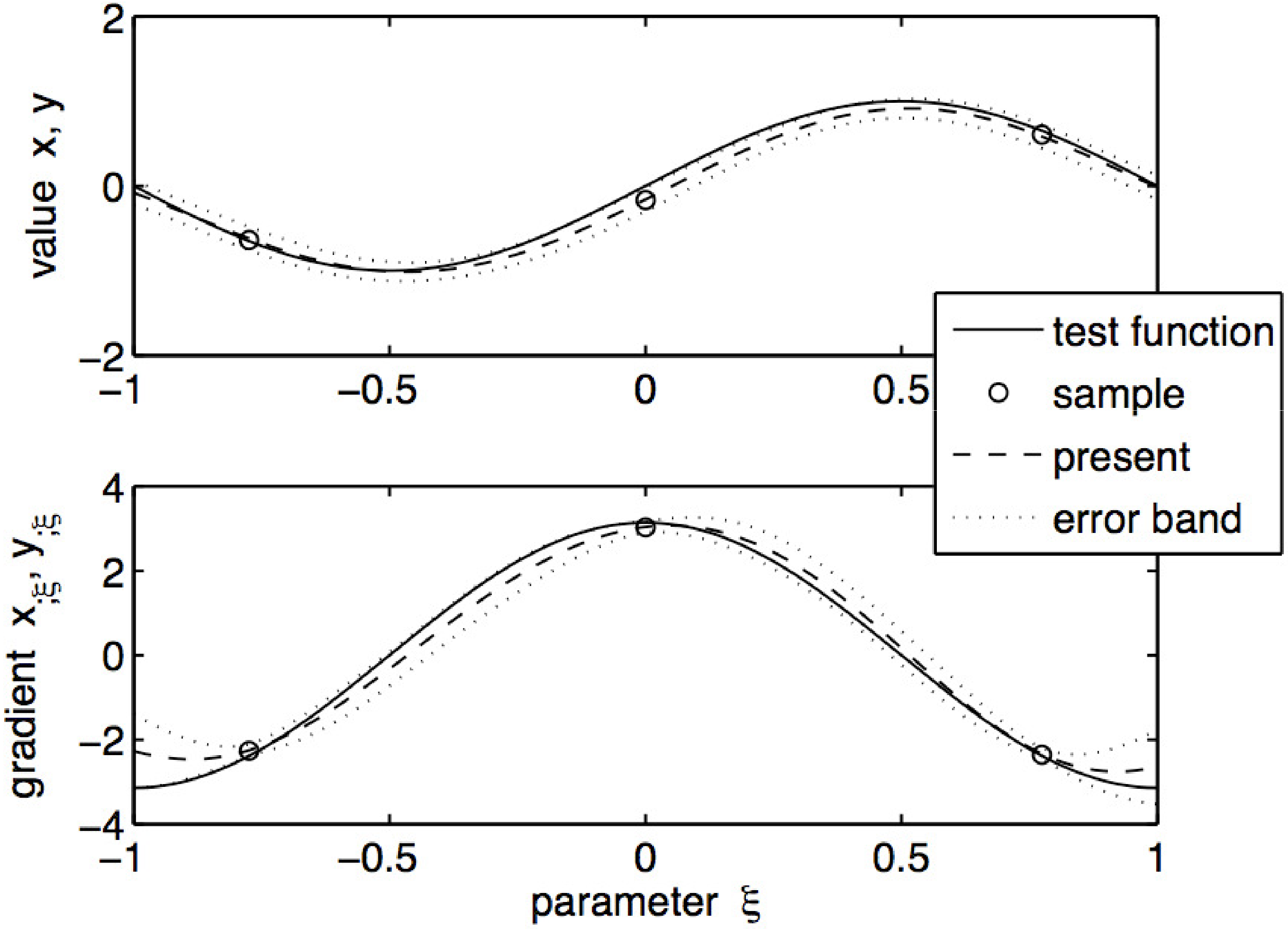
Exactly as in previous case

$$X_c | Y_c \sim \mathcal{N}(\cdot, \cdot)$$

$$\mathbb{E}(X_c | Y_c) = \mu_c + K_c(y_c - H_c \mu_c)$$

$$K := \Sigma_c H_c^T (R_c + H_c \Sigma_c H_c^t)^{-1}$$

Posterior = Kriging surface



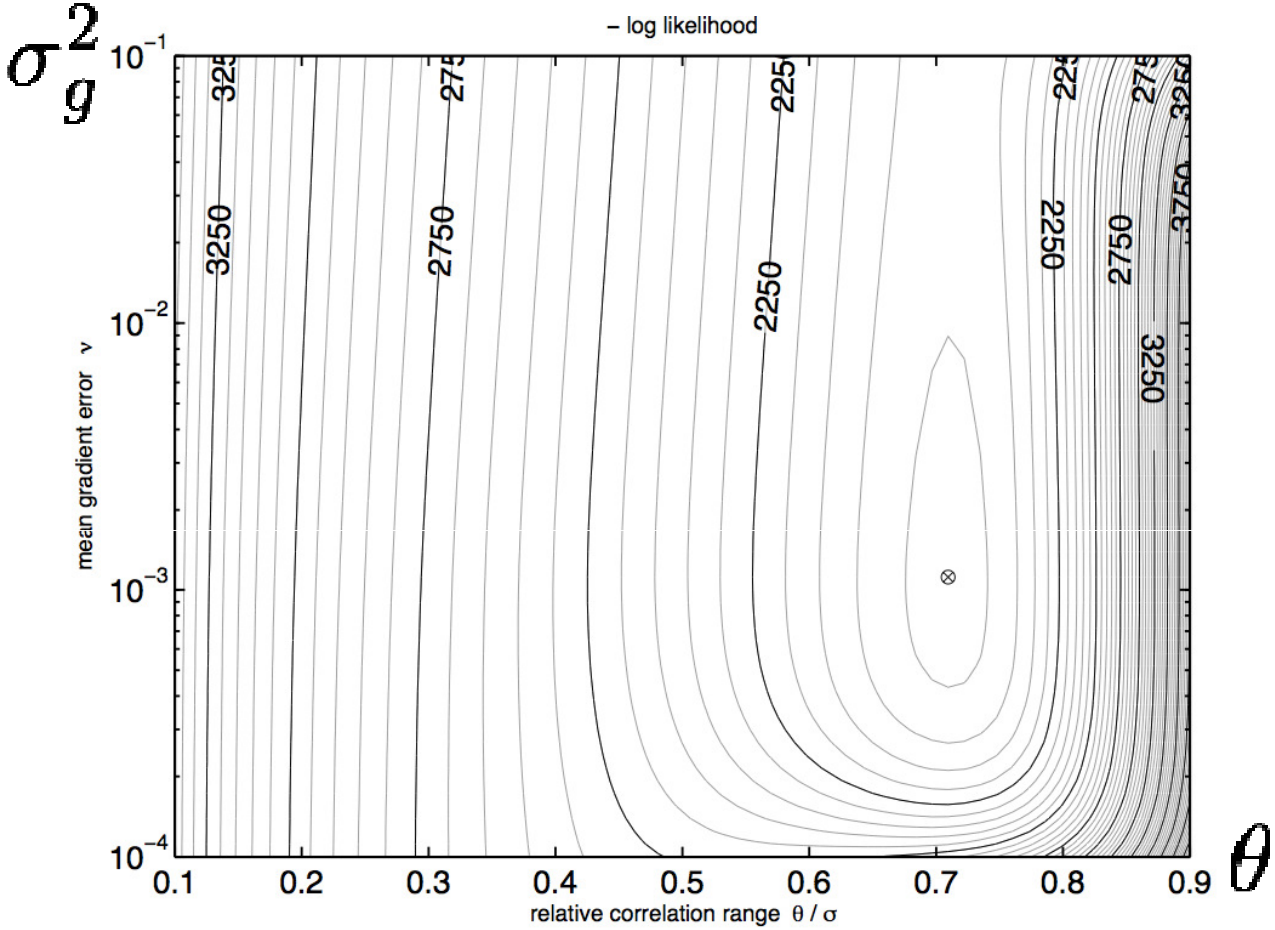
Estimating gradient error

MLE for gradient error

- We don’t know the gradient error.
- What do we do when we don’t know something?
- We apply Bayes and hope that what we **do** know informs what we **don’t** know!

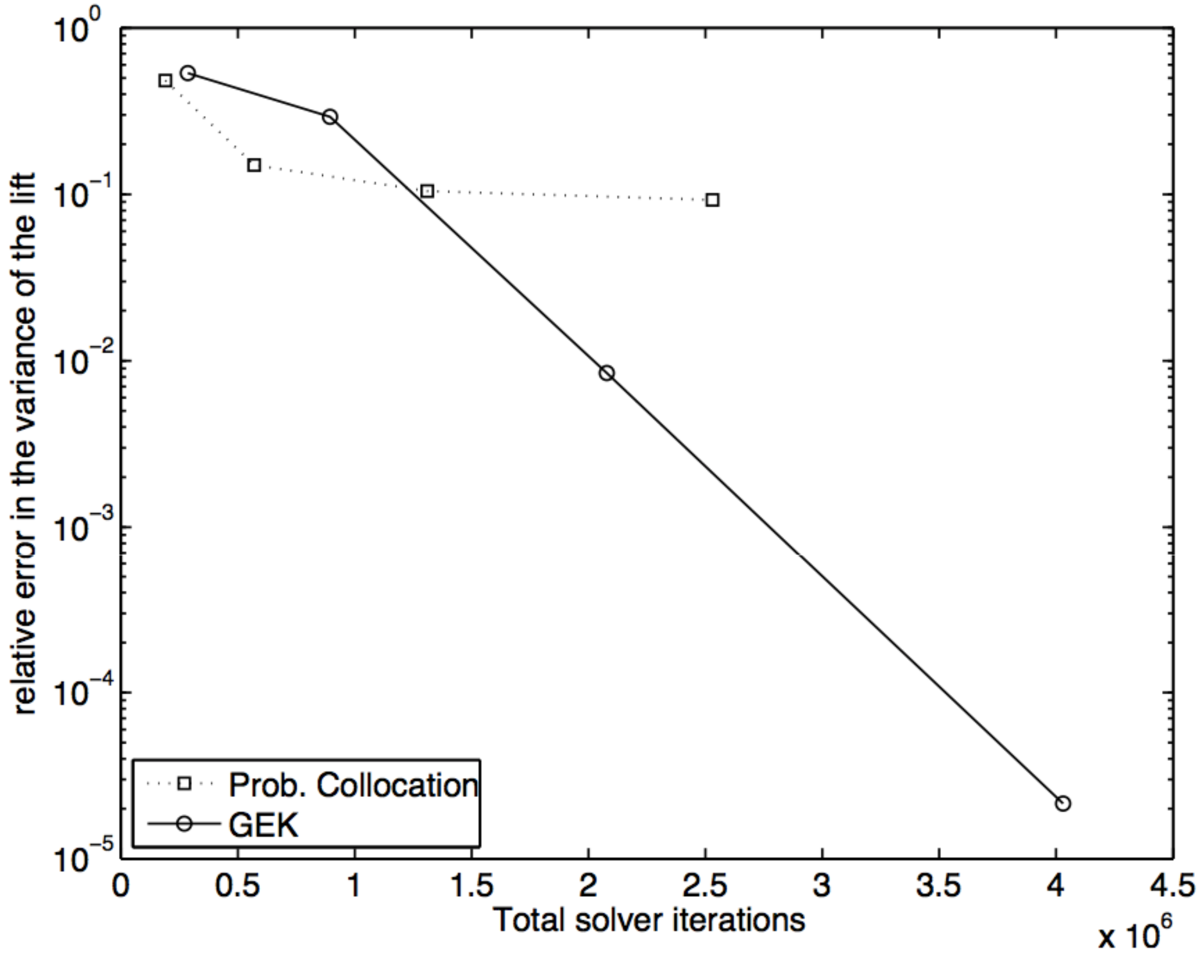
$$\rho(x_c, \sigma_g^2, \theta | y_c) \propto \rho(y_c | x_c, \sigma_g^2, \theta | y) \rho_0(x_c) \rho_0(\sigma_g^2) \rho_0(\theta)$$

Estimating gradient error



de Baar, Dwight, Bijl (2012), ***Improvements to Gradient-Enhanced Kriging using a Bayesian Interpretation.*** *Technometrics (submitted)*.

Uncertainty quantification accuracy



de Baar, Dwight, Bijl (2012), ***Improvements to Gradient-Enhanced Kriging using a Bayesian Interpretation.*** *Technometrics (submitted)*.

Thank you

for your attention