TOWARDS MULTIDISCIPLINARY ADJOINT OPTIMIZATION OF TURBOMACHINERY COMPONENTS

Marc Schwalbach\(^1\) and Tom Verstraete\(^2\)

\(^1\)von Karman Institute for Fluid Dynamics
72, Chaussee de Waterloo, B1640 Rhode Saint Genese, Belgium
e-mail: marc.schwalbach@vki.ac.be

\(^2\)Queen Mary University Of London
E1 4NS Mile End Road, London, UK
e-mail: t.verstraete@qmul.ac.uk

Keywords: adjoints, algorithmic differentiation, turbomachinery, CAD, multidisciplinary optimization, fluid-structure interaction

Abstract. The current state-of-the-art adjoint design optimizations for turbomachinery components focus solely on aerodynamic cost functions and constraints, yet disregard structural feasibility during the optimization procedure. This paper presents the first steps taken towards including structural constraints in a multidisciplinary adjoint optimization design chain for turbomachinery components. Particularly in turbomachinery, deformations arise due to centrifugal and pressure load during running conditions, which lead to a coupled fluid-structure interaction problem. While most optimization methods treat the fluid and structure domains separately in a single-disciplinary fashion, we seek to directly include the coupled fluid-structure interaction within the adjoint optimization. To this end, a cold-to-hot transformation tool that deforms a CAD geometry based on FEM displacements is implemented and differentiated using adjoint algorithmic differentiation to compute the required transformation sensitivities.
1 INTRODUCTION

Current adjoint optimization methods for turbomachinery components focus mainly on aero-
dynamic cost functions and constraints, e.g. [3], [10], [18]. The structural feasibility of the
resulting optimized shape is usually tested a posteriori by conducting a stress analysis. As a
result, an aerodynamically optimized shape may exceed defined stress tolerance levels and a
new shape will have to be designed. This can lead to several costly design iterations. Multidis-
ciplinary adjoint optimization seeks to take structural constraints into consideration during the
optimization process by coupling the fluid and structural disciplines. This will avoid unneces-
sary design iterations and directly compute an aerodynamically optimized shape, which is also
structurally feasible.

Turbomachinery components pose an additional challenge due to the deformations which
occur at running conditions. These are caused by rotation-induced centrifugal forces, gas pres-
sure loads, as well as thermal expansions due to temperature changes. The deformation process
is known as the cold-to-hot transformation [5], which can be seen as a fluid-structure interac-
tion problem. The cold state defines the at-rest state at which a component is manufactured.
The hot state defines the deformed state of a component during running conditions. Typically,
these deformations are not taken into account during the optimization process itself, but rather
a shape is optimized in its hot state and transformed a posteriori to its cold shape for manufac-
turing. In this case, the parametrized CAD geometry used in the optimization, also named the
master CAD geometry, is defined in the hot state [6]. However, this method has the downside
that the master CAD geometry used in the optimization was generated for one specific design
point. Different design points, each with their own centrifugal and pressure loads, would result
in different deformed hot geometries. Nevertheless, often the same master CAD geometry is
used to compute other design points. Additionally, the cold geometry that is generated after
the optimization may no longer fulfill manufacturing constraints, e.g. flank milling of radial
machines requires ruled surfaces.

As opposed to a hot state optimization and an a posteriori hot-to-cold transformation, we pro-
pose setting the cold state CAD parameters as the design parameters for the optimization with
the cold-to-hot deformation being computed during the simulation. As a result, the optimization
would directly compute a manufacturable cold state geometry of a component optimized
for a chosen design point. Since the final goal is an adjoint multidisciplinary optimization, the
differentiation of an entire cold-to-hot chain is required. In this paper, we present first steps
towards this alternative by implementing an adjoint version of a cold-to-hot transformation tool
to compute the required sensitivities.

First, we will motivate our goal of adjoint multidisciplinary optimization in section 2. Sec-
tions 3 and 4 outline the cold-to-hot algorithm and its differentiation with AD, respectively. The
results are presented in section 5 with concluding remarks in section 6.

2 ADJOINT MULTIDISCIPLINARY OPTIMIZATION

This section introduces the motivation of adjoint multidisciplinary optimization in turbo-
machinery. With the goal of optimizing the shape of a turbomachinery component, a metric
referred to as the cost function $J(\mathbf{x}) \in \mathbb{R}^c$, e.g. efficiency, has to be optimized, i.e. maxi-
mized. The parameters which can be varied to minimize the cost function are referred to as the *design parameters* $\mathbf{x} \in \mathbb{R}^d$. The cost function is minimized using an optimization method, which mainly falls into one of the two categories of *gradient-free* or *gradient-based* optimization methods.

Gradient-free optimization methods can be used to compute improved designs using only cost function evaluations, but finding the true optimal design is not guaranteed. Additionally, converging towards an optimum can require a large number of evaluations and the cost of the optimization is dependent on the number of design parameters $d$. Gradient-based optimization methods, on the other hand, use sensitivity information to compute a minimum with less iterations than gradient-free methods. However, these methods compute a local minimum and cannot guarantee that the computed minimum is indeed the global minimum, unless the optimization problem is convex. Furthermore, the sensitivities of the cost function with respect to the design variables

$$\frac{\partial J}{\partial \mathbf{x}} \in \mathbb{R}^{c \times d}$$

have to be computed. Typical approximation methods, such as finite differences (FD), compute the gradient (1) at a cost proportional to the number of design parameters $d$, which can lead to high computational costs.

Using the *adjoint* approach, first introduced by Pironneau [15] and later in the application of aerodynamic design optimization by Jameson [8], [16], the gradient (1) can be computed at a cost proportional to the size $c$ of the cost function $J$. In the context of aerospace engineering, one is typically interested in few parameters such as efficiency, mass flow rate, pressure ratio, and maximum stresses. Thus, the size of the cost function is usually much smaller than the size of the design space, such that $c \ll d$. Thus, the adjoint method offers a significant performance advantage for gradient-based optimization methods over gradient-free methods, given a smooth design space. Notably, the cost of computing the gradient is independent of the number of design variables $d$, allowing a much greater design space. The adjoint approach requires the evaluation of the *adjoint model*

$$\mathbf{x}_{(1)} = \frac{\partial J(\mathbf{x})}{\partial \mathbf{x}} J_{(1)}$$

of the cost function. First deriving a continuous form of the adjoint model (2), then discretizing it, is referred to as the *continuous adjoint* [9] approach. However, deriving a continuous adjoint model of a complex system of equations can be extremely difficult and error prone. Alternatively, the *discrete adjoint* [11] approach discretizes the system of equations first, then derives an adjoint model of the discretized system of equations. By using the code transformation method *algorithmic differentiation* (AD) [7], [13], also known as *automatic differentiation*, an adjoint model of a computer program code can be easily generated. In total, this makes the discrete adjoint optimization method an efficient and precise alternative to gradient-free optimization methods.

Gradient-free optimization methods have been applied in the context of CAD-based multidisciplinary optimization of turbomachinery components e.g. by [12], [17]. In this project, we seek to apply the discrete adjoint approach. We especially want to focus on the fluid-structure interaction (FSI) problem portrayed in the introduction (section 1). An essential aspect of FSI within a CAD-based optimization is the interface between the fluid, solid, and CAD disciplines. The
cold-to-hot transformation (section 3) provides one such interface. Within an adjoint optimization, a discrete adjoint model of this interface is required, which is achieved by differentiating the cold-to-hot transformation using AD (section 4).

3 COLD-TO-HOT TRANSFORMATION

The cold-to-hot transformation implemented in this project is based on the method detailed in [5]. Given an FEM mesh that is generated from a cold state CAD geometry, an FEM linear elastic computation is carried out to compute the displacements generated by given centrifugal forces. For now, we are focusing on deformations caused only by centrifugal forces. By perturbing the B-spline surface control points, the computed FEM displacements are matched and the CAD geometry is transformed to its hot state. In this section, the key steps of the transformation will be outlined. The interested reader is referred to [5] for a more detailed explanation of the algorithm and to [4], [14] for an in-depth discussion of computational geometry.

A B-spline surface

\[ S(u, v) = \sum_{i=0}^{n-1} \sum_{j=0}^{m-1} N_p^i(u) N_q^j(v) p_{i,j}^0, \]  

is given with basis functions \( N_p^i(u) \) and \( N_q^j(v) \) of order \( p \) and \( q \), respectively, foot points \( 0 \leq u \leq 1 \) and \( 0 \leq v \leq 1 \), and control points \( p_{i,j}^0 \in \mathbb{R}^3 \). The position of FEM nodes \( d_k^0 \in \mathbb{R}^3 \) and their respective displacements \( \Delta d_k \in \mathbb{R}^3 \) are given as well. \( k = 0, \ldots, s - 1 \), where \( s \) defines the number of FEM nodes on the surface.

3.1 FEM node projection

The FEM nodes \( d_k^0 \) are projected onto surface foot points \((u_k, v_k)\) by solving a point inversion problem, which can be solved using a suitable algorithm from [14]. As depicted in figure 1, the displacements \( \Delta d_k \) are imposed onto the foot points \((u_k, v_k)\), to compute the displaced node points

\[ d_k = S(u_k, v_k) + \Delta d_k. \]  

Figure 1: Displacing surface point \( S(u_k, v_k) \) with FEM displacement \( \Delta d_k \)
Using the assumption that the displaced node points \( \mathbf{d}_k \) must pass through the deformed B-spline surface \( \hat{\mathcal{S}}(u_k, v_k) \), the following system of equations can be derived:

\[
\sum_{i=0}^{n-1} \sum_{j=0}^{m-1} N_i^p(u_k) N_j^q(v_k) \Delta \mathbf{p}_{i,j} = \Delta \mathbf{d}_k, \quad k = 0, \ldots, s - 1,
\]

where \( \Delta \mathbf{p}_{i,j} \) represents the change to control point \( \mathbf{p}_{i,j} \). Typically, this system is overdetermined since the number of FEM nodes \( s \) is usually greater than the number of control points \( n \cdot m \). This system can be solved using a least-squares method.

### 3.2 Solve for outer control point displacements

While the solution of the system (5) can deliver the required control point displacements \( \Delta \mathbf{p}_{i,j} \) for a single surface \( \hat{\mathcal{S}}(u, v) \), multiple surfaces with adjacent edges are mostly involved in a complex geometry. An explicit treatment of the surface edges is required to avoid different solutions of \( \Delta \mathbf{p}_{i,j} \) along adjacent edges, which can lead to undesirable kinks. The control points along the surface edges are thus solved first by treating the edges as B-spline curves. This results in the overdetermined system of equations

\[
\sum_{i=0}^{n-1} N_i^p(u_k) \Delta \mathbf{p}_i = \Delta \mathbf{d}_k, \quad k \in \mathbf{K}_{\text{edge}},
\]

with \( \mathbf{K}_{\text{edge}} \) defining the set of FEM nodes along the edge.

### 3.3 Solve for inner control point displacements

Once the control point displacements along the edges \( \Delta \mathbf{p}_{i,0}, \Delta \mathbf{p}_{i,m-1}, \Delta \mathbf{p}_{0,j}, \Delta \mathbf{p}_{n-1,j} \) are determined, the known values of (5) can be moved to the right-hand side to solve for the remaining inner control point displacements

\[
\sum_{i=1}^{n-2} \sum_{j=1}^{m-2} N_i^p(u_k) N_j^q(v_k) \Delta \mathbf{p}_{i,j} = \Delta \mathbf{d}_k - \sum_{i=0}^{n-1} N_i^p(u_k) [N_0^q(v_k) \Delta \mathbf{p}_{i,0} + N_{n-1}^q(v_k) \Delta \mathbf{p}_{i,m-1}] + \sum_{j=1}^{m-2} N_j^q(v_k) [N_0^p(u_k) \Delta \mathbf{p}_{0,j} + N_{n-1}^p(u_k) \Delta \mathbf{p}_{n-1,j}],
\]

\( k \in \mathbf{K}_{\text{inner}} \),

with \( \mathbf{K}_{\text{inner}} \) defining the set of remaining inner FEM nodes. Afterwards, the control points can be updated by

\[
\mathbf{p} = \mathbf{p}^0 + \Delta \mathbf{p}
\]

and continuity correction methods can be used to ensure continuity at the common edges of deformed surfaces. Results of the implemented algorithm are presented in section 5.1.

### 4 SENSITIVITY COMPUTATION

For the computation of the sensitivities of the structural quantity of interest, e.g. the maximum von Mises stress \( \sigma_{\text{max}} \in \mathbb{R} \), with respect to the design parameters \( \mathbf{\alpha} \in \mathbb{R}^d \), the sensitivities of an entire chain of operations have to be considered. While the inclusion of pressure loads from the CFD solution is required to fully couple the fluid and structural disciplines, for now
we regard only centrifugal loads. Consider figure 2 for the chain of computations required to compute the structural quantity $\sigma_{max}$.

One starts off with design parameters $\alpha$, which can define initial control points of the cold geometry, as well as application specific parameters such as inlet- and outlet radius. $\alpha$ is then plugged into the CAD kernel which generates a cold B-spline surface $S^i_c$, which is used to generate the initial solid mesh $\Omega^i_s$. The solid mesh, along with the known centrifugal forces $F_c$, is then used in the computational structural mechanics (CSM) solver to compute the FEM displacements $d^i_{+1}$. The solution of the CSM solver automatically gives us an updated solid mesh $\Omega^i_{s+1}$. The FEM displacements $\Delta d^i_{+1}$ are then plugged into the cold-to-hot transformation algorithm to compute the updated control points $p^i_{+1}$, which can be used to generate an updated geometry $S^{i+1}$. Once this deformation loop, which is boxed within a dashed line in figure 2, has converged, the maximum von Mises stress $\sigma_{max}$ can be computed.

Performing a structurally constrained gradient-based optimization would require differentiating the entire described chain to compute the gradient

$$\frac{\partial \sigma_{max}}{\partial \alpha} \in \mathbb{R}^{1 \times d}.$$

Typically, it would hold that $d \gg 1$, which gives the opportunity of a cheap gradient computation using adjoints. However, this would involve computing the adjoints of the entire cold-to-hot chain. A first step towards computing the adjoints of this entire chain is made by computing the adjoints of the cold-to-hot transformation

$$p(\Delta d) = p^0 + \Delta p(\Delta d).$$

The corresponding 1st-order adjoint model is given by

$$\Delta d^{(1)} = \frac{\partial \Delta p}{\partial \Delta d}^T p^{(1)},$$

which will yield the gradient

$$\frac{\partial \Delta p(\Delta d)}{\partial \Delta d} = \frac{\partial p(\Delta d)}{\partial \Delta d} \in \mathbb{R}^{n \times m \times s}.$$
by performing \( n \cdot m \) runs of the adjoint model \( (11) \). Each adjoint run \( i \) is seeded with the \( i \)-th unit vector \( \vec{i} \in \mathbb{R}^{n \cdot m} \), where \( \vec{i} \) is a vector of 0s, except for the \( i \)-th variable, which is a 1:

\[
p^{(1)} = \vec{i} = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}.
\] (13)

Analogously, the 1st order tangent model is given by

\[
p^{(1)} = \frac{\partial \Delta p}{\partial \Delta d} \Delta d^{(1)},
\] (14)

which allows the computation of the gradient \( (12) \) by performing \( s \) runs with each run \( j \) seeded with a unit vector \( \vec{j} \in \mathbb{R}^{4} \):

\[
\Delta d^{(1)} = \vec{j}.
\] (15)

For the differentiation with AD, the cold-to-hot algorithm is not treated as a black box. Notably, the first step of the cold-to-hot transformation, the projection of FEM nodes onto the surface (section 3.1), does not need to be differentiated. This is because the projection is performed only to determine the foot points \((u_k, v_k)\) of a corresponding FEM node \(d_k\). Differentiating equations \( (6) \) and \( (7) \) is sufficient to compute the gradient \( (12) \). This is achieved using the open-source AD tool CoDiPack \( (1) \) developed by the Chair for Scientific Computing at TU Kaiserslautern, Germany. The results of the adjoint sensitivity computation compared against finite difference (FD) approximations are presented in section 5.2.

5 RESULTS

In this section, test results of the implemented cold-to-hot transformation and its sensitivity computation are presented. The selected test case is an axial fan blade geometry, which has been used as a baseline geometry for an optimization in \( [2] \).

5.1 Cold-to-hot transformation results

The results of the cold-to-hot transformation of the axial fan are quantified by computing the distance error between the deformed B-spline surface \( \hat{S}(u, v) \) and the displaced points \( d \). The error is computed in the same manner as in \( [5] \):

\[
\Delta_k = \min_{u, v \in [0, 1]} \| d_k - \hat{S}(u, v) \|^2, \quad k = 0, ..., s - 1
\] (16)

The maximum and mean distance errors are computed as

\[
\Delta_{\text{max}} = \max_k \Delta_k, \quad \Delta_{\text{mean}} = \frac{1}{s} \sum_{k=0}^{s-1} \Delta_k.
\] (17)

The deformations are visualized in figures 3 and 4, which show that the deformed blade geometry in figures 3(b) and 4(b) follows the displacements given by the vectors in figures 3(a) and 4(a). Distance error results are summarized in table 1.
Marc Schwalbach and Tom Verstraete

(a) displacements vectors of FEM results
(b) deformed blade: dark. original blade: light

Figure 3: axial fan cold-to-hot transformation

(a) displacements vectors of FEM results
(b) deformed blade: dark. original blade: light

Figure 4: axial fan cold-to-hot transformation

<table>
<thead>
<tr>
<th></th>
<th>Suction Side</th>
<th>Pressure Side</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \Delta_{\text{max}} ) [m]</td>
<td>( 2.83 \times 10^{-3} )</td>
<td>( 2.49 \times 10^{-3} )</td>
</tr>
<tr>
<td>( \Delta_{\text{mean}} ) [m]</td>
<td>( 7.36 \times 10^{-5} )</td>
<td>( 7.53 \times 10^{-5} )</td>
</tr>
</tbody>
</table>

Table 1: Maximum distance error \( \Delta_{\text{max}} \) and mean distance error \( \Delta_{\text{mean}} \) of the axial fan blade

5.2 Sensitivity computation results

The cold-to-hot transformation was differentiated using CoDiPack in both forward and reverse mode to compute the 1st order sensitivities [12]. The input variables are the FEM node
displacements $\Delta d \in \mathbb{R}^s$ and the output variables are the updated control points $p \in \mathbb{R}^{n \cdot m}$. The number of control points $n \cdot m$ is less than the number of FEM nodes $s$, with

$$n \cdot m = 1053, \quad s = 8149.$$  \hspace{1cm} (18)

A finite-difference (FD) approximation of (12) would require $2 \cdot s$ runs of (10) and a forward AD evaluation of the gradient requires $s$ runs of the tangent model (14). On the other hand, computing (12) using reverse AD would require only $n \cdot m$ runs of the adjoint model (11). Thus, a faster run time is expected and also observed for the adjoint AD computation compared to the FD approximations and tangent AD runs (see table 2). The relative run time results are compared to the measured cost of the primal computation of the control point displacements, excluding the FEM node projection, $\text{cost}(P) \approx 4.52$ seconds. Note that in the final application of the adjoint cold-to-hot chain, only a single adjoint computation of the cold-to-hot transformation is necessary for each output parameter of the entire chain.

<table>
<thead>
<tr>
<th></th>
<th>FD</th>
<th>forward AD</th>
<th>reverse AD</th>
</tr>
</thead>
<tbody>
<tr>
<td>absolute run time [min]</td>
<td>83.57</td>
<td>98.06</td>
<td>11.94</td>
</tr>
<tr>
<td>relative run time</td>
<td>1110.42 $\cdot$ cost($P$)</td>
<td>1303.04 $\cdot$ cost($P$)</td>
<td>158.61 $\cdot$ cost($P$)</td>
</tr>
</tbody>
</table>

Table 2: run time comparison of FD vs forward AD vs reverse AD for the 1st order sensitivities (12), with $\text{cost}(P) \approx 4.52$ seconds

Figures 5(a) and 5(b) present a comparison between gradients computed using FD, forward AD, and reverse AD for selected rows $r$ and columns $k$ of $\frac{\partial p_r}{\partial \Delta d_k}$. The plots show a good agreement between the FD and AD computed values with discrepancies between FD and AD at values of around $O(10^{-10})$ in figure 5(a). A maximum discrepancy of $O(10^{-6})$ and a mean discrepancy of $O(10^{-13})$ between FD and AD was computed for the entire matrix (12). A visualization of the 1st-order sensitivities of a single control point with respect to the FEM displacements $\Delta d$ is shown in figure 6.

![Comparison of FD vs forward AD vs reverse AD on a semi-logarithmic plot row of $\frac{\partial p_r}{\partial \Delta d_k}$](image)

Figure 5: Comparison of FD vs forward AD vs reverse AD on a semi-logarithmic plot row of $\frac{\partial p_r}{\partial \Delta d_k}$
6 CONCLUSIONS

With the implementation and adjoint differentiation of the cold-to-hot transformation, an important component of the cold-to-hot chain was addressed. The AD computed sensitivities agree well with the FD computed sensitivities, paving the way towards an adjoint multidisciplinary optimization for turbomachinery components. Further work could involve a local error reduction in the cold-to-hot transformation, as well as the inclusion of pressure forces in the cold-to-hot transformation to couple the fluid and solid computations. An adjoint differentiation of a CSM solver is the next key component of the cold-to-hot chain that is to be implemented.

7 ACKNOWLEDGMENTS

The authors would like to acknowledge the financial support of the IODA (Industrial Optimal Design using Adjoint CFD) project, which led to the work presented in this paper. The IODA project is funded by the European Commission within the Horizon 2020 program.
8 NOMENCLATURE

\[ J \in \mathbb{R}^c \] cost function
\[ J_{(1)} \in \mathbb{R}^c \] adjoint cost function
\[ \alpha, \alpha \in \mathbb{R}^d \] design parameters
\[ \alpha_{(1)} \in \mathbb{R}^d \] adjoint design parameters
\[ S, S_c, \hat{S} \in \mathbb{R} \] B-spline surface (initial, cold state, deformed)
\[ N^p, N^q \in \mathbb{R} \] B-spline basis functions
\[ p, q \in \mathbb{N} \] B-spline basis function orders
\[ u, v \in \mathbb{R} \] B-spline foot points
\[ p^0_{i,j}, p_{i,j} \in \mathbb{R}^3 \] control points (initial, displaced)
\[ p^{(1),i,j} \in \mathbb{R}^3 \] tangent control points
\[ p^{(1),i,j} \in \mathbb{R}^3 \] adjoint control points
\[ \Delta p_{i,j} \in \mathbb{R}^3 \] control point displacements
\[ n \cdot m \in \mathbb{N} \] number of control points
\[ d_k^0, d_k \in \mathbb{R}^3 \] FEM nodes (initial, displaced)
\[ \Delta d_k \in \mathbb{R}^3 \] FEM node displacements
\[ \Delta d^{(1),i,k} \in \mathbb{R}^3 \] tangent FEM node displacements
\[ \Delta d^{(1),i,k} \in \mathbb{R}^3 \] adjoint FEM node displacements
\[ s \in \mathbb{N} \] number of FEM nodes
\[ \mathcal{K}_{\text{edge}} \] set of FEM nodes along surface edge
\[ \mathcal{K}_{\text{inner}} \] set of inner FEM nodes
\[ \sigma_{\text{max}} \in \mathbb{R} \] maximum von Mises stress
\[ \Omega_s \] solid mesh
\[ F_c \in \mathbb{R}^3 \] centrifugal forces
\[ \mathbf{i} \in \mathbb{R}^{n \cdot m} \] \( i \)-th unit vector
\[ \mathbf{j} \in \mathbb{R}^s \] \( j \)-th unit vector
\[ \Delta_k, \Delta_{\text{max}}, \Delta_{\text{mean}} \in \mathbb{R} \] distance error (\( k \)-th, maximum, mean)

REFERENCES


