An adjoint method for gradient-based optimization of stellarator coil shapes

Elizabeth Paul\textsuperscript{1}
Dr. Matt Landreman\textsuperscript{1}
Dr. Aaron Bader\textsuperscript{2}
Dr. William Dorland\textsuperscript{1}

\textsuperscript{1}University of Maryland, College Park
\textsuperscript{2}University of Wisconsin, Madison

PPPL Theory Seminar
March 15\textsuperscript{th}, 2018
1. **Introduction**

2. **Current potential approach for coil optimization**
   a) NESCOIL
   b) REGCOIL

3. **Nonlinear optimization of the winding surface**
   a) Objective function
   b) Optimization constraints

4. **Adjoint method for gradient computation**
   a) Examples from CFD and electron gun optimization
   b) Linear adjoint method

5. **Applications**
   a) Optimization of W7-X and HSX winding surfaces

6. **Local sensitivity analysis**
1. Introduction
2. Current potential approach for coil optimization
   a) NESCOIL
   b) REGCOIL
3. Nonlinear optimization of the winding surface
   a) Objective function
   b) Optimization constraints
4. Adjoint method for gradient computation
   a) Examples from CFD and electron gun optimization
   b) Linear adjoint method
5. Applications
   a) Optimization of W7-X and HSX winding surfaces
6. Local sensitivity analysis
Designing & constructing 3-D coils is challenging

The construction of the National Compact Stellarator Experiment (NCSX) was never completed, party due to increasing costs and small tolerances in producing the magnetic field.

How can we more successfully design stellarator coils?

- Stellarators are fully 3-D, so design space is extremely large
- Large non-linear optimizations are now used in the design process
- First step in optimization: identify plasma shape and configuration with good physics properties (MHD stability, neoclassical & turbulent transport)
- Coils chosen to minimize error in producing magnetic surfaces desired
- Many desirable engineering considerations
  - Sufficient space between coils and plasma to allow for neutron shielding and blanket
  - Sufficient coil-coil spacing for maintenance, diagnostics, and neutral beams
  - Small curvature to allow for finite thickness of conducting material
  - Small bending ($\mathbf{J} \times \mathbf{B}$) forces on coils (less support required)
- Coil design is closely tied to cost and size of stellarator
Coil optimization tools

- Nonlinear optimization codes – not guaranteed to find global minimum, requires tuning of weighting parameters
  - COILOPT++
    - Coils represented by splines on winding surface
  - FOCUS
    - Coils represented by 3D space curves (no winding surface)
- Current potential methods – robust and fast, but do not include finite coil effects
  - NESCOIL
  - REGCOIL
- Coil optimization is difficult - **important to have multiple tools**

**COILOPT++ allows ‘straightening’ of modular coils** [14]

**W7-X modular coils optimized with FOCUS** [15]
Outline

1. Introduction
2. **Current potential approach for coil optimization**
   a) NESCOIL
   b) REGCOIL
3. Nonlinear optimization of the winding surface
   a) Objective function
   b) Optimization constraints
4. Adjoint method for gradient computation
   a) Examples from CFD and electron gun optimization
   b) Linear adjoint method
5. Applications
   a) Optimization of W7-X and HSX winding surfaces
6. Local sensitivity analysis
Finding coil shapes on a winding surface

Given a fixed winding surface and plasma surface, how should I design my coils?

- One of the earliest methods (NESCOIL [4]) used for stellarator coil design assumed all coils to lie on a winding surface (toroidal surface enclosing plasma)
- Plasma surface is fixed based on physics optimization
- Divergentless current density on a surface can be computed from current potential, $\Phi$
- Contours of $\Phi$ give coil shapes (streamlines of $K$)

\[ K = n \times \nabla \Phi \]

Unit normal on coil surface

Current potential is composed of secular terms, determined by net poloidal and toroidal surface currents, and single-valued (periodic) piece which we optimize

\[ \Phi = \Phi_{sv} + \frac{G\zeta}{2\pi} + \frac{I\theta}{2\pi} \]

Determined by net currents linking plasma

W7-X plasma surface (red) and coil-winding surface (green) [5]
Example – Current potential for W7-X modular coil set

Total current potential

Corresponding coil set (1/2 period)
REGCOIL [5] approach

- REGCOIL [5] is a current potential method which applies **Tikhonov regularization**
- Solves ill-posed problem: very different current distributions can give nearly identical magnetic surfaces
- Improves condition number of linear least-squares problem
- Simultaneously **improves engineering properties** of coils

**Tikhonov regularization**

$$\min_x \|Ax - b\| + \|\Gamma x\|$$

Analogous to polynomial fitting – overfitting amplifies noise

**REGCOIL regularization**

$$\min_{\Phi} \phi_B^2 + \lambda \phi_K^2$$

Fidelity in reproducing plasma surface

$$\phi_B^2 = \int_{\text{plasma}} d^2 A B_n^2$$

If plasma surface were exactly reproduced, $$\phi_B^2 = 0$$

Increased coil-coil spacing

$$\phi_K^2 = \int_{\text{coil}} d^2 A K^2$$
REGCOIL [5] approach

\[ \Phi_{sv} = \sum_{m,n} \Phi_{mn} \sin(m \theta - n \zeta) \]

- \( \Phi_{sv} \) expanded in sine series (assuming stellarator symmetry)
- For NESCOIL (\( \lambda = 0 \)), only regularization provided by truncation of Fourier series

Our task: \( \min_{\Phi} \chi^2 = \chi_B^2 + \lambda \chi_K^2 \)

Linear least-squares problem

\[ A\Phi = b \]

- Tradeoff between low field error and low current density
- Regularization (\( \lambda \)) chosen to meet some prescribed tolerance
  - e.g., \( K_{max} \) (coil-coil spacing)
  - Obtained from numerical root-finding algorithm
The current potential method has been widely applied

- **Robust**: guaranteed to find a minimum
- **Fast**: requires inverting a matrix of size $\lesssim 144 \times 144$
- For these reasons, often used in initial stages of coil optimization process

Contours of the current potential and coil set obtained with NESCOIL in initial NCSX design [6]

W7-X coils obtained from NESCOIL with modified winding surface [12]
1. Introduction
2. Current potential approach for coil optimization
   a) NESCOIL
   b) REGCOIL
3. Nonlinear optimization of the winding surface
   a) Objective function
   b) Optimization constraints
4. Adjoint method for gradient computation
   a) Examples from CFD and electron gun optimization
   b) Linear adjoint method
5. Applications
   a) Optimization of W7-X and HSX winding surfaces
6. Local sensitivity analysis
Optimization of the winding surface

- Compute the current distribution on fixed winding surface using REGCOIL
- We’d like to find the optimum **coil shapes in 3D space** rather than on a fixed surface – this amounts to optimizing the winding surface
- Plasma surface is fixed
- Properties we’d like in a winding surface:
  - Fidelity in **producing desired plasma surface**
  - **Maximize space** between coil surface and plasma surface
  - Allows for **simple coils**

Optimization parameters: \[ \Omega = \{ r^c_{m,n}, z^s_{m,n} \} \]

- \[ x = \sum_{m,n} r^c_{m,n} \cos (m\theta - n\zeta) \cos \zeta \]
- \[ y = \sum_{m,n} r^c_{m,n} \cos (m\theta - n\zeta) \sin \zeta \]
- \[ z = \sum_{m,n} z^s_{m,n} \sin (m\theta - n\zeta) \]
Current density is a proxy for coil complexity

- Increased maximum current density
- Increased coil length
- Increased toroidal extent
- Increased curvature
- Decreased coil-coil spacing

Coils computed on same winding surface (only \( \chi \) varying)

\[ \| \mathbf{K} \|_2 = 2.2 \text{ MA/m} \]
\[ \| \mathbf{K} \|_2 = 2.7 \text{ MA/m} \]
\[ \| \mathbf{K} \|_2 = 3.2 \text{ MA/m} \]
Objectifying our optimization

Our task: \( \min_{\Omega} f \)

\[
f(\Omega, \Phi(\Omega)) = \chi_B^2(\Omega, \Phi(\Omega)) - \alpha_V V_{\text{coil}}^{1/3}(\Omega) + \alpha_S S_p(\Omega) + \alpha_K \|K\|_2(\Omega, \Phi(\Omega))
\]

- **Fidelity in reproducing plasma surface**
- **Increase coil-plasma distance**
- **Represent surface with a condensed Fourier series**

We want to use a **gradient-based optimization** method to minimize \( f \)

- \( V_{\text{coil}} \) and \( S_p \) are only functions of geometry, so can be explicitly differentiated with respect to \( \Omega \)
- Gradient of \( \chi_B^2 \) could be computed by finite differencing
  - This could be prohibitively expensive – requires \( N_\Omega + 1 \) \((\approx 100)\) calls to REGCOIL
- Alternatively, use the **adjoint method** - only requires 2 calls to REGCOIL

Elizabeth Paul  
PPPL Theory Seminar  
March 15th, 2018
Outline

1. Introduction
2. Current potential approach for coil optimization
   a) NESCOIL
   b) REGCOIL
3. Nonlinear optimization of the winding surface
   a) Objective function
   b) Optimization constraints
4. **Adjoint method for gradient computation**
   a) Examples from CFD and electron gun optimization
   b) Linear adjoint method
5. Applications
   a) Optimization of W7-X and HSX winding surfaces
6. Local sensitivity analysis
What is an adjoint method?

- Allow for efficient computation of gradients of output quantities of a linear or non-linear solve with respect to input parameters
- Solve an adjoint (“backward in time”) equation in addition to original (“forward”) equation
- Developed in 1970s for analysis of drag and flow dynamics
- Widely used in computational fluid dynamics and aerodynamic engineering
- Applications
  - Gradient-based optimization
  - Uncertainty quantification in scientific computing
  - Surface sensitivity maps

**CFD Shape optimization**

How does the total drag depend on the shape of the mirror on the VW Passat? [7]

**Electron gun sensitivity**

What is the sensitivity of the displacement of the beam to the properties of the anode? [8]

How does the total drag depend on the shape of the mirror on the VW Passat? [7]
**Question:** How do small changes in position or potential of anode affect the properties of the beam leaving the gun?

**Adjoint problem:** perturb electron positions and run backwards in time, obtain $\nabla \hat{\Phi}$

\[ \delta x \propto \int_{\text{anode}} d^2 A \left( \mathbf{n} \cdot \nabla \hat{\Phi} \right) \delta \Phi \]

- Displacement of beam
- 'Sensitivity function' (from backward problem)
- Local change in electric potential

**Vector plot of the sensitivity function of the beam displacement to local changes in electric potential**
Adjoint method for a linear system

We’d like to compute the gradient of some function of input and output parameters

\[ \frac{\partial f(\Omega_i, x(\Omega_i))}{\partial \Omega_i} \bigg|_{A x = b} = \frac{\partial f}{\partial \Omega_i} \bigg|_x + \left( \frac{\partial f}{\partial x} \right) \cdot \frac{\partial x}{\partial \Omega_i} \bigg|_{A x = b} \]

State variable, \( x \), satisfies linear “forward” equation and depends on many input parameters, \( \Omega = \{\Omega_i\} \)

\[ A(\Omega_i)x = b(\Omega_i) \]

Differentiate with respect to \( \Omega_i \)

\[ \frac{\partial A}{\partial \Omega_i} x + A \frac{\partial x}{\partial \Omega_i} = \frac{\partial b}{\partial \Omega_i} \rightarrow \frac{\partial x}{\partial \Omega_i} \bigg|_{A x = b} = A^{-1} \left( \frac{\partial b}{\partial \Omega_i} - \frac{\partial A}{\partial \Omega_i} x \right) \]

Insert the expression for \( \frac{\partial x}{\partial \Omega_i} \bigg|_{A x = b} \)

This requires inverting \( A \) for each \( \Omega_i \) we’d like to differentiate with respect to!
We’d like to compute the derivative for many $\Omega_i$:

$$\frac{\partial f(\Omega_i, x(\Omega_i))}{\partial \Omega_i} \bigg|_{Ax=b} = \frac{\partial f}{\partial \Omega_i} \bigg|_x + \left( \frac{\partial f}{\partial x} \right) \cdot \left[ A^{-1} c(x, \Omega_i) \right]$$

We’ll exploit the adjoint property for this inner product:

$$\frac{\partial f(\Omega_i, x(\Omega_i))}{\partial \Omega_i} \bigg|_{Ax=b} = \frac{\partial f}{\partial \Omega_i} \bigg|_x + \left( A^T \right)^{-1} \left( \frac{\partial f}{\partial x} \right) \cdot c(x, \Omega_i)$$

We can instead invert the adjoint operator once:

$$A^T q = \frac{\partial f}{\partial x}$$

We take an inner product for each $\Omega_i$ we’d like to differentiate with respect to:

$$\frac{\partial f(\Omega_i, x(\Omega_i))}{\partial \Omega_i} \bigg|_{Ax=b} = \frac{\partial f}{\partial \Omega_i} \bigg|_x + q \cdot c(x, \Omega_i)$$

- Assuming that inverting $A$ is the computational bottleneck, the adjoint method scales independently of the number of input parameters you are differentiating with respect to (requires 2 linear solves).
- Can be generalized to non-linear systems and general inner product spaces.
We use the linear adjoint method to obtain

\[
\left\{ \frac{\partial \chi_B^2}{\partial \Omega}, \frac{\partial \|K\|_2}{\partial \Omega} \right\}
\]

**Forward equation**

\[ A\Phi = b \]

**Adjoint equation**

\[ A^T q = \frac{\partial \chi_B^2}{\partial \Phi} \]

---

**Benchmark with finite difference derivatives**

**Computational time scaling**

Elizabeth Paul

PPPL Theory Seminar

March 15th, 2018
Optimization constraints

Inequality constraint on minimum coil-plasma distance

\[ d_{\text{min}} = \min_{\theta, \zeta} (d_{\text{coil-plasma}}) \]

\[ d_{\text{min}} \geq d_{\text{min}}^{\text{target}} \]

Regularization, \( \lambda \), chosen to fix maximum current density

\[ \chi^2 = \chi_B^2 + \lambda \chi_K^2 \]

Corresponds to fixed coil-coil spacing

Elizabeth Paul

PPPL Theory Seminar

March 15th, 2018
Gradient-based optimization of objective function

\[ f(\Omega, \Phi(\Omega)) = \chi_B^2 + \alpha_K \| K \|_2 - \alpha_V V_{\text{coil}}^{1/3} + \alpha_S S_p \]

- Gradients can be computed with the linear adjoint method
- Analytically differentiate

Remaining challenge:
How do we choose \( \alpha_V, \alpha_S, \) and \( \alpha_K \)?

Can we design ‘better’ winding surfaces?

- We use the sequential linear-squares quadratic programming (SLSQP) for constrained optimization
  - Implementation in NLOPT package
  - For demonstration, begin with actual HSX and W7-X winding surfaces

Elizabeth Paul
PPPL Theory Seminar
March 15th, 2018
1. Introduction
2. Current potential approach for coil optimization
   a) NESCOIL
   b) REGCOIL
3. Nonlinear optimization of the winding surface
   a) Objective function
   b) Optimization constraints
4. Adjoint method for gradient computation
   a) Examples from CFD and electron gun optimization
   b) Linear adjoint method
5. Applications
   a) Optimization of W7-X and HSX winding surfaces
6. Local sensitivity analysis
Trends with optimization parameters

\[ f = \chi_B^2 + \alpha_S S_p \]

- As \( \alpha_S \) increases, winding surfaces approaches cylindrical cross section (minimal Fourier spectrum)
- At moderate \( \alpha_S \), approaches uniform offset from plasma surface (only minimizing \( \chi_B^2 \))
- Moderate \( \alpha_S \) needed to eliminate zero-gradient direction (non-unique parameterization)

\[ S_p = \sum_{m,n} m^p \left( (r_{mn}^c)^2 + (z_{mn}^s)^2 \right) \]
Trends with optimization parameters

\[ f = \chi_B^2 - \alpha_V V_{coil}^{1/3} + \alpha_S S_p \]

- Location of closest approach stays fixed.
- As \( \alpha_V \) increases, \( V_{coil} \) increases.
- At small \( \alpha_V \), approaches uniform offset from plasma surface.
- Minimum coil-plasma distance remains fixed – needed to produce plasma surface in concave region.
- In the convex, low field region, coils can be further away.
- Need to find a balance between the physics (small \( \chi_B^2 \)) and engineering objectives.
Optimizing the W7-X winding surface

\[ \zeta = 0.0 \quad \zeta = 0.25 \quad \zeta = 0.5 \quad \zeta = 0.75 \]

Actual W7-X winding surface (initial surface for optimization)

Optimum obtained

- \( \chi_B^2 \) decreased 90%
- \( V_{\text{coil}} \) increased 22%

\[ \alpha_V = 0.5, \quad \alpha_S = 0.24, \quad \alpha_K = 1.6 \times 10^{-6} \]
Optimized W7-X Coils

<table>
<thead>
<tr>
<th></th>
<th>Initial</th>
<th>Optimized</th>
</tr>
</thead>
<tbody>
<tr>
<td>Min. coil-coil distance [m]</td>
<td>0.223</td>
<td>0.271</td>
</tr>
<tr>
<td>Max curvature [m⁻¹]</td>
<td>9.01</td>
<td>4.84</td>
</tr>
<tr>
<td>Max toroidal extent [rad.]</td>
<td>0.222</td>
<td>0.197</td>
</tr>
</tbody>
</table>
Optimizing the HSX winding surface

Actual HSX winding surface
(initial surface for optimization)

Optimum obtained

\( \zeta = 0.0 \frac{2\pi}{N_p} \) \hspace{1cm} \( \zeta = 0.25 \frac{2\pi}{N_p} \)

\( \zeta = 0.5 \frac{2\pi}{N_p} \) \hspace{1cm} \( \zeta = 0.75 \frac{2\pi}{N_p} \)

\( \alpha_V = 3.13 \times 10^{-4} \), \( \alpha_S = 0 \), \( \alpha_K = 3 \times 10^{-10} \)

Elizabeth Paul
PPPL Theory Seminar
March 15th, 2018

- \( \chi_B^2 \) decreased 33%
- \( V_{coil} \) increased 30%
Optimized HSX Coils

<table>
<thead>
<tr>
<th></th>
<th>Initial</th>
<th>Optimized</th>
</tr>
</thead>
<tbody>
<tr>
<td>Min. coil-coil distance [m]</td>
<td>0.0850</td>
<td>0.0853</td>
</tr>
<tr>
<td>Max curvature [m⁻¹]</td>
<td>33.4</td>
<td>25.8</td>
</tr>
<tr>
<td>Max toroidal extent [rad.]</td>
<td>0.530</td>
<td>0.505</td>
</tr>
</tbody>
</table>
1. Introduction
2. Current potential approach for coil optimization
   a) NESCOIL
   b) REGCOIL
3. Nonlinear optimization of the winding surface
   a) Objective function
   b) Optimization constraints
4. Adjoint method for gradient computation
   a) Examples from CFD and electron gun optimization
   b) Linear adjoint method
5. Applications
   a) Optimization of W7-X and HSX winding surfaces
6. Local sensitivity analysis
Computing local sensitivity

“Fourier derivatives” computed using adjoint method

\[ \partial \chi_B^2 / \partial r_{mn} \]

Sensitivity in real space?
Computing local sensitivity with shape gradients

Consider a functional of the shape of some domain, $f(\Gamma)$

\[ \Gamma_\varepsilon = \{ \mathbf{r}_0 + \varepsilon \delta \mathbf{r}(\mathbf{r}_0) : \mathbf{r}_0 \in \Gamma \} \]

\[ \delta f(\Gamma, \delta \mathbf{r}) = \lim_{\varepsilon \to 0} \frac{f(\Gamma_\varepsilon) - f(\Gamma)}{\varepsilon} \]

For many functionals, can be written in the ‘Hadamard form’

\[ \delta f(\Gamma, \delta \mathbf{r}) = \int_{\partial \Gamma} d^2 A \, S \delta \mathbf{r} \cdot \mathbf{n} \]

Shape gradient of drag on a Volkswagen from Navier-Stokes calculations [7]

red: inwards for smaller drag
blue: outwards
Computing local sensitivity with shape gradients [13]

\[
\delta \chi_B^2(\Gamma_{\text{coil}}, \delta \mathbf{r}) = \int_{\partial \Gamma_{\text{coil}}} d^2 A S_{\chi_B^2} \delta \mathbf{r} \cdot \mathbf{n}
\]

\[
S_{\chi_B^2} = \sum_j S_j \cos(m_j \theta - n_j N_p \zeta)
\]

\[
\frac{\partial \chi_B^2}{\partial \Omega_i} = \int_{\partial \Gamma_{\text{coil}}} d^2 A S_{\chi_B^2} \left( \frac{\partial \mathbf{r}}{\partial \Omega_i} \right) \cdot \mathbf{n}
\]

Linear system (generally not square) can be solved with Moore-Penrose pseudoinverse to obtain \( S_{\chi_B^2} \)

\[
\frac{\partial \chi_B^2}{\partial \Omega_i} = \sum_j D_{ij} S_j
\]
W7-X winding surface shape gradient

Shape gradient of $\mathcal{X}_B^2$

Second principal curvature of plasma surface (negative = concave)

$$\delta \mathcal{X}_B^2 = \int_{\text{coil}} d^2A \mathcal{S}_\mathcal{X}_B^2 \delta \mathbf{r} \cdot \mathbf{n}$$
W7-X winding surface shape gradient

\[
\delta \|K\|_2 = \int_{\text{coil}} d^2A \, S_{\|K\|_2} \, \delta r \cdot n
\]
Concluding thoughts

- We have demonstrated the first application of adjoint methods to stellarator coil optimization
  - Efficient computation of gradients (reduces required function evaluations by a factor of $\approx 50$)
- We have obtained winding surfaces for W7-X and HSX which simultaneously reproduce the desired plasma surfaces with better fidelity, improve engineering properties of coils, and increase the coil-plasma distance, allowing for more experimental flexibility
- We have developed tools to compute sensitivity to local perturbations of the winding surface

Thank you for your attention!

Elizabeth Paul
PPPL Theory Seminar
March 15th, 2018
References